ECHO: The Epidemiology of HPV Infection in Oropharyngeal, Oral Cavity, and Laryngeal Cancer in Ireland

A thesis submitted to Trinity College, The University of Dublin

For the degree of Doctor of Philosophy, Histopathology and Morbid Anatomy in 2019

by

Imogen Sharkey Ochoa BA (Hons) MPhil (Hons)
Molecular Pathology Research Group and CERVIVA
The Coombe Women and Infants' University Hospital
Trinity College Dublin

Under the supervision of Dr. Cara Martin and Professor John O'Leary

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

I agree to deposit this thesis in the University's open access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

I do not consent to the examiner retaining a copy of the thesis beyond the examining period, should they so wish (EU GDPR May 2018).

Imogen Sharkey Ochoa

To my mum, Katrina, and my late grandmother, Bridie

This third generation would not have made it this far without you

Acknowledgements

The English language is not expansive enough to extend the appropriate thanks due to my Supervisor, Dr. Cara Martin. Thank you for giving me the opportunity to take this project on as your Supervisee in 2016 and dedicating so much of your time and energy to the ECHO study no matter the burden of all other responsibilities. Most importantly, thank you for believing in me and supporting me in developing beyond the bounds of academia, constantly providing an inspiration to aspire to both professionally and personally. I, and the ECHO study, would not be where we are today without your unwavering leadership, support, encouragement, and kindness.

To my Supervisor and Principal Investigator, Professor John O'Leary, thank you for making the decision to take me on as a part of your team three years ago. Your constant encouragement, optimism, and enthusiasm about the ECHO study and belief in my ability to "take on the world" have made me eternally grateful to have you as a leader and have forever changed me as a person.

To my colleagues, and now certainly close friends, in the Molecular Pathology Laboratory of the Coombe Women and Infants' Hospital and within the CERVIVA group, I would never have made it through with your generous donation of time and patience to answer all of my questions, train me in new techniques, and simply make day-to-day lab life a real pleasure, even in the most frustrating of times. To the original "bold" member, Prerna Tewari, the sky-diving queen, Christine White, the cell-culture oracle, Cathy Spillane, the INNOLiPA expert, Helen Keegan, "not-today Satan" Mark Ward, "looks 20 years younger than she is" Tanya Kelly, "*not* funny" Laura Edgerton, and all the members of our lab team, my gratitude is never-ending.

My deepest thanks are also extended to all of the participating hospitals, and the pathologists and lab scientists who made such immense efforts to allow the acquisition of samples for the ECHO study to proceed and reviewed thousands of slides without any qualms or questions. This project was nothing without its treasured "blocks" and their assessment, and your support was irreplaceable. Importantly, to Dr. Esther O'Regan, thank you for being the biggest supporter of this project from all sides, and for being the most kind-hearted and inspirational boss-lady I could have ever hoped to look up to.

I extend my eternal gratefulness to the Histology Departments in both the Coombe and St. James'. Jacqui Barry-O'Crowley, Niamh Kernan, and the rest of the infinitely giving team, thank you for dedicating your machinery and expert skills so kindly to helping this project progress.

Lastly, to my mum, Katrina, thank you for your unfaltering support of my every endeavor, and mostly for being you, the most accomplished person I know who never fails to still be the kindest and most loving.

Table of Contents

List of Figures

List of Tables

Summary

Human Papillomavirus (HPV) infection has been identified as a significant etiologic agent in the development of head and neck squamous cell carcinoma (HNSCC), particularly those of the oral region of the head and neck. There is mounting evidence from North America and Europe that HPV-related HNSCC is becoming the principal driver of increasing incidence trends amongst all HNSCC, as HPV-unrelated cases have seen a plateau or a decline in the last 20 to 30 years. This is especially true of HPV-related trends of oropharyngeal squamous cell carcinoma (OPSCC). HPV-related HNSCC has also consistently been shown to present at younger age and later stage, and have better survival and prognosis than its HPV-unrelated counterparts. No data on the epidemiology of HPV infection in oropharyngeal, oral cavity, and laryngeal SCC in Ireland currently exists in the literature.

The ECHO study thus investigated the epidemiology of HPV infection in oropharyngeal, oral cavity, and laryngeal cancer in Ireland, acquiring and assessing primary oropharyngeal, oral cavity, and laryngeal squamous cell carcinoma (SCC) diagnosed between 1994 and 2013 in Ireland and recorded at the National Cancer Registry of Ireland (NCRI). Cases were tested for HPV DNA using Multiplex PCR Luminex technology based in and sanctioned by the International Agency for Research on Cancer (IARC).

Samples were obtained from 7 hospitals spanning both the West and East parts of the country. The study found a significant disparity between the length of time needed to organize the review of pathology reports and retrieval of included samples from hospital and private storage [273.67 days (CI: 121.04, 426.30)] and the length of time needed to execute report review and retrieve blocks [7.17 days (CI: 1.35, 12.99)]. A preliminary pilot investigation also revealed the significant impact of technology type and sterility protocols used for determining HPV prevalence.

In total, 861 cases were included in the ECHO study, showing an overall HPV DNA prevalence of 17.1% (CI: 14.6, 19.6). The oropharyngeal sub-site saw the largest prevalence of 41.1% (CI: 34.5, 47.8), followed by the oral cavity with a prevalence of 10.9% (CI: 7.5, 14.2)
, followed by the larynx with a prevalence of 7.8% (CI: 4.9, 10.7). The tonsillar sub-site within the oropharynx saw the highest prevalence of all (60.0%). HPV16 was the overwhelmingly dominant genotype amongst all cases regardless of sub-site, representing over 80% if not 90% of infected cases. Two LR HPV6 cases were detected in the larynx. HPV-related and HPV-unrelated oropharyngeal, oral cavity, and laryngeal SCC saw significantly increased average annual percentage

change in the time-period, but the highest average annual percentage change of any sub-site or HPV status was seen in HPV-related oropharyngeal cases at 16.4% ($p<0.0001$).

Significant predictors of HPV positivity amongst oropharyngeal, oral cavity, and laryngeal SCC were younger age, oropharyngeal sub-site, and never- and ex-smoking status. HPV positive cases presented disproportionately at later TNM stage and with higher extent of nodal involvement. The only sub-site within which there were significant predictors was the oropharynx, and these mirrored overall findings with younger age and ex- and never-smoking status predicting HPV positivity. The only sub-site for which a relationship existed between HPV positivity and stage classifications was the oropharynx, with HPV positive tumours seeing greater extent of nodal involvement (N stage).

Both overall and cancer-specific survival were significantly improved amongst HPV positive SCC patients, though this relationship emanated strictly from oropharyngeal cases. Improved survival for HPV positive cases in the larynx and oral cavity was not observed. HPV positive tumours were more likely to be treated with surgery/radiotherapy/chemotherapy whereas HPV negative tumours were more likely to be treated with radiotherapy or surgery alone, likely due to their greater extent of nodal involvement and thus later TNM stage. For HPV positive OPSCC, surgery alone saw the best overall and cancer-specific survival, followed closely by surgery/radiotherapy and surgery/radiotherapy/chemotherapy. Significant predictors of survival in all cases together and OPSCC alone included HPV status, TNM stage, and age, amongst other characteristics. HPV status did not predict survival in the larynx or in the oral cavity.

The study highlighted the oropharyngeal sub-site as the only region of the head and neck to which carcinogenic HPV is relevant, though indicated the possible emergence of “hybrid” cases, involving both smoking history and the virus. It also emphasized the urgent need for imminently achievable and required prevention for HNSCC in the form of vaccination-based prophylaxis and head and neck cancer screening tools possibly involving HPV as a triage mechanism. It stressed the relevance of the intended introduction of the nona-valent Gardasil vaccine and the inclusion of boys into the Irish vaccination program in September 2019. Importantly, it alluded to the potential for de-escalation of treatment in HPV-related OPSCC in particular, to both maximize survival and minimize long-lasting side effects for patients. The study was funded by the Health Research Board and the Coombe Women and Infants’ University Hospital.

Awards, Publications, Press, Presentations, and Posters

Awards

EuroLife Summer School Scholarship (€700)

EuroLife Summer School 2018: "Molecular Mechanisms in Cancer – Translating Discoveries into Personalised Therapies"

Leiden University Medical Center, Leiden, the Netherlands

July 8-13 2018

Trinity Trust Travel Grant (€200)

Trinity Trust Travel Grant Scheme awarded for research conducted in the International Agency from Research on Cancer (IARC)

IARC, Lyon, France

July 2018

Publications

My Experience of the HPV Vaccine and the Epidemiology of HPV Infection in Oral Cancer in Ireland

Sharkey Ochoa I

thejournal.ie

thejournal.ie HQ, Dublin, Ireland

August 2017

Press

Radio Interview with NEWSTALK: Experience of the HPV Vaccine and the Epidemiology of HPV Infection in Oral Cancer in Ireland

Sharkey Ochoa I

NEWSTALK

NEWSTALK Studios, Dublin, Ireland

August 2017

Radio Interview with TODAY FM: Experience of the HPV Vaccine and the Epidemiology of HPV Infection in Oral Cancer in Ireland

Sharkey Ochoa I

HPV Alliance Press Launch and TODAY FM Studios

Irish Cancer Society, Dublin, Ireland

August 2017

Presentations

CERVIVA HPV Primary Screening Pilot Study: evaluation of triage strategies for HPV-positive women

White C, Reynolds S, Murphy K, Naik P, O' Brien R, Keegan H, Pilkington L, Sharkey Ochoa I, Powles C, Wright F, Bolger N, Barry-O'Crowley J, Tewari P, O'Toole S, Normand C, Sharp L, Flannelly G, Martin CM, O'Leary JJ

EUROGIN International Multidisciplinary HPV Congress

Congress Center, Lisbon, Portugal

December 2018

Comparison of partial HPV genotyping using the Cobas 4800 HPV test and the Aptima HPV 16 18/45 Genotype assay

White C, Reynolds S, Murphy K, Naik P, O' Brien R, Keegan H, Pilkington L, Sharkey Ochoa I, Powles C, Wright F, Bolger N, Barry-O'Crowley J, Tewari P, O'Toole S, Normand C, Sharp L, Flannelly G, Martin CM, O'Leary JJ

EUROGIN International Multidisciplinary HPV Congress

Congress Center, Lisbon, Portugal

December 2018

HPV Primary Screening Pilot Study: molecular testing of potential triage strategies for HPV-positive women

White C, Reynolds S, Naik P, O'Brien R, Pham T, Pilkington L, Sharkey Ochoa I, Powles C, Wright F, Bolger N, Barry-O'Crowley J, Tewari P, O'Toole S, Normand C, Sharp L, Flannelly G, Martin CM, O'Leary JJ

British Society for Colposcopy and Cervical Pathology

Exchange Hall, Manchester, UK

April 2018

ECHO: The Epidemiology of HPV Infection in Oral Cancer in Ireland

Sharkey Ochoa I, O'Murchu E, Woods R, Tewari P, O'Regan E, Toner M, Feeley K, Timon C, Sharp L, Martin CM, and O'Leary JJ

Trinity Translational Medicine Institute Conference

Trinity College Dublin, Dublin, Ireland

March 2018

ECHO: The Epidemiology of HPV Infection in Oral Cancer in Ireland

Sharkey Ochoa I, O'Murchu E, Woods R, Tewari P, O'Regan E, Toner M, Feeley K, Timon C, Sharp L, Martin CM, and O'Leary JJ

Thesis in 3 Minutes

Trinity College Dublin, Dublin, Ireland

March 2018

ECHO: The Epidemiology of HPV Infection in Oral Cancer in Ireland

Sharkey Ochoa I, O'Murchu E, Woods R, Tewari P, O'Regan E, Toner M, Feeley K, Timon C, Sharp L, Martin CM, and O'Leary JJ

Thesis in 3 Minutes (in French)

Embassy of France and Alliance Française, Dublin, Ireland

March 2018

Experience of the HPV Vaccine and ECHO: The Epidemiology of HPV Infection in Oral Cancer in Ireland

Sharkey Ochoa I, O'Murchu E, Woods R, Tewari P, O'Regan E, Toner M, Feeley K, Timon C, Sharp L, Martin CM, and O'Leary JJ

CERVIVA Cervical Cancer Prevention Workshop

Trinity College Dublin, Dublin, Ireland

January 2018

HPV Primary Screening Pilot Study: molecular testing of potential triage strategies for HPV-positive women

White C, Reynolds S, Naik P, O' Brien R, Pham T, Pilkington L, Sharkey Ochoa I, Powles C, Wright F, Bolger N, Barry-O'Crowley J, Tewari P, O'Toole S, Normand C, Sharp L, Flannelly G, O'Leary JJ, Martin CM

EUROGIN International Multidisciplinary HPV Congress

RAI Exhibition and Convention Center, Amsterdam, the Netherlands

October 2017

Experience of the HPV Vaccine and ECHO: The Epidemiology of HPV Infection in Oral Cancer in Ireland

Sharkey Ochoa I, O'Murchu E, Woods R, Tewari P, O'Regan E, Toner M, Feeley K, Timon C, Sharp L, Martin CM, and O'Leary JJ

HPV Alliance Press Launch

Irish Cancer Society, Dublin, Ireland

August 2017

ECHO: The Epidemiology of HPV Infection in Oral Cancer in Ireland

Sharkey Ochoa I, O'Murchu E, Woods R, Tewari P, O'Regan E, Toner M, Feeley K, Timon C, Sharp L, Martin CM, and O'Leary JJ

Dublin Dental Hospital Conference Day

Dublin Dental University Hospital, Dublin, Ireland

June 2017

HPV Primary Screening Pilot Study: molecular testing of potential triage strategies for HPV-positive women

White C, Reynolds S, Naik P, O' Brien R, Pham T, Pilkington L, Sharkey Ochoa I, Powles C, Wright F, Bolger N, Barry-O'Crowley J, Tewari P, O'Toole S, Normand C, Sharp L, Flannelly G, O'Leary JJ, Martin CM

British Society for Colposcopy and Cervical Pathology

Cardiff City Stadium, Cardiff, UK

May 2017

HPV Primary Screening Pilot Study: A comparison of HPV DNA and HPV mRNA Assays in a Primary Screening Population

Reynolds S, White C, Naik P, O'Brien R, Pham T, Sharkey Ochoa I, Powles C, Bolger N, Barry O'Crowley J, Tewari P, O'Toole S, Normand C, Sharp L, Flannelly G, O'Leary JJ, Martin CM

United States and Canadian Association of Pathology Annual Meeting

Henry B González Convention Center, San Antonio, Texas, USA

March 2017

ECHO: The Epidemiology of HPV Infection in Oral Cancer in Ireland

Sharkey Ochoa I, O'Murchu E, Woods R, Tewari P, O'Regan E, Toner M, Feeley K, Timon C, Sharp L, Martin CM, and O'Leary JJ

CERVIVA National HPV Awareness Symposium

Trinity Biomedical Sciences Institute, Dublin, Ireland

January 2017

HPV Primary Screening Pilot Study

White C, Reynolds S, Naik P, O' Brien R, Pham T, Sharkey Ochoa I, Bolger N, Barry-O'Crowley J, Tewari P, O'Toole S, Normand C, Sharp L, Flannelly G, O'Leary JJ, Martin CM

Trinity College Dublin 10th International Cancer Conference

Trinity Biomedical Sciences Institute, Dublin, Ireland

October 2016

Posters

Evaluation of Prognostic Impact of p53 Mutant Expression in HPV-Positive and HPV-Negative Head & Neck Squamous cell Carcinomas

Tewari P, Raguraman P, Woods R, Kernan N, Sharkey Ochoa I, Barry-O'Crowley J, O'Regan E, Martin CM, O'Leary JJ

The British Society for Oral and Maxillofacial Pathology Meeting

Ballsbridge Hotel, Dublin, Ireland

April 2019

ECHO: The Epidemiology of HPV Infection in Oral Cancer in Ireland

Sharkey Ochoa I, Woods R, O'Regan E, Barry-O'Crowley J, Kernan N, Kay E, Buckley C, Faul P, Toner M, Mullen D, Tewari P, Keegan H, Sharp L, O'Keane CJ, Timon C, Gheit T, Tommasino M, Martin CM, O'Leary JJ

The British Society for Oral and Maxillofacial Pathology Meeting

Ballsbridge Hotel, Dublin, Ireland

April 2019

Evaluation of Prognostic Impact of p53 Mutant Expression in HPV-Positive and HPV-Negative Head & Neck Squamous cell Carcinomas

Tewari P, Raguraman P, Woods R, Kernan N, Sharkey Ochoa I, Barry-O'Crowley J, O'Regan E, Martin CM, O'Leary JJ

United States and Canadian Academy of Pathology Annual Meeting

National Harbour, Maryland, USA

March 2019

ECHO: The Epidemiology of HPV Infection in Oral Cancer in Ireland

Sharkey Ochoa I, O'Murchu E, Woods R, Tewari P, O'Regan E, Toner M, Feeley K, Timon C, Sharp L, Martin CM, and O'Leary JJ

TCD International Postgraduate Research Conference

Trinity Biomedical Sciences Institute, Dublin, Ireland

March 2019

ECHO: The Epidemiology of HPV Infection in Oral Cancer in Ireland

Sharkey Ochoa I, O'Murchu E, Woods R, Tewari P, O'Regan E, Toner M, Feeley K, Timon C, Sharp L, Martin CM, and O'Leary JJ

Trinity College Dublin Multidisciplinary Research Showcase

Trinity Biomedical Sciences Institute, Dublin, Ireland

March 2019

ECHO: The Epidemiology of HPV Infection in Oral Cancer in Ireland

Sharkey Ochoa I, Woods R, O'Regan E, Barry-O'Crowley J, Kernan N, Kay E, Buckley C, Faul P, Toner M, Mullen D, Tewari P, Keegan H, Sharp L, O'Keane CJ, Timon C, Gheit T, Tommasino M, Martin CM, O'Leary JJ

Irish Association for Cancer Research 55th Annual Conference

Europa Hotel, Belfast, Northern Ireland

February 2019

ECHO: The Epidemiology of HPV Infection in Oral Cancer in Ireland

Sharkey Ochoa I, O'Murchu E, Woods R, Tewari P, O'Regan E, Toner M, Feeley K, Timon C, Sharp L, Martin CM, and O'Leary JJ

TCD International Postgraduate Research Conference

Trinity Biomedical Sciences Institute, Dublin, Ireland

March 2018

ECHO: The Epidemiology of HPV Infection in Oral Cancer in Ireland

Sharkey Ochoa I, O'Murchu E, Woods R, Tewari P, O'Regan E, Toner M, Feeley K, Timon C, Sharp L, Martin CM, and O'Leary JJ

Trinity College Dublin Multidisciplinary Research Showcase

Trinity Biomedical Sciences Institute, Dublin, Ireland

March 2018

ECHO: The Epidemiology of HPV Infection in Oral Cancer in Ireland

Sharkey Ochoa I, O'Murchu E, Woods R, Tewari P, O'Regan E, Toner M, Feeley K, Timon C, Sharp L, Martin CM, and O'Leary JJ

The Coombe Health and Social Worker Day Conference

Coombe Women and Infants University Hospital, Dublin, Ireland

February 2018

ECHO: The Epidemiology of HPV Infection in Oral Cancer in Ireland

Sharkey Ochoa I, O'Murchu E, Woods R, Tewari P, O'Regan E, Toner M, Feeley K, Timon C, Sharp L, Martin CM, and O'Leary JJ

TCD International Postgraduate Research Conference

Trinity Biomedical Sciences Institute, Dublin, Ireland

May 2017

ECHO: The Epidemiology of HPV Infection in Oral Cancer in Ireland

Sharkey Ochoa I, O'Murchu E, Woods R, Tewari P, O'Regan E, Toner M, Feeley K, Timon C, Sharp L, Martin CM, and O'Leary JJ

Trinity College Dublin Multidisciplinary Research Showcase

Trinity Biomedical Sciences Institute, Dublin, Ireland

March 2017

ECHO: The Epidemiology of HPV Infection in Oral Cancer in Ireland

Sharkey Ochoa I, O'Murchu E, O'Regan E, Toner M, Feeley K, Timon C, Sharp L, Martin CM, and O'Leary JJ

CERVIVA National HPV Awareness Symposium

Trinity Biomedical Sciences Institute, Dublin, Ireland

January 2017

ECHO: The Epidemiology of HPV Infection in Oral Cancer in Ireland

Sharkey Ochoa I, O'Murchu E, O'Regan E, Toner M, Feeley K, Timon C, Sharp L, Martin CM, and O'Leary JJ

Trinity College Dublin 10th International Cancer Conference

Trinity Biomedical Sciences Institute, Dublin, Ireland

October 2016

Abbreviations

μ	Micro
AJCC	American Joint Cancer Committee
CAP	College of American Pathologists
CDK	Cyclin dependent kinase
CERVIVA	The Irish HPV-related disease consortium
CCND1	Cyclin D1 gene
CWIUH	Coombe Women and Infants' University Hospital
DNA	Deoxyribonucleic acid
E6AP	E6 associated protein
ECHO	Epidemiology of HPV Infection in Oropharyngeal, Oral Cavity, and Laryngeal Cancer in Ireland
EGFR	Epidermal growth factor receptor
FFPE	Formalin-fixed paraffin-embedded
GDPR	General Data Protection Regulation
hc2	hybrid capture 2
HIV	Human Immunodeficiency Virus
HNC	Head and neck cancer
HNSCC	Head and neck squamous cell carcinoma
HIPE	Hospital Inpatient Episode Statistics
HPV	Human Papillomavirus
HR	Hazard ratio
HR HPV	High-risk HPV
HSE	Health Service Executive
hTERT	Telomerase reverse transcriptase
IARC	International Agency for Research on Cancer
ICD10	International Classification of Diseases – 10th Revision
IFN-α	Interferon-α
IFN-γ	Interferon-γ
IL-1B	Interleukin-1B

IL-6	Interleukin-6
IL-10	Interleukin-10
ISO	International Organization for Standardization
KPNC	Kaiser Permanente Northern California
LCR	Long control region
LR HPV	Low-risk HPV
LSCC	Laryngeal squamous cell carcinoma
mRNA	Messenger RNA
MALT	Mucosa-associated lymphoid tissue
MCHA	Colorimetric Hybridization Assay
NCRI	National Cancer Registry of Ireland
OPSCC	Oropharyngeal squamous cell carcinoma
OR	Odds ratio
OSCC	Oral cavity squamous cell carcinoma
PCR	Polymerase chain reaction
PD-1	Programmed cell death protein 1
PD-L1	Immune checkpoint ligand programmed cell death 1
pRb	Retinoblastoma gene/protein
RNA	Ribonucleic acid
SCC	Squamous cell carcinoma
SCJ	Cervical squamocolumnar junction
SIN	Squamous Intraepithelial Neoplasia
STI	Sexually transmitted infection
TNF-α	Tumour necrosis factor α
TOR	Transoral resection
TNM	TNM Classification of Malignant Tumours
UICC	Union for International Cancer Control
URR	Upstream regulatory region
WHO	World Health Organization

CHAPTER 1

Introduction

1 CHAPTER 1: INTRODUCTION

1.1 Head and Neck Cancers

1.1.1 Physiology of Head and Neck Cancers

Head and neck cancers (HNC) include those oncogenic growths of the oral cavity, pharynx, oropharynx, nasopharynx, hypopharynx, and larynx (Figure 1.1).

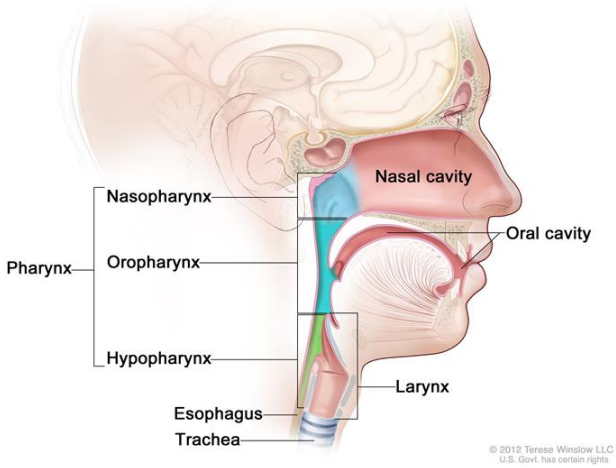

Figure 1.1 Anatomical sub-sites of the head and neck[1].

There are a vast number of types of malignancies that arise in the head and neck, with diverse pathologies and genetic profiles. Squamous Cell Carcinoma (SCC) growths comprise almost 95% of cancers of the head and neck, making head and neck squamous cell carcinoma (HNSCC) a valuable cancer group for epidemiological investigation[2].

1.1.2 Epidemiology of HNSCC

HNSCC is the sixth most common type of cancer worldwide, accounting for an estimated 633,000 new cases diagnosed annually accompanied by 355,000 deaths[3]. Over the last two decades, there has been a downward trend in HNSCC incidence and mortality, with variability based on anatomical site and geographic location or region[4].

Those variations based on geography tend to follow trends in tobacco use. For instance, the general incidence of HNSCC in the United States has varied and ultimately declined unwaveringly in tandem with changes in smoking activity (Figure 1.2)[5]. Similarly, the age-standardized incidence of HNSCC has increased in the last several decades in areas of Asia due to high rates of tobacco use[6].

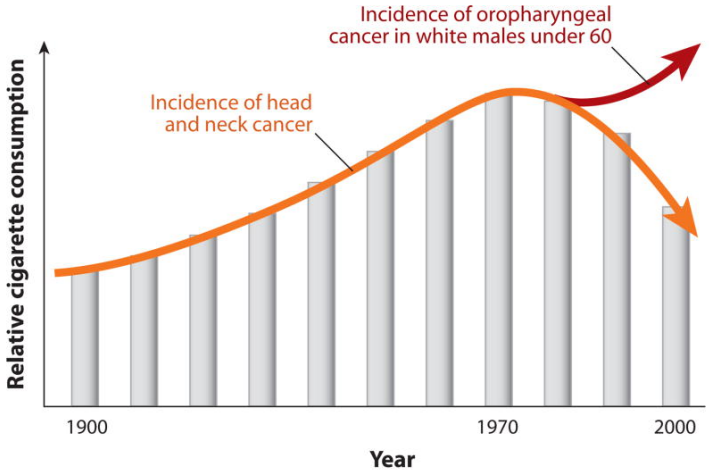

Figure 1.2 Incidence of HNCs in the United States as a function of smoking trends[5].

According to the National Cancer Registry of Ireland (NCRI), HNSCC, including the 5% of HNCs not diagnosed as SCC, is the ninth most common cancer in Ireland[7]. It accounts for approximately 1.7% of all invasive cancers in women, and 3.9% of all invasive cancers in men[4,7]. Comparatively, Ireland's estimated incidence of HNSCC is 8.3 per 100,000, as exhibited in Figure 1.3, where comparisons are shown with other countries and a difference in incidence between males and females is evident[7].

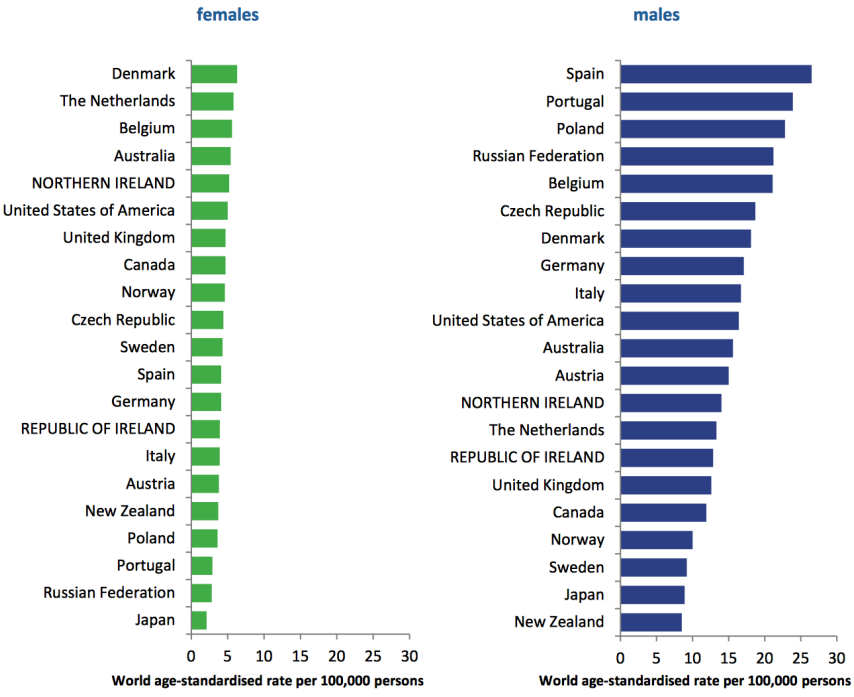

Figure 1.3 Estimated age-standardized incidence of HNCs per 100,000 persons for various countries around the world[7].

1.1.3 Histology of HNSCC

HNSCCs arise in most cases from preneoplastic lesions grouped under the term dysplasia. Microscopically, dysplasia is defined as architectural and cytological change of the epithelium, without evidence of invasion. In dysplastic lesions, the epithelium presents with irregular stratification, loss of polarity of basal cells, drop-shaped rete ridges, increased number of mitotic figures, abnormal superficial mitoses, premature keratinization in single cells (dyskeratosis), and keratin pearls within rete pegs[8]. Cytological changes include abnormal variation in nuclear or cell size and shape, increased nuclear to cytoplasmic ratio, increased nuclear size, atypical mitotic figures, increased number and size of nucleoli, and hyperchromasia[8].

The spectrum of dysplasia in HNSCC is divided into mild, or Squamous Intraepithelial Neoplasia 1 (SIN 1), moderate (SIN 2), and severe (SIN 3), based on the extent of architectural disturbances and cytological atypia. Once dysplasia surpasses severe, it becomes carcinoma in-situ presenting with full thickness or almost full thickness architectural abnormalities accompanied by cytological atypia.

HNSCC is characterized by invasive growth with disruption of the basement membrane, and squamous differentiation, often called keratinization, sometimes with keratin pearl formation. Extension into the underlying tissue is often accompanied by desmoplastic stromal reaction and a dense inflammatory infiltrate, mainly comprised of lymphocytes and plasma cells[8]. Angiolymphatic and perineural invasion may be seen.

SCC is graded into well-, moderately-, and poorly-differentiated classifications. Well-differentiated SCC resembles normal squamous mucosa whereas moderately-differentiated SCC displays nuclear pleomorphism, mitoses, and less keratinization. Poorly-differentiated SCC are composed predominantly of immature cells with typical and atypical mitoses, minimal keratinization, and sometimes necrosis. Most SCCs are moderately-differentiated[9].

Other noted epithelial markers for HNSCCs include cytokeratins, which are usually adequate for the diagnosis of well-differentiated tumours. In less well-differentiated lesions, further immunohistochemistry is useful for discerning grade, including staining for cytokeratin cocktails, AE1/AE3, pancytokeratin, CK5/CK6 and p63[10].

1.2 Oropharyngeal, Laryngeal, and Oral Cavity SCC

1.2.1 Physiology of Oropharyngeal, Laryngeal, and Oral Cavity SCC

HNSCC includes tumours from a number of subsites, three of the largest of which are the oropharynx, the larynx, and the oral cavity (Figure 1.1). Oropharyngeal SCC (OPSCC) accounts for approximately 10%[11] of HNSCC, while laryngeal SCC (LSCC) and oral cavity SCC (OSCC) account for approximately 33%[7,12] of HNSCC each.

The oropharynx begins at the soft palate, and descends through the base of the tongue (back 1/3 of the tongue), the lateral pharyngeal walls, the posterior wall of the throat, and the palatine tonsils (Figure 1.4). It is composed of three main anatomical sites: The palatine tonsils, the lingual tonsils at base of the tongue, and the soft palate. These are all flanked at the base by the larynx, and at the top by the oral cavity and the nasopharynx (Figures 1.4).

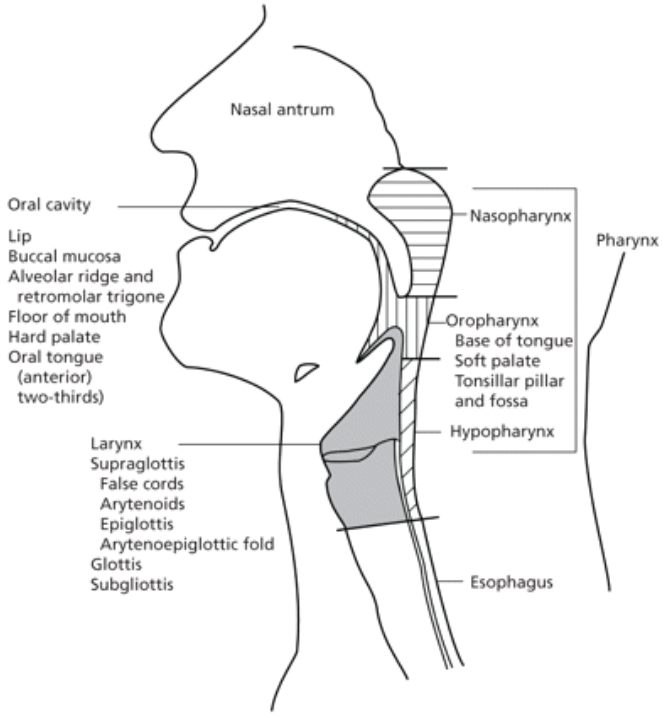

Figure 1.4 Detailed anatomical sub-sites of the head and neck[13].

The palatine tonsils are a pair of soft tissue masses located at the rear of the throat, composed entirely of lymphatic tissue. Similarly, the lingual tonsils are two small mounds of lymphatic tissue that are vertically adjacent to the soft palate and behind the circumvallate papilla taste buds which demarcate the base of the tongue from the oral tongue. The soft palate is the soft tissue constituting the back of the roof of the mouth, composed of muscle and connective tissue.

Embryologically, the lingual and palatine tonsils are respectively derived from the first and second pharyngeal pouch endoderm, as they become aggregated with mesodermal cells from the third and fourth pharyngeal pouches[14]. They contain four lymphoid compartments that influence immune functions, namely the reticular crypt epithelium, the extrafollicular area, the mantle zones of lymphoid follicles, and the follicular germinal centers[15]. They are part in parcel of the Waldeyer's ring, the cumulative name for the four pairs of tonsils in the head and neck (Figure 1.5).

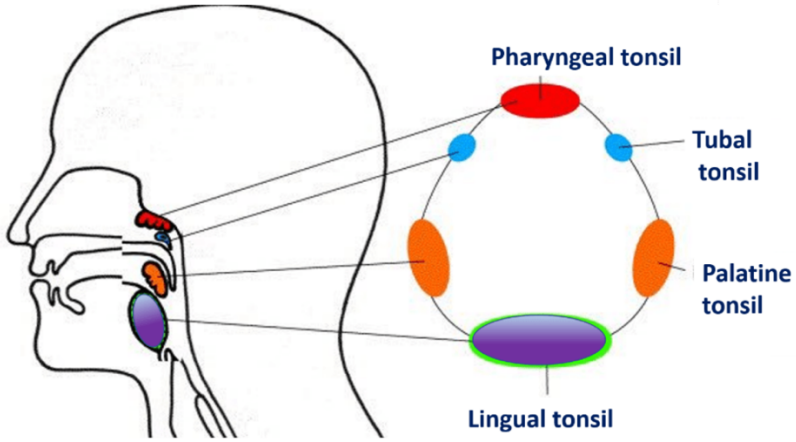

Figure 1.5 Physiology of Waldeyer's tonsillar ring in the head and neck[16].

In more detail, according to the American Joint Committee on Cancer (AJCC), the oral cavity site extends from the skin-vermilion junction of the lips to the junction of the hard and soft palate above and to the line of circumvallate papillae below[17]. It is divided into the mucosal lip, the buccal mucosa, the lower alveolar ridge, the upper alveolar ridge, the retromolar trigone, the floor of the mouth, the hard palate, and the anterior two thirds of the tongue. The oropharynx is the portion of the continuity of the pharynx extending from the plane of the inferior surface of the soft palate to the plane of the superior surface of the hyoid bone (or floor of the vallecula) and includes the base of tongue, the inferior surface of the soft palate and the uvula, the anterior and posterior tonsillar pillars, and the glossotonsillar sulci[17]. By contrast, the anterior limit of the larynx is composed of the anterior or lingual surface of the suprahyoid epiglottis, the thyrohyoid membrane, the anterior commissure, and the anterior wall of the subglottic region. The posterior and lateral limits include the laryngeal aspect of aryepiglottic folds, the arytenoid region, the interarytenoid space, and the posterior surface of the subglottic space, represented by the mucous membrane covering the surface of the cricoid cartilage[17]. The superolateral limits are composed of the tip and the lateral borders of the epiglottis. The inferior limits are made up of the plane passing through the inferior edge of the cricoid cartilage.

The oropharyngeal subsites, the larynx, and the oral cavity are all at least partially if not fully lined with mucosa-associated lymphoid tissue (MALT). The term MALT was first coined to emphasize that solitary organized mucosa-associated B-cell follicles and larger lymphoid aggregates have common features and are the origin of cells that traffic to mucosal effector sites[18,19]. The MALT mucosa is populated by lymphocytes such as T cells and B cells, as well as plasma cells and macrophages[20] and shows many histological similarities to the cervical mucosa. Given the position of the oropharynx, the larynx, and the oral cavity, their MALT is well positioned to encounter antigens passing through the mucosa's non-keratinising squamous epithelium.

The mucosa is particularly unique in the palatine and lingual tonsils, and the base of tongue (back 1/3 of tongue). Parts of the mucosal linings contain folds, or dips, often referred to as tonsillar and lingual "crypts" that extend through the full thickness of the tonsils (Figure 1.6). Crypts appear at around three to six months of age, with the capsule developing around the fifth month and germinal centres developing after birth.

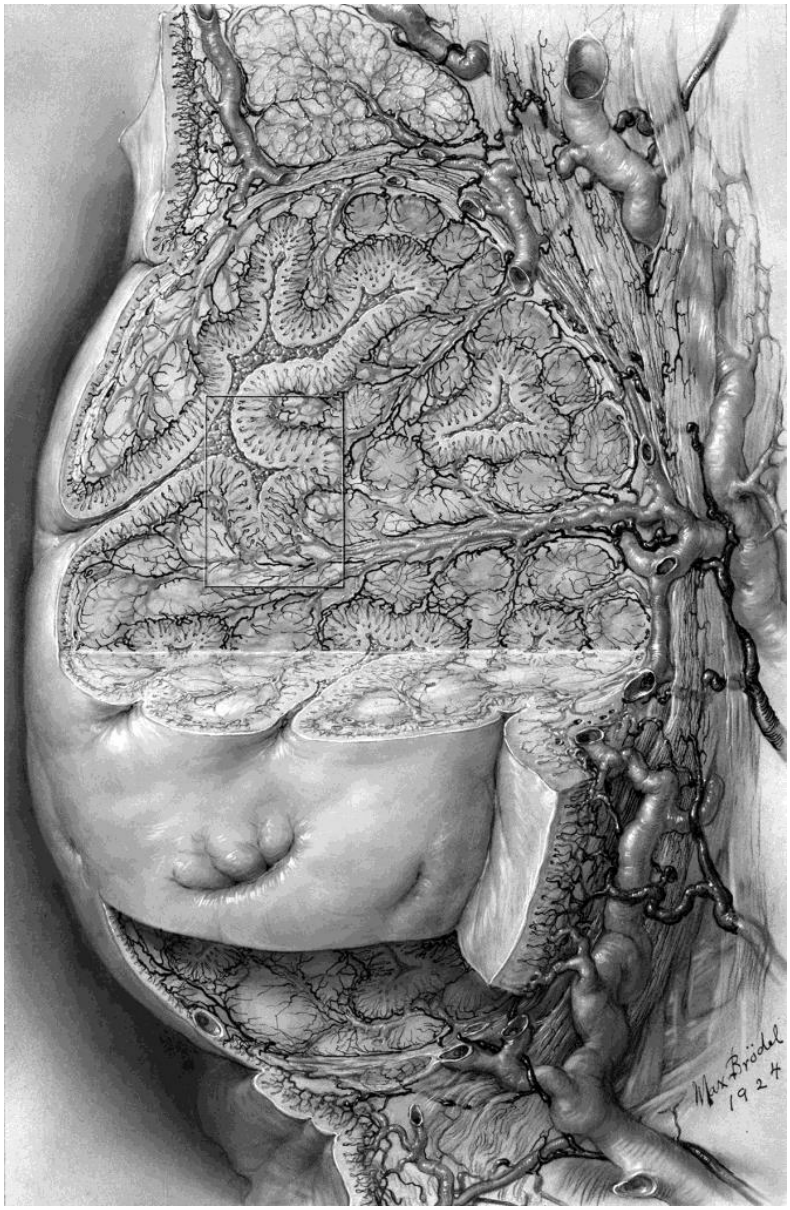

Figure 1.6 Microanatomy of the palatine tonsil with evident folds/dips of crypts, a portion of which are highlighted with a box. Drawing by Max Brödel[21].

Functionally, the crypts serve to greatly increase the total surface area of the oropharyngeal regions, dramatically increasing the efficiency of their antigen capture and immune surveillance[10]. Each palatine tonsil, for instance, has about 20 crypts which increase their contact surface area to approximately 295 cm^3, a 700% increase from the equivalent but unfolded region[5,22].

Unlike the rest of the oropharynx however, the crypt mucosa is lined with a squamous reticular crypt epithelium composed of three layers. The basal layer and membrane is disrupted and porous, thus allowing for the direct passage of lymphocytes, mononuclear phagocytic cells, plasma cells, and antigenic transport cells known as M (membranous) cells[5]. The intermediate layer is permeated by

lymphocytes and antigen-presenting cells[5]. The superficial squamous layer is thin and fragile. Complete desquamation of the superficial cells exposes the internal environment of the tonsil to external pathogens. Nonetheless, the crypts are a highly specialised region for antigen recognition.

It is from this squamous cell epithelium, most often in its unique reticulated state in the crypts, from which SCC originates. 44% to 51% of OPSCC in particular arise in the squamous lining of the palatine tonsil, while 35% to 55% originate in the base of the tongue, with another 2% to 12% developing in other squamous cell oropharyngeal sites[23,24].

1.2.2 Epidemiology of OPSCC, LSCC, and OSCC

While the incidence of other HNSCCs has decreased over the past three decades, correlating with decreased tobacco use, the age-standardized incidence of OPSCCs has been increasing over the same period across North America and Europe[23,25]. This is especially true of the SCCs of the base of the tongue and tonsillar region[26]. The incidence of these increased by 2% to 3% annually from 1974 to 2001, and then by 5.22% from 2000 to 2004 in the United States[27]. Most of Europe has had a similar experience, with Sweden seeing over a three-fold increase in tonsillar SCC, and the Netherlands observing an almost 2.7% annual increase in OPSCC between 1989 and 2011[12,26,28–31]. In the United Kingdom, statistically significant increases in incidence of OPSCC of 18% and 30% were seen in males and females respectively between 1990 to 1999[32]. A 51% increase of OPSCC in men was then observed from 7 per 100,000 to 11 per 100,000 between 1989 and 2006[33]. Danish data is even more resounding showcasing a constant linear increase in age-standardized incidence in tonsil cancers from 1980 to 2014[34,35].

Ireland is no different. Between 1996 and 2007, age-standardized oropharyngeal cancer incidence (excluding the tonsil) wavered between 0.23 cases per 100,000 per year to 0.1 cases per 100,00 per year[36]. A constant and annual increase after 2007 yielded an incidence of 0.98 cases per 100,000 per year by 2015. In fact, the Irish age-standardized incidence rate for cancer of the tonsil increased by 4.9% annually for women and 4.3% annually for men between the periods 1994 to 1998 and 2004 to 2008[4]. This consistent increase in age-standardized incidence continued after 2008 yielding an enormous and linear trend exemplified most succinctly in men from 1.2 cases per 100,000 per year in 1994 to 2.5 cases per 100,000 per year in 2015[36]. The base of tongue was the only other region of the oropharynx to experience a similar increase with the remaining regions seeing no change or in fact a decrease in the comparable age-standardized incidences.

Worldwide, the incidence of oropharyngeal cancer in white males under the age of 60 has since the 1980s been increasing directly in opposition to the trend of other HNSCCs (see Figure 1.2). As a consequence, OSSC is now the most common HNC in the world with an estimated 85,000 new cases of OPSCC diagnosed every year[37].

The larynx shows a more classic HNSCC incidence trend. Incidence was relatively constant in the United States between 1975 and 1990 at 5 per 100,000, and began descending after 1990 in an almost linear fashion[38,39]. Since 2015, incidence has hovered between 2 and 3 per 100,000[38]. Either the same decrease or a plateau is true of most European countries with variation mainly due to cultural differences between nations relating to smoking, alcohol, diet, and exposure to other carcinogens[30,35,40–43]. In Ireland, however, there has been no significant change in laryngeal cancer incidence regardless of gender or age since 1994 when data was first collected on the subject[36].

Cancers of the oral cavity are relatively difficult to epidemiologically track as they are often grouped into different subsets of the "tongue" and "mouth." That said, all of those groupings excluding the base of tongue have seen generally decreasing incidence trends across Europe and North America in the last three decades[31,40,44–46]. In Ireland, cancers of the floor of the mouth and palate have seen fluctuations in age-standardized incidence, but these have rested between 0.5 cases per 100,000 per year to 1 case per 100,000 per year[36]. Incidence of cancers of the mouth in women have actually seen an increase in the last decade, likely due to a resurgence in smoking behaviours amongst younger Irish people or the involvement of other carcinogens[36].

As the most common malignancy of the head and neck, that LSCC follows declines in smoking behaviour is unsurprising. Similarly, no widespread increase in cancers of the oral cavity suggests that behavioural and attitudinal changes towards health are having concrete effects on SCC incidence, though variations remain by country[47]. However, LSCC and OSCC resistance to unequivocal decline in smoking behaviour in some European countries, and more specifically in Ireland, suggest other possible carcinogens being involved. Furthermore, the larynx and oral cavity's proximity to the physiologically vulnerable parts of the head and neck in the oropharynx (see Section 1.6.4 for further details) and their similar MALT lining make them a necessary inclusion in all epidemiological assessments.

Thus, OPSCC incidence trends in particular have defied those seen across Europe and North America with respect to other HNCs, suggesting that unique carcinogens and risk factors are increasing the

incidence of oncogenesis in this particular physiological region. Several carcinogenic risk factors have been suggested to explain these epidemiological trends, the most significant of which is Human Papilloma Virus (HPV).

1.3 Human Papilloma Virus

1.3.1 The Virus

HPV is the most common sexually transmitted infection (STI)[48]. It is an epitheliotropic, non-enveloped DNA virus. Within its 55nm diameter, it carries a single molecule of circular double-stranded DNA, consisting of about 8,000 kilo-base pairs[49].

The HPV genome, as represented by Figure 1.7 depicting HPV genotype 16, has three regions: The early (E) region, the late (L) region, and the long control region (LCR), otherwise known as the upstream regulatory region (URR). The E region is comprised of 8 genes, where the L region is comprised of 2. The HPV genome encodes for 6 early proteins and 2 late proteins, called L1 and L2[50]. Each gene can be translated theoretically into three different proteins, depending on the site where transcription begins[50]. The LCR with E2 binding sites and the origin of replication of the virus contains the highest degree of variation in the viral genome[50].

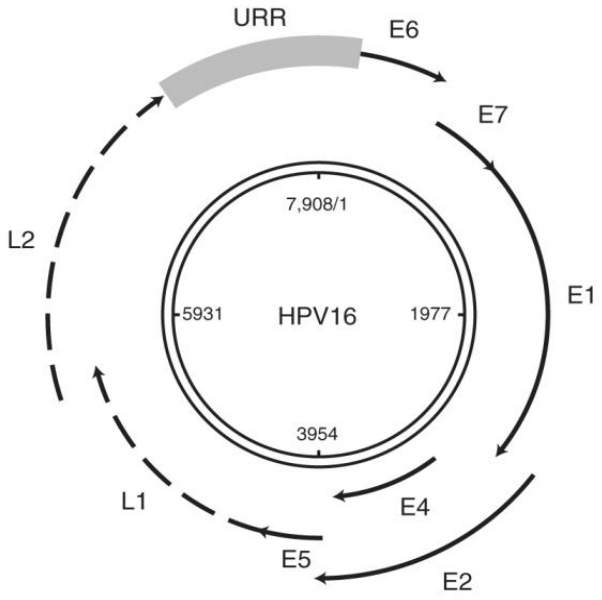

Figure 1.7 The structure and organisation of the HPV16 genome[49].

Generally, the early genes code for proteins responsible for virus replication, transcription, cell cycle, cell signalling and apoptosis control, immune modulation, and structural modification of the infected cell. Most of these proteins are expressed throughout the infectious cycle perhaps with reduced

expression at late times[51]. The late genes code for structural proteins, including the virus capsid required for virus transmission, spread, and survival in the environment[51]. Early proteins E3 and E8 have only recently been described in a few HPV types but their functions are relatively unknown[50]. Table 1.1 summarizes the functions of all of the proteins associated with the genes of the same name in the HPV genome including those genes most strongly associated with carcinogenesis.

Table 1.1 Roles and functions of each protein encoded by genes of the same name in HPV.

Oncoproteins implicated in carcinogenesis are indicated as such.

Protein	Role in the virus life cycle
E1	○ Genome replication: ATP-dependent DNA helicase.
E2	○ Genome replication, transcription, segregation, encapsidation. ○ Regulation of cellular gene expression. ○ Cell cycle and apoptosis regulation.
E3	○ Function relatively unknown.
E4	○ Remodels cytokeratin network. ○ Cell cycle arrest. ○ Virion assembly.
E5	○ Control of cell growth and differentiation. ○ Immune modulation.
E6	○ Oncoprotein. ○ Inhibits apoptosis and differentiation. ○ Regulates cell shape, polarity, mobility, and signalling.
E7	○ Oncoprotein. ○ Cell cycle control. ○ Controls centrosome duplication.
E8	○ Function relatively unknown.
L1	○ Major capsid protein.
L2	○ Minor capsid protein. ○ Recruits L1. ○ Virus assembly.

Though the general genomic structure of the virus is common to all of its types, there are over 200 different genotypes of papillomaviruses characterised by at least 10% nucleotide divergence in capsid gene (L1)[52]. The genotypes are classified in three ways. First, they are classified based on similarities in their DNA sequences. Second, they are grouped into mucosal-, mostly those of the alpha genus, or cutaneous-, mostly those of the beta genus, types. Third, they are most often classified into low-risk (LR) and high-risk (HR) types based on virulence and ability to promote malignant transformation in host cells.

HPV 16, 18, 31, 33, 34, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68, 73, and 82 are some of those classified as HR viruses, detectable in high grade squamous intraepithelial lesions in the cervix or in invasive cancer. HPV 6, 11, 40, 42, 43, 44, 54, 61, 72, 81, and 89 are designated LR and can be isolated from low grade epithelial lesions of the cervix[10]. HPV 6 and 11 are responsible for 90% of cases of genital warts[53].

Not all HPV types have been classified under the HR and LR categorization, as some have currently unknown oncogenic potential. In addition, there is a degree of intratypic variation, which may also relate to pathogenesis, as well as geographic variation in genotype prevalence[54–59].

1.3.2 HPV as a Carcinogen

Though HPV is highly ubiquitous in humans, only a small percentage of infections result in HPV-related cancers. In 1977, zur Hausen[60] first suggested a link between HPV and the pathogenesis of SCCs. By 1995, the International Agency for Research on Cancer (IARC) recognised that HR HPV types 16 and 18 were carcinogenic in humans[61]. Since, the popular and research focus has been on the role HPV plays in cervical cancer, now well-described with nearly all cervical cancer cases being caused by HPV[62]. HR HPV types are now also linked with other ano-genital tumours, including those originating in the vulva, vagina, penis, and anus, along with cutaneous, esophageal, and HNSCCs[63,64].

It is estimated that the virus currently accounts for 4.8% to 5.2% of the total global cancer burden, making it the most powerful carcinogenic virus in the world[65,66].

1.4 HPV in HNSCC, OPSCC, LSCC, and OSCC

HPV was first suggested and subsequently identified as a major etiologic factor in HNSCCs in 1983 and 1985 respectively[67,68]. Later its particular relevance to the subset of HNSCCs that arise mainly in the oropharynx was clarified[69,70].

Ever since, OPSCCs have been classified as a distinct clinicopathological entity in comparison to the traditional smoking- and alcohol-related HNSCCs, mostly given a unique genetic signature that HPV E6 and E7 proteins induce in contrast to the more genetically diverse tobacco-associated HNSCCs[10]. As a result, HPV is less heavily implicated in the carcinogenesis of LSCC and OSCC, but it still remains a carcinogen of interest given the varying epidemiology of LSCC and OSCC in many European countries.

1.4.1 The Epidemiology and Genotype Distribution of HPV in HNSCC, OPSCC, LSCC, and OSCC

The epidemiological data strongly supports the involvement of HPV as a carcinogen in HNSCC, specifically in OPSCC. The population-level incidence of HPV-positive OPSCC in the United States increased by 225% between 1988 and 2004, with a concomitant decline of 50% for HPV-negative OPSCC (Figure 1.8)[71]. More recent modelling data using the SEER dataset confirms these diverging trends in incidence[39].

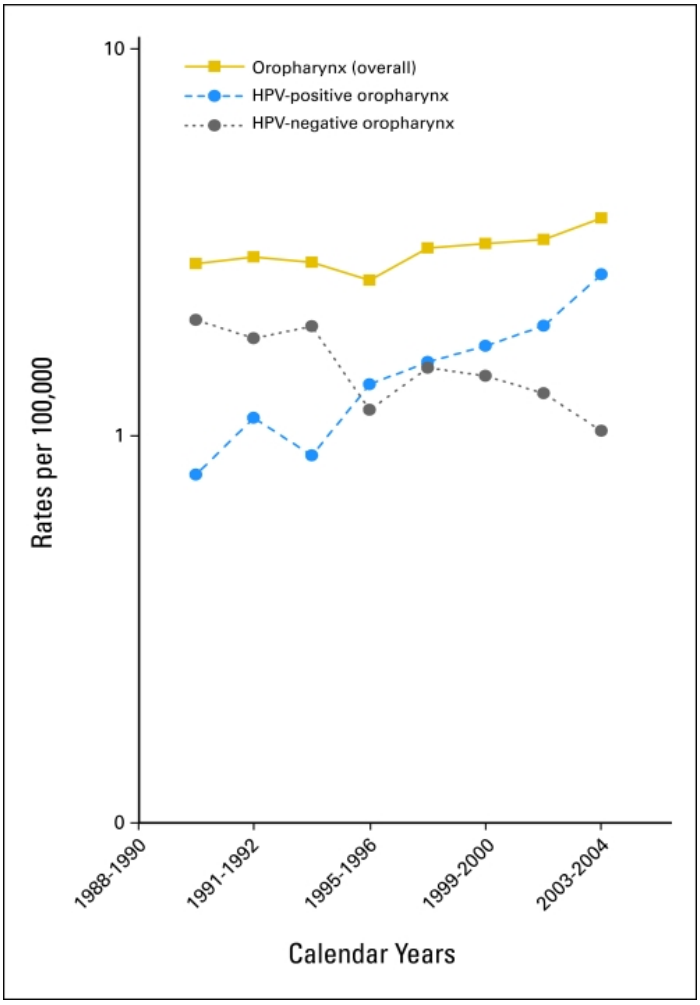

Figure 1.8 Incidence of oropharyngeal cancers per 100,000 between 1988 to 2004 in the United States overall and further subdivided by HPV status[71].

The Australian data showcases precisely the same trends with increasing incidence of HPV-associated oropharyngeal cancers between 1982 and 2005[26]. Canadian data also highlights that the age-standardized incidence of HPV-associated oropharyngeal cancer rose significantly from 1.6 per 100,000 in 1992 to 2.6 per 100,000 in 2009[72].

In Europe, similar figures are found in Denmark[73], and more specifically in Sweden where the incidence of HPV-positive tonsillar tumours almost doubled each decade between 1970 to 2007, indicating a 7-fold increase over the whole period, in conjunction with a parallel decline of HPV-negative tonsillar tumour incidence[29]. During the last 2 years of this Swedish study, 93% of all tonsillar cancer was HPV-positive[29]. Furthermore, Sweden has seen an overall increase in the incidence of base of tongue cancer from 0.15 per 100,000 between 1970 to 1974 to 0.47 per 100,000 between 2005 to 2007[74]. The prevalence of HPV in base of tongue cancer in Stockholm

county alone increased from 58% between 1998 and 2001 to 84% between 2004 and 2007, suggesting a link between the virus and base of tongue cancers[74].

There is limited data from other European countries regarding the epidemiology of HPV in HNSCC. Modelling work from Germany has shown without actual testing of archival samples that HPV-related HNSCC incidence has been increasing over the last two decades and represents a significant proportion of the overall HPV-related disease burden[75,76]. Preliminary French data also shows an undeniable role for HPV in oropharyngeal and oral cavity cancers[77].

This said, in the United Kingdom, despite significantly increased incidence of cancers of the oral cavity and oropharynx, the proportion of HPV-positive tumours in the region has hovered around 51% since 2002, suggesting that in the British context, HPV is not the only trend-driving carcinogen[78]. Data from northern Spain also shows relatively low HPV involvement in oropharyngeal and oral cavity cancer[79].

Nonetheless, the fact that HPV-negative OPSCC incidence has declined whilst overall OPSCC incidence has increased in the majority of North America and several European countries during the same time period suggests the significant impact of HPV-positive OPSCCs in driving the overall increasing incidence trend.

Prevalence statistics of OPSCC reflect this impact. A particularly steep rise of over 70% has been reported for prevalence of HPV-related OPSCCs in the past decade, with prevalence in Europe increasing at a faster rather than in North America[80,81]. With the rise in HPV-related OPSCC coupled with the decline of HPV-related cervical SCC, it has been suggested that the annual numbers of HPV-related OPSCC cases could soon surpass that of cervical cancer[82,83].

The long-term epidemiology of HPV in LSCC and OSCC is less clear. Since the late 1990s, it has been long-established that HPV is absent in the majority of laryngeal carcinomas[84,85]. Nonetheless, prevalence data from across the globe suggests a significant role for the virus in certain geographic regions. In the United States, 21% of invasive laryngeal cancers are HPV positive[86], a figure that plummets in Turkish patients to 7%[87], and skyrockets in Northeast China to 62%[88]. Systematic review finds a more representative figure of 24% LSCC HPV positivity in North America and Europe[89]. Comparisons of benign and malignant lesions of the larynx also suggest that LSCC appears to be characterized by an active HR HPV infection.

Several publications have found that HPV DNA is present in 10% to 25% of tumours of the oral cavity[90]. Frequencies vary between studies but the floor of the mouth and the tongue have been documented as infected by HPV in anywhere between 8% to 42% of cases[90]. Systematic review suggests an OSCC HPV prevalence very similar to that in LSCC at 24%[89]. However, the very definition of what the 'oral cavity' comprises in most studies is uncertain, which likely expands the range of available prevalence data and makes it more difficult to discern HPV's role in the anatomical region.

14 different oncogenic HPV types have been detected in cervical cancer, with HPV16 being the most common type followed in order of prevalence by HPV18, 58, 33, 45, 31, 52, 35, 59, 39, 51 and 56[91]. The genotype distribution of HPV in OPSCC, LSCC, and OSCC is similar with respect to detected types but completely different in terms of proportional representation across almost all North American and European populations analysed to date. Where, for instance HPV16 is identified in 60% of cervical cancers[91], it is overwhelmingly present in OPSCC, comprising at least 80% if not over 90% of cases in a large proportion of studies[28,70,77,89,92]. The same remains true in HPV-positive LSCC and OSCC[89,93].

Furthermore, unlike in the cervix where HPV18 prevalence is as high as 16%[91], the prevalence and incidence of HPV18 in HNSCC lags significantly behind that of HPV16, often representing less than 5% of HPV-positive cases. In OPSCC, HPV18 accounts for less than 3%[34] of HPV-positive cases, a figure that often drops to 1%, dependent upon the population[72]. LSCC and OPSCC also reflect this drastically diminished role for HPV18, though prevalence of the genotype tends to be higher than in OPSCC, between 4% and 8%[89]. In the developing world, HPV18 may play a larger role in HNSCC with evidence of HPV18 in oral lesions rising to 34% in Sudan[94]. Other genotypes that feature in OPSCC, LSCC, and OSCC are types 6, 26, 30, 31, 33, 35, 44, 45, 52, 58, 61, 67, 69, 82, and 91[14,95], the majority of which are HR types.

HPV's role in the carcinogenesis of other cancers is also becoming apparent[91]. The virus is increasingly implicated in anogenital[96–102], vulvar[103–106], vaginal[103,104], and penile cancers[104,106–109], all of which show varying HPV genotype distributions. Interestingly, those cancers more closely associated with men including anogenital and penile growths tend to see extreme HPV16 predominance, mimicking both the sex and genotype most related to HNSCC. The crystallizing epidemiological evidence thus suggests a crucial role for the preventative HPV vaccine and the development of further screening tools for HNSCC and others.

1.4.2 Histology of HPV-related HNSCC, OPSCC, LSCC, and OSCC

The pathological features of HPV-positive HNSCCs deviate from the moderately differentiated keratinizing morphology that typifies most HNSCCs. The latter are associated with dysplasia of the surface epithelium and are keratinizing. HPV-positive HNSCCs are consistently associated with lack of dysplasia of the surface epithelium and are poorly differentiated[110]. They exhibit lobular growth, and are permeated by infiltrating lymphocytes. They often involve central necrosis and cystic degeneration. They also lack significant keratinization, and demonstrate prominent basaloid morphology[110]. Lymph node involvement is frequent and commonly cystic.

The differential histological morphology of HPV-related HNSCC gives rise to three factors that cause diagnostic ambiguity in the field. The first is variation between the three subtypes in question. Those HPV-positive characteristics are most typical of OPSCC, and are less common and more heterogeneous in LSCC and OSCC given that much of their aetiology is still heavily influenced by the effects of smoking. The second and third factors are microscopic features of HPV-related HNSCCs[111]. The first of these is that HPV-related HNSCC is often mistaken as poorly differentiated carcinoma based on the immature appearance of tumour cells. The appearance of the tumour cells closely emulates the appearance of the reticulated epithelium of the region which is the specialized epithelium lining the tonsillar crypts and base of tongue from which the grand majority of HPV-related cancers arise[112]. Thus, most OPSCC in particular are in fact well-differentiated and not poorly differentiated as is often presumed. Second, the term “basaloid” is confusing for the way it invites an erroneous connection with basaloid squamous cell carcinoma, a subtype of HNSCC notorious for its aggressive clinical behaviour[111].

1.4.3 TNM Staging of HPV-related and HPV-unrelated HNSCC, OPSCC, LSCC, and OSCC

The TNM Classification of Malignant Tumours (TNM) is a cancer staging system that is used to describe the stage of a particular cancer when it originates with a solid tumour. The Union for International Cancer Control (UICC) TNM classification for OPSCC is identical to that of the American Joint Cancer Committee (AJCC).

According to the classification system, T refers to the size of the primary tumour and whether it has invaded nearby tissue or not. N refers to the degree of spread to regional lymph nodes. M refers to distant metastatic spread of the cancer to another site in the body[113].

The 8th edition of TNM staging was recently published[114]. However, the 5th edition[17] of the TNM staging recommendations is most relevant to the current endeavour given that this was the most up-to-date

staging system available at the time of diagnoses for the study between 1994 and 2013. It should be noted that in recent years, the staging for OPSCC specifically has been updated to reflect HPV and/or p16 status, showcasing the significance of the virus and its associated biomarkers in its carcinogenesis and improved prognosis[114] (see Section 1.8 for further details).

The TNM staging for oral cavity and oropharyngeal cancers is shown in Table 1.2 and that for laryngeal cancers is shown in Table 1.3. The stage groupings are then summarized for all of these cancers in Table 1.4.

Table 1.2 TNM staging classifications according to the AJCC 5th edition[17] for oral cavity and oropharyngeal cancers.

PRIMARY TUMOUR (T)	
TX	Primary tumour cannot be assessed
T0	No evidence of primary tumour
Tis	Carcinoma in situ
T1	Tumour 2cm or less in greatest dimension
T2	Tumour more than 2cm but not more than 4cm in greatest dimension
T3	Tumour more than 4cm in greatest dimension
T4	Tumour invades adjacent structures. In the case of the oral cavity this includes invasion through cortical bone, into deep (extrinsic) muscle of tongue, maxillary sinus, or skin. Superficial erosion alone of bone/tooth socket by gingival primary is not sufficient to classify as T4. In the case of the oropharynx including the base of tongue this includes the pterygoid muscle(s), mandible, hard palate, deep muscle of tongue, and larynx.
REGIONAL LYMPH NODES (N)	
NX	Regional lymph nodes cannot be assessed
N0	No regional lymph nodes metastasis
N1	Metastasis in a single ipsilateral lymph node, 3cm or less in greatest dimension
N2	Metastasis in a single ipsilateral lymph node, more than 3cm but not more than 6cm in greatest dimension; or in multiple ipsilateral lymph nodes, none more than 6cm in greatest dimension; or in bilateral or contralateral lymph nodes, none more than 6cm in greatest dimension
N2a	Metastasis in single ipsilateral lymph node more than 3cm but not more than 6cm in greatest dimension
N2b	Metastasis in multiple ipsilateral lymph nodes, none more than 6cm in greatest dimension
N2c	Metastasis in bilateral or contralateral lymph nodes, none more than 6cm in greatest dimension
N3	Metastasis in a lymph node more than 6cm in greatest dimension
DISTANT METASTASIS (M)	
MX	Distant metastasis cannot be assessed
M0	No distant metastasis
M1	Distant metastasis

Table 1.3 TNM staging classifications according to the AJCC 5th edition[17] for laryngeal cancers.

PRIMARY TUMOUR (T)	
TX	Primary tumour cannot be assessed
T0	No evidence of primary tumour
Tis	Carcinoma in situ
Supraglottis	
T1	Tumour limited to one subsite of supraglottis with normal vocal cord mobility
T2	Tumour invades mucosa of more than one adjacent subsite of supraglottis or glottis or region outside the supraglottis (e.g. mucosa of base of tongue, vallecula, medial wall of pyriform sinus) without fixation of the larynx
T3	Tumour limited to larynx with vocal cord fixation and/or invades any of the following: postcricoid area, pre-epiglottic tissues
T4	Tumour invades through the thyroid cartilage, and/or extends into soft tissues of the neck, thyroid, and/or esophagus
Glottis	
T1	Tumour limited to the vocal cord(s) (may involve anterior or posterior commissure) with normal mobility
T1a	Tumour limited to one vocal cord
T1b	Tumour involves both vocal cords
T2	Tumour extends to supraglottis and/or subglottis, and/or with impaired vocal cord mobility
T3	Tumour limited to the larynx with vocal cord fixation
T4	Tumour invades through the thyroid cartilage and/or to the other tissues beyond the larynx (e.g. trachea, soft tissues of neck, including thyroid, and pharynx)
Subglottis	
T1	Tumour limited to the subglottis
T2	Tumour extends to vocal cord(s) with normal or impaired mobility
T3	Tumour imited to larynx with vocal cord fixation
T4	Tumour invades through cricoid or thyroid cartilage and/or extends to other tissues beyond the larynx (e.g. trachea, soft tissues of neck, including thyroid and esophagus)
REGIONAL LYMPH NODES (N)	
NX	Regional lymph nodes cannot be assessed
N0	No regional lymph nodes metastasis
N1	Metastasis in a single ipsilateral lymph node, 3cm or less in greatest dimension
N2	Metastasis in a single ipsilateral lymph node, more than 3cm but not more than 6cm in greatest dimension; or in multiple ipsilateral lymph nodes, none more than 6cm in greatest dimension; or in bilateral or contralateral lymph nodes, none more than 6cm in greatest dimension
N2a	Metastasis in single ipsilateral lymph node more than 3cm but not more than 6cm in greatest dimension
N2b	Metastasis in multiple ipsilateral lymph nodes, none more than 6cm in greatest dimension
N2c	Metastasis in bilateral or contralateral lymph nodes, none more than 6cm in greatest dimension
N3	Metastasis in a lymph node more than 6cm in greatest dimension
DISTANT METASTASIS (M)	
MX	Distant metastasis cannot be assessed
M0	No distant metastasis
M1	Distant metastasis

Table 1.4 TNM stage groupings for oral cavity, oropharyngeal, and laryngeal cancers according to the AJCC 5th edition[17].

STAGE GROUPINGS			
Stage 0	**Tis**	**N0**	**M0**
Stage I	**T1**	**N0**	**M0**
Stage II	**T2**	**N0**	**M0**
Stage III	**T3**	**N0**	**M0**
	T1	**N1**	**M0**
	T2	**N1**	**M0**
	T3	**N1**	**M0**
Stage IVA	**T4**	**N0**	**M0**
	T4	**N1**	**M0**
	Any T	**N2**	**M0**
Stage IVB	**Any T**	**N3**	**M0**
Stage IVC	**Any T**	**Any N**	**M1**

1.5 Risk Factors in HNSCC

Most HNSCCs have traditionally been associated with patients with a long history of heavy smoking and alcohol consumption[115], poor oral hygiene[116,117], a diet low in fruit and vegetable consumption[118,119], and chronic inflammatory disease in the oral cavity[120,121], with age of onset falling in the seventh decade of life. HNSCC has also been associated with exposure to radiation[122], occupational exposures like leather dust or asbestos[123], and underlying genetic factors like Fanconi anemia[124]. Those HNSCCs, especially those of the oropharynx, that are HPV-related however, have their own risk factor signature, distinct from all other HNSCCs.

1.5.1 Age

HNSCC generally present in patients over the age of 50, with the majority of LSCC and OSCC being diagnosed from the age of 60 onwards[10,14,42,44,125]. HPV-positive HNSCCs however present at a younger age, averaging a few years lower than HPV-negative tumours[10,14,82,99,106,126,127]. HNSCCs developing in these younger patients show different genetic signatures than those in older patients, with distinct germline and somatic mutations[128–130]. Patients under 55 see up to a 3.4-fold higher risk of infection with carcinogenic HPV[131], and a strong association has been demonstrated with HPV16 infection and tonsillar cancer in males under the age of 40[132]. In the United States there has been an increasing

incidence in OPSCC in particular in those aged under 60, with a steep rise seen between the ages of 50 to 59[31].

1.5.2 Sex

HNSCCs show a male predominance ratio of approximately 3:1[10]. The NCRI reports an incidence rate for HNSCC of under 5 per 100,000 in women, and almost 15 per 100,000 in men[4]. American data shows however that for tobacco- and alcohol-related HNSCC, the gender difference has converged in tandem with smoking trends, with 43% of men and 30% of women smoking in 1974 compared to 26% of men and 21% of women smoking in 2000[133].

Male predominance remains amongst HPV-related HNSCC, and cannot be fully explained by factors like sexual behaviours. The literature also reports that oral HPV infection is more common in men than in women[92,134–137]. This suggests that innate biological differences between men and women render men particularly susceptible to the onset of HNSCC, including HPV-related HNSCC[82]. Some characteristics that may preferentially predispose men to cancer of the oropharynx are hormonal differences[45,138,139] and the potential protective immunity from seroconversion in response to cervical HPV infections among women[45,140,141]. It is also possible that the transmissibility of oral HPV may be higher for men performing oral sex on women, possibly due to a higher HPV copy number in the vagina and cervix[142].

1.5.3 Socio-Economic Status, Region, and Race

HNSCCs have historically been associated with patients from low socio-economic groups given the higher rates of smoking and alcohol consumption amongst this group and the slower rates of decline in these behaviours[143]. Indeed, recent data from North America indicates that socio-economic status still plays a determining role in HNSCC patient outcomes, with lower socio-economic status being associated with later stage presentation and worse survival[144–146]. LSCC and OSCC are also persistently associated with smoking and alcohol behaviours and socioeconomic deprivation[125,147–149]. However, in the case of HPV-related HNSCCs, it is patients from higher socio-economic groups who have a better baseline performance status that are at higher risk[150,151].

White males seem to be particularly at risk of HPV-related HNSCC, with a rise in incidence reported in this group alone, mostly due to the influence of OPSCC[31,71,152]. HPV-positive HNSCC has a lower incidence and prevalence in African Americans than in other racial groups, with poorer survival in this racial group given that a higher proportion of HNSCC in this group is related to tobacco and alcohol

exposure[153,154]. LSCC and OSCC are common in African Americans, reflecting their relationship to tobacco and alcohol[155,156].

Given the significant role that smoking and alcohol still play in LSCC and OSCC, their relative contribution to HNSCC prevalence overall in the developing and developed worlds is less distinct than that of OPSCC. The developing world has a relatively low proportion of OPSCC, comprising about 1% to 10% of HNSCC, while the incidence rate of OPSCC has steadily increased in most developed countries[46,157]. The developed world features a relatively high proportion of OPSCC, comprising about 15% to 30% of all HNSCCs[158]. These varying proportions might be attributed to an ever-increasing number of HNSCC cases in the developing world caused by increasing smoking and alcohol consumption rates. Where the latter trend is the opposite in the developed world, non-smoking and non-alcohol HNSCC carcinogens like HPV account for a larger proportion of the burden.

1.5.4 HPV Serology

There is a strong association between serologic evidence of HPV infection and HNSCC risk, even after adjustment for other HNSCC risk factors[159]. One highly-cited study has even revealed a distinct temporal association between HPV serology and HNSCC development. Pre-diagnostic serum samples from ten year periods in this investigation that were positive for HPV16 capsid antibodies conferred an increased risk of OPSCC of 14.4[160]. An additional serum-based study also showed that patients with pre-diagnostic E6 seropositivity had a significantly higher risk of oropharyngeal cancer[161].

1.5.5 Sexual Behaviour

Sexual behaviours have not typically been deemed causal in HNSCC, but sexual behaviour and smoking and alcohol consumption are often related to one another in risk analysis. With the emerging relevance of HPV, HPV-positive HNSCCs have been strongly associated with number of lifetime sexual partners, number of vaginal, oral, and anal sex partners, young age at first intercourse/earlier sexual contact, and history of sexually transmitted diseases, including genital warts[126,131,162,163]. The greater the number of partners and historical sexually transmitted diseases, no matter their classifications, the larger the risk of HPV-positive HNSCC. In fact, even after adjusting for HPV16 serology, the associations in case-control studies were no longer significant[162]. This finding reveals that sexual behaviours may be considered a surrogate for HPV16 exposure.

Further data from Europe and North America indicates that markers of HR sexual behaviours, such as earlier ages of sexual debut, practice of premarital sex, average number of lifetime partners, and

practice of oral sex, have all increased among recent birth cohorts, a trend that very clearly follows the patterns of incidence of HPV-related OPSCCs[14,142,164].

These behavioural observations are sensical given that a greater number of partners and a larger medical history suggests increased exposure to HPV or other immunocompromising pathogens. Increased sexual exposure as a risk-factor for HNSCC is especially relevant to HPV-related HNSCC as the virus is independently the most globally common sexually transmitted infection, suggesting a potentially exponential risk increase in relation to particular sexual behaviours.

1.5.6 Oral HPV Infection

Oral HPV is most often acquired via sexual transmission. HPV acquisition increases around age of sexual debut with oral HPV prevalence of 1.5% in 12 to 15 year-olds, 3.3% in 16 to 20 year-olds, and 4.5% to 6.9% in healthy adults[135,136,165,166]. Higher oral HPV prevalence has been reported in women with cervical HPV infection[167,168], and people infected with Human Immunodeficiency Virus (HIV)[167,169]. Some studies have posited that HPV transmission is possible through kissing[166,170], as well as intrapartum transmission[171], and transmission during laser surgery[172].

Previous studies and reviews have established that oral HPV16 infection is a strong risk factor for oropharyngeal cancer, while the relationship has not always been clear for OPSCCs specifically[14,159,173]. More recent data suggests however that persistent oral HPV infection risk is significantly higher in patients with HNSCC than in those with no malignancies[174]. This is especially true in men whose of risk OPSCC is significantly increased with detected oral HPV infection, specifically of the HPV16 variety[137].

The relationship between oral HPV infection genotype and HPV-related HNSCC, and thus the exact natural history of HPV infection in the oral cavity, is as yet still an undetermined one. Where it is well known that HPV16 plays the largest role in HPV-related HNSCC carcinogenesis, oral HPV genotype distribution does not always correspond to HNSCC HPV genotype distribution[175,176]. However, this may simply be a testament to the unique persistence of HPV16 in the oral cavity after the onset of infection. Recent data has shown that where, for instance, HPV18 is more frequently detected in the oral cavity than in developing HNSCC, it is one of the most easily cleared HPV genotypes[177].

In comparison to the cervix, HPV prevalence is lower than cervical HPV prevalence at an incidence of 3.5% to 3.7%[135,136]. This may be due to a lower proportion of oral-genital than genital-genital partners[138]. Additionally, though type-specific concordance is low, HPV infection of the cervix and oral

cavity are not independent[178] and so cervical HPV infection could be considered a risk factor for oral cavity HPV infection.

1.5.7 Immunodeficiency

Immunodeficiency is a risk factor for all HNSCC[122,179,180]. However, given the viral origins of carcinogenesis in HPV-related HNSCC, immunodeficient patients are at particular risk of developing this subset of cancers. Patients infected with HIV for instance have a 2 to 6 times increased risk of HPV-related HNSCC[181–183]. These patients are also at even greater risk of ano-genital SCCs[184]. This said, amongst HIV-positive patients, risk factors for the development of HNSCC are similar to the general population, including both HPV-related and tobacco/alcohol-related HNSCC[185].

In cervical cancer patients, immunosuppression leads to HPV persistence and disease progression[186]. This may translate to the relationship between immunosuppression and HPV-related HSNCC development, and may also partly explain any potential association with tobacco exposure due to the immunosuppressive effects of smoking[187], including reduced antibody response in smokers[188].

1.5.8 Tobacco and Alcohol Exposure

Tobacco and alcohol exposure are the two main carcinogens in the majority of HNSCC, especially those of the larynx and oral cavity[10,26,38,43,44,147,155]. This is clear in the literature and has been well-documented since the 1970s.

The roles of tobacco and alcohol exposure in HPV-related HNSCC and in oral HPV infection are by contrast uncertain based on the literature[189]. Some studies suggest a positive association by way of smoking-induced immunosuppression, leaving patients more vulnerable to HPV infection or an inability to clear the virus before persistent infection transforms cells. Others report a possible additive and/or multiplicative role for smoking and alcohol in the genetic transformation of cells into malignancies in HPV-infected patients[131,189–192]. Recent investigations show a clear role for the synergistic effects of smoking on increased HPV-related HNSCC risk in tandem with other risk factors including oral HPV infection, age, and gender[137,175].

Many other studies report no association, especially when accounting for the confounding elements that associate increased risk of HPV infection with smoking and alcohol given particular sexual behaviours[69,126,162,176,193,194]. It is possible that tobacco exposure potentiates the effects of HPV carcinogenesis, but the resounding evidence does not yet exist in the literature[195].

The latter is reflected in the fact that the role for tobacco smoking in cervical cancer becomes a weak one after adjustment for sexual and reproductive factors[196]. Furthermore, in comparison to traditional HNSCCs, patients with HPV-related HNSCC are less likely to have excessive tobacco exposure and alcohol use[70,197,198], though HPV-related HNSCC does also occur in both smokers and alcohol users.

1.6 Carcinogenesis in HNSCC

1.6.1 Carcinogenesis in HNSCC

Carcinogenesis in HNSCC can be summarized with reference to the hallmarks of cancer. Importantly, genomic instability, resistance to cell death, evading growth suppressors, sustained proliferative signalling are crucial to disease progression and are underpinned by exposure to carcinogens over time leading to genetic and epigenetic changes that accumulate, resulting in the transformation of cells due to altered transcriptional and translational activity to premalignant and eventually malignant lesions. This said, HNSCC is a heterogeneous subset of cancers with various subtypes described, based on histological appearance, and supported by different gene expression profiles[199,200]. Deep-sequencing studies on the HNSCC oncogenome have demonstrated a vast number of diverse genetic alterations affecting multiple different molecular mechanisms and cancer hallmarks[201].

As synthesized by Woods[14], these alterations coincide in four targetable hallmark molecular pathways. The convergence of HNSCC genomic alterations highlights the fact that genes affected by smoking and alcohol in carcinogenesis are for the most part distinct[202]. The four targetable pathways involved are summarized here:

1. ***Mitogenic signalling***, in particular the amplification or up-regulation of Epidermal Growth Factor Receptor (EGFR) and the downstream pathway of phosphoinositide 3-kinase (PI3K)/mTOR as well as PTEN inactivation, each leading to pathways involving proliferation, DNA repair, survival, and spread.
2. ***Defective differentiation*** involving NOTCH signalling alterations.
3. ***Cell cycle de-regulation*** involving inactivation of CDKN2a tumour suppressor gene encoding p16, and the Cyclin D1 gene (CCND1) which encodes Cyclin D1 amplification.
4. ***Genomic instability*** involving loss of TP53 and other genes related to DNA damage recognition and repair which is the most common genomic alteration in all HNSCC.

Field cancerisation is a hallmark in the carcinogenesis of most HNSCC. First denoted in 1953[203], field cancerisation is the wide distribution of genetic alterations throughout mucosal lining in the aerodigestive tract despite the absence of overt malignant histopathological changes[112]. Leukoplakia and erythroplakia represent a minority of precancerous fields in the oral cavity[204]. A further minority of patients presenting with either, between 6% to 36%, go onto develop oropharyngeal, oral cavity, and laryngeal SCC[14,205]. The accumulation of further genetic changes in these precancerous fields leads to the development of HNSCC, with the presence of field change leading to a higher risk of multiple synchronous or metachronous primary tumours[14]. However, there is no strong evidence for a field effect in HPV-related SCC[206] and the risk of second primary OPSCC has markedly decreased over time[207]. It is also well-established that the mutation rate of HPV-positive tumours is less than that of HPV-negative tumours, with some studies citing a rate of half that found in HPV-negative HNSCC[208,209].

Nonetheless, chromosomal alterations and gene transcription result in all HNSCC, though there are noted differences identified between HPV- and non-HPV related HNSCCs[210–213]. TP53 mutations, loss of chromosome locus 9p21, hypermethylation of 14-2-2 sigma protein and RASSF1A promotions, and overexpression of Cyclin D1 are all common in HPV-unrelated OPSCCs while pRb levels are normal and p16 expression is often decreased[214,215].

With respect to induction of angiogenesis and reprogramming of energy metabolism in HNSCC, these are induced by hypoxia, stimulating the production of vascular endothelial growth factor, platelet-derived growth factor, fibroblast growth factors, and interleukin-8 (IL-8)[216]. High vascular endothelial growth factor has been documented in HNSCC with the factor or its receptors increased in 50% of premalignant and 75% of malignant oral and laryngeal lesions[217,218]. The factor also shows high positive correlation for IL-8 overexpression. Tumours that overexpress these angiogenic elements are more aggressive and have shorter disease-free intervals[217]. The role of tumour-promoting inflammation is less clear in HNSCC, especially when divided by HPV status. High density of immune cells in some tumours indicates better survival in HPV positive cases, where it seems to promote spread in HPV negative cases[219–221]. The same is true of immune evasion, which is discussed further in Section 1.6.2. Furthermore, E-cadherin forms key protein complexes with catenins which link neighbouring cells, and their under-expression in HNSCC is associated with lymph node metastases[222]. Integrins, transmembrane proteins that allow attachment to the extra-cellular matrix and have a key role in migration, are often irregularly expressed in HNSCC causing disordered cellular attachment and local tumour spread[223].

The role of mutational load in HNSCC, both HPV positive and negative, is as yet unconfirmed. Indeed, some studies suggest an overall lower level of mutational loads in HPV positive HNSCCs, but others observe a comparable level of mutational burden or frequency, albeit with differing profiles[224,225]. Mutational loads in HPV negative cases exhibit a mutational profile consistent with tobacco exposure, while HPV positive cases rarely show mutations in these more classic HNSCC genes including TP53 and CDKN2A, and PTEN[226]. For all HNSCC, up to 30% of cases harbour mutations in genes that regulate squamous differentiation (NOTCH1, IRF6, and TP63), implicating its dysregulation as a major driver of HNSCC carcinogenesis[225].

The epigenome also contributes to the evolution of HNSCC, hypermethylation of various genes being the most studied element of this area. Hypermethylated genes in HNSCC fall strictly into several of the key hallmarks of cancer including apoptosis, cell cycle progression, DNA repair, inflammation, protein glycosylation, and invasion and metastasis. In HPV negative cases, RASSF1, STAT5, CDKN2A, CHFR, MGMT, SPDEF, and ESR2 are often hypermethylated, where the same is true of CCNA1, TUSC3, JAK3, CADM1, CDH11, IGSF4, TIMP3, and GRB7 in HPV positive cases[227]. Recent studies have also identified unique DNA methylation signatures in HPV-positive HNSCC, showing novel differentially methylated CpGs and regions associated with viral infection that are independent of anatomic site. Most hypomethylated regions appear to be characterized by a marked loss of CpG island boundaries[228]. E6 and E7 oncoproteins have also been show to affect histone acetylation and methylation pattern by interacting with histone acetyltransferases, histone deacetylases, histone methyltransferases, and histone demethylases. Non-coding RNA are also of biologic relevance to HNSCC due to their role in modulating gene expression[227]. They can interact with histone-modifying complexes and DNA methyl transferases. In fact, using different cohorts of OPSCC, an miRNA panel that differentiated HPV positive from HPV negative tumours has been recently identified. Strong upregulation of miR-9 for instance has been observed in HPV positive cases compared to HPV negative cases in which miR-9 expression has been found to be silenced by promoter methylation[229,230]. Behavioural factors, most notably smoking and alcohol consumption, affect all of these epigenetic factors however, making HNSCC a diverse set of disease regardless of HPV status.

Ultimately, it is exposure to carcinogens that brings about these field changes and genetic alterations in all HNSCC, one of the most powerful of which is HPV.

1.6.2 HPV-Related Carcinogenesis in HNSCC: Persistent Infection, Immune Evasion, Viral Integration, and Viral Load

HPV-related carcinogenesis in HNSCC follows the same general progression as those in other HNSCCs, starting with consistent exposure to a carcinogen over time. Drawing from the cervical experience, exposure to HPV is generally temporary. HPV infections are almost always cleared by the immune system with some studies reporting that 90% of infections are cleared within two years[231,232]. However, HR HPV genotypes tend to persist longer than LR types[231,232]. It is this persistence, a direct result of immune evasion by the virus, that ultimately leads to carcinogenesis.

HPV evades the immune system in a variety of different ways related to both its natural life-cycle and its specific oncogenic activity led by oncogenes E6 and E7[233]. HPV manipulates gene expression, protein function, antigen processing, and the extracellular environment to evade the immune system in an almost infinite number of ways[234–238]. There are several key features of the virus' activity however that highlight its evasive potential.

First, HPV's normal life cycle is conducive to immune evasion itself. The replication of viral DNA in undifferentiated cells and subsequent particle formation in external epithelial cells reduce its exposure to the host immune system[238]. HPV does not cross the basement membrane of these epithelial cells where immune cells are far more abundant, and instead deposits on the basement membrane (see Figure 1.11 for further detail). That its life cycle is non-lytic means that it does not elicit any pro-inflammatory signals that activate dendritic cells. This also means that there is no blood-borne phase of the HPV life cycle, again minimizing the virus' exposure to the immune system.

Second, in initial infection, the virus keeps gene expression at low levels, controlled by protein E2 hindering the detection of infected cells by local immune cells[237]. These E6 and E7 proteins, though up-regulated upon integration of the virus, are non-secreted proteins and remain within the nucleus of cells, reducing the risk of immune detection. Furthermore, the virus has been implicated in alterations of the gene expression of chemokines, cytokines, and adhesion molecules that facilitate its immune evasion[237].

Third, HPV mediates the dysregulation of antigen processing machinery and the production of the very antigens that are processed and presented to the adaptive immune system[233]. The virus down-regulates the peptide MHC-complexes on the surface of infected cells, protecting them from immune detection and destruction. In fact, HPV-related tumours exhibit loss of MHC Class I[237,238]. E6 and E7

immunogenic peptides are also not efficiently processed and presented. E6 and E7 from HPV16 and HPV18 in particular actually repress promoters for MHC Class I heavy chain expression[239], while E7 actively disrupts antigen presentation.

Fourth, HPV dysregulates and avoids interferon type I effects. There is evidence to suggest that interferon-α (IFN-α) does not inhibit the transcription of E6 and E7 oncoproteins, though interferon-γ (IFN-γ) does[238]. HPV16 and HPV18 E7 can bind to the transactivation domain of some interferon response factors which are downregulated as a result of E7 expression[237]. E6-mediated down-regulation of E-cadherin in keratinocytes may also limit the presentation of viral antigens by dendritic cells to the immune system[233,238].

Finally, HPV evades the immune system through the manipulation of extracellular strategies affecting antigen presenting cells, T-cells, tumour-associated macrophages, myeloid-derived suppressor cells, and regulatory T-cells[237]. Particularly, HPV E6 and E7 oncoproteins alter the recruitment and location of epidermal dendritic cells, while soluble regulatory factors derived from HPV-induced hyperplastic epithelium change dendritic development and influence the initiation of specific cellular immune responses[240]. Even more importantly, HPV interferes with the recruitment of T helper cells through suppression of associated cytokines including interleukin-6 (IL-6) and interleukin-10 (IL-10)[237].

Once HPV establishes immune evasion through these extensive and intricate mechanisms, integration of HPV DNA into the cellular genome is likely the next critical step for carcinogenic progression. This has been noted especially amongst patients whose HPV infections affect their tonsils[112]. In fact, anatomically, HPV DNA integration is consistently centred on the tonsillar crypt epithelium[112].

HPV's integration into the host genome in particular cells is thought to be mostly a random process, though it is very likely that along with systemic risk factors, there are certain cellular characteristics that make some cells more at risk of infection and integration than others. Nonetheless, integration results in the clonal selection of aggressively expanding cells that display altered gene expression of integrated HPV genomes and potential perturbations of cellular genes at or near viral integration sites[241]. Furthermore, integration immortalizes host cells with potential for malignant transformation through increased proliferative capacity, upregulation of anti-apoptotic pathways, and increased genomic instability[14,242]. Viral integration can also lead to loss of E2-mediated inhibition of viral oncoprotein expression, further potentiating carcinogenesis[243].

It should be noted however that transcription of E6 and E7 viral oncogenes can occur when HPV is still episomal. Some studies suggest that episomal HPV alone contributes to the development of most HPV-related HNSCC in contrast to SCCs of the cervix[244,245]. This said, in cervical SCC, alteration of E2 on integration facilitates increased expression of E6 and E7 oncogenes[244], and some investigations have concluded that viral integration in the tonsillar crypts is crucial to carcinogenesis[246,247].

Once integrated and replicating, a greater HPV viral load should logically result from an immune evasive HPV infection. It has been suggested that high HPV viral load (at least one HPV copy per tumour cell) in HNSCC predicts active HPV infection[248–250], but the cervical case has shown that it is impossible to predict tumour progression based on viral load alone[251]. In fact, the proportion of HPV-positive SCCs with high viral load varies dependent upon studies which report high viral load prevalence between 33% to 77.5%[71,248]. It is possible that low viral load is representative of a transient HPV infection that simply happens to be present in the tumour at the time of testing. However, gene and oncogene expression in HNSCC varies greatly. Thus, the simple existence, rather than a high expression of viral oncogenes may be sufficient to spur HPV-related HNSCC carcinogenesis[245].

1.6.3 HPV-Related Carcinogenesis: Oncogenes and Genome Alterations

Following HPV's persistent infection, immune evasion, and likely, integration, it is the oncogenes of the virus, in particular E6 and E7, that are the driving force of HPV carcinogenesis[252]. The E6 and E7 genes produce E6 and E7 oncoproteins that respectively inhibit the activity of tumour-suppressing p53 and retinoblastoma (pRb) genes and proteins.

The E6 oncoprotein's main contribution to HPV-related HNSCC carcinogenesis is the promotion of ubiquitin-mediated proteolysis of the p53 protein through an E6 associated protein (E6AP) (Figure 1.9). When HPV E6 binds to the E6AP, the complex then binds to p53, ultimately leading to its degradation (Figure 1.9). Under normal circumstances, p53 facilitates repair to damaged host DNA by arresting cells in the G1 phase or inducing apoptosis. E6-expressing cells therefore face increased mitotic stress and genomic instability[253]. p53's degradation is accompanied by E6's interaction with c-myc which results in increased proliferation through the consequential up-regulated transcription of telomerase (hTERT in Figure 1.9)[254]. Furthermore, under normal circumstances, p53 induces protein p21 activity, whose overexpression results in cell cycle arrest and the maintenance of DNA repair (Figure 1.9). Degradation of p53 therefore decreases p21 expression and cells proliferate freely[50]. Finally, along with targeting the Wnt and NOTCH signalling pathways[255], HR HPV E6 proteins bind to PDZ domains containing proteins with tumour suppressor activity[255,256].

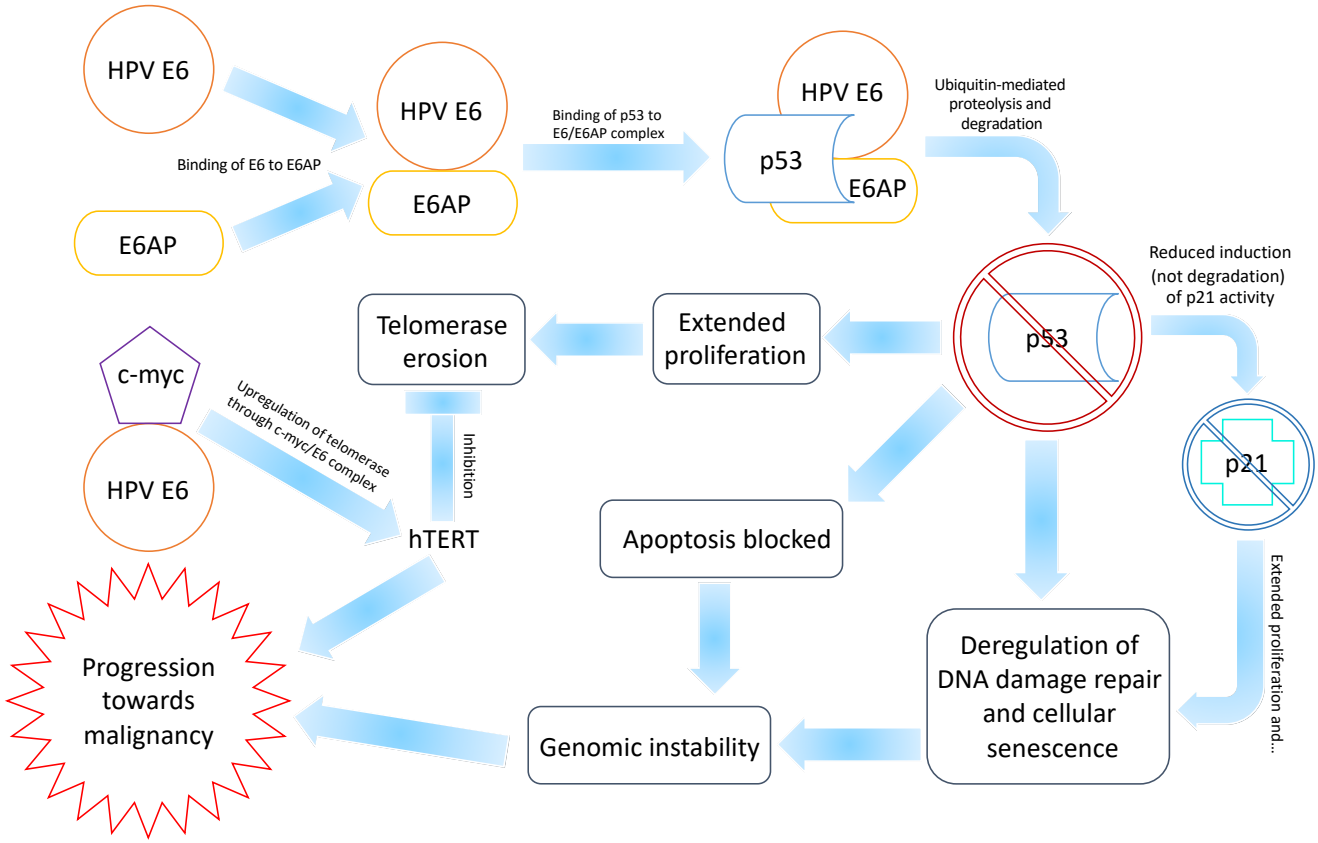

Figure 1.9 Schematic diagram of HPV oncoprotein E6's main involvement in the carcinogenesis of HPV-related HNSCC.

The E6AP binds to the E6 oncoprotein, a complex which then binds tumour suppressor protein p53. This leads to ubiquitin-mediated proteolysis and degradation of p53. Under normal circumstances, p53 induces the activity of protein 21, but when it degrades, p21 is less active and cell proliferation and DNA damage occur. E6 also interacts with c-myc, causing the upregulated expression of telomerase through the transcription of the hTERT gene. Overall, the downstream effects of E6's genomic alterations include extended cellular proliferation, the arrest of apoptosis, deregulation of DNA damage repair, and cellular senescence.

The E7 oncoprotein is the most important driver of cell cycle deregulation in HPV-related HNSCC carcinogenesis. It causes cell cycle disruption through its interactions with the pRb tumour suppressor protein and thus increases the downstream activity of histone deacetylases, cyclins, and cyclin-dependent kinases (CDK) (Figure 1.10). It does this primarily by mimicking the phosphorylation of pRb, which cyclins and CDKs normally induce, resulting in the release of E2F transcription factors (Figure 1.10). E2F factors are necessary for the transcription of genes involved in proliferation and cell cycle progression[257]. Cells therefore abnormally enter the S-phase from the G1 phase in the cell cycle (Figure 1.10). E7 also binds to cyclin and CDK inhibitors, most importantly protein p21 (Figure 1.10). It also binds and degrades the pRb protein itself, further releasing E2F and destabilizing regulation of cellular senescence[258]. This functional inactivation of pRb results in overexpression of the p16 tumour suppressor protein, which is a CDK4A inhibitor (Figure 1.10). This renders HPV a somewhat masochistic carcinogen but also allows the use of p16 as a surrogate marker for HPV-related oncogenesis[127,259–262].

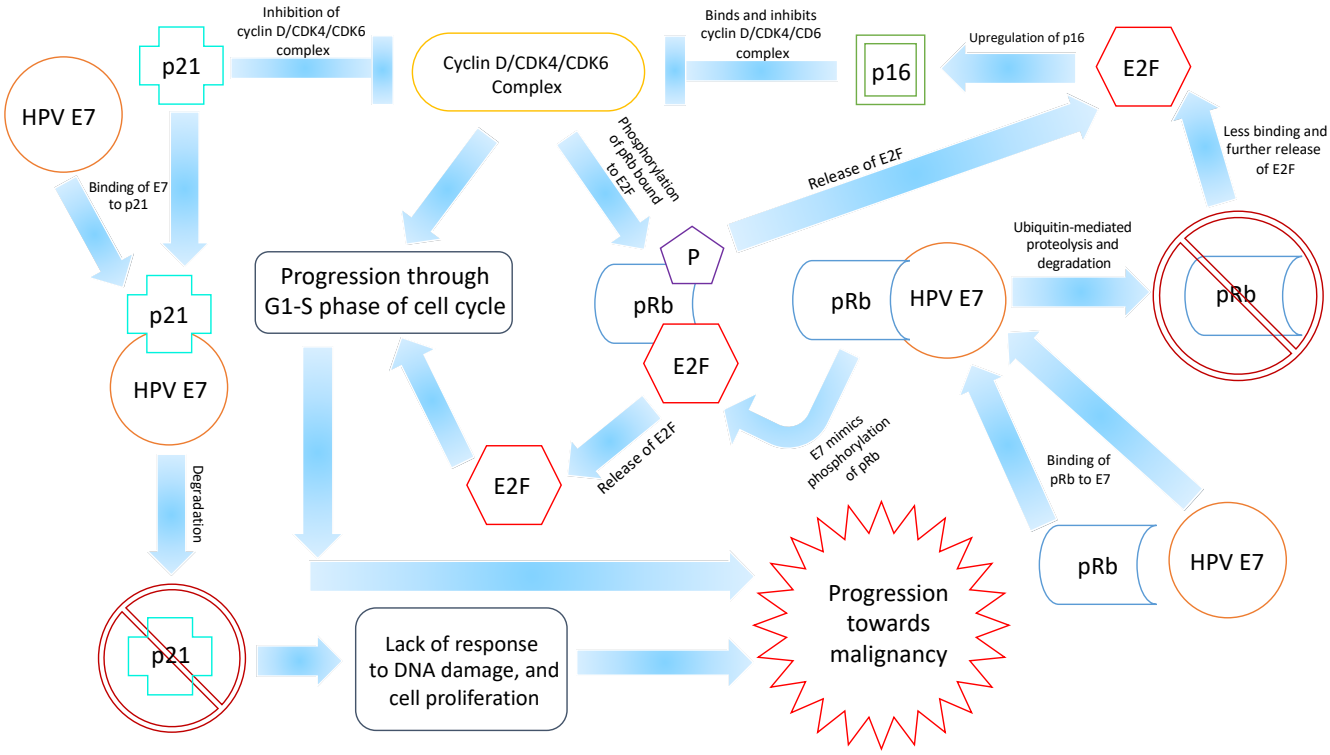

Figure 1.10 Schematic diagram of HPV oncoprotein E7's main involvement in the carcinogenesis of HPV-related HNSCC.

E7 targets the pRb tumour suppressor pathway and protein, and deregulates the cell cycle during the G1 phase. Under normal circumstances, mitogenic signals activate cyclins which promote progression though the G1-S phase of the cell cycle in a manner dependent on cyclin-dependent kinases CDK4 and CDK6. pRb in its hypo-phosphorylated form is associated with E2F molecules. E2F transcription factors allow cell cycle progression through the G1 to S phases. Cyclin D/CDK4/CDK6 complexes cause the phosphorylation (represented by P in a purple pentagon in the diagram) of pRb bound to E2F, which in turn results in the release of E2F. Thus, cells

progress through the G1 and S phases. During HPV carcinogenesis, E7 mimics the phosphorylation of pRb which leads to the release of E2F, progression into the S phase, and proliferation. E7 also interacts with the p21 protein leading to its degradation. This is significant because p21 plays a critical role in the cellular response to DNA damage and its overexpression results in cell cycle arrest. p21 also inhibits the cyclin D/CDK4/CDK6 complexes and its degradation prevents this inhibition. Finally, E7 binds to pRb and induces its ubiquitin-mediated proteolysis and degradation. This furthers the lack of binding of pRb to E2F transcription factors which does indeed yield increased progression through the G1-S phase of the cell cycle. However, the release of E2F through E7 actions also causes the up-regulation of p16. p16 is a tumour suppressor and binds to cyclin D/CDK4/CDK6 complexes, inhibiting progression through the G1-S phase of the cell cycle and preventing phosphorylation of pRb bound to E2F transcriptions factors.

The E5 oncoprotein cooperates with E6 and E7 to promote proliferation of infected cells and probably facilitates malignant progression[263]. E5 primarily modulates EGFR signalling pathways and delays the internalization and degradation of EGFR, leading to increased levels of EGFR on the surface of E5-expressing keratinocytes through interaction with the 16kD subunit of the vacuolar ATPase, which normally acts to acidify the endosomal environment[50]. HPV-related HNSCC, in particular OPSCC, is strongly associated with low EGFR expression. In tandem, viral integration frequently leads to loss of E5 expression. Thus, it is logical to conclude that E5's oncogenic activity likely occurs in the early stages of carcinogenesis[264].

Genomic instability underpins the development of dysplasia, malignancy, invasion, and metastasis in cancers[265]. Indeed, the aberrant proliferation induced by E7 described and pictured in Figure 1.10 is facilitated by suppression of apoptosis by E6 mechanisms described and pictured in Figure 1.9. However, it is the additional functions of E6 and E7, and to an extent E5, by multiple mechanisms that result in chromosomal mutations and genomic instability. These chromosomal abnormalities include centrosome irregularities and spindle checkpoint failure, both of which give rise to polyploidy, aneuploidy, and chromosomal rearrangement[266,267], DNA damage[268,269], disruption of checkpoint control mechanisms, variation in the Fanconi anemia DNA repair pathway[14], and induction of the ATM-ATR DNA damage repair pathway[258]. Tobacco and alcohol consumption also cause genomic instability[270]. This may be used as further evidence for a role for tobacco especially in the potentiation of HPV-related HNSCC in the context of the effects of E5, E6, and E7.

With respect to DNA methylation, diverging patterns have been observed between HPV-related and HPV-unrelated HNSCCs. HPV-related HNSCC methylation patterns are far more similar to those demonstrated in cervical SCC than those in HPV-unrelated HNSCC[271]. The same is true of the miRNA

profiles of HPV-related HNSCCs which are similar to those of cervical SCCs[272]. Excess DNA methylation could be a mechanism used by the virus itself to evade the immune system, or it may by a defence tactic utilized by the host cell[271].

To note is that there are also differences in DNA methylation rates between HNSCCs in tobacco-users versus non-users as well as specific mRNA and miRNA clusters[273]. This suggests that in tandem with differential methylation patterns for HPV-related HNSCC, there may be a distinct set of cancers with unique genetic and molecular profiles that result from a combination of viral and behavioural origins. There is evidence to suggest that smoking and HPV status for instance, when combined in survival analysis, may segregate patients into distinct categories[10,14,274].

It should be noted that despite differential carcinogenic mechanisms, the effects on downstream pathways of HPV- and smoking-related cancers are often the same. This includes, for instance, mTOR inhibition which results from TP53 mutations in tobacco related cases and from E6 induced degradation of p53 in HPV-related cases[201]. That said, the specific viral origins of HPV-related HNSCC have very evident clinical implications for patient prognosis, survival, and treatment (see Section 1.8). Thus, the early effects of HPV in carcinogenesis are significant and not to be dismissed simply because their downstream effects converge with those of smoking- and alcohol-related cases. It is also important to highlight that E6 and E7 proteins expressed in LR HPV types do not induce the same genetic changes that HR HPV do in HNSCCs. These may therefore exist as latent passenger viruses with no transcriptional activity[275,276].

In any case, there is a distinct group of HPV-related head and neck tumours arising from the epithelium of lymphoid tissue characterised by viral oncoprotein expression. Much like in the cervix, HPV infection is likely not only an early factor in carcinogenesis, but the origin of carcinogenesis in these cases. The viral oncoproteins E6 and E7 of HR HPV genotypes, particularly HPV16, are the primary drivers of cellular transformation in HPV-related HNSCC[252]. Their downstream effects result in the onset of tumours in specific parts of the head and neck with particular genetic and molecular profiles.

1.6.4 HPV-Related Carcinogenesis in HNSCC: Anatomical and Functional Vulnerabilities

HPV is implicated in the carcinogenesis of cancers in a defined region of the head and neck, including the oropharynx, the larynx, and the oral cavity. This suggests that there are particular anatomical and functional characteristics of the region that render it especially vulnerable to persistent HPV infection that evades the immune system and ultimately transforms cells.

The oropharynx, larynx, and oral cavity are full of lymphatic tissue and MALT whose role is to enact and support immune function. In fact, it is the palatine and lingual tonsils, protected only by the immune properties of the saliva, that act as a first line of immune defence against pathogens and carcinogens for the rest of the body (Figures 1.4 and 1.5). Positionally, the location of the oropharynx and the palatine and lingual tonsils in particular, enhances their ability to immunologically defend against pathogens entering through the mouth (Figures 1.4 and 1.5). However, these functional and structural properties, though optimal for immunological activity, result in the sacrificial vulnerability of the region, disproportionately exposing it to pathogens and carcinogens. This is in direct contrast to for instance, the pharyngeal tonsils which are physically and biochemically protected from direct pathogenic contact by the uvula and the major cavity of the nasopharynx (Figures 1.4 and 1.5). For this reason, or perhaps despite it, the oropharynx, larynx, and oral cavity, though most especially the tonsils, are known to harbour pathogenic viruses such as Epstein Barr virus, adenoviruses, and herpes simplex virus[277].

Much the same functional and structural vulnerability is present at the cellular level, as showcased in Figure 1.11. HPV requires an exposed basement membrane to depose itself and become persistent. In the cervix, the basal cell layer of the three-layered mucosa is not naturally porous or disrupted. Thus, HPV infection not only requires desquamation of the superficial epithelium but also demands trauma to the basal cell layer with subsequent deposition of the virus onto an exposed basement membrane[278]. In the reticulated epithelium of the oropharynx and to some extent in the MALT of the larynx and oral cavity, the natural state of the basal cell layer is disrupted (Figure 1.11). Thus, superficial epithelial disruption and unrest leave the basement membrane exposed to viral deposition without the need for mucosal trauma (Figure 1.11). The normal cellular structure of the tonsillar crypts therefore uncovers the basement membrane of the tonsil to the precise pathogens, including HPV, whose successful persistent infection necessitate its exposure (Figure 1.11).

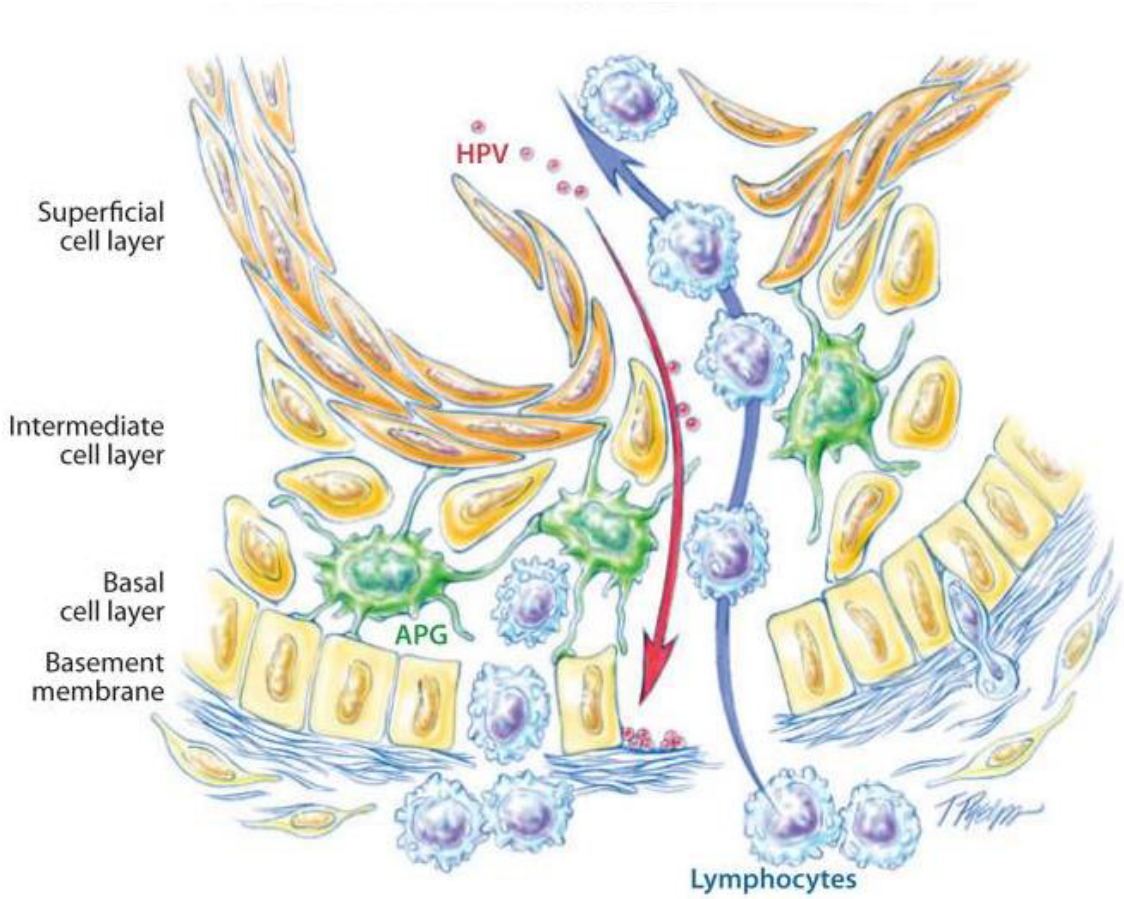

Figure 1.11 The reticular epithelium of a tonsillar crypt, drawing by T. Phelps[5].

The drawing shows the three cell layers of the reticulated epithelium of the tonsillar crypts, highlighting the fact that the basal cell layer exists in a naturally disrupted and porous state. Under normal circumstances, this allows for the efficient transfer of lymphocytes across the basement membrane to support the best possible immune function. However, the drawing also indicates that this natural porosity results in the exposure of the basement membrane after partial or complete superficial cell desquamation. HPV, to become persistent in infection, must be able to deposit itself on an exposed basement membrane. Thus, the natural exposure of the basement membrane in the tonsillar crypt is the precise anatomical environment required for HPV to become carcinogenic.

Viral entry itself may be facilitated by M-cells lining the crypt epithelium[50], as with other viruses[279,280]. Increased cytokines related to nearby lymphoid tissue may additionally influence the promotion of carcinogenic HPV infection specifically[281]. Furthermore, the permissive nature of the reticular epithelium to cellular migration may contribute to the early metastasis of HPV-positive OPSCC[5].

Functionally, the basal cells of the crypts are the only dividing cells in the epithelium, indicating that their transcriptional and translational machinery, involved in cellular division, are uniquely and

continuously operational. HPV's genomic alterations to a basal cell may therefore not need to be as significant as in the other epithelial cells of the crypts, or other parts of the head and neck, for a carcinogenic pathway to develop.

That the vast majority of HPV-positive HNSCCs localize to the tonsillar crypts of the lingual and palatine tonsils and head and neck sites lined with MALT is thus unsurprising. It is clear that the structural and functional properties of the region that normally optimize its immunological functions ultimately increase its risk for originating HPV-related HNSCC.

1.6.5 A Comprehensive Model for HPV-Related Carcinogenesis in HNSCC

Based on the cervical model, a small number of patients with persistent HR HPV infection, immune evasion, viral integration, and significant viral load will ultimately see genetic changes which lead to precancerous lesions. A further fraction of these patients will eventually develop cancer many years after the original infection. Indeed, HPV-related precursor lesions in the oral cavity have been identified[282]. This said, there have so far been no detectable HPV-related precancerous lesions in the oropharynx[82], despite SCC of this region being the most heavily associated with viral origins. This may be related to the technical difficulties in clinically assessing and sampling deep tonsillar crypts[283,284]. No dedicated HNC screening tools have been developed, nor have any accompanying screening systems, due to the relatively newly discovered significance of HPV-related HNSCC.

Those patients who do go on to develop HPV-related HNSCC however will never do so without the combination and cascade of events the foregoing discussion describes. In other words, none of the genetic, molecular, or anatomical features of HPV-related HNSCC will work in isolation from one another throughout carcinogenesis. For instance, the vulnerability of the oropharynx's anatomy to infection only becomes a risk when mechanisms of immune evasion allow persistent infection at these sites, which is one of the main reasons why immunosuppressed individuals are particularly at risk.

Thus, based on the foregoing review of the unique epidemiology, physiology, and oncogenesis of HPV-related HNSCC in comparison to HPV-unrelated HNSCC, it is possible to construct a distinct model of HPV-related HNSCC carcinogenesis. The model is detailed and portrayed in Figure 1.12. It begins with active HPV infection in the vulnerable anatomical sites of the head and neck, particularly in the MALT tissue and tonsillar crypts, followed by viral immune evasion, the expression of viral oncogenes, E6/E7 mediated cellular transformation, progressive genetic alteration, eventual malignancy, and potential metastasis.

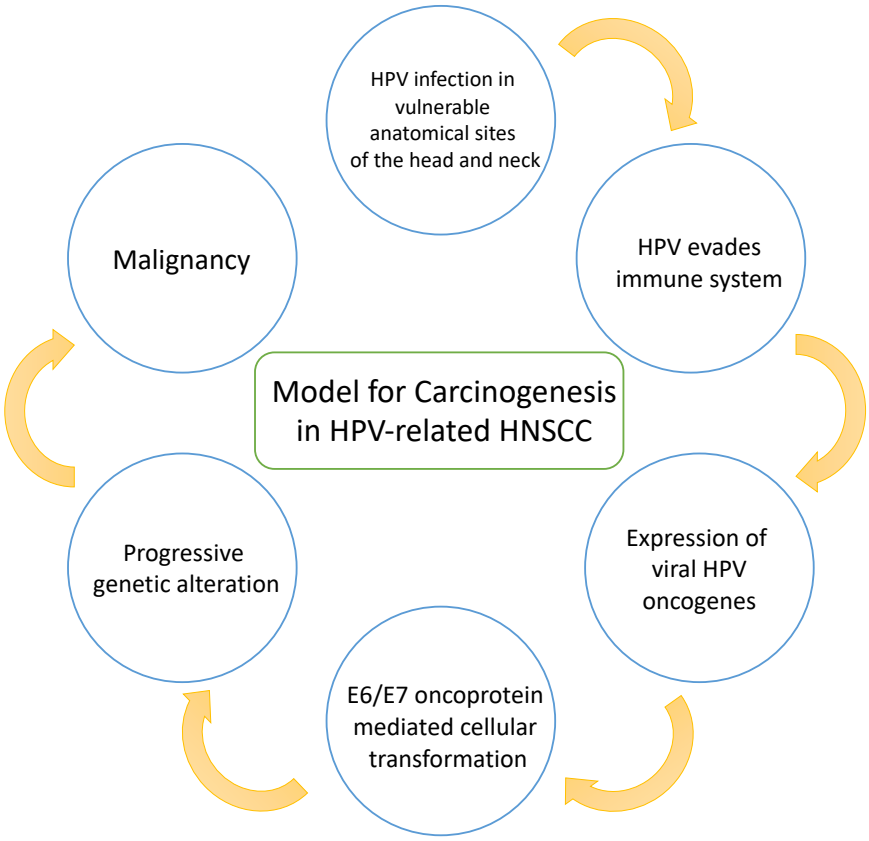

Figure 1.12 Proposed model for carcinogenesis of HPV-related HNSCC.

The model posits that carcinogenesis begins with HPV infection of vulnerable anatomical sites of the head and neck, including the oropharynx, oral cavity, and larynx. In particular, HPV infection in palatine and lingual tonsils, and base of tongue are suggestive of potential carcinogenesis given the partially desquamated reticulated epithelium of the region that allows HPV's deposition on the basement membrane. The virus then evades the immune response through various mechanisms, becoming persistent in the vulnerable anatomical sites. The virus, perhaps integrated or still episomal, then expresses its viral oncogenes including E5, E6, and E7. The latter two then subsequently mediate cellular transformation through their interactions with the p53 and pRb tumour suppressor pathways. This leads to progressive genetic alteration, and eventual malignancy.

1.7 Biomarkers in HNSCC

1.7.1 Defining 'Biomarkers'

The word biomarker refers to a broad subcategory of medical signs, objective indications of medical state observed from outside the patient, which can be measure accurately and reproducibly including molecules, genes, and alleles[285]. Medical signs stand in contrast to medical symptoms, which by comparison are limited to indications of health or illness perceived by the patients themselves[285].

The use of biomarkers in both basic and clinical research has become commonplace. Some biomarkers have been so well defined that many serve as reliable predictors, detectors, confirmers, and identifiers of clinical outcomes across a variety of diseases, treatments, outcomes, and populations.

Biomarkers can be pharmacodynamic based on their mechanistic action on tissue, prognostic based on their effect on clinical outcome, predictive based on an effect on clinical outcome from any intervention, and surrogate, acting as substitutes for clinical outcomes[14,286].

1.7.2 Biomarkers in HPV-related and HPV-unrelated HNSCC

It is difficult to identify clinically useful biomarkers in HNSCC due to the heterogeneity in HNSCC carcinogenesis, genome, and subsequent prognosis[287]. Nonetheless, as examined above, HPV-related HNSCC, especially those of the oropharynx and to an extent those of the oral cavity and larynx, show converging agreement for several biomarkers, some proving promising in treatment and prognosis predictions. The same is true for HPV-unrelated HNSCC as a comparative sub-category. Biomarkers therefore represent the opportunity to assess specific therapeutic approaches to dealing with these subsets of cancer based on their developing genetic signatures.

1.7.3 HPV Status

HPV-status has been designated a predictor of prognosis, survival, and outcome in HNSCC in many studies[288–290]. Those studies that find significance in statistical analyses find that HPV-positivity predicts improved response to treatment, prognosis, and outcome[287,291–293]. These conclusions hold true no matter the indicator of HPV-positivity including HPV DNA and HPV RNA[88,249,294,295]. However, HPV-positive OPSCC patients with high expression of the E6 oncoprotein have 5 times greater a risk of distant disease recurrence and significantly worse cancer-specific survival[296]. In addition to this, though the literature is almost unanimous in its determination of HPV status as a positive prognostic biomarker in HNSCC, one study did show poorer survival for HPV-positive OSCC[297]. Individual experience of HPV infections, sub-site, and geographic region thus nuance more positive outcomes documented for HPV-positive OPSCC patients.

1.7.4 p16

p16 has proven an extremely useful biomarker in HNSCC. As shown in Figure 1.10, the tumour suppressor protein p16 is upregulated in HPV-related carcinogenesis as a result of the downstream effects of the HPV E7 oncoprotein. p16 binds to the cyclin D/CDK4/CDK6 complex, stabilizing the pRb tumour suppressor protein in its active hypophosphorylated form (Figure 1.10). It is therefore

unsurprising that a huge number of studies observe a strong correlation between integrated HPV detection and p16 protein overexpression in HNSCC, so much so that many researchers suggest p16 as a surrogate biomarker for HPV-related HNSCC[127,259–262,288,293,298]. Pooled analysis shows that 46% to 98% of HPV-positive OPSCCs demonstrate p16 positivity[262].

By contrast, the upregulation of p16 is not evident in HPV-unrelated HNSCC, where downregulation or loss of the p16 protein expression is a common early event and is associated with worse prognosis, consistent with the tumour-suppressor role it has[265]. As a consequence of these diverging p16 profiles, p16 is becoming a more routine part of clinical assessment of HNSCC in terms of subsequent treatment options due to better survival in HPV-related cases, and its use has now expanded to the reframing of staging for OPSCC and OSCC[17,114].

Controversy remains however about p16 as a true surrogate for HPV-positivity. What is defined as HPV-positivity (e.g. HPV DNA positivity, HPV mRNA positivity) differs in most studies, and which indicator equivocates to p16 as a biomarker changes accordingly. False positivity for HPV based on p16 may be higher outside the oropharynx where the prevalence of HPV-related SCC is lower. p16 expression can also occur in 5% to 8% of HPV-negative HNSCCs[299].

Furthermore, p16 can often be a confounding factor when trying to differentiate between HPV-related and smoking-related HNSCC. Many HPV-positive tumours are p16 negative. In fact, pooled analysis shows that between 2% and 54% overexpression of p16 is absent from a subgroup of HNSCC with an HPV infection[262,300]. Furthermore, HPV positivity and p16 negativity/methylation often correlate with patients who are smokers or have a history of smoking[130,301]. Conversely, some HPV-negative tumours are p16 positive, with dysregulation of epigenetic control or multiple transcription factors being mechanisms that lead to aberrant expression of p16[302]. This may suggest the emergence of new hybrid classes of HNSCC where either both HPV and smoking are involved in disease progression, or for some other reason, p16 is upregulated in HPV-unrelated cases.

Preliminary data already suggests that though HPV-positive/p16-positive cases show the best overall survival, p16-positive cases show better survival regardless of HPV status. Thus, several studies point out that p16 might be a better independent prognosticator, regardless of HPV status[262,303–306]. Others disagree based on the fact that overexpression of p16 has been found in normal tonsillar tissue[112,307].

The pairing of HPV DNA status and p16 status is one suggested diagnostic method that may help to overcome the uncertainty surrounding p16's use a surrogate biomarker and to improve its specificity. Since p16 likely represents the activity of viral oncogenes, specifically E7, the identification of HPV-positive/p16-positive cases may distinguish carcinogenic infections from transient ones, with likely HPV-positive/p16-negative results. This may actually explain why HPV-positive/p16-negative HNSCCs have a slightly worse prognosis than their dually positive counterparts[10,194,308].

Having said this, some studies have found significant heterogeneity in overall survival rates for p16-positive/HPV DNA-positive cases, despite the overall significance of a more favourable survival rate[288]. There are also suggestions that HPV E6/E7 mRNA expression may be just as if not more reliable and representative of an active oncogenic HPV infection as the combination of HPV DNA detection and p16 expression. The fact that HPV-negative HNSCC and HPV-positive but E6/E7 mRNA-negative HNSCC show similar survival curves justifies this proposition[249]. HPV mRNA use is still however mostly restricted to the laboratory as a result of its intensive labour and large costs[309] and its use in the clinic has not yet been fully validated. The same is true for p16 and HPV DNA, all of which continue to be evaluated for their comparability, validity, reliability, feasibility, and ultimate use in the clinic.

For the moment, the pairing of p16 immunohistochemistry and HPV DNA testing represents the most cost effective and valid way of using biomarkers for the benefit of patient outcomes. The current literature thus suggests that combining immunohistochemistry for p16, with an estimated cost of €25 per slide, with a test for HPV DNA, either using ISH with estimated cost €49 per slide or PCR with an estimated cost €38 per slide using SPF-10 primers, likely represents the best strategy for determination of clinically relevant HPV infection in HSCC[310]. New emerging technologies detecting HPV DNA that are more sensitive and specific than ISH and traditional PCR may come to replace them given the high-stakes-, patient outcome-nature of HPV DNA's clinical use.

1.7.5 EGFR

EGFR is the cell-surface receptor for members of the epidermal growth factor family of extracellular protein ligands[311]. EGFR's binding to its associated family of epidermal growth factors results in cellular proliferation, differentiation, and survival[125].

Across all HNSCC, EGFR expression is diverse. For the most part, EGFR overexpression is not hugely common in patients with head and neck cancer, with the promising results of targeted treatment in cell line models being less applicable in the clinical context[312]. Where it is overexpressed, it is

associated with poor survival in HNSCC in several studies[313–315], including HPV-related HNSCC[50,316]. In fact, many studies identify EGFR expression intensity as an independent predictor of response to induction chemotherapy, chemo/radiotherapy, overall survival, and disease-specific survival[10,125,317,318].

HPV-related HNSCC is characterized by low EGFR expression[14,310,317–319]. Much like in the case of their upregulation of tumour suppressor p16, low EGFR expression in another hallmark of the seemingly self-defeating nature of some HPV-related HNSCC, especially those of the oropharynx. In fact, low EGFR expression may partially explain the better response to treatment and better survival in these patients[313,314,317]. In combination with other biomarkers of HPV-related HNSCC, the relationship between low EGFR expression and better survival in these patients becomes incrementally stronger[125,317,318].

By contrast, HPV-unrelated HNSCC tend to show higher EGFR expression in tandem with an increasing pattern of expression between never-smokers, past smokers, and current smokers[318]. Given that smoking is the primary carcinogen implicated in HPV-unrelated carcinogenesis, smoking make be responsible for increased EGFR expression in these HNSCC through increased hypoxia in the pre-cancerous and cancerous tissue of smokers. This renders the tumours more aggressive, but also better targets for EGFR inhibitors including Cetuximab.

1.7.6 Cervical Squamocolumnar Junction Biomarkers

In the literature there are suggestions that cervical biomarkers might be translated to those for HPV-related HNSCC given the similarities in carcinogenesis between these tumour-types. Biomarkers specifically expressed in the cells of the cervical squamocolumnar junction (SCJ) have been discovered through laser capture microdissection by Affymetrix exon array[14]. The panel of biomarkers established (CK7, AGR2, MMP7, and GDA) share expression with over 90% of high grade squamous intraepithelial lesions and carcinomas. The biomarkers converge in a population of cuboidal epithelial cells of embryonic origin at the SCJ[320–323] which are the origin of the majority of cervical SCCs. Though tonsillar crypt epithelial cells have previously been described as cuboidal, the expression of this protein biomarker panel is not induced by HPV E6 or E7 in squamous epithelial cells in vitro, and their expression is lost if the SCJ is removed by cone biopsy or excision[324]. The SCJ-specific expression profile in cervical lesions and cancers is likely not a result of the transformation of cells but instead is simply a reflection of the already known embryonal origin of the cells.

1.7.7 Other and Combined Biomarkers

The construction of a biomarker library for HNSCC, and HPV-related HNSCC presents a valuable opportunity given the lack of any defined precancerous phase in HPV-related HNSCC especially. In the case of HPV-related HNSCC, it might be logical to assume that oral HPV infection would be indicative of eventual malignancy. However, given that the majority of HPV infections are cleared, it is unsurprising that oral fluid shows correlation but lacks sensitivity and specificity for clinically relevant HPV infections[242]. In this context, Raman spectroscopy may prove useful still, as may the potential for analysing tumour DNA in blood via circulating tumour DNA or tumour cells to facilitate the early detection and treatment surveillance of cancer[325].

Biomarkers related to treatment response, including radiation, chemotherapy, and immunotherapeutic agents have been studied with mixed results[319,326–328]. Suggested biomarker correlations are often contradictory and inconsistent, so much so that the limited success of individual markers to predict tumour behaviour has led to attempts to classify biomarker 'signatures' such as panels of RNA, miRNA, and protein expression alterations[10,14,286,329]. Some studies have investigated panels of predictive biomarkers in both HPV-related OPSCC and non HPV-related OPSCC in particular[330], but not many of these have been clinically validated[331].

Mutations of genes involved in unifying HNSCC carcinogenic pathways such as squamous differentiation (NOTCH), apoptosis (TP63, FAS/FASL), cell cycle control (CCND1), and DNA repair (TP53, TP73) may prove useful as biomarkers. Although HPV-related HNSCC tend to have far fewer genetic mutations than HPV-unrelated HNSCC, notably in these precise genes, the downstream effects of viral oncoproteins are still similar. For instance, the substitution of Pro for Arg in the transactivation domain of the p53 gene[332] may alter susceptibility of p53 to oncogenic HPV E6-mediated degradation[333]. Homozygous Arg/Arg genotype has been significantly associated with an increased risk of HPV associated cancer[333,334]. The same is true of 2 linked non-coding exon 2 polymorphisms of the p73 gene in the p53 pathway at positions 4 (Guanine to Adenine) and 14 (Cytosine to Tyrosine) which are thought to affect gene function by altering expression, and efficiency of translational initiation[335]. MDM2, also in the p53 pathway, downregulates p53 given certain SNPs at codon 72. SNP G2580T of MDM2 at nucleotide 309 in the promoter region has been shown to alter p53 expression levels with subsequent attenuation of the p53 pathway[336].

CCND1 downregulation is also strongly associated with HPV-positivity in tonsil carcinomas[337]. The gene promotes transition through the restriction point in the G1 phase of the cell cycle[338]. A polymorphism

in CCND1 exists at codon 242 within the conserved splice donor site of exon 4, modulating the splicing of CCND1 mRNA and causing two transcripts with different half-lives and functional activity to result[339]. Consequently, the reduced levels of CCND1 could facilitate the interaction of the HPV16 E7 protein with pRb, contributing then to p16 overexpression. Genetic polymorphisms of the FAS/FASL promoters have also been suggested to contribute to HPV-associated cancer risk by inducing differential apoptosis of immune cells in response to the micro-environment signals after HPV infection[340].

Some studies have identified 10 cytokeratin genes (CK19, KLK7, KLK8, KRT10, KRT75, KRTDAP, DMKN, SBSN, SPRR2A, SPRR2G, and SPRR4) that are down-regulated in HPV-positive tumours, which correlates well with the fact that HPV-positive HNSCCs are predominantly non-keratinizing, poorly differentiated, or basaloid carcinomas[310,341]. This downregulation may be attributed to the E6 and E7 oncoproteins that are known to modulate keratinocyte differentiation, suggesting that this HPV gene expression is intimately linked to differentiation states of the infected keratinocytes.

Other, and lesser-known markers, include EPS15, involved in clathrin-mediated endocytosis for HPV16 entry[342], MTMR14, an inhibitor of autophagy and HPV16 infectivity, DYRK1A, involved in stabilizing HPV16 E7 oncoprotein and increasing its transforming potential[343], and UBE3A, which interacts with HPV16 E6 oncoproteins resulting in p53 proteolysis[344]. Many of these markers, including CK19, are now considered markers of pre-malignancy and susceptibility to HPV-related OPSCC. What is becoming apparent however, is that the collective analysis of significant markers will prove very useful as determinants for prediction and prognosis when enough statistical analyses have been carried out to personalize probabilities to specific combinations of markers. Research of this type is ongoing, with the combination of EGFR expression, HPV16-positivity, and p16 expression accurately stratifying patients with the best and worst survival – low EGFR, high p16, and HPV positive tumours generally presenting the best survival statistics – and the pairing of low p53 and high Bcl-xL expression associated with poor prognosis[317,318].

1.7.8 Immune Markers

Chronic inflammation is well-known to be related to increased risk of cancer and of HNSCC in general. The role of inflammation, chronic or otherwise, in HPV infection and associated disease is complex, on the one hand preventing initial infections and clearing ongoing ones and on the other promoting the persistence and progression of related lesions[345]. Chronic inflammation has been implicated in the development of HPV-related cervical cancer[346].

Given that immune evasion is key to the pathogenesis of HPV-related cancer, pro-inflammatory cytokines may play roles in regulating the growth of HPV-infected cells[345]. That viral perspective, disease progression, and/or malignant transformation may involve escape of these cytokines could modify the efficiency of HPV clearance, especially since the etiology of HPV-induced cancer is triggered by persistent, repetitive, viral infection. Polymorphisms of a number of cytokine genes have been implicated as biomarkers in influencing susceptibility or resistance to cancers caused by HPV infection owing to their role in determining host immune response.

The IFN-γ gene plays a pivotal role in defence against viruses and intra-cellular pathogens through the induction of immune-mediate inflammatory responses[347]. IFN-γ is significantly decreased in HNSCC patients[348]. The T+874A SNP is located at the translation start site of the IFN-γ gene and coincides with a putative NF-KB binding site[349]. This SNP could play a fundamental role in the induction of constitutively high IFN-γ production. IL-10 also has a suppressive effect on cell-mediated immunity, which may be critical in the elimination of HPV-harbouring cells[350]. A number of polymorphisms exist at the -1082 position of the promoter region that play an important role in determining high, medium, and low production of IL-10[351]. The G/A SNP at position -1802 is associated with low (AA), medium (AG), and high (GG) cytokine production[351]. Interleukin-1B (IL-1B) and Tumour Necrosis Factor α (TNF-α) have also both been shown to influence cytokine expression, with the latter directly controlling HPV infection by induction of apoptosis in HPV-infected cells, as seen in cervical cells[352].

Furthermore, the immune checkpoint ligand programmed cell death 1 (PD-L1) is present in normal tonsillar crypts irrespective of HPV infection. Programmed cell death protein 1 (PD-1) expressing lymphocytes are also found in both chronic tonsillitis and HPV-related oropharyngeal tumours[353]. The PD-1:PD-L1 pathway may play a role in HNSCC, particularly in HPV-related oropharyngeal cases by encouraging HPV-related carcinogenesis in an immune-rich site[353,354].
In fact, that there is significantly decreased IFN-γ in HNSCC patients[348] may be caused by inhibition of T-cell regulation from increased expression of PD-1:PD-L1.

For these reasons, immune checkpoint blockades in the form of a monoclonal antibody that inhibits the PD-1 receptor has the potential to play a role in future therapies for HNSCC. The initiation of the anti-tumour response has already been documented in animal studies[355]. Pembrolizumab, a drug with this specific function, has demonstrated promising clinical responses in head and neck cancer patients[356,357]. Trials are ongoing to further assess the role of these antibodies in HNSCC.

1.8 Clinical Implications

1.8.1 Diagnosis, Treatment, and Prognosis of HNSCC

Prognosis and survival for HNSCC is generally poor. Approximately half of all patients with HNSCC have advanced stage disease at the time of diagnosis, with an expected 5-year survival rate between 10% to 40%[358]. This is mostly attributed to the fact that diagnosis of HNSCC is frequently delayed because symptoms for which patients will seek medical attention such as pain, dysphagia, and shortness of breath occur late in the stage of disease[359]. Despite treatments that may consist of multiple combinations of surgery, radiotherapy, and chemotherapy, overall long-term survival remains low due to persistent or recurrent disease[359]. In fact, the median overall survival for recurrent or metastatic HNC remains less than 1 year despite modern chemotherapy and targeted agents[360].

Palliative chemotherapy and cetuximab, the EGFR inhibitor, constitute the backbone of treatment for patients with recurrent and metastatic HNSCC[360]. Cetuximab in combination with chemotherapy has been shown to improve response rate from 20% to 36% in some trials though toxicity continues to be a problem[361]. Platinum chemotherapy in combination with 5-fluorouracil and cetuximab has resulted in the longest median overall survival for HNSCC[360]. Other phase III trials have been carried out on platinum doublets including cisplatin/5-FU, cisplatin/paclitaxel, and cisplatin/pemetrexed. Most of these have proven no consistent benefit over single agent therapy given that although they increase response rates and toxicity, they do not increase survival[362–365]. Patients treated with other EGFR inhibitors including Zalutumumab[366] and Panitumumab[363] also show no consistent improvements in survival. These kinds of treatments should therefore be reserved for patients who are symptomatic of the disease for whom the benefit of partial response may be worth the cost and increased treatment side effects[360]. Other drugs used as single agents for patients with low disease burden and few symptoms include docetaxel, paclitaxel, capecitabine, pemetrexed, and methotrexate[360].

Given poor performance status in most HNSCC, most studies on the topic focus on palliative care and the improvement of quality life for patients. Best supportive care, and not survival, are more often than not the focus of HNSCC late-stage treatment. It is known that palliative radiation therapy, for instance, is beneficial for treating symptomatic metastatic sites[360]. This said, mTOR inhibitors, IGF1R inhibitors, and anti-angiogenic agents are currently under investigation to monitor impact on survival. Most recently, mTOR inhibitors have been found to increase anti-tumour responses and reduce tumour growth in HNSCC[367,368].

1.8.2 Diagnosis, Treatment, and Prognosis of HPV-Related HNSCC

HPV-related HNSCC have a unique and distinct set of diagnostic, prognostic, and treatment-related characteristics. These tumours generally present with a more advanced clinical stage, with a higher nodal category[369–371], despite lower tumour extent[151,370], and have different tendencies for extracapsular spread and perineural invasion[372]. They often present with early lymph node metastases[69,373], which are sometimes confused with branchial cleft cysts[374]. However, tonsil SCCs in general are known to present with early lymph node metastases[375] and it may simply be that the anatomy of the site itself facilitates the early spread an depth of invasion[373].

Treatment utilized for most patients with high-risk, resected HNSCC, is adjuvant radiation therapy with high doses of cisplatin which appears to work well for HPV-positive tumours. 3-year overall survival rates are between 86% and 91% while 3-year recurrence free survival rates are between 82% and 84%[376,377]. Thus, despite more advanced presentation, improved survival, which is consistently higher than 30%[378], is evident in HPV-related HNSCC[373,379–381], irrespective of treatment modality[151,292,293,382–385]. The improved prognosis and response to treatment holds true for all indicators of HPV-positivity including seropositivity, mRNA, oncoprotein expression, and viral load and copy number[318]. It also remains salient in the case of HPV-positive HNSCC biomarkers, including p16, p53, EGFR, and Bcl-xL[289,318].

These response, survival, and prognosis statistics remain significant after adjusting for confounding factors including age, sex, smoking-status, and performance status. Nonetheless, HPV-positive HNSCC patients tend to be younger, and are less likely to have had significant exposure to tobacco, marijuana, alcohol, diabetes, chronic obstructive pulmonary disease, anxiety disorders, and major depression[10,95,127,376,386,387]. Immune function in non-diabetics and those unaffected by pulmonary ailments is also not symptomatically malfunctioning, and/or constantly under stress[376]. In all, this indicates that the most at-risk populations for HPV-related HNSCC are those with the best immune ability to combat it.

Furthermore, the viral nature of the origins of HPV-positive tumours, accompanied by their expression of viral oncoproteins and related HPV-positive tumour antigens at sites of huge immune and lymphatic activity likely attracts a more aggressive and specific immune response. This is especially relevant given that immune evasion allows HPV to infect cells, but does not seem to play a role after viral integration. HPV-positive tumours often show strong T-cell infiltration, especially CD8+ and CD3+ infiltration, compared to HPV-negative tumours[233,388]. Genes involved in cell-based immunity are also

more highly expressed in HPV-related OPSCC in particular[389]. Most HNSCC do not overexpress CD200, a protein regulating myeloid cell activity and inhibitory signalling for macrophage lineage[390]. HPV-positive tumours also contain a lower percentage of cells with CD44 and CD98 expression[391]. Three serological elements have additionally been shown to be correlated with HPV-related OPSCC: E1, NE2, and E6 antibody positivity[289]. These immune biomarkers are all in turn related to increased overall survival, better prognosis, low regional recurrences, and increased response to chemoradiation in vivo in HNSCC[388,389,391].

Genetically, given that HPV-related tumours have fewer genotype alterations than negative ones, they have increased sensitivity to DNA-damaging agents[392]. HPV-positive tumours often have TP53 mutations that do not confer chemo- and radio-resistance[214,393]. It is also well-established that the overexpression of p16 dramatically affects radiation sensitivity in HNSCC cells since it impairs the recruitment of RAD51 to the site of DNA damage by down-regulating cyclin D1 protein expression[376] (Figure 1.10). Thus, a direct pathway between the E7 HPV oncogene and increased sensitivity to radiation therapy emerges: Increased p16 expression correlates with decreased cyclin D1 expression; the resulting cell cycle dysregulation and impaired DNA repair correlates with higher cellular radiosensitivity; this results in an accumulation of cells in the G2/M phase, and a resulting increased responsiveness to treatment.

1.8.3 De-Escalation and Targeted Treatment in HPV-Related HNSCC

The very significant better responses to treatment amongst HPV-related HNSCC patients is extremely encouraging with respect to possible treatment de-escalation[151,386,394–397]. As with all therapies, there are many toxic, long-term, debilitating side effects associated with HNSCC treatments that affect morbidity and quality of life. These include but are not limited to trouble swallowing and breathing, difficulty speaking, and third degree burns as a result of radiation. The possibility of reducing the severity of these symptoms is something to be capitalized on if the evidence confirms survival is not compromised.

The population of mostly younger patients with improved prognosis is an encouraging population for transoral resection (TOR), the removal of the tumour completely with sound oncological margins, especially with emerging robotic techniques and laser microsurgery. Robotic TOR is particularly attractive as it often results in better functional outcomes for patients than other surgical approaches and radiotherapy[398–400]. TOR without adjuvant therapy has been shown to be adequate treatment for HPV-related HNSCC, with anywhere between 48% to 74% of patients not requiring chemotherapy[401–

[403]. There appears to be no significant difference in the success of TOR alone in local control of disease by HPV status for early stage OPSCC patients[404], but the landscape of outcomes for later stage patients by HPV status using TORs alone is less clear, especially given the newly updated 8th edition AJCC staging criteria. The better survival of HPV-related tumours regardless of stage suggests that TOR has a better chance of operating alone in these OPSCC patients, perhaps with de-escalated adjuvant radiotherapy.

There are also many Phase II trials ongoing with several Phase III trials underway examining the possible replacement of cisplatin with cetuximab, less aggressive radiation/chemoradiation regimens, and the removal of chemotherapy altogether. These are summarized in several extensive reviews[14,386] and can be found at www.clinicaltrials.gov but an overview is found in Table 1.5 below.

Table 1.5 Summary of key completed and ongoing trials on treatment and de-escalation for HPV-related HNSCC.

Trial	Phase	Inclusion	Arm 1	Arm 2	Outcomes
RTOG 1016	III	P16-positive locally advanced OPSCC	Radiation and concurrent chemotherapy	Radiation and concurrent cetuximab	Survival, toxicity, locoregional recurrence, and quality of life
ECOG E1308	II	Stage III-IVa HPV-positive OPSCC	Complete response to induction chemotherapy and reduced dose radiation with concurrent cetuximab	Incomplete response to induction chemotherapy and standard dose radiation with concurrent cetuximab	Survival, toxicity, response, quality of life, and biomarker correlation
De-ESCALATE HPV	III	Stage III-IVa HPV-positive OPSCC	Cetuximab and concurrent radiotherapy	Standard concurrent cisplatin chemoradiotherapy	Morbidity, quality of life, cost, survival, and recurrence
QUARTERBACK	III	Locally advanced HPV16 positive OPSCC, unknown primary SCC, or nasopharyngeal SCC showing complete or partial response to induction therapy	Reduced dose radiation with cetuximab and chemotherapy	Standard dose radiation with chemotherapy	Survival, locoregional control, toxicity, and biomarker correlation
LCCC 1120	II	HPV-positive and/or p16-positive low-risk OPSCC	Decreased dose of radiation and chemotherapy	Standard radiation and chemotherapy	Pathological response rate, locoregional control, survival and quality of life
NCT01221753	II	Locally advanced HPV-positive OPSCC	Docetaxel/Cisplatin/5-Fluorouracil induction chemotherapy followed by concurrent chemoradiation using modified radiation dose	N/A	Locoregional control, survival, toxicity
SIRS	II	Early to mid-stage HPV-positive OPSCC who receive transoral robotic surgery plus a neck dissection, where clinically indicated	Observation only	Low dose postoperative radiation only **Arm 3**: Chemoradiation	Rates of locoregional control, overall survival, and use of salvage chemoradiation in the observation group
TROG 12.01	III	HPV-positive OPSCC	Radiation and cetuximab	Radiation and cisplatin	Symptoms severity, swallowing, quality of life, toxicity, survival, locoregional recurrence
ADEPT	III	P16-positive OPSCC that has underground TOR with negative margins	Postoperative radiation alone	Postoperative radiation with cisplatin	Survival, locoregional control, toxicity, quality of life
NCT01088802	I/II	HPV-positive T1-3 OPSCC	De-escalation radiation from 70 Gy to 63 Gy with concurrent chemotherapy	De-escalated radiation from 58.1 Gy to 50.75 with concurrent chemotherapy	Toxicity, quality of life and adverse events
ECOG E3311	II	Stage III-IVa HPV-positive OPSCC after transoral surgery and neck dissection with negative margins, no extracapsular spread and less than 4 lymph nodes involved	TOR with standard radiation	TOR with low-dose radiation	Survival, surgical complications, toxicity, and swallowing

Most of these trials understandably focus on originally good responders as the risk of more aggressive progression is smaller in these patients. Low induction of chemotherapy followed by decreased radiation doses or volumes in these patients has so far shown extremely good survival rates[386], but these results are short-term and longer-time follow-up is necessary to make sure that these lower doses are actually controlling and suppressing tumour growth effectively. Results from chemoradiation with decreased dose of radiation and chemotherapy have also shown promising potential for radiation being sufficient for disease control in HPV-driven OPSCC patients[386]. Indeed, the removal of chemotherapy altogether may be possible in patients with stage III/IV disease and non-smokers.

With respect to EGFR, there have been conflicting reports on the benefit of replacing cisplatin with cetuximab, an EGFR inhibitor, in HPV-related OPSCC specifically. Some suggest that it improves survival for OPSCC patients[405], while others including the RTOG o522 and SPECTRUM trials disagree[378,406]. Given that low EGFR expression is a biomarker for HPV-positive tumours, population-wide success of cetuximab should be expectedly low, suggesting a targeted use of cetuximab for HPV-positive HNSCC with high EGFR expression and the majority of non-HPV related HNSCC with high EGFR expression. In fact, the most recent evidence from the De-ESCALATE trials shows no improvements in toxicity and expectedly and significantly worse outcomes amongst patients treated with cetuximab instead of cisplatin[407–409]. For this reason, pending results expected after 2020 of the ADEPT, SIRS, and ECOG E3311 (Table 1.5) trials are highly anticipated to determine whether or not de-escalation in the form of surgery with adjuvant low-dose radiation, or perhaps no radiation at all for some patients, is the best form of de-escalated treatment for HPV-related OPSCC.

Potential for de-escalation in HPV-related HNSCC must be approached very cautiously, however. Treatment de-escalation is really only conceivable in “low-risk” patients, in other words, not only those with HPV-positive disease in the oropharynx and perhaps more specifically in the tonsil and base of tongue, but also those who are non-smokers, young, and have little significant family history. In fact, the parameters that determine “low-risk” are not defined in the research literature and are by consequence not applied homogenously in the clinical context.

Furthermore, that there is no standard by which a tumour is definitively considered “HPV-positive” makes this identification of “low-risk” patients more difficult. HPV DNA is not sufficient to determine whether or not the virus is causal in the HNSCC given that the virus may be simply transient. That

many trials use p16 as a surrogate biomarker for HPV infection is thus logical. However, as previously indicated, p16 positivity is not a guarantee of an active HPV infection, and the pairing of p16 positive and HR HPV DNA detection shows better surviving patients than p16 and HPV alone. Even paired with surrogate biomarker p16, current HPV DNA detecting technologies used in the clinic are not as sensitive as those being developed in the laboratory. Thus, some patients with HPV-related HNSCC might be missed and relegated to more intense treatment schemes if their tumours only present as p16 positive due to shortcomings in current HPV molecular testing. More importantly, some patients could be under-treated on the basis of HPV-positivity alone.

Compromising patient safety when the parameters identifying "low-risk" patients are uncertain and not optimized is unacceptable. It has also been noted that the benefit/risk balance between decreased toxicity and cancer control is potentially narrow[386]. It is understandable that on this basis, patients and physicians are reluctant to risk the possibility of a worse outcomes in exchange for the possibility of improved morbidity and less severe side effects. One of the only studies assessing patient attitudes to de-escalation in its current state showed that nearly 70% of patients were not willing to risk a 5% or less drop in survival likelihood to switch from chemoradiation to radiation alone[410].

Given that patient-led rather than paternalistic treatment is becoming widespread in the clinic, it is of the utmost importance that patients have access to the most detailed and certain research regarding survival and side-effects. The need for more definitive "low-risk" and "high-risk" categorizations and definitive outcome data based on a scale of both HPV and p16 status is therefore urgent if patients are to not only survive disease, but thrive in its aftermath. That the AJCC has updated their 8th edition[114] staging guidelines to reflect the differential prognosis in HPV-related HNSCC is significant. It is the first step in adjusting reliance on severe chemotherapy and radiation treatments without compromising patient safety, but there is still justification for extending its use of only p16 to determine HPV status to both p16 and HR HPV.

For now however, de-escalation is not a mandated part of HPV-related HNSCC treatment. The current treatment of HPV-positive HNSCC of adjuvant radiation therapy with high doses of cisplatin is reliable and safe and is the best available way thus far to preserve survival and minimize toxicity[376,377]. Furthermore, around 20% of HPV positive patients do not show an improved prognosis or respond to treatment[411]. This may be due to the aforementioned synergistic involvement of tobacco use, which has been identified as an additional risk factor in stratifying outcomes in patients with HPV-related

cases, which was recently confirmed in clinical trials. Patients with large primary and lymph-node tumours have also been described as high-risk groups that might have poor outcomes[411].

1.8.4 Additional and Personalized Treatments for HPV-Related HNSCC

There are several novel therapies that may be useful for treating HPV-related HNSCC[376]. Everolimus, an inhibitor of the mammalian target of rapamycin, Sorafenib, and Sunitinib, two multi-tyrosine inhibitors, are three of the most promising. All three have shown anti-tumour effects against various tumour entities with significantly higher sensitivity to the drugs in HPV-related HNSCC compared to HPV-unrelated HNSCC. Each are administered orally with moderate side effects, high compliance, and decreased hospitalization rates[390].

Further, as the PD-1:PD-L1 pathway of immune evasion may play a role in HPV-related and non-HPV related HNSCCs, immunotherapy with antibodies, such as the anti-PD-1 inhibitors already approved for use in advanced melanoma cases, could also prove useful treatments in primary or refractory cases[14].

The tumour microenvironment may also have a role to play in regulating unique response to treatment and determining prognosis. Modulation of intercellular signalling in the tumour microenvironment in particular could be a valid and robust therapeutic modality[412]. It is well recognised for instance that high expression of the VEGF-A factor, which supports tumour vascularisation, is linked to poor prognosis[413]. Combination of anti-VEGF-A humanised monoclonal antibody with anti-EGF receptor antibody can be used for the treatment of recurrent and metastatic HNSCC[413]. Cytokines as prominent mediators of intercellular crosstalk include IL-2, IL-6, IL-u, and IFN α/γ have been subjected to clinical studies but with limited success[414]. Targeting PD-1 and PD-L1 both on natural killer cells and on T lymphocytes represents an excellent strategy for many sensitive tumours. Targeting numerous growth factors, cytokines, and chemokines produced by members of the extra-cellular matrix (including cancer-associated fibroblasts) may also be a promising therapeutic strategy for these cancers, though few HNSCC-specific experiments have been conducted. Those that have show no consistently conclusive results. Galectin-1, for instance, produced by cancer-associated fibroblasts, induces apoptosis in T lymphocytes and has an immunosuppressive effect[415]. However, galectin-1 also diminishes resistance of cancer cells to anoikis, indicating a pleiotropic effect of the factor in cancer biology[416].

1.8.5 The Use of HPV Vaccines to Prevent and Treat HNSCC

Epidemiological research is essential if the use of prophylactic HPV vaccines to prevent HNSCC is to be determined. There are three FDA-approved prophylactic HPV vaccines that have primarily been introduced for the prevention of cervical cancer in women. These are summarized, along with the genotypes of the virus against which they protect, in Table 1.6. All three of the vaccines are approved for a dosing schedule of 3 doses over 6 months, although more recent trials suggest that shorter dosing schedules are just as effective[417,418]. All of the vaccines are currently based on virus-like particles of the HPV L1 protein. In countries where the vaccines are available, girls and women between the ages of 9 to 26 can avail of it.

Table 1.6 HPV vaccines currently available and the genotypes against which they protect, respectively.

Vaccine Type	Genotypes Protected Against
Cervarix (bi-valent)	16, 18
Gardasil (quadri-valent)	6, 11, 16, 18
Gardasil-9 (nona-valent)	6, 11, 16, 18, 31, 33, 45, 52, 58

There is currently very little data evaluating the impact of the HPV vaccine on oral HPV infection or HPV-related HNSCC in particular[419]. This is mainly due to the fact that epidemiological data regarding the relationship between the virus and HNSCC is still emerging; the natural history of HPV in the oral cavity is to be confirmed; and as a consequence, clinical trials have not been designed for the study of oral HPV infection in the context of vaccines. Recent studies suggest an effective protection against the virus in the oral cavity using the vaccine in both men and women[420]. Furthermore, in the CVT trial on the Cervarix vaccine, oral rinses were taken at the 4-year follow-up mark after vaccination[421]. Overall and HPV16/HPV18 prevalence was lower in the vaccinated group compared to the control group, with an estimated efficacy of 93.3% for HPV16/18[421]. While this data is promising, there was no original oral rinse taken, and though HNSCC is most prevalent in men, this study only included women given its cervical origins.

There is however ample evidence from animal studies that vaccination protects against oral infection and the development of oral lesions from canine oral papillomavirus[422]. Predictive modelling studies also suggest that with a 50% vaccination uptake and 50% vaccine efficacy, the vaccination of young boys for the prevention of HPV-related OPSCC would be cost-effective[419,423,424]. Nonetheless, the effectiveness of any HPV vaccine in protecting against HNSCC would not be able to be determined

until 2060 should widespread vaccination be introduced immediately given the average age of onset of HPV-related HNSCC in the fifth decade of life[419].

This said, emerging epidemiological data plays an enormous role in predicting the potential impact of the vaccine on HNSCC. HPV16 appears to be the main culprit in HPV-related HNSCC, accounting for over 90% of HPV prevalence in these cancers[28,34,71,105,425]. All other genotypes, including the cervical SCC-implicated HPV18, trail far behind in HNSCC prevalence. Despite accounting for few HPV-positive cases, the rest of the HR HPV genotypes detected in HNSCC align with those in cervical cancer. Thus, though modelling the impact of the vaccine based on this genotyping data has yet to be published, preliminary logic would suggest that the currently available quadri- and/or nona-valent vaccines should protect against persistent HPV infection in the oral cavity given the systemic rather than local protection vaccines provide.

On the basis of this logic and epidemiological recommendations, the HPV vaccine, mostly the quadri- and nona-valent Gardasil vaccines are already offered to both boys and girls in about 20 countries including Australia, Austria, Brazil, Canada, Germany, Israel, Italy, New Zealand, Norway, and the USA. In the United Kingdom, boys are to be offered the vaccine in 2019 where it was previously only available to some men who have sex with men[426]. In these nations, the quadri-valent Gardasil vaccine is generally administered in boys and men between the ages of 9 and 26, with the nona-valent version recommended between the ages of 9 and 15. In Ireland, HIQA recommended the inclusion of boys into the national vaccination scheme in 2018, and boys will be included in the program from September 2019, which will also see the introduction of the nona-valent vaccine[427–431].

Vaccines may also have a role to play in therapeutic treatment of HNSCC. Therapeutic vaccines can be classified in to DNA, mRNA, peptide, viral/bacterial vector-based, and cellular vaccines[432]. The majority of these target the E6 and E7 oncoprotein antigens. The most effective of these appear to be DNA and mRNA vaccines which have been shown to induce tumour regression and improved T-cell mediated immune response in HPV-related HNSCC. For instance, the Listena-based HPV vaccine ADX-11-001 is currently being trialled in Phase II to determine if it induces curating and tumour infiltrating specific T-cell antigens in HPV16-positive OPSCC patients[433]. Other recent phase I and II clinical trials have also reported that vaccines in combination with chemotherapy may boost effectiveness[378,434]. The literature on these trials is expanding at such a rate that summaries of those completed and ongoing trials have already been published[432].

1.8.6 Future Directions and Screening Tools for HNSCC

There is no single standardized treatment for HNSCC, but before recommended management strategies are altered results from randomised controlled trials are needed to assess the efficacy of the different treatment modalities available for both HPV-positive and HPV-negative HNSCC[435], although recruitment of sufficient numbers remains difficult[372]. Given the current literature, it is likely that no single course of treatment will guarantee better prognosis for any one patient. Instead, the idea of personalized medicine determining best treatment options, including the specific analysis of tumour biomarkers independently and as groups, will predominate the future of HNSCC treatment.

Nonetheless, preventative medicine in the form of HPV-specific immune responses by prophylactic vaccination with recombinant HPV virus-like particles is likely essential to successful prevention of persistent HPV infection and its subsequent consequences. As such, bi-, quadri-, and nona-valent vaccines are now widely available and have shown efficacy in prevention of anal, cervical, vaginal, and vulvar pre-cancers in unexposed individuals[142,436,437].

It is not only vaccination that will lead to the effective prevention of HNSCC however. Other cancer types, specifically cervical cancer, rely on pairing vaccination with sensitive screening tools and programs for early detection. In Ireland, Cervical Check is responsible for the national screening program for pre-cancerous and cancerous cervical lesions[438]. HPV's causal role in almost all cervical cancers has already transformed other nations' screening programs to HPV-based ones[439]. The Irish government and the Health Service Executive (HSE) plan to introduce HPV-based screening imminently.

No such comparative screening tools exist for HNC in general, let alone those that are HPV-related. Most HNC and HNSCC are detected only when they become symptomatic or by external physical presentation during routine doctor visits, something that likely explains the late stage of presentation for all HNC[359,373].

HNSCC and HPV-related HNSCC present a particular challenge for screening tools given the difficult-to-access areas of the head and neck in which they originate[283,284]. It is problematic, for instance, to design a sampling tool to access the full scope of deep tonsillar crypts and the base of tongue down to the larynx. Furthermore, the role that HPV might play in this screening process is uncertain. It is well-documented that though HPV16 is the most common genotype detected in the oral cavity, the genotype distribution of oral HPV prevalence does not exactly correlate to that in HPV-related

HNSCC[176,177]. However, given that HPV16 is the least easily cleared by the immune system in the oral cavity[176,177], smear-like sampling and detection of HPV16 may be an extremely valuable predictor of eventual HNSCC. At the very least, it could be the basis for closer monitoring of patients with HPV16 infection detected from oropharyngeal sampling specifically. Though suggestions for screening have been previously proposed based on logical extrapolation from the cervical context[440], further clarity on the natural history of HPV16 infection in the oropharynx in particular is needed if a monitoring-based screening system is to be targeted at infections that are likely to become carcinogenic.

Public health schemes including vaccination and screening, much like in the cervical case, will likely underpin the successful eradication of HNSCC and HPV-related HNSCC in particular. The development of screening tools is therefore essential, along with the validation and further supporting evidence for the usefulness of the HPV vaccines for those genotypes most implicated in HNSCC.

1.9 Classification of HNSCC

The previous review determines that HPV-related HNSCC are a unique set of HNSCC with characteristics providing for a new focus for prevention and treatment in patients. The generalized differentiators of HPV-related HNSCC from other HNSCCs are summarized in Table 1.7 below.

Table 1.7 Generalized patient and clinical characteristics of HPV-related and other HNSCCs.

Characteristic	HPV-related HNSCC	Other HNSCCs
Physiology	MALT tissue, more often than not in oropharynx	All regions, less prevalent in oropharynx
Epidemiology	Increasing incidence in Europe and North America since the 1990s	Decreasing incidence in Europe and North America in tandem with smoking trends
Carcinogens	HPV	Smoking, Alcohol
Histology	Lobular growth, infiltrating lymphocytes, non-keratinizing, poorly differentiated unless misdiagnosed based on cellular immaturity	Moderately differentiated, keratinizing, few infiltrating lymphocytes
Risk Factors	Young, white, male, higher socioeconomic status, developed world, HR HPV infection	Older, smoker, heavy drinker, lower socioeconomic status, developing world
Biomarkers	HPV-DNA, HPV-E6/7 mRNA, HPV serology, p16 upregulation, EGFR downregulation, immune markers	EGFR upregulation, p16 downregulation, mutations associated with smoking- and alcohol-related cancers
Clinical Implications	Better prognosis and outcome than other HNSCC, de-escalation of treatment promising, still constructing biomarker profiles, heterogeneity still apparent	Poorer prognosis than HPV-related HNSCC, treatment more akin to smoking-related lung cancer, cetuximab more successful given EGFR, heterogeneity still apparent

1.10 The Justification for and Value of the "Epidemiology of HPV Infection in Oropharyngeal, Oral Cavity, and Laryngeal Cancer in Ireland" Study

CERVIVA is a multi-investigator collaboration led by Trinity College Dublin, encompassing researchers at several national and international academic institutions, health agencies, and commercial diagnostic and biotechnology companies. The purpose of the consortium is to provide relevant and necessary research evidence to support health policy and health services in the area of HPV-associated diseases. The "Epidemiology of HPV Infection in Oropharyngeal, Oral Cavity, and Laryngeal Cancer in

Ireland" (ECHO) study is being conducted within the CERVIVA consortium on the basis of the differential characteristics of HPV-related HNSCC emerging in the literature (Table 1.7).

Currently, the Irish epidemiological data regarding HPV's role in HNC does not exist in the literature. The foregoing analysis clearly defines why establishing the HPV prevalence, HPV genotype distribution, time-trends, risk factors, survival, and treatment data of HNSCC is crucial to creating the best preventative and curative mechanisms to eradicate the disease. The World Health Organization (WHO) is also currently conducting a worldwide meta-analysis on HPV-related HNSCC called the HPV-AHEAD study[441]. HPV-AHEAD uses standardized sectioning, extraction, and testing methods for which the ECHO study will provide the first ever Irish data.

The homogenous nature of the methodology used in the ECHO study as a part of the larger HPV-AHEAD study is extremely valuable given the extent of heterogeneity of technologies and principles used to test for HPV throughout the literature. The consequences of varying detection methods and differing definitions for what constitutes an 'HPV-positive' case for the validity, feasibility, and reliability of results are significant (see Chapter 2). In fact, not only does the ECHO study provide the opportunity to review the currently available technologies for detecting DNA (see Chapter 2), but the pilot study for the ECHO project provides an additional chance to assess the validity of two different HPV DNA-detecting methodologies. Currently, only a few comparisons between technologies optimized for cervical samples have been carried out[442,443], with no such studies existing for HNCs.

Finally, the size of the intended ECHO study makes it possible to evaluate and make constructive suggestions regarding the manner in which Irish clinical research is carried out. The significance of this lies in the impact that the current systems have on the patient and academic impact of Irish research. No appraisal of this nature has yet been conducted. Thus, the essential nature of the ECHO study for the current gaps in the literature is evident.

1.11 Aims

The aims of the ECHO study are three-tiered. The first tier is to conduct a population-based investigation of the epidemiology of HPV infection in oropharyngeal, oral cavity, and laryngeal cancer in Ireland. More specifically, the ECHO study intends to:

1 estimate the prevalence of HPV DNA positivity in archival tumour specimens from patients diagnosed with oropharyngeal, oral cavity, and laryngeal SCC in Ireland in the period between 1994 and 2013.
2 describe the genotype distribution in HPV positive tumours in this population.
3 estimate the raw incidence of HPV-positive and HPV-negative oropharyngeal, oral cavity, and laryngeal SCC.
4 identify patient (e.g. sex, age at diagnosis, smoking status) and clinical (e.g. stage, grade) factors associated with HPV-positivity for oropharyngeal, oral cavity, and laryngeal SCC.
5 compare treatment received and survival in patients with HPV-positive and HPV-negative tumours and identify significant predictors of survival.

The achievement of the above mentioned aims gives rise to the opportunity to assess the comparability of and implications of this for two different HPV DNA detecting methodologies/technologies. Thus, the pilot study for the ECHO project creates a second tier of aims. These are:

1 To compare HPV DNA detection between SPF10 PCR Gel Electrophoresis and Multiplex PCR xMAP® Luminex laser-based technologies.
2 To assess the effect of altering sterility protocols on the validity of results.
3 To determine which method is best suited for the epidemiological endeavours of the ECHO study as a whole.

Finally, the execution of the ECHO study allows for the evaluation of the current logistical manner in which Irish medical research is conducted on the basis of data collected whilst acquiring samples for the ECHO study. In particular, the study expects:

1 To establish the number of procedures and parties necessary to acquire pathology reports and retrieve sample FFPE blocks from hospital sites.
2 To determine the lengths of time (days) taken to organize and execute the review of pathology reports and the retrieval of FFPE blocks from hospital sites.
3 To pinpoint the steps of the pathology report review and FFPE block retrieval process contributing most to the attrition of cases from the study.
4 To identify the reasons for which attrition occurred at each step of the pathology report review and FFPE block retrieval process.

1.12 Hypothesis

HPV is unequivocally involved in the carcinogenesis of a significant portion of HNSCC. This is particularly true in North America and Europe, where the incidence of HPV-unrelated HNSCC has been decreasing tandem with trends in smoking and alcohol consumption[31,71,82]. The risk factors and positive clinical implications of HPV-related HNSCC summarized in Table 1.7 are well-established in Scandinavian[28,34], American[71], and a limited number of continental European[75–77] countries. Nonetheless, not all developed nations see precisely the same epidemiological patterns in HNSCC, the United Kingdom being a prime example of this due to more mixed smoking, alcohol, and sexual behaviours amongst younger people to this day[78]. That said, the United Kingdom has still seen increasing HPV-related HNSCC incidence trends.

Whether Ireland fits or diverges from the epidemiological landscape of HNSCC found in developed world has yet to be determined. Demographically and with respect to population size, Ireland is similar to Scandinavian countries, as are its cultural practices with reference to smoking[444] and sexual behaviours. That said, alcohol consumption is much higher per capita in Ireland and is not dissimilar to that of the United Kingdom[445]. It is therefore expected that HPV DNA positive cases will be prevalent in the Irish oropharyngeal, oral cavity, and laryngeal cancer population, from anywhere between 18% and 40% based on generalized prevalence data from across Europe.

Broken down by anatomical subsite, there is nothing to suggest that Irish people would be any less anatomically susceptible to persistent HPV infection in the oropharyngeal subsites including the tonsil and the base of tongue. There should be some laryngeal and oral cavity involvement given the vulnerability of MALT tissue, but the prevalence of HPV DNA in SCC of these regions is likely to be less significant given their stronger relationship to smoking and alcohol behaviours. Oropharyngeal cases would be expected to represent the majority of HPV DNA positive cases. HPV16 should also remain the most HR genotype for HNSCC with HPV18 and other HR and LR types being very poorly represented.

It would be expected that younger age, non- and/or ex-smokers, and higher socioeconomic status Irish patients would have a higher risk of HPV DNA-positive tumours. This said, smoking rates amongst young people in Ireland have not dropped as drastically as in other countries[444]. In fact, younger women have seen increasing smoking rate in the last 10 years[444]. The raw incidence numbers for HPV-related and HPV-unrelated cases may therefore not have increased and decreased in the 1994 to 2013

time-period as distinctly as in, for instance, the United States or Sweden. It is anticipated however that HPV positive OPSCC in particular will have better survival than HPV negative cancer as a whole, with these cancers potentially being treated quite severely due to the later stage at which they generally present as seen in the literature. HPV positivity would also be expected to be a significant predictor of better survival amongst oropharyngeal, oral cavity, and laryngeal SCC, and OPSCC especially.

Regarding the second tier of aims, there are no major comparative studies in the literature regarding HPV-detecting technologies which could be used to predict the outcomes of the ECHO pilot study. Some comparative studies do exist using cervical cases[442,443], but none involve xMAP® technologies. The reported differential sensitivities and specificities of the SPF10 PCR Gel Electrophoresis[446,447] and the Multiplex PCR xMAP® Luminex[448,449] alone would suggest that there may be differences in HPV DNA prevalence detected in the same population. Even more variation might result from altering the sterility protocols for two platforms with unequal sensitivities. This comparison will delineate a more valid technology for the ECHO study as a whole and have implications for the current standards of HPV detection in research and in the clinic.

Lastly, with respect to the third tier of aims, no studies exist in the literature regarding the evaluation of the logistical steps and sample attrition involved in conducting nationwide research in Ireland. The current decentralized state and public-private nature of the management of patients and samples in Ireland would insinuate that organizing sample acquisition for studies involving multiple hospital sites may cause delays. There may also be disconnects between documentation at different public hospitals, private storage sites, and national registries given the sheer number that are involved in the storing of patient data.

The Irish data regarding HPV infection in oropharyngeal, oral cavity, and laryngeal cancer is currently absent from the literature. The context of the HPV-AHEAD study provides the perfect opportunity to fill this gap by using a globally standardized and hyper-sensitive DNA-detecting technology. Obtaining data that is generated using precisely the same WHO protocols from around the world is powerful given the currently heterogeneous landscape of HPV-detecting technologies. This is particularly true given the already well-evidenced notion that HPV-related HNSCC have better survival in the clinic, but more importantly that they might be almost entirely preventable using the HPV vaccine and appropriately developed screening tools. Only when standardized, comparable, and valid data is generated from across the globe regarding these virally-caused cancers will patients in the clinic ultimately see the benefits of discovering the strangely self-destructive nature of their HPV-related

disease. Furthermore, it is only then that the already established power of preventative medicine will come into full force to eradicate not only an imminently treatable, but avoidable class of disease.

References

1. Winslow, T. Scientific and Medical Illustrations - Head and Neck Cancer Regions. *Terese Winslow LLC* (2012). Available at: http://www.teresewinslow.com/portshow.asp?nxt=13&sid=492DA908-CB0A-433F-BE09-0C3E5F6FC4B4&portfolioid=%7B4B56C61F-9C24-47C6-9F4D-9444E1D75BA2%7D. (Accessed: 7th January 2017)
2. Schiff, B. A. Oropharyngeal Squamous Cell Carcinoma - Ear, Nose, and Throat Disorders. *MSD Manual Professional Edition* (2016). Available at: http://www.msdmanuals.com/professional/ear,-nose,-and-throat-disorders/tumors-of-the-head-and-neck/oropharyngeal-squamous-cell-carcinoma. (Accessed: 7th January 2017)
3. Ferlay, J. *et al.* Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer* **127**, 2893–2917 (2010).
4. National Cancer Registry of Ireland. *Cancer Trends: Cancers of the Head and Neck*. (2011).
5. Pai, S. I. & Westra, W. H. Molecular pathology of head and neck cancer: implications for diagnosis, prognosis, and treatment. *Annu. Rev. Pathol.* **4**, 49–70 (2009).
6. Jemal, A. *et al.* Global Cancer Statistics. *CA Cancer J Clin* **61**, 69–90 (2011).
7. National Cancer Registry of Ireland/Northern Ireland Cancer Registry. *All Ireland Cancer Atlas 1995-2007*. (2011).
8. Rousseau, A. & Badoual, C. Head and Neck: Squamous Cell Carcinoma: An Overview. *Atlas of Genetics and Cytogenetics in Oncology and Haematology* (2011). Available at: http://atlasgeneticsoncology.org/Tumors/HeadNeckSCCID5078.html. (Accessed: 2nd February 2017)
9. Barnes, L., Eveson, J. W., Reichart, P. & Sidransky, D. *World Health Organization Classification of Tumours: Head and Neck Tumours*. (2005).
10. Woods, R. *et al.* Role of human papillomavirus in oropharyngeal squamous cell carcinoma: A review. *World J. Clin. cases* **2**, 172–93 (2014).
11. Gubanova, E. *et al.* Downregulation of SMG-1 in HPV-Positive Head and Neck Squamous Cell Carcinoma Due to Promoter Hypermethylation Correlates with Improved Survival. *Clin. Cancer Res.* **18**, 1257–1267 (2012).
12. National Cancer Registry of Ireland. *Cancer Trends: Cancers of the Head and Neck*. (2014).

13. Pfister, D. G. *et al.* Head and neck cancers. *J. Natl. Compr. Canc. Netw.* **9**, 596–650 (2011).
14. Woods, R. Human Papillomavirus-related Oropharyngeal Squamous Cell Carcinoma – Prevalence and Incidence in a Defined Population and Analysis of the Expression of Specific Cellular Biomarkers. (2016).
15. Kheirandish-Gozal, L. & Gozal, D. *Sleep Disordered Breathing in Children: A Comprehensive Clinical Guide to Evaluation and Treatment*. *Respiratory Medicine* (Humana Press, 2012). doi:10.3342/kjorl-hns.2013.56.6.391
16. Palatine Tonsil - The Components of Waldeyer's Tonsillar Ring. *Anatomy Questions and Answers* (2019). Available at: http://www.anatomyqa.com/anatomy/important-question-and-answers-on-head-and-neck-anatomy/palatine-tonsil/. (Accessed: 21st May 2019)
17. American Joint Committee on Cancer & American Cancer Society. Head and Neck Sites. in *AJCC Cancer Staging Manual* 24–59 (Lippincott-Raven Publishers, 1997).
18. Brandtzaeg, P., Kiyono, H., Pabst, R. & Russell, M. W. Terminology: nomenclature of mucosa-associated lymphoid tissue. *Mucosal Immunol.* **1**, 31–37 (2008).
19. McDermott, M. R. & Bienenstock, J. Evidence for a common mucosal immunologic system. I. Migration of B immunoblasts into intestinal, respiratory, and genital tissues. *J. Immunol.* **122**, 1892–8 (1979).
20. Janeway, C., Travers, P., Walport, M. & Shlomchik, M. *Immunobiology: The Immune System in Health and Disease*. (2001).
21. Howard, J. D. & Chung, C. H. Biology of Human Papillomavirus–Related Oropharyngeal Cancer. *Semin. Radiat. Oncol.* **22**, 187 (2012).
22. CTI Reviews. *Essentials of Anatomy and Physiology: Medicine, Internal Medicine.* (CTI Reviews, 2016).
23. Auluck, A. *et al.* Trends in oropharyngeal and oral cavity cancer incidence of human papillomavirus (HPV)-related and HPV-unrelated sites in a multicultural population. *Cancer* NA-NA (2010). doi:10.1002/cncr.25087
24. Bledsoe, T. J. *et al.* Oropharyngeal squamous cell carcinoma with known human papillomavirus status treated with definitive chemoradiotherapy: patterns of failure and toxicity outcomes. *Radiat. Oncol.* **8**, 174 (2013).
25. Ernster, J. A. *et al.* Rising Incidence of Oropharyngeal Cancer and the Role of Oncogenic Human Papilloma Virus. *Laryngoscope* **117**, 2115–2128 (2007).
26. Hocking, J. S. *et al.* Head and neck cancer in Australia between 1982 and 2005 show increasing incidence of potentially HPV-associated oropharyngeal cancers. *Br. J. Cancer* **104**, 886–891 (2011).

27. Shiboski, C. H., Schmidt, B. L. & Jordan, R. C. K. Tongue and Tonsil Carcinoma. *Cancer* **103**, 1843–1849 (2005).
28. Näsman, A. *et al.* Incidence of human papillomavirus (HPV) positive tonsillar carcinoma in Stockholm, Sweden: An epidemic of viral-induced carcinoma? *Int. J. Cancer* **125**, 362–366 (2009).
29. Ramqvist, T. & Dalianis, T. Oropharyngeal cancer epidemic and human papillomavirus. *Emerg. Infect. Dis.* **16**, 1671–7 (2010).
30. Braakhuis, B. J. M., Leemans, C. R. & Visser, O. Incidence and survival trends of head and neck squamous cell carcinoma in the Netherlands between 1989 and 2011. *Oral Oncol.* **50**, 670–675 (2014).
31. Chaturvedi, A. K. *et al.* Worldwide Trends in Incidence Rates for Oral Cavity and Oropharyngeal Cancers. *J. Clin. Oncol.* **31**, 4550–4559 (2013).
32. Conway, D. I., Stockton, D. L., Warnakulasuriya, K. A. A. S., Ogden, G. & Macpherson, L. M. D. Incidence of oral and oropharyngeal cancer in United Kingdom (1990–1999)—recent trends and regional variation. *Oral Oncol.* **42**, 586–592 (2006).
33. Mehanna, H., Jones, T. M., Gregoire, V. & Ang, K. K. Oropharyngeal carcinoma related to human papillomavirus. *BMJ* **340**, c1439 (2010).
34. Carlander, A.-L. F. *et al.* Continuing rise in oropharyngeal cancer in a high HPV prevalence area: A Danish population-based study from 2011 to 2014. *Eur. J. Cancer* **70**, 75–82 (2017).
35. Blomberg, M., Nielsen, A., Munk, C. & Kjaer, S. K. Trends in head and neck cancer incidence in Denmark, 1978-2007: Focus on human papillomavirus associated sites. *Int. J. Cancer* **129**, 733–741 (2011).
36. National Cancer Registry Ireland. National Cancer Registry Ireland. *National Cancer Registry* (2019). Available at: https://www.ncri.ie/. (Accessed: 30th April 2019)
37. de Martel, C. *et al.* Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. *Lancet Oncol.* **13**, 607–615 (2012).
38. National Cancer Institute, N. I. of H. Laryngeal Cancer - Cancer Stat Facts. *NIH Cancer Stat Facts* (2019). Available at: https://seer.cancer.gov/statfacts/html/laryn.html. (Accessed: 9th April 2019)
39. Mourad, M. *et al.* Epidemiological Trends of Head and Neck Cancer in the United States: A SEER Population Study. *J. Oral Maxillofac. Surg.* **75**, 2562–2572 (2017).
40. Diz, P. *et al.* Oral and pharyngeal cancer in Europe. *Transl. Res. Oral Oncol.* **2**, 2057178X1770151 (2017).
41. Chatenoud, L. *et al.* Laryngeal cancer mortality trends in European countries. *Int. J. Cancer*

138, 833–842 (2016).

42. Wierzbicka, M., Winiarski, P. & Osuch-Wójcikiewicz, E. The incidence of laryngeal cancer in Europe with special regard to Poland in last 2 decades. *Otolaryngol. Pol.* **70**, 16–21 (2016).
43. Markou, K. *et al.* Laryngeal cancer: epidemiological data from Northern Greece and review of the literature. *Hippokratia* **17**, 313–8 (2013).
44. Marur, S. & Forastiere, A. A. Head and Neck Squamous Cell Carcinoma: Update on Epidemiology, Diagnosis, and Treatment. *Mayo Clin. Proc.* **91**, 386–396 (2016).
45. Warnakulasuriya, S. Global epidemiology of oral and oropharyngeal cancer. *Oral Oncol.* **45**, 309–316 (2009).
46. McKean-Cowdin, R., Feigelson, H. S., Ross, R. K., Pike, M. C. & Henderson, B. E. Declining cancer rates in the 1990s. *J. Clin. Oncol.* **18**, 2258–68 (2000).
47. Diz, P. *et al.* Oral and pharyngeal cancer in Europe. *Transl. Res. Oral Oncol.* **2**, 2057178X1770151 (2017).
48. Center for Disease Control. STD Facts - Human papillomavirus (HPV). *Center for Disease Control* (2017). Available at: https://www.cdc.gov/std/hpv/stdfact-hpv.htm. (Accessed: 2nd February 2017)
49. Van Doorslaer, K. & Burk, R. D. Evolution of Human Papillomavirus Carcinogenicity. *Adv. Virus Res.* **77**, 41–62 (2010).
50. Rautava, J. & Syrjänen, S. Biology of human papillomavirus infections in head and neck carcinogenesis. *Head Neck Pathol.* **6 Suppl 1**, S3-15 (2012).
51. Graham, S. V. Human papillomavirus: gene expression, regulation and prospects for novel diagnostic methods and antiviral therapies. *Future Microbiol.* **5**, 1493–506 (2010).
52. de Villiers, E.-M., Fauquet, C., Broker, T. R., Bernard, H.-U. & zur Hausen, H. Classification of papillomaviruses. *Virology* **324**, 17–27 (2004).
53. Yanofsky, V. R., Patel, R. V & Goldenberg, G. Genital warts: a comprehensive review. *J. Clin. Aesthet. Dermatol.* **5**, 25–36 (2012).
54. Pande, S. *et al.* Human papillomavirus type 16 variant analysis of E6, E7, and L1 genes and long control region in biopsy samples from cervical cancer patients in north India. *J. Clin. Microbiol.* **46**, 1060–6 (2008).
55. Qmichou, Z. *et al.* Analysis of mutations in the E6 oncogene of human papillomavirus 16 in cervical cancer isolates from Moroccan women. *BMC Infect. Dis.* **13**, 378 (2013).
56. Pista, A. *et al.* Molecular variants of human papillomavirus type 16 and 18 and risk for cervical Neoplasia in Portugal. *J. Med. Virol.* **79**, 1889–1897 (2007).
57. Tornesello, M. L. *et al.* Analysis of human papillomavirus type-16 variants in Italian women

with cervical intraepithelial neoplasia and cervical cancer. *J. Med. Virol.* **74**, 117–126 (2004).

58. Xi, L. F. *et al.* Risk for High-Grade Cervical Intraepithelial Neoplasia Associated with Variants of Human Papillomavirus Types 16 and 18. *Cancer Epidemiol. Biomarkers Prev.* **16**, 4–10 (2007).
59. Cornet, I. *et al.* Human papillomavirus type 16 E6 variants in France and risk of viral persistence. *Infect. Agent. Cancer* **8**, 4 (2013).
60. zur Hausen, H. Human papillomaviruses and their possible role in squamous cell carcinomas. *Curr. Top. Microbiol. Immunol.* **78**, 1–30 (1977).
61. International Agency for Research on Cancer. *Human papillomaviruses: IARC monographs on the evaluation of carcinogenic risks to humans*. (1995).
62. Walboomers, J. M. M. *et al.* Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J. Pathol.* **189**, 12–19 (1999).
63. Liyanage, S. S. *et al.* The Aetiological Role of Human Papillomavirus in Oesophageal Squamous Cell Carcinoma: A Meta-Analysis. *PLoS One* **8**, e69238 (2013).
64. Aldabagh, B., Angeles, J. G. C., Cardones, A. R. & Arron, S. T. Cutaneous squamous cell carcinoma and human papillomavirus: is there an association? *Dermatol. Surg.* **39**, 1–23 (2013).
65. Bosch, F. X. *et al.* Comprehensive Control of Human Papillomavirus Infections and Related Diseases. *Vaccine* **31**, I1–I31 (2013).
66. Parkin, D. M. The global health burden of infection-associated cancers in the year 2002. *Int. J. Cancer* **118**, 3030–3044 (2006).
67. Syrjänen, K., Syrjänen, S., Lamberg, M., Pyrhönen, S. & Nuutinen, J. Morphological and immunohistochemical evidence suggesting human papillomavirus (HPV) involvement in oral squamous cell carcinogenesis. *Int. J. Oral Surg.* **12**, 418–24 (1983).
68. Löning, T. *et al.* Analysis of oral papillomas, leukoplakias, and invasive carcinomas for human papillomavirus type related DNA. *J. Invest. Dermatol.* **84**, 417–20 (1985).
69. Paz, I. B., Cook, N., Odom-Maryon, T., Xie, Y. & Wilczynski, S. P. Human papillomavirus (HPV) in head and neck cancer. An association of HPV 16 with squamous cell carcinoma of Waldeyer's tonsillar ring. *Cancer* **79**, 595–604 (1997).
70. Gillison, M. L. *et al.* Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. *J. Natl. Cancer Inst.* **92**, 709–20 (2000).
71. Chaturvedi, A. K. *et al.* Human Papillomavirus and Rising Oropharyngeal Cancer Incidence in the United States. *J. Clin. Oncol.* **29**, 4294–4301 (2011).
72. Forte, T., Niu, J., Lockwood, G. A. & Bryant, H. E. Incidence trends in head and neck cancers and human papillomavirus (HPV)-associated oropharyngeal cancer in Canada, 1992–2009.

Cancer Causes Control **23**, 1343–1348 (2012).

73. Svahn, M. F., Munk, C., von Buchwald, C., Frederiksen, K. & Kjaer, S. K. Burden and incidence of human papillomavirus-associated cancers and precancerous lesions in Denmark. *Scand. J. Public Health* **44**, 551–559 (2016).
74. Attner, P. *et al.* The role of human papillomavirus in the increased incidence of base of tongue cancer. *Int. J. Cancer* **126**, NA-NA (2010).
75. Wittekindt, C. *et al.* Increasing Incidence rates of Oropharyngeal Squamous Cell Carcinoma in Germany and Significance of Disease Burden Attributed to Human Papillomavirus. *Cancer Prev. Res.* (2019). doi:10.1158/1940-6207.CAPR-19-0098
76. Buttmann-Schweiger, N., Deleré, Y., Klug, S. J. & Kraywinkel, K. Cancer incidence in Germany attributable to human papillomavirus in 2013. *BMC Cancer* **17**, 682 (2017).
77. St Guily, J. L. *et al.* Human papillomavirus genotype distribution in oropharynx and oral cavity cancer in France—The EDiTH VI study. *J. Clin. Virol.* **51**, 100–104 (2011).
78. Schache, A. G. *et al.* HPV-Related Oropharynx Cancer in the United Kingdom: An Evolution in the Understanding of Disease Etiology. *Cancer Res.* **76**, 6598–6606 (2016).
79. Rodrigo, J. P. *et al.* Time trends in the prevalence of HPV in oropharyngeal squamous cell carcinomas in northern Spain. *Int. J. Cancer* **134**, 487–492 (2014).
80. Stein, A. P. *et al.* Prevalence of Human Papillomavirus in Oropharyngeal Cancer. *Cancer J.* **21**, 138–146 (2015).
81. Mehanna, H. *et al.* Prevalence of human papillomavirus in oropharyngeal and nonoropharyngeal head and neck cancer-systematic review and meta-analysis of trends by time and region. *Head Neck* **35**, 747–755 (2013).
82. Chaturvedi, A. K. Epidemiology and Clinical Aspects of HPV in Head and Neck Cancers. *Head Neck Pathol.* **6**, 16–24 (2012).
83. Center for Disease Control and Prevention. Cancers Associated with Human Papillomavirus, United States—2010–2014. *Center for Disease Control and Prevention* (2017). Available at: https://www.cdc.gov/cancer/uscs/about/data-briefs/no1-hpv-assoc-cancers-UnitedStates-2010-2014.htm. (Accessed: 16th September 2019)
84. Lindeberg, H. & Krogdahl, A. Laryngeal cancer and human papillomavirus: HPV is absent in the majority of laryngeal carcinomas. *Cancer Lett.* **146**, 9–13 (1999).
85. Matzow, T., Boysen, M., Kalantari, M., Johansson, B. & Hagmar, B. Low Detection Rate of HPV in Oral and Laryngeal Carcinomas. *Acta Oncol. (Madr).* **37**, 73–76 (1998).
86. Hernandez, B. Y. *et al.* Human papillomavirus prevalence in invasive laryngeal cancer in the United States. *PLoS One* **9**, e115931 (2014).

87. Gungor, A. *et al.* Human papilloma virus prevalence in laryngeal squamous cell carcinoma. *J. Laryngol. Otol.* **121**, 772–774 (2007).
88. Tong, F. *et al.* Prevalence and Prognostic Significance of HPV in Laryngeal Squamous Cell Carcinoma in Northeast China. *Cell. Physiol. Biochem.* **49**, 206–216 (2018).
89. Kreimer, A. R., Clifford, G. M., Boyle, P. & Franceschi, S. Human Papillomavirus Types in Head and Neck Squamous Cell Carcinomas Worldwide: A Systematic Review. *Cancer Epidemiol. Biomarkers Prev.* **14**, 467–475 (2005).
90. Hübbers, C. U. & Akgül, B. HPV and cancer of the oral cavity. *Virulence* **6**, 244–8 (2015).
91. Li, N., Franceschi, S., Howell-Jones, R., Snijders, P. J. F. & Clifford, G. M. Human papillomavirus type distribution in 30,848 invasive cervical cancers worldwide: Variation by geographical region, histological type and year of publication. *Int. J. Cancer* **128**, 927–935 (2011).
92. Kreimer, A. R. *et al.* The Epidemiology of Oral HPV Infection among a Multinational Sample of Healthy Men. *Cancer Epidemiol. Biomarkers Prev.* **20**, 172–182 (2011).
93. Hernandez, B. Y. *et al.* Human papillomavirus prevalence in invasive laryngeal cancer in the United States. *PLoS One* **9**, e115931 (2014).
94. Babiker, A. Y. *et al.* Screening for high risk human papilloma virus (HR-HPV) subtypes, among Sudanese patients with oral lesions. *Int. J. Clin. Exp. Med.* **6**, 275–81 (2013).
95. de Sanjose, S. *et al.* Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. *Lancet Oncol.* **11**, 1048–1056 (2010).
96. Mai, S. *et al.* Prognostic Relevance of HPV Infection and p16 Overexpression in Squamous Cell Anal Cancer. *Int. J. Radiat. Oncol.* **93**, 819–827 (2015).
97. Hoots, B. E., Palefsky, J. M., Pimenta, J. M. & Smith, J. S. Human papillomavirus type distribution in anal cancer and anal intraepithelial lesions. *Int. J. Cancer* **124**, 2375–2383 (2009).
98. Beckmann, A. M. *et al.* Human papillomavirus infection and anal cancer. *Int. J. Cancer* **43**, 1042–1049 (1989).
99. Ouhoummane, N. *et al.* Squamous anal cancer: Patient characteristics and HPV type distribution. *Cancer Epidemiol.* **37**, 807–812 (2013).
100. Palmer, J. G. *et al.* Anal cancer and human papillomaviruses. *Dis. Colon Rectum* **32**, 1016–1022 (1989).
101. Palefsky, J. M. *et al.* Detection of human papillomavirus DNA in anal intraepithelial neoplasia and anal cancer. *Cancer Res.* **51**, 1014–9 (1991).
102. Palefsky, J. M. *et al.* HPV Vaccine against Anal HPV Infection and Anal Intraepithelial Neoplasia. *N. Engl. J. Med.* **365**, 1576–1585 (2011).

103. Smith, J. S., Backes, D. M., Hoots, B. E., Kurman, R. J. & Pimenta, J. M. Human Papillomavirus Type-Distribution in Vulvar and Vaginal Cancers and Their Associated Precursors. *Obstet. Gynecol.* **113**, 917–924 (2009).
104. Parkin, D. M. & Bray, F. Chapter 2: The burden of HPV-related cancers. *Vaccine* **24**, S11–S25 (2006).
105. de Martel, C., Plummer, M., Vignat, J. & Franceschi, S. Worldwide burden of cancer attributable to HPV by site, country and HPV type. *Int. J. Cancer* **141**, 664–670 (2017).
106. Chaturvedi, A. K. Beyond Cervical Cancer: Burden of Other HPV-Related Cancers Among Men and Women. *J. Adolesc. Heal.* **46**, S20–S26 (2010).
107. Backes, D. M., Kurman, R. J., Pimenta, J. M. & Smith, J. S. Systematic review of human papillomavirus prevalence in invasive penile cancer. *Cancer Causes Control* **20**, 449–457 (2009).
108. Kelly L. Stratton, M. D. J. C. M. F. M. A Contemporary Review of HPV and Penile Cancer. (2016).
109. Miralles-Guri, C. *et al.* Human papillomavirus prevalence and type distribution in penile carcinoma. *J. Clin. Pathol.* **62**, 870–878 (2009).
110. Husain, N. & Neyaz, A. Human papillomavirus associated head and neck squamous cell carcinoma: Controversies and new concepts. *J. oral Biol. craniofacial Res.* **7**, 198–205 (2017).
111. Westra, W. H. The changing face of head and neck cancer in the 21st century: the impact of HPV on the epidemiology and pathology of oral cancer. *Head Neck Pathol.* **3**, 78–81 (2009).
112. Begum, S., Cao, D., Gillison, M., Zahurak, M. & Westra, W. H. Tissue distribution of human papillomavirus 16 DNA integration in patients with tonsillar carcinoma. *Clin. Cancer Res.* **11**, 5694–9 (2005).
113. Union for International Cancer Control. TNM. *UICC: Global Cancer Control* (2016). Available at: http://www.uicc.org/resources/tnm. (Accessed: 3rd February 2017)
114. American Joint Committee on Cancer & American Cancer Society. *AJCC Cancer Staging Manual*. (Springer Publishing, 2016).
115. Lee, Y.-C. A. *et al.* Smoking addiction and the risk of upper-aerodigestive-tract cancer in a multicenter case-control study. *Int. J. Cancer* **133**, n/a-n/a (2013).
116. Guha, N. *et al.* Oral Health and Risk of Squamous Cell Carcinoma of the Head and Neck and Esophagus: Results of Two Multicentric Case-Control Studies. *Am. J. Epidemiol.* **166**, 1159–1173 (2007).
117. Talamini, R. *et al.* Oral hygiene, dentition, sexual habits and risk of oral cancer. *Br. J. Cancer* **83**, 1238–42 (2000).

118. Franceschi, S. *et al.* Food groups, oils and butter, and cancer of the oral cavity and pharynx. *Br. J. Cancer* **80**, 614–620 (1999).
119. Fioretti, F., Bosetti, C., Tavani, A., Franceschi, S. & La Vecchia, C. Risk factors for oral and pharyngeal cancer in never smokers. *Oral Oncol.* **35**, 375–8 (1999).
120. Ahn, J., Segers, S. & Hayes, R. B. Periodontal disease, Porphyromonas gingivalis serum antibody levels and orodigestive cancer mortality. *Carcinogenesis* **33**, 1055–1058 (2012).
121. Tezal, M. *et al.* Chronic Periodontitis and the Incidence of Head and Neck Squamous Cell Carcinoma. *Cancer Epidemiol. Biomarkers Prev.* **18**, 2406–2412 (2009).
122. Deeken, J. F. *et al.* The Rising Challenge of Non-AIDS-Defining Cancers in HIV-Infected Patients. *Clin. Infect. Dis.* **55**, 1228–1235 (2012).
123. Langevin, S. M. *et al.* Occupational dust exposure and head and neck squamous cell carcinoma risk in a population-based case-control study conducted in the greater Boston area. *Cancer Med.* **2**, 978–986 (2013).
124. Kutler, D. I. *et al.* High incidence of head and neck squamous cell carcinoma in patients with Fanconi anemia. *Arch. Otolaryngol. Head. Neck Surg.* **129**, 106–12 (2003).
125. Khalil, D. *et al.* Does Socioeconomic Status Affect Stage at Presentation for Larynx Cancer in Canada's Universal Health Care System? *Otolaryngol. Neck Surg.* **160**, 488–493 (2019).
126. Gillison, M. L. *et al.* Distinct Risk Factor Profiles for Human Papillomavirus Type 16-Positive and Human Papillomavirus Type 16-Negative Head and Neck Cancers. *JNCI J. Natl. Cancer Inst.* **100**, 407–420 (2008).
127. Marur, S., D'Souza, G., Westra, W. H. & Forastiere, A. A. HPV-associated head and neck cancer: a virus-related cancer epidemic. *Lancet Oncol.* **11**, 781–789 (2010).
128. O'Regan, E. M. *et al.* Distinct array comparative genomic hybridization profiles in oral squamous cell carcinoma occurring in young patients. *Head Neck* **28**, 330–338 (2006).
129. Kutler, D. I. *et al.* High incidence of head and neck squamous cell carcinoma in patients with Fanconi anemia. *Arch. Otolaryngol. Head. Neck Surg.* **129**, 106–12 (2003).
130. O'Regan, E. M. *et al.* p16INK4A genetic and epigenetic profiles differ in relation to age and site in head and neck squamous cell carcinomas. *Hum. Pathol.* **39**, 452–458 (2008).
131. Smith, E. M. *et al.* Age, sexual behavior and human papillomavirus infection in oral cavity and oropharyngeal cancers. *Int. J. Cancer* **108**, 766–772 (2004).
132. El-Mofty, S. K. & Lu, D. W. Prevalence of human papillomavirus type 16 DNA in squamous cell carcinoma of the palatine tonsil, and not the oral cavity, in young patients: a distinct clinicopathologic and molecular disease entity. *Am. J. Surg. Pathol.* **27**, 1463–70 (2003).
133. American Lung Association. *Trends in Tobacco Use*. (2011).

134. Miller, C. S. & White, D. K. Human papillomavirus expression in oral mucosa, premalignant conditions, and squamous cell carcinoma: a retrospective review of the literature. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod.* **82**, 57–68 (1996).
135. Gillison, M. L. *et al.* Prevalence of Oral HPV Infection in the United States, 2009-2010. *JAMA* **307**, 693 (2012).
136. Kreimer, A. R. *et al.* Oral Human Papillomavirus in Healthy Individuals: A Systematic Review of the Literature. *Sex. Transm. Dis.* **37**, 1 (2010).
137. D'Souza, G., McNeel, T. S. & Fakhry, C. Understanding personal risk of oropharyngeal cancer: risk-groups for oncogenic oral HPV infection and oropharyngeal cancer. *Ann. Oncol.* **28**, 3065–3069 (2017).
138. Gillison, M. L. *et al.* Eurogin Roadmap: Comparative epidemiology of HPV infection and associated cancers of the head and neck and cervix. *Int. J. Cancer* **134**, 497–507 (2014).
139. Shatalova, E. G., Klein-Szanto, A. J. P., Devarajan, K., Cukierman, E. & Clapper, M. L. Estrogen and Cytochrome P450 1B1 Contribute to Both Early- and Late-Stage Head and Neck Carcinogenesis. *Cancer Prev. Res.* **4**, 107–115 (2011).
140. Markowitz, L. E., Sternberg, M., Dunne, E. F., Mcquillan, G. & Unger, E. R. Seroprevalence of Human Papillomavirus Types 6, 11, 16, and 18 in the United States: National Health and Nutrition Examination Survey 2003–2004. *J. Infect. Dis.* **200**, (2009).
141. Safaeian, M. *et al.* Epidemiological study of anti-HPV16/18 seropositivity and subsequent risk of HPV16 and -18 infections. *J. Natl. Cancer Inst.* **102**, 1653–62 (2010).
142. D'Souza, G. & Dempsey, A. The role of HPV in head and neck cancer and review of the HPV vaccine. *Prev. Med. (Baltim).* **53**, S5–S11 (2011).
143. O'Hanlon, S., Forster, D. P. & Lowry, R. J. Oral cancer in the North-East of England: incidence, mortality trends and the link with material deprivation. *Community Dent. Oral Epidemiol.* **25**, 371–6 (1997).
144. Megwalu, U. C. Impact of County-Level Socioeconomic Status on Oropharyngeal Cancer Survival in the United States. *Otolaryngol. Neck Surg.* **156**, 665–670 (2017).
145. Adeyemi, B. F., Olusanya, A. A. & Lawoyin, J. O. Oral squamous cell carcinoma, socioeconomic status and history of exposure to alcohol and tobacco. *J. Natl. Med. Assoc.* **103**, 498–502 (2011).
146. Choi, S. H. *et al.* Socioeconomic and Other Demographic Disparities Predicting Survival among Head and Neck Cancer Patients. *PLoS One* **11**, (2016).
147. Ramsey, T. *et al.* Laryngeal cancer: Global socioeconomic trends in disease burden and smoking habits. *Laryngoscope* **128**, 2039–2053 (2018).

148. Groome, P. A. *et al.* Explaining Socioeconomic Status Effects in Laryngeal Cancer. *Clin. Oncol.* **18**, 283–292 (2006).

149. Auluck, A. *et al.* Socio-economic deprivation: a significant determinant affecting stage of oral cancer diagnosis and survival. *BMC Cancer* **16**, 569 (2016).

150. Boscolo-Rizzo, P. *et al.* New insights into human papillomavirus-associated head and neck squamous cell carcinoma. *Acta Otorhinolaryngol. Ital.* **33**, 77–87 (2013).

151. Fakhry, C. *et al.* Improved Survival of Patients With Human Papillomavirus-Positive Head and Neck Squamous Cell Carcinoma in a Prospective Clinical Trial. *JNCI J. Natl. Cancer Inst.* **100**, 261–269 (2008).

152. Colevas, A. D. Population-based evaluation of incidence trends in oropharyngeal cancer focusing on socioeconomic status, sex, and race/ethnicity. *Head Neck* **36**, 34–42 (2014).

153. Jiron, J. *et al.* Racial disparities in Human Papillomavirus (HPV) associated head and neck cancer. *Am. J. Otolaryngol.* **35**, 147–53 (2014).

154. Settle, K. *et al.* Racial Survival Disparity in Head and Neck Cancer Results from Low Prevalence of Human Papillomavirus Infection in Black Oropharyngeal Cancer Patients. *Cancer Prev. Res.* **2**, 776–781 (2009).

155. American Cancer Society. Laryngeal Cancer Risk Factors | Throat Cancer Risk Factors. *American Cancer Society* (2017). Available at: https://www.cancer.org/cancer/laryngeal-and-hypopharyngeal-cancer/causes-risks-prevention/risk-factors.html. (Accessed: 14th April 2019)

156. Ernani, V. & Saba, N. F. Oral Cavity Cancer: Risk Factors, Pathology, and Management. *Oncology* **89**, 187–195 (2015).

157. van Monsjou, H. S., Balm, A. J. M., van den Brekel, M. M. & Wreesmann, V. B. Oropharyngeal squamous cell carcinoma: A unique disease on the rise? *Oral Oncol.* **46**, 780–785 (2010).

158. Elrefaey, S., Massaro, M. A., Chiocca, S., Chiesa, F. & Ansarin, M. HPV in oropharyngeal cancer: the basics to know in clinical practice. *Acta Otorhinolaryngol. Ital.* **34**, 299–309 (2014).

159. Gillison, M. L. Human papillomavirus-related diseases: oropharynx cancers and potential implications for adolescent HPV vaccination. *J. Adolesc. Health* **43**, S52-60 (2008).

160. Mork, J. *et al.* Human Papillomavirus Infection as a Risk Factor for Squamous-Cell Carcinoma of the Head and Neck. *N. Engl. J. Med.* **344**, 1125–1131 (2001).

161. Kreimer, A. R. *et al.* Evaluation of Human Papillomavirus Antibodies and Risk of Subsequent Head and Neck Cancer. *J. Clin. Oncol.* **31**, 2708–2715 (2013).

162. D'Souza, G. *et al.* Case–Control Study of Human Papillomavirus and Oropharyngeal Cancer. *N. Engl. J. Med.* **356**, 1944–1956 (2007).

163. Syrjanen, S. *et al.* Oral HPV infection: current strategies for prevention and therapy. *Curr. Pharm. Des.* **18**, 5452–69 (2012).

164. Chaturvedi, A. K. *et al.* NHANES 2009-2012 Findings: Association of Sexual Behaviors with Higher Prevalence of Oral Oncogenic Human Papillomavirus Infections in U.S. Men. *Cancer Res.* **75**, 2468–2477 (2015).

165. Smith, E. M. *et al.* Human papillomavirus seropositivity and risks of head and neck cancer. *Int. J. Cancer* **120**, 825–832 (2007).

166. D'Souza, G., Agrawal, Y., Halpern, J., Bodison, S. & Gillison, M. L. Oral Sexual Behaviors Associated with Prevalent Oral Human Papillomavirus Infection. *J. Infect. Dis.* **199**, 1263–1269 (2009).

167. Fakhry, C. *et al.* Relationship between prevalent oral and cervical human papillomavirus infections in human immunodeficiency virus-positive and -negative women. *J. Clin. Microbiol.* **44**, 4479–85 (2006).

168. Termine, N. *et al.* Oral human papillomavirus infection in women with cervical HPV infection: New data from an Italian cohort and a metanalysis of the literature. *Oral Oncol.* **47**, 244–250 (2011).

169. Kreimer, A. R. *et al.* Oral Human Papillomavirus Infection in Adults Is Associated with Sexual Behavior and HIV Serostatus. *J. Infect. Dis.* **189**, 686–698 (2004).

170. Pickard, R. K. L., Xiao, W., Broutian, T. R., He, X. & Gillison, M. L. The Prevalence and Incidence of Oral Human Papillomavirus Infection Among Young Men and Women, Aged 18–30 Years. *Sex. Transm. Dis.* **39**, 559–566 (2012).

171. Rautava, J. & Syrjänen, S. Human papillomavirus infections in the oral mucosa. *J. Am. Dent. Assoc.* **142**, 905–14 (2011).

172. Rioux, M., Garland, A., Webster, D. & Reardon, E. HPV positive tonsillar cancer in two laser surgeons: case reports. *J. Otolaryngol. - Head Neck Surg.* **42**, 54 (2013).

173. Lingen, M. W. *et al.* Low etiologic fraction for high-risk human papillomavirus in oral cavity squamous cell carcinomas. *Oral Oncol.* **49**, 1–8 (2013).

174. Morbini, P. *et al.* Oral HPV infection and persistence in patients with head and neck cancer. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol.* **116**, 474–484 (2013).

175. Gillison, M. L. *et al.* Prevalence of oral HPV infection in the United States, 2009-2010. *JAMA* **307**, 693–703 (2012).

176. Lupato, V. *et al.* Prevalence and Determinants of Oral Human Papillomavirus Infection in 500 Young Adults from Italy. *PLoS One* **12**, e0170091 (2017).

177. Louvanto, K. *et al.* Genotype-specific incidence and clearance of human papillomavirus in oral

mucosa of women: a six-year follow-up study. *PLoS One* **8**, e53413 (2013).

178. Steinau, M. *et al.* Prevalence of Cervical and Oral Human Papillomavirus Infections Among US Women. *J. Infect. Dis.* **209**, 1739–1743 (2014).
179. Patel, P. *et al.* Incidence of types of cancer among HIV-infected persons compared with the general population in the United States, 1992-2003. *Ann. Intern. Med.* **148**, 728–36 (2008).
180. Rabinovics, N. *et al.* Cancer of the head and neck region in solid organ transplant recipients. *Head Neck* **36**, 181–186 (2014).
181. Engels, E. A. *et al.* Cancer risk in people infected with human immunodeficiency virus in the United States. *Int. J. Cancer* **123**, 187–194 (2008).
182. Gillison, M. L. Oropharyngeal cancer: a potential consequence of concomitant HPV and HIV infection. *Curr. Opin. Oncol.* **21**, 439–444 (2009).
183. Beachler, D. C. & D'Souza, G. Oral human papillomavirus infection and head and neck cancers in HIV-infected individuals. *Curr. Opin. Oncol.* **25**, 503–10 (2013).
184. Beachler, D. C., D'Souza, G., Sugar, E. A., Xiao, W. & Gillison, M. L. Natural history of anal vs oral HPV infection in HIV-infected men and women. *J. Infect. Dis.* **208**, 330–9 (2013).
185. D'souza, G. *et al.* Epidemiology of head and neck squamous cell cancer among HIV-infected patients. *J. Acquir. Immune Defic. Syndr.* **65**, 603–10 (2014).
186. Garcia-Pineres, A. J. *et al.* Persistent Human Papillomavirus Infection Is Associated with a Generalized Decrease in Immune Responsiveness in Older Women. *Cancer Res.* **66**, 11070–11076 (2006).
187. Arnson, Y., Shoenfeld, Y. & Amital, H. Effects of tobacco smoke on immunity, inflammation and autoimmunity. *J. Autoimmun.* **34**, J258–J265 (2010).
188. Simen-Kapeu, A. *et al.* Smoking impairs human papillomavirus (HPV) type 16 and 18 capsids antibody response following natural HPV infection. *Scand. J. Infect. Dis.* **40**, 745–751 (2008).
189. Asvadi Kermani, I. *et al.* Human papilloma virus in head and neck squamous cell cancer. *Iran. J. cancer Prev.* **5**, 21–6 (2012).
190. Esquenazi, D., Bussoloti Filho, I., Carvalho, M. da G. da C. & Barros, F. S. de. The frequency of human papillomavirus findings in normal oral mucosa of healthy people by PCR. *Braz. J. Otorhinolaryngol.* **76**, 78–84 (2010).
191. Silva, K. C. *et al.* Risk factors associated with human papillomavirus infection in two populations from Rio de Janeiro, Brazil. *Mem. Inst. Oswaldo Cruz* **104**, 885–91 (2009).
192. Alam, S., Conway, M. J., Chen, H.-S. & Meyers, C. The Cigarette Smoke Carcinogen Benzo[a]pyrene Enhances Human Papillomavirus Synthesis. *J. Virol.* **82**, 1053–1058 (2008).
193. Snijders, P. J. F. *et al.* Prevalence of mucosotropic human papillomaviruses in squamous-cell

carcinomas of the head and neck. *Int. J. Cancer* **66**, 464–469 (1996).

194. Hong, A. M. *et al.* Human papillomavirus, smoking status and outcomes in tonsillar squamous cell carcinoma. *Int. J. Cancer* **132**, 2748–2754 (2013).

195. Sinha, P., Logan, H. L. & Mendenhall, W. M. Human papillomavirus, smoking, and head and neck cancer. *Am. J. Otolaryngol.* **33**, 130–6 (2012).

196. International Collaboration of Epidemiological Studies of Cervical Cancer *et al.* Carcinoma of the cervix and tobacco smoking: Collaborative reanalysis of individual data on 13,541 women with carcinoma of the cervix and 23,017 women without carcinoma of the cervix from 23 epidemiological studies. *Int. J. Cancer* **118**, 1481–1495 (2006).

197. Ritchie, J. M. *et al.* Human papillomavirus infection as a prognostic factor in carcinomas of the oral cavity and oropharynx. *Int. J. Cancer* **104**, 336–344 (2003).

198. Lindel, K., Beer, K. T., Laissue, J., Greiner, R. H. & Aebersold, D. M. Human papillomavirus positive squamous cell carcinoma of the oropharynx: a radiosensitive subgroup of head and neck carcinoma. *Cancer* **92**, 805–13 (2001).

199. Chung, C. H. *et al.* Molecular classification of head and neck squamous cell carcinomas using patterns of gene expression. *Cancer Cell* **5**, 489–500 (2004).

200. Walter, V. *et al.* Molecular Subtypes in Head and Neck Cancer Exhibit Distinct Patterns of Chromosomal Gain and Loss of Canonical Cancer Genes. *PLoS One* **8**, e56823 (2013).

201. Iglesias-Bartolome, R., Martin, D. & Gutkind, J. S. Exploiting the Head and Neck Cancer Oncogenome: Widespread PI3K-mTOR Pathway Alterations and Novel Molecular Targets. *Cancer Discov.* **3**, 722–725 (2013).

202. Urashima, M. *et al.* Distinct Effects of Alcohol Consumption and Smoking on Genetic Alterations in Head and Neck Carcinoma. *PLoS One* **8**, e80828 (2013).

203. Slaughter, D. P., Southwick, H. W. & Smejkal, W. Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. *Cancer* **6**, 963–8 (1953).

204. Graveland, A. P. *et al.* Molecular screening of oral precancer. *Oral Oncol.* **49**, 1129–1135 (2013).

205. Lucio, M., Andrea, G., Bartolomeo, G. D., Fabio, C. & Dora, S. Between-lesion discrepancies in terms of dysplasia, cell turnover and diagnosis in patients with multiple potentially malignant oral lesions. *Open Dent. J.* **7**, 169–74 (2013).

206. Rietbergen, M. M. *et al.* No evidence for active human papillomavirus (HPV) in fields surrounding HPV-positive oropharyngeal tumors. *J. Oral Pathol. Med.* **43**, 137–142 (2014).

207. Jain, K. S., Sikora, A. G., Baxi, S. S. & Morris, L. G. T. Synchronous cancers in patients with head and neck cancer. *Cancer* **119**, 1832–1837 (2013).

208. Agrawal, N. *et al.* Exome Sequencing of Head and Neck Squamous Cell Carcinoma Reveals Inactivating Mutations in NOTCH1. *Science (80-.).* **333**, 1154–1157 (2011).

209. Stransky, N. *et al.* The Mutational Landscape of Head and Neck Squamous Cell Carcinoma. *Science (80-.).* **333**, 1157–1160 (2011).

210. Pyeon, D. *et al.* Fundamental Differences in Cell Cycle Deregulation in Human Papillomavirus-Positive and Human Papillomavirus-Negative Head/Neck and Cervical Cancers. *Cancer Res.* **67**, 4605–4619 (2007).

211. Slebos, R. J. C. *et al.* Gene Expression Differences Associated with Human Papillomavirus Status in Head and Neck Squamous Cell Carcinoma. *Clin. Cancer Res.* **12**, 701–709 (2006).

212. Smeets, S. J. *et al.* Genome-wide DNA copy number alterations in head and neck squamous cell carcinomas with or without oncogene-expressing human papillomavirus. *Oncogene* **25**, 2558–2564 (2006).

213. Wilting, S. M. *et al.* Genomic profiling identifies common HPV-associated chromosomal alterations in squamous cell carcinomas of cervix and head and neck. *BMC Med. Genomics* **2**, 32 (2009).

214. Perrone, F. *et al.* Molecular and Cytogenetic Subgroups of Oropharyngeal Squamous Cell Carcinoma. *Clin. Cancer Res.* **12**, 6643–6651 (2006).

215. Strati, K. & Lambert, P. F. HUMAN PAPILLOMAVIRUS ASSOCIATION WITH HEAD AND NECK CANCERS: UNDERSTANDING VIRUS BIOLOGY AND USING IT IN THE DEVELOPMENT OF CANCER DIAGNOSTICS. *Expert Opin. Med. Diagn.* **2**, 11–20 (2008).

216. Bernstein, J. M., Bernstein, C. R., West, C. M. L. & Homer, J. J. Molecular and cellular processes underlying the hallmarks of head and neck cancer. *Eur. Arch. Oto-Rhino-Laryngology* **270**, 2585–2593 (2013).

217. Eisma, R. J., Spiro, J. D. & Kreutzer, D. L. Role of angiogenic factors: Coexpression of interleukin-8 and vascular endothelial growth factor in patients with head and neck squamous carcinoma. *Laryngoscope* **109**, 687–693 (1999).

218. Caponigro, F., Formato, R., Caraglia, M., Normanno, N. & Iaffaioli, R. V. Monoclonal antibodies targeting epidermal growth factor receptor and vascular endothelial growth factor with a focus on head and neck tumors. *Curr. Opin. Oncol.* **17**, 212–7 (2005).

219. Friemel, J. *et al.* Pretreatment oral hygiene habits and survival of head and neck squamous cell carcinoma (HNSCC) patients. *BMC Oral Health* **16**, 33 (2016).

220. Gallo, O. *et al.* Prognostic significance of cyclooxygenase-2 pathway and angiogenesis in head and neck squamous cell carcinoma. *Hum. Pathol.* **33**, 708–714 (2002).

221. Grivennikov, S. I., Greten, F. R. & Karin, M. Immunity, Inflammation, and Cancer. *Cell* **140**,

883–899 (2010).

222. Tanaka, N., Odajima, T., Ogi, K., Ikeda, T. & Satoh, M. Expression of E-cadherin, α-catenin, and β-catenin in the process of lymph node metastasis in oral squamous cell carcinoma. *Br. J. Cancer* **89**, 557 (2003).

223. Georgolios, A. K., Batistatou, A. & Charalabopoulos, K. Integrins in Head and Neck Squamous Cell Carcinoma (HNSCC): A Review of the Current Literature. *Cell Commun. Adhes.* **12**, 1–8 (2005).

224. Hanna, G. J. *et al.* Frameshift events predict anti–PD-1/L1 response in head and neck cancer. *JCI Insight* **3**, (2018).

225. Stransky, N. *et al.* The Mutational Landscape of Head and Neck Squamous Cell Carcinoma. *Science (80-.).* **333**, 1157–1160 (2011).

226. Canning, M. *et al.* Heterogeneity of the Head and Neck Squamous Cell Carcinoma Immune Landscape and Its Impact on Immunotherapy. *Front. Cell Dev. Biol.* **7**, 52 (2019).

227. Boscolo-Rizzo, P., Furlan, C., Lupato, V., Polesel, J. & Fratta, E. Novel insights into epigenetic drivers of oropharyngeal squamous cell carcinoma: role of HPV and lifestyle factors. *Clin. Epigenetics* **9**, 124 (2017).

228. Degli Esposti, D. *et al.* Unique DNA methylation signature in HPV-positive head and neck squamous cell carcinomas. *Genome Med.* **9**, 33 (2017).

229. Miller, D. L. *et al.* Identification of a Human Papillomavirus–Associated Oncogenic miRNA Panel in Human Oropharyngeal Squamous Cell Carcinoma Validated by Bioinformatics Analysis of The Cancer Genome Atlas. *Am. J. Pathol.* **185**, 679–692 (2015).

230. Minor, J. *et al.* Methylation of microRNA-9 is a specific and sensitive biomarker for oral and oropharyngeal squamous cell carcinomas. *Oral Oncol.* **48**, 73–78 (2012).

231. Franco, E. L. *et al.* Epidemiology of Acquisition and Clearance of Cervical Human Papillomavirus Infection in Women from a High-Risk Area for Cervical Cancer. *J. Infect. Dis.* **180**, 1415–1423 (1999).

232. Wheeler, C. M. Natural History of Human Papillomavirus Infections, Cytologic and Histologic Abnormalities, and Cancer. *Obstet. Gynecol. Clin. North Am.* **35**, 519–536 (2008).

233. Wang, H.-F. *et al.* The Double-Edged Sword-How Human Papillomaviruses Interact With Immunity in Head and Neck Cancer. *Front. Immunol.* **10**, 653 (2019).

234. Bordignon, V. *et al.* How Human Papillomavirus Replication and Immune Evasion Strategies Take Advantage of the Host DNA Damage Repair Machinery. *Viruses* **9**, (2017).

235. Amador-Molina, A., Hernández-Valencia, J. F., Lamoyi, E., Contreras-Paredes, A. & Lizano, M. Role of innate immunity against human papillomavirus (HPV) infections and effect of

adjuvants in promoting specific immune response. *Viruses* **5**, 2624–42 (2013).

236. Smola, S. Immunopathogenesis of HPV-Associated Cancers and Prospects for Immunotherapy. *Viruses* **9**, (2017).
237. Steinbach, A. & Riemer, A. B. Immune evasion mechanisms of human papillomavirus: An update. *Int. J. Cancer* **142**, 224–229 (2018).
238. Kanodia, S., Fahey, L. M. & Kast, W. M. Mechanisms used by human papillomaviruses to escape the host immune response. *Curr. Cancer Drug Targets* **7**, 79–89 (2007).
239. Georgopoulos, N. T., Proffitt, J. L. & Blair, G. E. Transcriptional regulation of the major histocompatibility complex (MHC) class I heavy chain, TAP1 and LMP2 genes by the human papillomavirus (HPV) type 6b, 16 and 18 E7 oncoproteins. *Oncogene* **19**, 4930–4935 (2000).
240. Bashaw, A. A., Leggatt, G. R., Chandra, J., Tuong, Z. K. & Frazer, I. H. Modulation of antigen presenting cell functions during chronic HPV infection. *Papillomavirus Res. (Amsterdam, Netherlands)* **4**, 58–65 (2017).
241. Lace, M. J. *et al.* Human Papillomavirus Type 16 (HPV-16) Genomes Integrated in Head and Neck Cancers and in HPV-16-Immortalized Human Keratinocyte Clones Express Chimeric Virus-Cell mRNAs Similar to Those Found in Cervical Cancers. *J. Virol.* **85**, 1645–1654 (2011).
242. Chai, R. C., Lambie, D., Verma, M. & Punyadeera, C. Current trends in the etiology and diagnosis of HPV-related head and neck cancers. *Cancer Med.* **4**, 596–607 (2015).
243. Park, I.-S. *et al.* Characterization of the Methylation Patterns in Human Papillomavirus Type 16 Viral DNA in Head and Neck Cancers. *Cancer Prev. Res.* **4**, 207–217 (2011).
244. Gao, G. *et al.* Mate pair sequencing of oropharyngeal squamous cell carcinomas reveals that HPV integration occurs much less frequently than in cervical cancer. *J. Clin. Virol.* **59**, 195–200 (2014).
245. Olthof, N. C. *et al.* Comprehensive Analysis of HPV16 Integration in OPSCC Reveals No Significant Impact of Physical Status on Viral Oncogene and Virally Disrupted Human Gene Expression. *PLoS One* **9**, e88718 (2014).
246. Kim, S.-H. *et al.* HPV integration begins in the tonsillar crypt and leads to the alteration of p16, EGFR and c-myc during tumor formation. *Int. J. Cancer* **120**, 1418–1425 (2007).
247. Pannone, G. *et al.* Evaluation of a combined triple method to detect causative HPV in oral and oropharyngeal squamous cell carcinomas: p16 Immunohistochemistry, Consensus PCR HPV-DNA, and In Situ Hybridization. *Infect. Agent. Cancer* **7**, 4 (2012).
248. Holzinger, D. *et al.* Viral RNA Patterns and High Viral Load Reliably Define Oropharynx Carcinomas with Active HPV16 Involvement. *Cancer Res.* **72**, 4993–5003 (2012).
249. Jung, A. C. *et al.* Biological and clinical relevance of transcriptionally active human

papillomavirus (HPV) infection in oropharynx squamous cell carcinoma. *Int. J. Cancer* **126**, NA-NA (2010).

250. Smeets, S. J. *et al.* A novel algorithm for reliable detection of human papillomavirus in paraffin embedded head and neck cancer specimen. *Int. J. Cancer* **121**, 2465–2472 (2007).
251. Schmitt, M. *et al.* Multiple human papillomavirus infections with high viral loads are associated with cervical lesions but do not differentiate grades of cervical abnormalities. *J. Clin. Microbiol.* **51**, 1458–64 (2013).
252. Rampias, T., Sasaki, C., Weinberger, P. & Psyrri, A. E6 and E7 Gene Silencing and Transformed Phenotype of Human Papillomavirus 16-Positive Oropharyngeal Cancer Cells. *JNCI J. Natl. Cancer Inst.* **101**, 412–423 (2009).
253. Thomas, M., Pim, D. & Banks, L. The role of the E6-p53 interaction in the molecular pathogenesis of HPV. *Oncogene* **18**, 7690–7700 (1999).
254. McMurray, H. R. & McCance, D. J. Human papillomavirus type 16 E6 activates TERT gene transcription through induction of c-Myc and release of USF-mediated repression. *J. Virol.* **77**, 9852–61 (2003).
255. Rampias, T., Sasaki, C. & Psyrri, A. Molecular mechanisms of HPV induced carcinogenesis in head and neck. *Oral Oncol.* **50**, 356–363 (2014).
256. Mischo, A., Ohlenschläger, O., Hortschansky, P., Ramachandran, R. & Görlach, M. Structural Insights into a Wildtype Domain of the Oncoprotein E6 and Its Interaction with a PDZ Domain. *PLoS One* **8**, e62584 (2013).
257. Liu, X., Clements, A., Zhao, K. & Marmorstein, R. Structure of the Human Papillomavirus E7 Oncoprotein and Its Mechanism for Inactivation of the Retinoblastoma Tumor Suppressor. *J. Biol. Chem.* **281**, 578–586 (2006).
258. Moody, C. A. & Laimins, L. A. Human papillomavirus oncoproteins: pathways to transformation. *Nat. Rev. Cancer* **10**, 550–560 (2010).
259. Bose, S., Evans, H., Lantzy, L., Scharre, K. & Youssef, E. p16INK4A is a surrogate biomarker for a subset of human papilloma virus-associated dysplasias of the uterine cervix as determined on the Pap smear. *Diagn. Cytopathol.* **32**, 21–24 (2005).
260. El-Naggar, A. K. & Westra, W. H. p16 expression as a surrogate marker for HPV-related oropharyngeal carcinoma: A guide for interpretative relevance and consistency. *Head Neck* **34**, 459–461 (2012).
261. Schlecht, N. F. *et al.* A comparison of clinically utilized human papillomavirus detection methods in head and neck cancer. *Mod. Pathol.* **24**, 1295–1305 (2011).
262. Wang, H., Sun, R., Lin, H. & Hu, W. P16 ^{INK4A} as a surrogate biomarker for human

papillomavirus-associated oropharyngeal carcinoma: Consideration of some aspects. *Cancer Sci.* **104**, 1553–1559 (2013).

263. Kang, J.-H. *et al.* Molecular genetic characterization of p53 mutated oropharyngeal squamous cell carcinoma cells transformed with human papillomavirus E6 and E7 oncogenes. *Int. J. Oncol.* **43**, 383–93 (2013).
264. Duensing, S. & Münger, K. Mechanisms of genomic instability in human cancer: Insights from studies with human papillomavirus oncoproteins. *Int. J. Cancer* **109**, 157–162 (2004).
265. Jenkins, G., O'Byrne, K. J., Panizza, B. & Richard, D. J. Genome Stability Pathways in Head and Neck Cancers. *Int. J. Genomics* **2013**, 1–19 (2013).
266. Duensing, S. *et al.* The human papillomavirus type 16 E6 and E7 oncoproteins cooperate to induce mitotic defects and genomic instability by uncoupling centrosome duplication from the cell division cycle. *Proc. Natl. Acad. Sci.* **97**, 10002–10007 (2000).
267. Patel, D., Incassati, A., Wang, N. & McCance, D. J. Human papillomavirus type 16 E6 and E7 cause polyploidy in human keratinocytes and up-regulation of G2-M-phase proteins. *Cancer Res.* **64**, 1299–306 (2004).
268. Duensing, S. & Münger, K. The human papillomavirus type 16 E6 and E7 oncoproteins independently induce numerical and structural chromosome instability. *Cancer Res.* **62**, 7075–82 (2002).
269. Akagi, K. *et al.* Genome-wide analysis of HPV integration in human cancers reveals recurrent, focal genomic instability. *Genome Res.* **24**, 185–99 (2014).
270. Strati, K., Pitot, H. C. & Lambert, P. F. Identification of biomarkers that distinguish human papillomavirus (HPV)-positive versus HPV-negative head and neck cancers in a mouse model. *Proc. Natl. Acad. Sci.* **103**, 14152–14157 (2006).
271. Sartor, M. A. *et al.* Genome-wide methylation and expression differences in HPV(+) and HPV(-) squamous cell carcinoma cell lines are consistent with divergent mechanisms of carcinogenesis. *Epigenetics* **6**, 777–87 (2011).
272. Lajer, C. B. *et al.* The role of miRNAs in human papilloma virus (HPV)-associated cancers: bridging between HPV-related head and neck cancer and cervical cancer. *Br. J. Cancer* **106**, 1526–34 (2012).
273. Pickering, C. R. *et al.* Integrative Genomic Characterization of Oral Squamous Cell Carcinoma Identifies Frequent Somatic Drivers. *Cancer Discov.* **3**, 770–781 (2013).
274. Fonmarty, D. *et al.* Study of the concordance between p16 immunohistochemistry and HPV-PCR genotyping for the viral diagnosis of oropharyngeal squamous cell carcinoma. *Eur. Ann. Otorhinolaryngol. Head Neck Dis.* **132**, 135–9 (2015).

275. Hafkamp, H. C., Manni, J. J. & Speel, E. J. M. Role of human papillomavirus in the development of head and neck squamous cell carcinomas. *Acta Otolaryngol.* **124**, 520–6 (2004).

276. Liang, C. *et al.* Biomarkers of HPV in Head and Neck Squamous Cell Carcinoma. *Cancer Res.* **72**, 5004–5013 (2012).

277. Brook, I. The clinical microbiology of Waldeyer's ring. *Otolaryngol. Clin. North Am.* **20**, 259–72 (1987).

278. Roberts, J. N. *et al.* Genital transmission of HPV in a mouse model is potentiated by nonoxynol-9 and inhibited by carrageenan. *Nat. Med.* **13**, 857–861 (2007).

279. Fujimura, Y. *et al.* The role of M cells of human nasopharyngeal lymphoid tissue in influenza virus sampling. *Virchows Arch.* **444**, 36–42 (2004).

280. Moutsopoulos, N. M. *et al.* Regulation of the tonsil cytokine milieu favors HIV susceptibility. *J. Leukoc. Biol.* **80**, 1145–55 (2006).

281. Hafkamp, H. C. *et al.* A subset of head and neck squamous cell carcinomas exhibits integration of HPV 16/18 DNA and overexpression of p16INK4A and p53 in the absence of mutations in p53 exons 5-8. *Int. J. Cancer* **107**, 394–400 (2003).

282. Woo, S.-B., Cashman, E. C. & Lerman, M. A. Human papillomavirus-associated oral intraepithelial neoplasia. *Mod. Pathol.* **26**, 1288–1297 (2013).

283. Fakhry, C., Rosenthal, B. T., Clark, D. P. & Gillison, M. L. Associations between Oral HPV16 Infection and Cytopathology: Evaluation of an Oropharyngeal "Pap-Test Equivalent" in High-Risk Populations. *Cancer Prev. Res.* **4**, 1378–1384 (2011).

284. Kreimer, A. R. & Chaturvedi, A. K. HPV-associated Oropharyngeal Cancers--Are They Preventable? *Cancer Prev. Res.* **4**, 1346–1349 (2011).

285. Strimbu, K. & Tavel, J. A. What are biomarkers? *Curr. Opin. HIV AIDS* **5**, 463–6 (2010).

286. Lucs, A. V., Saltman, B., Chung, C. H., Steinberg, B. M. & Schwartz, D. L. Opportunities and challenges facing biomarker development for personalized head and neck cancer treatment. *Head Neck* **35**, 294–306 (2013).

287. Rainsbury, J. W. *et al.* Prognostic biomarkers of survival in oropharyngeal squamous cell carcinoma: Systematic review and meta-analysis. *Head Neck* **35**, 1048–1055 (2013).

288. Sedghizadeh, P. P. *et al.* Is p16-positive oropharyngeal squamous cell carcinoma associated with favorable prognosis? A systematic review and meta-analysis. *Oral Oncol.* **54**, 15–27 (2016).

289. Dahlstrom, K. R. *et al.* HPV Serum Antibodies as Predictors of Survival and Disease Progression in Patients with HPV-Positive Squamous Cell Carcinoma of the Oropharynx. *Clin.*

Cancer Res. **21**, 2861–9 (2015).

290. Kimple, R. J. & Harari, P. M. The prognostic value of HPV in head and neck cancer patients undergoing postoperative chemoradiotherapy. *Ann. Transl. Med.* **3**, S14 (2015).
291. Liu, S. Z., Zandberg, D. P., Schumaker, L. M., Papadimitriou, J. C. & Cullen, K. J. Correlation of p16 expression and HPV type with survival in oropharyngeal squamous cell cancer. *Oral Oncol.* **51**, 862–9 (2015).
292. Ang, K. K. *et al.* Human papillomavirus and survival of patients with oropharyngeal cancer. *N. Engl. J. Med.* **363**, 24–35 (2010).
293. Rischin, D. *et al.* Prognostic Significance of p16INK4A and Human Papillomavirus in Patients With Oropharyngeal Cancer Treated on TROG 02.02 Phase III Trial. *J. Clin. Oncol.* **28**, 4142–4148 (2010).
294. Augustin, J. *et al.* HPV RNA CISH score identifies two prognostic groups in a p16 positive oropharyngeal squamous cell carcinoma population. *Mod. Pathol.* **31**, 1645–1652 (2018).
295. Vermorken, J. B. *et al.* Impact of tumor HPV status on outcome in patients with recurrent and/or metastatic squamous cell carcinoma of the head and neck receiving chemotherapy with or without cetuximab: retrospective analysis of the phase III EXTREME trial. *Ann. Oncol.* **25**, 801–7 (2014).
296. Khwaja, S. S. *et al.* High E6 Gene Expression Predicts for Distant Metastasis and Poor Survival in Patients With HPV-Positive Oropharyngeal Squamous Cell Carcinoma. *Int. J. Radiat. Oncol. Biol. Phys.* **95**, 1132–41 (2016).
297. Duray, A. *et al.* Human papillomavirus DNA strongly correlates with a poorer prognosis in oral cavity carcinoma. *Laryngoscope* **122**, 1558–1565 (2012).
298. Klussmann, J. P. *et al.* Expression of p16 Protein Identifies a Distinct Entity of Tonsillar Carcinomas Associated with Human Papillomavirus. *Am. J. Pathol.* **162**, 747–753 (2003).
299. Vokes, E. E., Agrawal, N. & Seiwert, T. Y. HPV-Associated Head and Neck Cancer. *J. Natl. Cancer Inst.* **107**, djv344 (2015).
300. Hoffmann, M. *et al.* HPV DNA, E6*I-mRNA expression and p16INK4A immunohistochemistry in head and neck cancer – How valid is p16INK4A as surrogate marker? *Cancer Lett.* **323**, 88–96 (2012).
301. Strzelczyk, J. K., Krakowczyk, Ł. & Owczarek, A. J. Aberrant DNA methylation of the p16, APC, MGMT, TIMP3 and CDH1 gene promoters in tumours and the surgical margins of patients with oral cavity cancer. *J. Cancer* **9**, 1896–1904 (2018).
302. Rayess, H., Wang, M. B. & Srivatsan, E. S. Cellular senescence and tumor suppressor gene p16. *Int. J. Cancer* **130**, 1715–1725 (2012).

303. Lewis, J. S. *et al.* p16 Positive Oropharyngeal Squamous Cell Carcinoma:An Entity With a Favorable Prognosis Regardless of Tumor HPV Status. *Am. J. Surg. Pathol.* **34**, 1088–1096 (2010).

304. Park, K. *et al.* p16 immunohistochemistry alone is a better prognosticator in tonsil cancer than human papillomavirus in situ hybridization with or without p16 immunohistochemistry. *Acta Otolaryngol.* **133**, 297–304 (2013).

305. Zhao, N. *et al.* Different cellular p16INK4a localisation may signal different survival outcomes in head and neck cancer. *Br. J. Cancer* **107**, 482–490 (2012).

306. Stephen, J. K. *et al.* Significance of p16 in site-specific HPV positive and HPV negative HNSCC. *Cancer Clin. Oncol.* **2**, 51–61 (2012).

307. Klingenberg, B. *et al.* p16INK4A overexpression is frequently detected in tumour-free tonsil tissue without association with HPV. *Histopathology* **56**, 957–967 (2010).

308. Heath, S. *et al.* Clinically Significant Human Papilloma Virus in Squamous Cell Carcinoma of the Head and Neck in UK Practice. *Clin. Oncol.* **24**, e18–e23 (2012).

309. Bishop, J. A. *et al.* Detection of Transcriptionally Active High-risk HPV in Patients With Head and Neck Squamous Cell Carcinoma as Visualized by a Novel E6/E7 mRNA In Situ Hybridization Method. *Am. J. Surg. Pathol.* **36**, 1874–1882 (2012).

310. Mirghani, H. *et al.* A predictive transcriptomic signature of oropharyngeal cancer according to HPV16 status exclusively. *Oral Oncol.* (2014). doi:10.1016/j.oraloncology.2014.07.019

311. Herbst, R. S. Review of epidermal growth factor receptor biology. *Int. J. Radiat. Oncol.* **59**, S21–S26 (2004).

312. Khaznadar, S. S. *et al.* EGFR overexpression is not common in patients with head and neck cancer. Cell lines are not representative for the clinical situation in this indication. *Oncotarget* **9**, 28965–28975 (2018).

313. Chung, C. H. *et al.* Increased Epidermal Growth Factor Receptor Gene Copy Number Is Associated With Poor Prognosis in Head and Neck Squamous Cell Carcinomas. *J. Clin. Oncol.* **24**, 4170–4176 (2006).

314. Hama, T. *et al.* Prognostic Significance of Epidermal Growth Factor Receptor Phosphorylation and Mutation in Head and Neck Squamous Cell Carcinoma. *Oncologist* **14**, 900–908 (2009).

315. Temam, S. *et al.* Epidermal Growth Factor Receptor Copy Number Alterations Correlate With Poor Clinical Outcome in Patients With Head and Neck Squamous Cancer. *J. Clin. Oncol.* **25**, 2164–2170 (2007).

316. Rampias, T. *et al.* Molecular profile of head and neck squamous cell carcinomas bearing p16 high phenotype. *Ann. Oncol.* **24**, 2124–2131 (2013).

317. Reimers, N. *et al.* Combined analysis of HPV-DNA, p16 and EGFR expression to predict prognosis in oropharyngeal cancer. *Int. J. Cancer* **120**, 1731–1738 (2007).
318. Kumar, B. *et al.* EGFR, p16, HPV Titer, Bcl-xL and p53, sex, and smoking as indicators of response to therapy and survival in oropharyngeal cancer. *J. Clin. Oncol.* **26**, 3128–37 (2008).
319. Jedlinski, A., Ansell, A., Johansson, A.-C. & Roberg, K. EGFR status and EGFR ligand expression influence the treatment response of head and neck cancer cell lines. *J. Oral Pathol. Med.* **42**, 26–36 (2013).
320. Herfs, M. *et al.* A novel blueprint for 'top down' differentiation defines the cervical squamocolumnar junction during development, reproductive life, and neoplasia. *J. Pathol.* **229**, 460–468 (2013).
321. Herfs, M. *et al.* Cervical Squamocolumnar Junction–specific Markers Define Distinct, Clinically Relevant Subsets of Low-grade Squamous Intraepithelial Lesions. *Am. J. Surg. Pathol.* **37**, 1311–1318 (2013).
322. Herfs, M. *et al.* A discrete population of squamocolumnar junction cells implicated in the pathogenesis of cervical cancer. *Proc. Natl. Acad. Sci.* **109**, 10516–10521 (2012).
323. Yang, E. J. *et al.* Microanatomy of the cervical and anorectal squamocolumnar junctions: a proposed model for anatomical differences in HPV-related cancer risk. *Mod. Pathol.* **28**, 994–1000 (2015).
324. Perry, M. E. The specialised structure of crypt epithelium in the human palatine tonsil and its functional significance. *J. Anat.* 111–27 (1994).
325. Kang, S. Y. C. *et al.* Characterization of Epithelial Progenitors in Normal Human Palatine Tonsils and Their HPV16 E6/E7-Induced Perturbation. *Stem cell reports* **5**, 1210–25 (2015).
326. Ansell, A. *et al.* Matrix metalloproteinase-7 and -13 expression associate to cisplatin resistance in head and neck cancer cell lines. *Oral Oncol.* **45**, 866–71 (2009).
327. Farnebo, L. *et al.* Strong expression of survivin is associated with positive response to radiotherapy and improved overall survival in head and neck squamous cell carcinoma patients. *Int. J. Cancer* **133**, 1994–2003 (2013).
328. Johansson, A.-C. *et al.* Cancer-Associated Fibroblasts Induce Matrix Metalloproteinase-Mediated Cetuximab Resistance in Head and Neck Squamous Cell Carcinoma Cells. *Mol. Cancer Res.* **10**, 1158–1168 (2012).
329. Nagadia, R., Pandit, P., Coman, W. B., Cooper-White, J. & Punyadeera, C. miRNAs in head and neck cancer revisited. *Cell. Oncol.* **36**, 1–7 (2013).
330. Wu, Y. *et al.* Novel biomarker panel predicts prognosis in human papillomavirus-negative oropharyngeal cancer: an analysis of the TAX 324 trial. *Cancer* **118**, 1811–7 (2012).

331. Langer, C. J. Exploring biomarkers in head and neck cancer. *Cancer* **118**, 3882–3892 (2012).
332. Kawajiri, K., Nakachi, K., Imai, K., Watanabe, J. & Hayashi, S. Germ line polymorphisms of p53 and CYP1A1 genes involved in human lung cancer. *Carcinogenesis* **14**, 1085–9 (1993).
333. Matlashewski, G. *et al.* Role of a p53 polymorphism in the development of human papillomavirus-associated cancer. *Nature* **393**, 229–234 (1998).
334. Nagpal, J. K., Sahni, S. & Das, B. R. p53 codon 72 polymorphism and susceptibility to development of human papilloma virus-associated cervical cancer in Indian women. *Eur. J. Clin. Invest.* **32**, 943–948 (2002).
335. Kaghad, M. *et al.* Monoallelically expressed gene related to p53 at 1p36, a region frequently deleted in neuroblastoma and other human cancers. *Cell* **90**, 809–19 (1997).
336. Bond, G. L. *et al.* A Single Nucleotide Polymorphism in the MDM2 Promoter Attenuates the p53 Tumor Suppressor Pathway and Accelerates Tumor Formation in Humans. *Cell* **119**, 591–602 (2004).
337. Li, W. *et al.* The expression of key cell cycle markers and presence of human papillomavirus in squamous cell carcinoma of the tonsil. *Head Neck* **26**, 1–9 (2004).
338. Sherr, C. J. D-type cyclins. *Trends Biochem. Sci.* **20**, 187–90 (1995).
339. Hosokawa, Y. *et al.* Cyclin D1 (PRAD1) alternative transcript b: full-length cDNA cloning and expression in breast cancers. *Cancer Lett.* **113**, 123–30 (1997).
340. Lai, H.-C. *et al.* Single nucleotide polymorphism at Fas promoter is associated with cervical carcinogenesis. *Int. J. Cancer* **103**, 221–225 (2003).
341. Santoro, A. *et al.* Relationship between CK19 expression, deregulation of normal keratinocyte differentiation pattern and high risk-human papilloma virus infection in oral and oropharyngeal squamous cell carcinoma. *Infect. Agent. Cancer* **10**, 46 (2015).
342. Bousarghin, L., Touzé, A., Sizaret, P.-Y. & Coursaget, P. Human papillomavirus types 16, 31, and 58 use different endocytosis pathways to enter cells. *J. Virol.* **77**, 3846–50 (2003).
343. Liang, Y.-J., Chang, H.-S., Wang, C.-Y. & Yu, W. C. Y. DYRK1A stabilizes HPV16E7 oncoprotein through phosphorylation of the threonine 5 and threonine 7 residues. *Int. J. Biochem. Cell Biol.* **40**, 2431–41 (2008).
344. Howley, P. M. Warts, cancer and ubiquitylation: lessons from the papillomaviruses. *Trans. Am. Clin. Climatol. Assoc.* **117**, 113–26; discussion 126-7 (2006).
345. Boccardo, E., Lepique, A. P. & Villa, L. L. The role of inflammation in HPV carcinogenesis. *Carcinogenesis* **31**, 1905–1912 (2010).
346. Fernandes, J. V. *et al.* Link between chronic inflammation and human papillomavirus-induced carcinogenesis (Review). *Oncol. Lett.* **9**, 1015–1026 (2015).

347. Billiau, A., Heremans, H., Vermeire, K. & Matthys, P. Immunomodulatory Properties of Interferon-gamma: An Updatea. *Ann. N. Y. Acad. Sci.* **856**, 22–32 (1998).

348. Kaskas, N. M. *et al.* Serum Biomarkers in Head and Neck Squamous Cell Cancer. *JAMA Otolaryngol. Neck Surg.* **140**, 5 (2014).

349. Pravica, V., Perrey, C., Stevens, A., Lee, J. H. & Hutchinson, I. V. A single nucleotide polymorphism in the first intron of the human IFN-gamma gene: absolute correlation with a polymorphic CA microsatellite marker of high IFN-gamma production. *Hum. Immunol.* **61**, 863–6 (2000).

350. Zoodsma, M. *et al.* Interleukin-10 and Fas polymorphisms and susceptibility for (pre)neoplastic cervical disease. *Int. J. Gynecol. Cancer* **15**, 282–290 (2005).

351. Turner, D. M. *et al.* AN INVESTIGATION OF POLYMORPHISM IN THE INTERLEUKIN-10 GENE PROMOTER. *Eur. J. Immunogenet.* **24**, 1–8 (1997).

352. Basile, J. R., Zacny, V. & Münger, K. The cytokines tumor necrosis factor-alpha (TNF-alpha) and TNF-related apoptosis-inducing ligand differentially modulate proliferation and apoptotic pathways in human keratinocytes expressing the human papillomavirus-16 E7 oncoprotein. *J. Biol. Chem.* **276**, 22522–8 (2001).

353. Lyford-Pike, S. *et al.* Evidence for a Role of the PD-1:PD-L1 Pathway in Immune Resistance of HPV-Associated Head and Neck Squamous Cell Carcinoma. *Cancer Res.* **73**, 1733–1741 (2013).

354. Ukpo, O. C., Thorstad, W. L. & Lewis, J. S. B7-H1 Expression Model for Immune Evasion in Human Papillomavirus-Related Oropharyngeal Squamous Cell Carcinoma. *Head Neck Pathol.* **7**, 113–121 (2013).

355. Badoual, C. *et al.* PD-1-Expressing Tumor-Infiltrating T Cells Are a Favorable Prognostic Biomarker in HPV-Associated Head and Neck Cancer. *Cancer Res.* **73**, 128–138 (2013).

356. Salama, A. K. & Gangadhar, T. Clinical applications of PD-1-based therapy: a focus on pembrolizumab (MK-3475) in the management of melanoma and other tumor types. *Onco. Targets. Ther.* **8**, 929 (2015).

357. Seiwert, T. Y. *et al.* 2015 ASCO Annual Meeting: Antitumor activity and safety of pembrolizumab in patients (pts) with advanced squamous cell carcinoma of the head and neck (SCCHN): Preliminary results from KEYNOTE-012 expansion cohort. *J Clin Oncol* **33**, (2015).

358. Goldberg, H. I., Lockwood, S. A., Wyatt, S. W. & Crossett, L. S. Trends and differentials in mortality from cancers of the oral cavity and pharynx in the United States, 1973-1987. *Cancer* **74**, 565–72 (1994).

359. Thomas, G. R., Nadiminti, H. & Regalado, J. Molecular predictors of clinical outcome in

patients with head and neck squamous cell carcinoma. *Int. J. Exp. Pathol.* **86**, 347–63 (2005).

360. Price, K. A. R. & Cohen, E. E. Current Treatment Options for Metastatic Head and Neck Cancer. *Curr. Treat. Options Oncol.* **13**, 35–46 (2012).
361. Vermorken, J. B. *et al.* Platinum-Based Chemotherapy plus Cetuximab in Head and Neck Cancer. *N. Engl. J. Med.* **359**, 1116–1127 (2008).
362. Fury, M. G. & Pfister, D. G. Current recommendations for systemic therapy of recurrent and/or metastatic head and neck squamous cell cancer. *J. Natl. Compr. Canc. Netw.* **9**, 681–9 (2011).
363. Vermorken, J. B. *et al.* An analysis of safety in patients (pts) with recurrent and/or metastatic squamous cell carcinoma of the head and neck (R/M SCCHN) receiving chemotherapy (CT) with or without panitumumab (pmab) in a phase III clinical trial (SPECTRUM). *J. Clin. Oncol.* **27**, 6050 (2009).
364. Gibson, M. K. *et al.* Randomized Phase III Evaluation of Cisplatin Plus Fluorouracil Versus Cisplatin Plus Paclitaxel in Advanced Head and Neck Cancer (E1395): An Intergroup Trial of the Eastern Cooperative Oncology Group. *J. Clin. Oncol.* **23**, 3562–3567 (2005).
365. Urba, S. *et al.* Pemetrexed in combination with cisplatin versus cisplatin monotherapy in patients with recurrent or metastatic head and neck cancer. *Cancer* **118**, 4694–4705 (2012).
366. Machiels, J.-P. *et al.* Zalutumumab plus best supportive care versus best supportive care alone in patients with recurrent or metastatic squamous-cell carcinoma of the head and neck after failure of platinum-based chemotherapy: an open-label, randomised phase 3 trial. *Lancet Oncol.* **12**, 333–343 (2011).
367. Wang, Z., Valera, J. C., Zhao, X., Chen, Q. & Silvio Gutkind, J. mTOR co-targeting strategies for head and neck cancer therapy. *Cancer Metastasis Rev.* **36**, 491–502 (2017).
368. Day, T. A. *et al.* Inhibition of mTOR Signaling and Clinical Activity of Rapamycin in Head and Neck Cancer in a Window of Opportunity Trial. *Clin. Cancer Res.* **25**, 1156–1164 (2019).
369. Hafkamp, H. C. *et al.* Marked differences in survival rate between smokers and nonsmokers with HPV 16-associated tonsillar carcinomas. *Int. J. Cancer* **122**, 2656–2664 (2008).
370. Krane, J. F. Role of Cytology in the Diagnosis and Management of HPV-Associated Head and Neck Carcinoma. *Acta Cytol.* **57**, 117–126 (2013).
371. Stenmark, M. H. *et al.* Influence of human papillomavirus on the clinical presentation of oropharyngeal carcinoma in the United States. *Laryngoscope* **127**, 2270–2278 (2017).
372. Mroz, E. A., Forastiere, A. A. & Rocco, J. W. Implications of the Oropharyngeal Cancer Epidemic. *J. Clin. Oncol.* **29**, 4222–4223 (2011).
373. Joo, Y.-H. *et al.* High-risk human papillomavirus and cervical lymph node metastasis in

patients with oropharyngeal cancer. *Head Neck* **34**, 10–14 (2012).

374. McHugh, J. B. Association of cystic neck metastases and human papillomavirus-positive oropharyngeal squamous cell carcinoma. *Arch. Pathol. Lab. Med.* **133**, 1798–803 (2009).

375. Thompson, L. D. & Heffner, D. K. The clinical importance of cystic squamous cell carcinomas in the neck: a study of 136 cases. *Cancer* **82**, 944–56 (1998).

376. Wierzbicka, M., Szyfter, K., Milecki, P., Składowski, K. & Ramlau, R. The rationale for HPV-related oropharyngeal cancer de-escalation treatment strategies. *Contemp. Oncol. (Poznan, Poland)* **19**, 313–22 (2015).

377. Geiger, J. L. *et al.* Adjuvant chemoradiation therapy with high-dose versus weekly cisplatin for resected, locally-advanced HPV/p16-positive and negative head and neck squamous cell carcinoma. *Oral Oncol.* **50**, 311–318 (2014).

378. Psyrri, A., Sasaki, C., Vassilakopoulou, M., Dimitriadis, G. & Rampias, T. Future directions in research, treatment and prevention of HPV-related squamous cell carcinoma of the head and neck. *Head Neck Pathol.* **6 Suppl 1**, S121-8 (2012).

379. Chien, C.-Y. *et al.* Lower prevalence but favorable survival for human papillomavirus-related squamous cell carcinoma of tonsil in Taiwan. *Oral Oncol.* **44**, 174–179 (2008).

380. Fischer, C. A. *et al.* p16 expression in oropharyngeal cancer: its impact on staging and prognosis compared with the conventional clinical staging parameters. *Ann. Oncol.* **21**, 1961–1966 (2010).

381. Marur, S. & Forastiere, A. A. Head and Neck Cancer: Changing Epidemiology, Diagnosis, and Treatment. *Mayo Clin. Proc.* **83**, 489–501 (2008).

382. Gillison, M. L. HPV and prognosis for patients with oropharynx cancer. *Eur. J. Cancer* **45**, 383–385 (2009).

383. Lassen, P. *et al.* The influence of HPV-associated p16-expression on accelerated fractionated radiotherapy in head and neck cancer: Evaluation of the randomised DAHANCA 6&7 trial. *Radiother. Oncol.* **100**, 49–55 (2011).

384. Posner, M. R. *et al.* Survival and human papillomavirus in oropharynx cancer in TAX 324: a subset analysis from an international phase III trial. *Ann. Oncol.* **22**, 1071–1077 (2011).

385. Shaw, R. & Robinson, M. The increasing clinical relevance of human papillomavirus type 16 (HPV-16) infection in oropharyngeal cancer. *Br. J. Oral Maxillofac. Surg.* **49**, 423–429 (2011).

386. Mirghani, H. & Blanchard, P. Treatment de-escalation for HPV-driven oropharyngeal cancer: Where do we stand? *Clin. Transl. Radiat. Oncol.* **8**, 4–11 (2018).

387. Albers, A. E., Qian, X., Kaufmann, A. M. & Coordes, A. Meta analysis: HPV and p16 pattern determines survival in patients with HNSCC and identifies potential new biologic subtype. *Sci.*

Rep. **7**, 16715 (2017).

388. Balermpas, P. *et al.* Tumour-infiltrating lymphocytes predict response to definitive chemoradiotherapy in head and neck cancer. *Br. J. Cancer* **110**, 501–509 (2014).
389. Jung, A. C. *et al.* CD8-alpha T-cell infiltration in human papillomavirus-related oropharyngeal carcinoma correlates with improved patient prognosis. *Int. J. Cancer* **132**, E26–E36 (2013).
390. Jung, Y.-S. *et al.* CD200: association with cancer stem cell features and response to chemoradiation in head and neck squamous cell carcinoma. *Head Neck* **37**, 327–35 (2015).
391. Rietbergen, M. M. *et al.* Cancer stem cell enrichment marker CD98: A prognostic factor for survival in patients with human papillomavirus-positive oropharyngeal cancer. *Eur. J. Cancer* **50**, 765–773 (2014).
392. Braakhuis, B. J. M. *et al.* Expression profiling and prediction of distant metastases in head and neck squamous cell carcinoma. *J. Clin. Pathol.* **59**, 1254–1260 (2006).
393. Zhou, G., Liu, Z. & Myers, J. N. TP53 Mutations in Head and Neck Squamous Cell Carcinoma and Their Impact on Disease Progression and Treatment Response. *J. Cell. Biochem.* **117**, 2682–2692 (2016).
394. Kimple, R. J. *et al.* Enhanced Radiation Sensitivity in HPV-Positive Head and Neck Cancer. *Cancer Res.* **73**, 4791–4800 (2013).
395. Rieckmann, T. *et al.* HNSCC cell lines positive for HPV and p16 possess higher cellular radiosensitivity due to an impaired DSB repair capacity. *Radiother. Oncol.* **107**, 242–246 (2013).
396. Vu, H. L., Sikora, A. G., Fu, S. & Kao, J. HPV-induced oropharyngeal cancer, immune response and response to therapy. *Cancer Lett.* **288**, 149–155 (2010).
397. An, Y., Holsinger, F. C. & Husain, Z. A. De-intensification of adjuvant therapy in human papillomavirus-associated oropharyngeal cancer. *Cancers Head Neck* **1**, 18 (2016).
398. Golusiński, W. Functional Organ Preservation Surgery in Head and Neck Cancer: Transoral Robotic Surgery and Beyond. *Front. Oncol.* **9**, 293 (2019).
399. Mahmoud, O., Sung, K., Civantos, F. J., Thomas, G. R. & Samuels, M. A. Transoral robotic surgery for oropharyngeal squamous cell carcinoma in the era of human papillomavirus. *Head Neck* **40**, 710–721 (2018).
400. Weinstein, G. S. *et al.* Transoral Robotic Surgery Alone for Oropharyngeal Cancer. *Arch. Otolaryngol. Neck Surg.* **138**, 628 (2012).
401. Genden, E. M. *et al.* Transoral robotic resection and reconstruction for head and neck cancer. *Laryngoscope* **121**, 1668–1674 (2011).
402. White, H. N. *et al.* Transoral Robotic-Assisted Surgery for Head and Neck Squamous Cell

Carcinoma. *Arch. Otolaryngol. Neck Surg.* **136**, 1248 (2010).

403. Iseli, T. A. *et al.* Functional Outcomes after Transoral Robotic Surgery for Head and Neck Cancer. *Otolaryngol. Neck Surg.* **141**, 166–171 (2009).
404. Baskin, R. M. *et al.* Transoral robotic surgery for oropharyngeal cancer: patient selection and special considerations. *Cancer Manag. Res.* **10**, 839–846 (2018).
405. Bonner, J. A. *et al.* Radiotherapy plus cetuximab for locoregionally advanced head and neck cancer: 5-year survival data from a phase 3 randomised trial, and relation between cetuximab-induced rash and survival. *Lancet Oncol.* **11**, 21–28 (2010).
406. Koutcher, L. *et al.* Concurrent Cisplatin and Radiation Versus Cetuximab and Radiation for Locally Advanced Head-and-Neck Cancer. *Int. J. Radiat. Oncol.* **81**, 915–922 (2011).
407. Economopoulou, P. & Psyrri, A. Comparing cisplatin-Chemoradiotherapy to Cetuximab-radiotherapy in HPV+ “low-risk” locally advanced oropharyngeal squamous cell carcinoma: lessons from De-escalate study. *Cancers Head Neck* **4**, 1 (2019).
408. Oosthuizen, J. C. & Doody, J. De-intensified treatment in human papillomavirus-positive oropharyngeal cancer. *Lancet (London, England)* **393**, 5–7 (2019).
409. Mehanna, H. *et al.* Radiotherapy plus cisplatin or cetuximab in low-risk human papillomavirus-positive oropharyngeal cancer (De-ESCALaTE HPV): an open-label randomised controlled phase 3 trial. *Lancet (London, England)* **393**, 51–60 (2019).
410. Brotherston, D. C. *et al.* Patient preferences for oropharyngeal cancer treatment de-escalation. *Head Neck* **35**, 151–159 (2013).
411. Mesía, R. & Taberna, M. HPV-related oropharyngeal carcinoma de-escalation protocols. *Lancet Oncol.* **18**, 704–705 (2017).
412. Plzák, J. *et al.* The Head and Neck Squamous Cell Carcinoma Microenvironment as a Potential Target for Cancer Therapy. *Cancers (Basel).* **11**, (2019).
413. Argiris, A. *et al.* Cetuximab and bevacizumab: preclinical data and phase II trial in recurrent or metastatic squamous cell carcinoma of the head and neck. *Ann. Oncol.* **24**, 220–225 (2013).
414. Schuler, P. J., Laban, S., Doescher, J., Bullinger, L. & Hoffmann, T. K. Novel Treatment Options in Head and Neck Cancer. *Oncol. Res. Treat.* **40**, 342–346 (2017).
415. Deák, M. *et al.* Novel role for galectin-1 in T-cells under physiological and pathological conditions. *Immunobiology* **220**, 483–489 (2015).
416. Sanchez-Ruderisch, H. *et al.* Galectin-1 sensitizes carcinoma cells to anoikis via the fibronectin receptor α5β1-integrin. *Cell Death Differ.* **18**, 806–16 (2011).
417. Krajden, M. *et al.* Human Papillomavirus 16 (HPV 16) and HPV 18 Antibody Responses Measured by Pseudovirus Neutralization and Competitive Luminex Assays in a Two- versus

Three-Dose HPV Vaccine Trial. *Clin. Vaccine Immunol.* **18**, 418–423 (2011).

418. Romanowski, B. *et al.* Immunogenicity and safety of the HPV-16/18 AS04-adjuvanted vaccine administered as a 2-dose schedule compared to the licensed 3-dose schedule. *Hum. Vaccin.* **7**, 1374–1386 (2011).
419. Guo, T., Eisele, D. W. & Fakhry, C. The potential impact of prophylactic human papillomavirus vaccination on oropharyngeal cancer. *Cancer* **122**, 2313–23 (2016).
420. Chaturvedi, A. K. *et al.* Effect of Prophylactic Human Papillomavirus (HPV) Vaccination on Oral HPV Infections Among Young Adults in the United States. *J. Clin. Oncol.* **36**, 262–267 (2018).
421. Herrero, R. *et al.* Reduced Prevalence of Oral Human Papillomavirus (HPV) 4 Years after Bivalent HPV Vaccination in a Randomized Clinical Trial in Costa Rica. *PLoS One* **8**, e68329 (2013).
422. Suzich, J. A. *et al.* Systemic immunization with papillomavirus L1 protein completely prevents the development of viral mucosal papillomas. *Proc. Natl. Acad. Sci.* **92**, 11553–11557 (1995).
423. Graham, D. M. *et al.* A cost-effectiveness analysis of human papillomavirus vaccination of boys for the prevention of oropharyngeal cancer. *Cancer* **121**, 1785–1792 (2015).
424. Ward, G., Mehta, V. & Moore, M. Morbidity, mortality and cost from HPV-related oropharyngeal cancer: Impact of 2-, 4- and 9-valent vaccines. *Hum. Vaccin. Immunother.* **12**, 1343–1347 (2016).
425. LeConte, B. A. *et al.* Differences in the viral genome between HPV-positive cervical and oropharyngeal cancer. *PLoS One* **13**, e0203403 (2018).
426. Kmietowicz, Z. Boys in England to get HPV vaccine from next year. *BMJ* **362**, k3237 (2018).
427. Hilliard, M. HPV vaccination programme for boys to proceed in September. *The Irish Times* (2019). Available at: https://www.irishtimes.com/news/health/hpv-vaccination-programme-for-boys-to-proceed-in-september-1.3933725. (Accessed: 21st July 2019)
428. Libreri, S. HPV vaccine to be extended to boys. *RTE News* (2018). Available at: https://www.rte.ie/news/health/2018/1207/1015750-hiqa-hpv/. (Accessed: 21st July 2019)
429. HSE. About the HPV Vaccine - HSE.ie. *HSE* (2019). Available at: https://www.hse.ie/eng/health/immunisation/pubinfo/schoolprog/hpv/about/. (Accessed: 21st July 2019)
430. Health Information and Quality Authority (HIQA). PRESS RELEASE: HIQA advises changing to a more effective HPV vaccine and extending the vaccine to boys. *HIQA News Updates* (2018). Available at: https://www.hiqa.ie/hiqa-news-updates/hiqa-advises-changing-more-effective-hpv-vaccine-and-extending-vaccine-boys. (Accessed: 27th February 2019)
431. Health Information and Quality Authority. *HTA of extending the HPV vaccination to boys*.

(2018).

432. Wang, C. *et al.* Targeting Head and Neck Cancer by Vaccination. *Front. Immunol.* **9**, 830 (2018).

433. Miles, B., Safran, H. P. & Monk, B. J. Therapeutic options for treatment of human papillomavirus-associated cancers - novel immunologic vaccines: ADXS11–001. *Gynecol. Oncol. Res. Pract.* **4**, 10 (2017).

434. Lui, V. W. Y. & Grandis, J. R. Primary Chemotherapy and Radiation as a Treatment Strategy for HPV-Positive Oropharyngeal Cancer. *Head Neck Pathol.* **6**, 91–97 (2012).

435. Mehanna, H., Olaleye, O. & Licitra, L. Oropharyngeal cancer – is it time to change management according to human papilloma virus status? *Curr. Opin. Otolaryngol. Head Neck Surg.* **20**, 120–124 (2012).

436. Garland, S. M. & Smith, J. S. Human Papillomavirus Vaccines. *Drugs* **70**, 1079–1098 (2010).

437. Lu, B., Kumar, A., Castellsagué, X. & Giuliano, A. R. Efficacy and Safety of Prophylactic Vaccines against Cervical HPV Infection and Diseases among Women: A Systematic Review & Meta-Analysis. *BMC Infect. Dis.* **11**, 13 (2011).

438. Health Service Executive. Cervical Check. *Health Service Executive* (2019). Available at: https://www.hse.ie/eng/cervicalcheck/. (Accessed: 19th May 2019)

439. Wentzensen, N., Schiffman, M., Palmer, T. & Arbyn, M. Triage of HPV positive women in cervical cancer screening. *J. Clin. Virol.* **76**, S49–S55 (2016).

440. Taberna, M. *et al.* Human papillomavirus-related oropharyngeal cancer. *Ann. Oncol.* **28**, 2386–2398 (2017).

441. International Agency for Research on Cancer & World Health Organization. HPV-AHEAD Study. *International Agency for Research on Cancer* (2019). Available at: http://hpv-ahead.iarc.fr/about/index.php. (Accessed: 27th April 2019)

442. Safaeian, M. *et al.* Comparison of the SPF10-LiPA system to the Hybrid Capture 2 Assay for detection of carcinogenic human papillomavirus genotypes among 5,683 young women in Guanacaste, Costa Rica. *J. Clin. Microbiol.* **45**, 1447–54 (2007).

443. Hesselink, A. T. *et al.* Comparison of GP5+/6+-PCR and SPF10-Line Blot Assays for Detection of High-Risk Human Papillomavirus in Samples from Women with Normal Cytology Results Who Develop Grade 3 Cervical Intraepithelial Neoplasia. *J. Clin. Microbiol.* **46**, 3215–3221 (2008).

444. WHO | Prevalence of tobacco smoking. *WHO* (2016).

445. Ritchie, H. & Roser, M. Alcohol Consumption. *OurWorldData* (2019). Available at: https://ourworldindata.org/alcohol-consumption.

446. Kleter, B. *et al.* Novel Short-Fragment PCR Assay for Highly Sensitive Broad-Spectrum

Detection of Anogenital Human Papillomaviruses. *Am. J. Pathol.* **153**, 1731–1739 (1998).

447. Kleter, B. *et al.* Development and clinical evaluation of a highly sensitive PCR-reverse hybridization line probe assay for detection and identification of anogenital human papillomavirus. *J. Clin. Microbiol.* **37**, 2508–17 (1999).

448. Schmitt, M. *et al.* Bead-based multiplex genotyping of human papillomaviruses. *J. Clin. Microbiol.* **44**, 504–12 (2006).

449. Gheit, T. *et al.* Development of a sensitive and specific assay combining multiplex PCR and DNA microarray primer extension to detect high-risk mucosal human papillomavirus types. *J. Clin. Microbiol.* **44**, 2025–31 (2006).

CHAPTER 2

REVIEW OF THE METHODS FOR HPV DETECTION: RE-DEFINING 'HPV POSITIVITY' AND A CASE STUDY OF HEAD AND NECK CANCERS

2 CHAPTER 2: REVIEW OF THE METHODS FOR HPV DETECTION: RE-DEFINING 'HPV POSITIVITY' AND A CASE STUDY OF HEAD AND NECK CANCERS

2.1 Introduction

HPV is the most common sexually transmitted infection[1]. It is an epitheliotropic, non-enveloped DNA virus consisting of 8,000 kilo-base pairs[2]. The virus is a leading cause of cancer worldwide[1]. HPV16, 18, 31, 33, 34, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68, 73, and 82 are classified as HR carcinogens[3]. HPV6, 11, 40, 42, 43, 44, 54, 61, 72, 81, and 89 are designated as LR carcinogens[3].

HPV's origins as a carcinogen were first unveiled in its relationship to cervical cancer[4], the oncogenesis of which is well documented[5]. However, HPV is now understood to be a carcinogen in many other types of disease including anal, vulval, and vaginal cancers[6,7]. HNCs, first linked to the virus in the 1980s[8,9], are emerging as the second most common form of HPV-related cancer.

Despite its ever-clarified role in the etiology of non-cervical cancers, the majority of the technologies available for detecting HPV in human tissue have been optimized for the clinical analysis of fresh cervical samples for the purposes of triage. Particularly, techniques have been developed for the testing of cytology-based sampling. Indeed, HPV's role in cervical cancer screening is revolutionizing the entire screening process itself[10]. However, this population of cervical patients is becoming an increasingly narrower proportion of the rapidly expanding population of HPV-related cancers[11]. The majority of the available technologies are therefore not developed for testing samples, fresh or otherwise (e.g. FFPE, frozen), from the ever-emerging HPV-related disease sites including the head and neck, and other ano-genital regions.

Furthermore, the methods available for detecting HPV in non-cervical, non-fresh samples are entirely heterogeneous. This is firstly due to the multiple indicators (e.g. DNA, mRNA, viral load, integration, p16) of HPV's presence in a cell. Second, there are numerous technologies developed to detect each individual indicator. Each of these technologies is

also based on different theoretical assumptions, principles, and chemistries. Third, the presentation of results for each technology is different. For some, an HPV positive case is simply defined by the lowest limit of detection of the technology in a binary manner[12–14], and for others, results are quantitatively determined with cut-offs set to distinguish negative from positive cases[15].

A review of the literature showcases this heterogeneity. In some studies, HPV DNA detection is tantamount to designating a case HPV positive for HPV DNA using PCR and gel electrophoresis[3,12,13]. In others, the detection of HPV E6/E7 mRNA, or HPV DNA and other surrogate biomarkers including p16 classifies an HPV positive case[16], either individually or paired together.

What is also important to note in the case of emerging HPV-related diseases, HNC in particular, is that technologies available for detecting HPV in all sample types have varying implications for the significance of virus detection. If HPV DNA is detected in an HNC sample, for instance, there is no guarantee that this infection has a causal relationship to the HNC in question. It may just be a transient infection that happens to be present at the time of sample collection. This is especially significant for HNCs in the clinical context because distinguishing transient from clinical infections is equivalent to identifying patients with better predicted overall response to treatment[3,17,18].

Thus, the majority of available HPV-detecting technologies have been validated using different clinical end points for the most part in cervical cytology samples. What constitutes a HPV positive case outside the cervical clinical context has consequently never been explicitly defined. This has large implications for the value of HPV testing in non-cervical, non-fresh, and non-clinical scenarios. For example, determining what constitutes a HPV positive case for future screening tools developed for HNC in tandem with developed vaccines will be tantamount to creating an HPV-based triage and monitoring system to prevent these cancers in the first place. Conversely, defining what constitutes a HPV positive case outside the clinic in the research context is crucial if reflective epidemiological indicators like prevalence, incidence, and risk factors are to be established.

On the basis of the heterogeneous landscape of HPV-detecting technologies and emerging non-cervical HPV-related cancers, especially HPV-related HNCs, this Chapter explores three topics. In the first instance, it reviews the different methods and technologies currently used to test for HPV in cervical samples and other HPV-related diseases. In the second instance, it deconstructs the heterogeneous state of the definition of 'HPV positivity' in the literature and suggests the revolutionizing of the current approach. It posits that there should not be a singular definition for 'HPV positivity' across all research and clinical contexts. Instead, it states that there should be heterogeneous definitions of what constitutes an HPV 'positive' case. However, these definitions should vary based on the contexts that require them and, within each context, they should be homogenous. The article thus conceives of a necessary systematic structure for determining what standardized definition of 'HPV positivity' is applicable to particular contexts.

In the third and final instance, the Chapter uses HNC as a case study for the application of this systematic structure. It ultimately concludes that this structure, or one based on similar logic, is crucial should the emerging uniqueness of HPV-related disease, especially that of HNC with respect to survival and potential prevention, be capitalized on, and the maximal number of lives be saved.

2.2 HPV Indicators and the Technologies That Detect Them

As a virus, HPV can be detected using several different indicators and sample types. These indicators include HPV DNA, HPV mRNA, HPV viral load, and HPV integration into the host cell genome. Each indicator is also associated with various different technologies, based on different chemical and biological principles. Figure 2.1 is a summary of the majority of the available HPV-detecting technologies, the principles upon which they are based, the indicators that they detect, and the sample types for which they are used.

HPV Indicator	Principle	Technology Type/Name	Optimized for…
HPV DNA			
	Nucleic acid hybridization	Southern blot	Cytology samples, FFPE tissue, frozen tissue
		In-situ hybridization	FFPE tissue, frozen tissue
		Dot-blot hybridization	Cytology samples, FFPE tissue, frozen samples
	Signal amplification	Digene® HPV (hybrid capture 2)*	Cytology samples
		CareHPV*	Cytology samples
		Cervista® 16/18*, Cervista® HPV HR*	Cytology samples
	Nucleic acid amplification	PCR: Many types including pure PCR and real-time PCR, key examples of which are detailed in the below rows	Cytology samples, FFPE tissue, frozen samples
		Gel electrophoresis, Microarray analysis (post-PCR)	Cytology samples, FFPE tissue, frozen samples
		Papillo Check®*	Cytology samples
		COBAS® 4800 HPV *	Cytology samples
		Abbott RealTime High Risk*	Cytology samples
		BD Onlcarity HPV* Assay	Cytology samples
		Xpert HPV	Cytology samples
		CLART® HPV 2	Cytology samples, FFPE tissue
		INNOLiPA®	Cytology samples, FFPE tissue, frozen tissue
		Linear Array®	Cytology samples, FFPE tissue, frozen tissue
		Clinical Arrays® HPV	Cytology samples, FFPE tissue, frozen tissue
		Colorimetric Hybridization Assay (MCHA)	FFPE tissue, frozen tissue
		Genome sequencing (dideoxy-fluoro-integrated Sanger sequencing)	Cytology samples, FFPE tissue, frozen tissue
HPV mRNA			
	Hybridization	mRNA In-situ hybridization	FFPE tissue, frozen tissue
	Transcription-mediated amplification	PreTect® HPV Proofer	Cytology samples
		APTIMA® HPV E6/E7	Cytology samples
		HPV E*6 Assay	FFPE tissue
	Monoclonal antibodies	AVantage® HPV E6	Cytology samples
HPV Viral Load			
	Amplification	Various PCR tests including real-time PCR	Cytology samples, FFPE tissue, frozen tissue
HPV Integration			
	Amplification	Various PCR tests including pure and real-time PCR	Cytology samples, FFPE tissue, frozen tissue
	Molecular methods	Fluorescence in-situ hybridization	Cytology samples, FFPE tissue, frozen tissue

Figure 2.1 Summary of the types of HPV indicators (HPV DNA, HPV mRNA, HPV viral load, and HPV integration), the principles used to detect them, and the technologies that are currently available on the basis of these principles.

*Sample types for which these technologies are optimized are also listed. 'Cytology samples' refers to liquid-based cervical cytology samples and/or swabs. * indicates a test that is clinically validated.*

The coming section elaborates on the indicators of HPV and their associated detection platforms summarized in Figure 2.1.

2.2.1 HPV DNA

As showcased in Figure 2.1, HPV DNA is the most widely used target for identifying the virus in samples. There are three central approaches for detecting HPV DNA for which there are multiple different technologies. These three techniques are nucleic acid hybridization, signal amplification, and nucleic acid amplification[19]. The majority of these are optimized for cytology samples (swabs included), and all clinically validated tests for the virus fall within this DNA category for cervical cytology samples alone.

2.2.1.1 Nucleic Acid Hybridization and Signal Amplification

Nucleic acid hybridization assays are all based on nucleic probe technology. They include in-situ hybridization and dot-blot hybridization. These can be used to test for HPV DNA in fresh cervical cytology samples, FFPE tissue, and frozen tissue. Signal amplification is another molecular method for detecting HPV DNA. It forms the basis of the Digene® HPV platform with hybrid capture 2 (hc2) technology (Qiagen Ltd, Manchester, UK)[20] and the Cervista® HPV HR assay (Hologic®, MA, USA)[21]. These are both optimized for cervical cytology samples and are clinically validated.

The Digene® HPV test is an in-vitro microplate assay based on signal-amplified nucleic acid hybridization, whereby an antibody enzyme conjugate and antibody are paired with a chemo-luminescent substrate should hybridized viral DNA be present[20,22]. Chemo luminescence is indicative of a positive result. hc2 is optimized for detecting 13 HR and 5 LR HPV types including HR HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68, and LR HPV6, 11, 42, 43, and 44[20]. It functions as a cervical cancer-screening test as it is highly sensitive and gives a high negative predictive value[22]. However, the test has a low specificity and it can only be used for testing fresh cervical samples. Additionally, it cannot identify the

HPV type or whether one or more HPV types are present in one sample. It also does not assign a quantitative value to the level of luminescence intensity. This is particularly important in the context of cervical screening given that HPV16 and 18 determine a risk of pre-cancerous lesion over 5 times higher than the risk of all other HR types combined[19].

The Cervista® HPV HR (and Cervista® HPV 16/18) is another signal amplification method that is analytically and clinically validated as an in-vitro diagnostic test for the qualitative detection of 14 HR HPV types in cervical specimens including HPV16, 18, 31, 33, 34, 39, 45, 51, 52, 56, 58, 59, 66, and 68[21]. Invader chemistry forms the basis of the Cervista® tests which uses two types of isothermal reactions that occur simultaneously[22]. The first reaction occurs on the targeted DNA sequence and a secondary reaction produces a fluorescent signal. The instrument itself has an internal control that reduces false negatives produced by a low of number cells[22]. The Cervista® platforms give results that are highly reproducible and sensitive, and the internal control to confirm sample quality makes it a good competitor in the clinical context[22]. However, like the hc2, the technology cannot determine specific HPV type, nor can it be used on any other sample type but those of the fresh cervical variety. Furthermore, cross-reactivity of two HPV types of unknown risk and the false negative risk in cases with low levels of infection or sampling error are significant disadvantages of the platforms.

2.2.1.2 Nucleic Acid Amplification

Nucleic acid amplification methods are based on the extraction of DNA from relevant tissue, followed by an optimized PCR. The PCR product can then be treated in different ways. Often, the PCR product is subjected to gel electrophoresis alongside positive and negative controls to determine HPV status. Microarray analysis is an alternative that uses probe amplification and further hybridization of the PCR product onto a chip that is visualized with a DNA chip scanner[23]. The chip method shows increased sensitivity and specificity in comparison to gel electrophoresis, and gives better results than DNA sequencing[24]. Other amplification methods are pure PCR, PCR-RFLP, real-time PCR, Abbott real-time PCR, Papillo Check® (Greiner Bio-One, Kremsmünster, Austria)[25], and the newer COBAS® 4800 HPV Test (Roche Diagnostics, Basel, Switzerland)[26]. Pure PCR, PCR-RFLP, real-time PCR, and Abbott

real-time PCR can be used for cytology samples, FFPE tissue, and frozen tissue. Papillo Check® and the COBAS® 4800 HPV Test are both optimized for fresh cytology samples and clinically validated.

PCR-based techniques are generally highly sensitive, using consensus primers including PGMY09/PGMY1 and GP5+/GP6+ which amplify a large number of HPV genotypes in a single reaction[19]. These mainly target highly conserved parts of the HPV genome like the L1 capsid gene. However, it has been remarked that in samples where copy number is very low, PCR's reliance on other technologies, including the aforementioned microarray analysis and gel electrophoresis, for the visualization of enough amplified DNA may underestimate positivity. Furthermore, in samples where there are multiple HPV infections, competition for PCR reagents may result in inaccurate assessments of present genotypes[19]. There may in fact be much stronger amplification of one of the HPV genotype sequences present. The range of genotypes that can be detected is also varying, and their specific identification depends on the PCR technique.

Papillo Check® detects 24 HPV types including HPV6, 11, 16, 18, 31, 33, 34, 39, 40, 42, 43, 44, 45, 51, 52, 53, 55, 56, 58, 59, 66, 68, 70, 73, and 82[25]. The assay is a Multiplex PCR using fluorescent primers to amplify a fragment of the E1 gene, comprising 28 probes, each in 5 replicate spots fixed on a DNA chip[25]. It distinguishes between HR and LR genotypes and is reliable in the case of multiple infections.

The COBAS® 4800 is one of the most widely used HPV DNA tests in the clinic for liquid-based cervical cytology samples. It is a combination of automated sample preparation with real-time PCR technology to detect 14 HR HPV types including 12 types (HPV31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68) and to specifically report on HPV16 and 18. Beta-globin is used as the control for extraction and amplification adequacy, and data is extremely reproducible[27]. The COBAS® 4800 fulfills international guidelines for validated screening, and is approved for ASC-US triage[28]. The system is particularly efficient, carrying out 96 tests in approximately five hours[26,29]. There is also no cross-reaction with non-carcinogenic genotypes. Furthermore, the operator has minimal contact with the samples, preventing contamination and reducing processing and work time, and errors due to fatigue[22].

Limitations mentioned by the manufacturer include that testing needs to be done by personnel with experience in PCR techniques and with the COBAS® HPV test system itself. The test can also only be used with liquid-based cervical cytology samples, and the presence of PCR inhibitors, as well as low virus copy numbers, may cause false negatives or invalid results[22].

The Abbott RealTime HR HPV assay (Abbott Molecular, IL, USA)[30] detects 14 HR HPV types (HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, and 68) and reports on HPV16 and 18 separately from other HR types. Like the COBAS® 4800, this system is an automated in vitro qualitative process that is clinically validated according to international requirements for use in screening of fresh cervical cytology samples in women starting at age 30[22]. Internally, the system consists of an m2000sp instrument that prepares the nucleic acid and an m200rt analyzer that carries out real-time PCR using a mixture of multiple primers and probes for amplification and detection of HR HPV DNA. The beta-globin gene is used as an internal quality control[30]. Slightly slower than the COBAS® 4800, it processes 96 samples in six to eight hours depending on the DNA extraction method used[30].

Two Real-Time PCR-based tests target the HR HPV E6/E7 oncogenes. The BD HPV Assay (Becton, Dickinson, and Company, MD, USA)[31] is clinically validated for fresh cervical cytology samples and provides information for six HPV types (16, 18, 31, 45, 51, and 52) as well as detection of 14 HR HPV types. The system, which also utilizes an internal quality control, is automated and can process 1 to 30 samples per run, giving 120 results per patient per day including genotype. The test performs as well as other tests approved by the FDA including hc2[22]. The Xpert HPV test (Cepheid, CA, USA)[32] detects DNA encoding for E6/E7 oncoproteins of 14 HPV types (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, and 68) in liquid-based cervical cytology samples. It has the capacity to process one test at a time, with up to 80 tests in one hour[32].

2.2.1.2.1 Genotyping Technologies

Several nucleic acid amplification technologies have been developed to provide information on a wide scope of individual HPV DNA genotypes present in a sample. HPV genome

sequencing is one which relies on dideoxy, fluoro-integrated Sanger sequencing and the technological advances that have led to throughput sequencing[33,34]. Though it is not yet validated for clinical use, genome sequencing is less expensive than other methods and is also more efficient, reading out a sequence rather than projecting a fluorescent signal that must then be converted to a sequence. This financial efficiency results from the approach's sequence-by-synthesis process where a DNA sequence is read in real-time and synthesized by the addition of inexpensive and unlabeled nucleotides[35]. Furthermore, genome sequencing methods can be used on a wide variety of sample types including fresh cervical cytology samples, FFPE tissue, and frozen tissue, along with blood, saliva, cell lines, plasma, serum, and whole genome-amplified DNA.

The CLART® HPV 2 (Genomica, Madrid, Spain)[36] is another platform that details HPV genotypes present in a sample. It uses biotinylated primers that amplify a fragment within the L1 region of the HPV genome. Amplicons are then detected by hybridization in a low-density microarray which has probes for 35 HPV genotypes including HPV6, 11, 16, 18, 26, 31, 33, 35, 39, 40, 42, 43, 44, 45, 51, 52, 53, 54, 56, 58, 59, 61, 62, 66, 68, 70, 71, 72, 73, 81, 82, 83, 84, 85, and 89[36]. The platform can be used with fresh cytology samples and has also been employed with FFPE tissue.

The INNOLiPA® assay (Fujirebio, Dublin, Ireland)[37] genotypes based on the amplification of a 65bp region of the L1 gene using biotinylated SPF10 primers. It does not however detail HPV35, 39, 52, 56, and 66, and appears to be ineffective for genotyping HPV42 and 59[37]. Genotypes are visualized on strips and assessed based on the associated INNOLiPA® results guide. FFPE tissue, frozen tissue, and liquid-based cytology samples can be used with the INNOLiPA®. It is particularly good at dealing with the complexities of processing FFPE tissue as it is a highly sensitive technology. However, due to its expensive nature and laborious protocol, it is often used after binary HPV status (positive or negative) has already been determined by another method. Its higher sensitivity can therefore be undercut as a result.

The Linear Array® Genotyping (Roche Diagnostics, Basel, Switzerland)[38] platform is another PCR-based assay coupled with a reverse line blot hybridization. It discriminates between 36 types of HPV including HR HPV16, 18, 31, 33, 35, 36, 39, 45, 51, 52, 53, 56, 58, 59, 66, 68, 73,

and 82, along with LR types HPV6, 11, 40, 42, 54, 61, 70, 72, and 81[38]. It also discriminates between other risk-as-yet-undetermined genotypes HPV55, 62, 63, 67, 69, 72, 83, and 83. It makes use of biotinylated PGMY09/11 primers to amplify a portion of the L1 region[38]. Similar to the INNOLiPA®, these are then visualized by coloured signals on strips assessed according to the associated results guide. The platform can be used with liquid-based cytology samples, FFPE tissue, and frozen tissue.

The Clinical Arrays® Assay (Genomica, Madrid, Spain)[39] uses absorption columns, employing biotinylated primers to define a sequence of 451 nucleotides within the L1 region. It detects HR HPV16, 18, 26, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 68, 70, 73, 82, and 85 as well as LR HPV6, 11, 40, 42, 43, 44, 54, 61, 62, 71, 72, 81, 83, 84, and 89[39]. The Clinical Arrays® is less widely employed than its alternatives due to lower sensitivity. This said, it can be used with a wide variety of tissue types including fresh cytology samples, FFPE tissue, and frozen tissue.

Finally, Microplate Colorimetric Hybridization Assays (MCHA) can identify HR HPV16, 18, 31, 33, 39, and 45. It uses PCR to amplify a fragment within the L1 region using GP5+/6+ primers accompanied by colorimetric hybridization on to type-specific probes on microwell plates[19]. It appears to show higher sensitivity than other technologies, specifically Papillo Check® for HPV16 and 18. However, it is generally optimized dependent upon the tissue type used (FFPE or frozen) and sensitivities and specificities can vary.

2.2.1.3 Challenges for All HPV DNA Platforms

Though each HPV DNA-detecting platform described above has its advantages and disadvantages, all of them are united by several faults. None of these technologies have been evaluated in HPV-vaccinated individuals. Their detection of HPV DNA depends on the number of viral copies in the specimen, and false positives can also occur with LR HPV.

In addition to this, the aforementioned technologies are for the most part optimized for HPV detection in fresh cervical cytology samples, with the exception of some nucleic acid hybridization and amplification techniques. PCR methods in particular have been re-

optimized to deal with FFPE and frozen tissue for research, especially for the head and neck[12,13]. However, sensitivities and specificities vary based on tissue type, and none of these have been clinically validated.

Indeed, the field of HPV-related disease outside the cervix is still an emerging one, and one mostly confined to research. This said, in-situ hybridization has already been in use in the clinic in the preliminary assessment of HNC patients, for instance. Enough evidence has therefore accrued to support HPV's role in the carcinogenesis of HNCs. However, HPV DNA technologies have not yet been developed to: Fulfill the clinic's need to distinguish between clinically relevant and transient HPV infections; assume their role in potential HPV-based screening for HNCs; and appropriately deal with non-fresh sample types like FFPE or frozen tissue in this context.

2.2.2 HPV mRNA

HPV's proteins E6 and E7 are known as "oncoproteins", as they are central to the carcinogenic properties of the HR virus types. The detection of associated mRNA transcripts is significant as it indicates a transforming HPV infection if HR HPV genotypes are identified. This is the key difference between HPV DNA and mRNA tests. Where the former is only proof of the presence of the virus in tissue, mRNA is indicative of an HPV infection that plays a causal role in the onset of the associated cancer.

With respect to HPV mRNA detection, technologies for fresh cervical clinical samples focus on E6/E7 mRNA and can be considered a marker for diagnosis in pre-cancerous lesions by HPV[40]. The chemistry of most mRNA tests is based on transcription-mediated amplifications of full-length E6/E7 transcripts pre-empted by target capture[19]. The PreTect® HPV Proofer (PreTect, Klokkarstua, Norway)[41] and the APTIMA® assay (Hologic, MA, USA)[42] are the two main technologies used for mRNA detection and are both optimized for fresh cervical cytology samples.

Though based on similar principles, the APTIMA® assay is becoming a more widespread technology in research than the PreTect®, given that its limit of detection is lower than the

limits reported for other tests[19], and because the PreTect® predates it and detects 5, rather than 14, HR HPV types. The APTIMA® assay detects HPV E6/E7 mRNA for HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, and 68[42]. Further genotyping for HPV16, 18, and 45 can be achieved using the APTIMA® Genotyping Assay. It is a fully automated processing system that involves three main steps, all of which take place in a single tube. These steps begin with target capture, followed by target amplification, and detection of the amplification products by the hybridization protection assay[42]. The assay incorporates an internal control to monitor all steps, as well as operator or instrument error. The system is efficient, and can carry out up to 250 tests in approximately five hours[42]. It was also approved by the FDA in 2011 for cervical screening in women starting at age 30, in combination with Pap smears[22].

Several mRNA assays have been developed for non-cervical samples that may function with both FFPE tissue and frozen tissue. The most sensitive of these is the E*6 mRNA assay developed by Halec et al.[43], optimized for FFPE tissue of HNCs. Though not yet tested, it follows that the assay may work with tissue from other anatomical regions and could also be optimized for use with frozen tissue.

Much like all DNA tests, all mRNA detection platforms have several limitations, including that: The tests have not been evaluated in HPV-vaccinated individuals; detection of HR HPV mRNA depends on the number of copies in the specimen; and false positives can occur with LR HPV. This said, the APTIMA® in particular has been shown to have a higher specificity than most DNA-based tests. It is however still unclear if mRNA testing alone is in any way an improvement upon HPV DNA detection for liquid-based cervical samples in HPV-based cervical screening, let alone for any other kind of HPV-related disease for which screening tools and approaches are far less developed.

For HNCs in particular, mRNA detection is not clinically validated given that it is still an emerging indicator of HPV-positivity in this disease type. The relationship between HPV mRNA status and clinical outcomes is less well established in the literature than for HPV DNA, p16, and the two combined. It is also known that some cases of HNC have characteristics of both HPV positive and smoking-related carcinogens, so the detection of

HPV mRNA may not be a foolproof way to determine the origins and best treatment protocols for a particular HNC.

2.2.3 HPV Integration

Part in parcel of HPV's life cycle is the integration of the virus, and thus viral DNA, into the human host cell genome[44]. Detection of an integrated virus, rather than one that is simply episomal, may be more indicative of a transforming infection if the genotype detected is HR. This is especially true since integration only occurs after successful active infection.

The main methods used for HPV integration detection are PCR, real-time PCR, and fluorescence in-situ hybridization, all of which can be used with fresh samples, FFPE tissue, and frozen tissue. Real-time PCR allows for calculation of the ratios between the levels of E2, and E6/E7 genes. When there is HPV integration, the viral genome shows a 1:1 ratio[45].

PCR is a widely available tool and is a relatively efficient method for detecting integrated viral DNA, even in the context of available sequencing. However, PCR provides no mechanism for determining the site of integration, or the unknown carcinogenic implications of that integration. In addition to this, it does not allow the distinguishing of extrachromosomal or episomal forms of HPV, and the pure integrated form of the virus[46].

The use of integration as indicative of a carcinogenic HPV infection is also conceptually misleading. The assumption underlying detecting integrated HPV DNA is that all integrated HPV is also successfully replicating, and more specifically, replicating viral oncogenes. These assumptions are not necessarily true and make integration a risky test in the clinical diagnostic and screening contexts. It is also documented in many studies that for HNC in particular, integration has no relationship to carcinogenic activity and that instead, most clinically relevant HPV infections are episomal[47,48]. Thus, where integration may not be appropriate for clinical use, it is also an inappropriate test for population-level studies given that epidemiological endeavors are not just interested in integrated DNA.

2.2.4 HPV Viral Load

Once a virus is integrated and replicating, it is logical to assume that assessments of HPV viral load would indicate the extent or success of an active HR infection and its carcinogenic potential. Methods for assessing viral load are generally limited to a variant of PCR, which must be quantitative[49,50]. Indeed, some studies using PCR show that viral load declines in response to therapy, and provides an acceptable detection mechanism for fresh samples, FFPE tissue, and frozen tissue[51].

However, many others show no added value over pure cytology in the case of fresh cervical samples, and that testing for high-load levels may not be clinically useful except in the case of HPV16[51–53]. Even in the cervical case, there is little evidence to suggest that there are persistent and graduated differences between viral load in CIN I, II, and III lesions[51]. Most studies conclude that though increases in copy-number are associated with an increased risk of abnormality, "a single measurement of viral load made at an indeterminate point during the natural history of HPV infection does not reliably predict the risk of acquiring cervical neoplasia"[50].

The same is true in the emerging field of HPV-related HNCs, where viral load does not seem to correlate with an active HPV infection[54,55]. This may indicate that the simple fact, rather than the quantity, of the virus' presence may be enough to implicate carcinogenic activity. Thus, the indicator's value for research and clinical purposes overall is likely not as great as that of HPV DNA and mRNA.

2.3 Re-defining 'HPV Positivity'

2.3.1 Three-tiered Heterogeneity of the Current Definitions for 'HPV positivity'

The preceding review of the methods for detecting HPV reveal three tiers of heterogeneity in defining an HPV positive case in any given study in the current literature. The first tier is the type of HPV material targeted: DNA, mRNA, integrated DNA, or viral load. The second is the technology used to detect that single type of HPV indicator, as there are several platforms that can be used to detect the same material. These platforms use heterogeneous

principles and mechanics. They also have drastically different sensitivities and specificities, making each one optimal for different detection purposes. Third, the presentation of results varies. Some platforms give quantitative results for which cut offs must be set. Others define positivity in a more binary manner, identifying a positive case as one containing any extent of HPV material. An HPV positive case is therefore determined solely by the lowest limit of detection of the given technology. Thus, what constitutes 'HPV positivity' is currently a moving target.

2.3.2 The Consequences of Three-tiered Heterogeneity

This three-tiered variation is problematic for a multiplicity of reasons. The first is a resulting inability to meaningfully compare the results of studies. If there is no standard manner with which to determine an HPV positive case, then the seemingly agreeing or powerfully diverging results across the literature are not as impactful as they could be, or even comparable in the first place. For example, in the case of HPV-related HNCs, the prevalence of the virus in seemingly similar European and North American populations can vary from anywhere between 18% to 90%[54,56–62]. Whether this range is due to the heterogeneity in technologies used to define an HPV positive case, or genuine population-level variance is unclear. However, the difference between sensitivities of technologies, and the HPV indicators used to define 'HPV positivity' undoubtedly contribute to this enormous scope.

Second, that the majority of detection technologies were developed and validated using different clinical end points and for the most part with liquid-based cervical cytology samples makes it difficult to distinguish between clinically relevant and transient infections in other sample types from other anatomical regions. This is especially pertinent in the case of HNCs where the determination of a carcinogenic infection in surgically removed tissue in FFPE blocks or frozen samples could mean the difference between life and death, or a life with or without long-term side-effects for patients. As HPV is emerging as a predictor of survival in HNC[17,18,63–65], patients may also suffer from unnecessary anxiety regarding their cancer's origins, or false hope regarding their outcome. Furthermore, the potential for screening tools to prevent HNC may require another definition of 'HPV positivity' more akin

to that used in screening for cervical pre-cancer, especially given the differing sampling technique that this could involve.

Third, enormous variation in what constitutes an HPV positive case means that conclusions drawn from research cannot always be translated to the clinical context with certainty, and vice versa. One emerging example of this in the case of HNCs is the possibility of the de-escalation of treatment for HPV-related HNSCC given their overall improved survival compared to HPV-unrelated HNSCC. On the basis of the evident role of the virus in HNSCC, many trials are exploring what indicators of HPV-positivity can robustly identify a carcinogenic infection in an HNSCC that will, as a result, respond just as well to lower doses of harsh treatments (see Table 1.5)[65–67]. Results are not resounding for all types of HPV indicators or any other associated biomarkers, most notably p16. Whether or not evidence from the research context can actually be successfully applied in the clinic is thus still unclear.

Finally, the enormous variation in HPV detecting technologies not only puts into question what platforms should be used in particular contexts, but also highlights that the majority of those optimized for non-cervical samples are not particularly cost effective. This is especially true in research, meaning that their current potential for use in the clinic *en masse* is in jeopardy.

2.3.3 Re-defining 'HPV Positivity': Further Diversification Before Standardization

At first glance, it is logical to assume that the currently wide range of available HPV-detecting technologies should be narrowed down to standardize the definition of an HPV positive specimen and to subsequently eliminate the consequences of heterogeneity discussed in Section 2.3.2. This is not the case. There should be **heterogeneous** definitions of HPV positivity based on varying technology types. These definitions should be defined by the **contexts** that require them (e.g. cervical screening, penile cancer epidemiology, HNC triage), which are becoming more diverse every day. Further diversification of the definitions for 'HPV positivity' should therefore be expected as the number of contexts demanding them grows.

However, **within** these present and multiplying **contexts**, there should only be one **standardized** and **homogenous** definition of an HPV positive case. In this way, the challenges associated with heterogeneity (Section 2.3.2) are minimized given that definitions are optimized for particular contexts but are entirely comparable and reliable within them. This is already the current protocol for the developed screening programs for cervical pre-cancer, where only particular technologies have been validated for country-wide screening programs that have recently added HPV-testing to their triage. For instance, in the United States, the COBAS® 4800 is the standard approved technology used for this primary HPV cervical screening **context**.

Standardizing definitions of 'HPV positivity' outside the liquid-based cervical cytology context is difficult given that the currently available technologies are not themselves fully optimized and are still developing. This is simply a product of time and resources, where up until the last few years, cervical cancer has been the main target of HPV-related research. Platforms that are increasingly sensitive are still emerging, as is the role of biomarkers in addition to HPV-detection. This said, it is possible to identify the kinds of technologies that would be optimal for different contexts, taking into account sample types, anatomical sites from which samples are taken, and the research and/or clinical goals of the required test.

2.3.4 Re-defining 'HPV Positivity': A Suggested Systematic Structure for Delineating the Contexts and Standardizing the Definitions of 'HPV Positivity'

Figure 2.2 is a posited systematic structure of the way in which definitions of 'HPV positivity' can be diversified to fit particular contexts, and subsequently standardized to ensure comparability and reliability within these contexts. The figure references no particular technologies, but simply defines current and potential methods for deciding upon the appropriate definitions of 'HPV positivity'. It is by no means a summary of the current state of HPV's use in research and in the clinic but is instead a suggested method for delineating the contexts of 'HPV positivity' and what characteristics would be necessary for the singular technologies defining an HPV positive case within these contexts. These suggestions include

the use of biomarkers discussed in Section 2.3.5, many of which are not yet validated for use in the clinic.

Figure 2.2 posits that the emerging contexts demanding definitions of 'HPV positivity' can be stratified by two crucial characteristics. The first level of stratification is that of cancer site (orange bubbles in Figure 2.2). Though the viral origins of HPV-related cervical, head and neck, and ano-genital cancers are the same, their eventual carcinogenesis are extremely different[68–73]. As a consequence, they have distinctive molecular profiles and signatures resulting in singular optimal treatment methods, and survival rates. The greatest amount of variation in HPV-related cancers is thus between sites, rather than within them. What constitutes an HPV positive case therefore varies in the first instance by the site of tumour emergence.

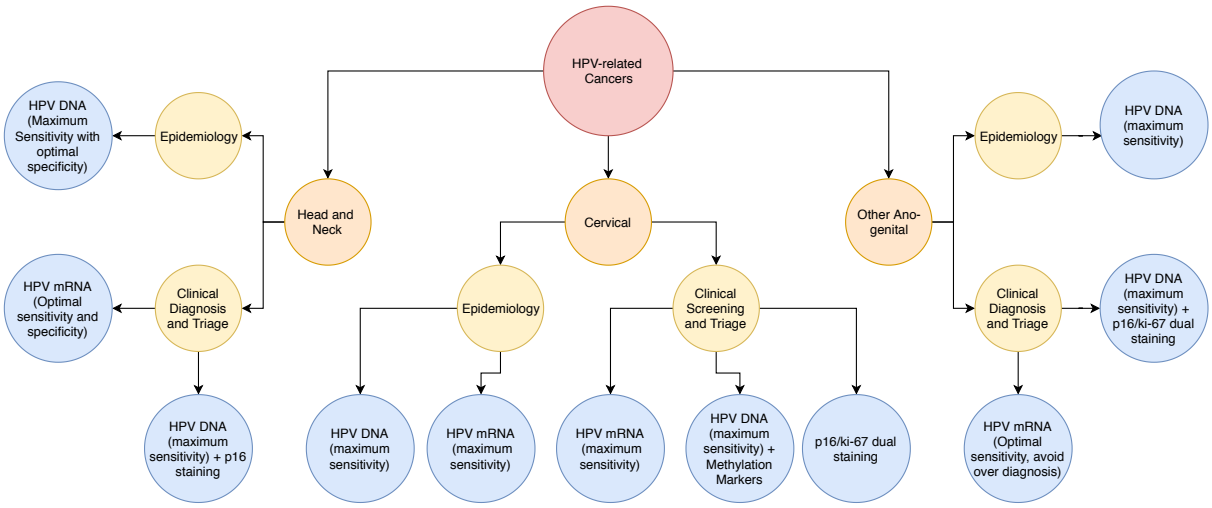

Figure 2.2 Posited systematic structure for delineating the contexts requiring standardized definitions for 'HPV positivity', and the suggested technologies and/or characteristics of technologies that should be used to create these definitions.

The first level of stratification is that of cancer sub-site, represented by the orange bubbles. The second level is that of the goals and constraints of the context itself, represented by the yellow bubbles. Goals include the hypothesis made, and the information and conclusions being sought. Constraints include the type of tissue being assessed. The resulting standardized definition for a context is represented by the blue bubbles. It should be noted that this schematic is a suggested method for identifying and/or developing the appropriate technologies for HPV detection in each context. It is not a finalized stratification for standardizing definitions of 'HPV positivity' but simply a visualization of what this might eventually look like when the spectrum of technologies evolves to allow it. Some proposed methods are thus purely projection and have not yet been validated for use in the clinic. For instance, methylation markers are still under evaluation as a triage mechanism to add to HPV-based cervical screening. p16/ki-67 dual staining is also emerging as a method just as valuable as cytology for assessing risk of pre-cancerous cervical lesions in screened women but has evidently not been clinically introduced.

The second level of stratification is defined by the goals and constraints of the context itself (yellow bubbles in Figure 2.2). The goals of the context include the hypothesis made, and the information and conclusions being sought. This may change the optimal sensitivities and specificities required of technologies and the kind of indicator that would achieve these goals best. The constraints are limited to the status of the tissue (e.g. fresh, frozen, FFPE). Both goals and constraints designate the optimal principles and technologies for a given HPV test. These circumstantial factors therefore determine the ultimate homogenous definition of 'HPV positivity' for one particular context (blue bubbles in Figure 2.2).

For example, if an HPV test is required for research purposes to determine HNC (first level of stratification: Orange bubble in Figure 2.2) epidemiological characteristics (second level of stratification: Yellow bubble in Figure 2.2), a highly sensitive test is needed to ascertain the presence of HPV in the population, regardless of the virus' true relationship to the cancer. It would be prudent to use HPV DNA as an indicator given that it designates the presence of the virus and does not define whether or not it is active and replicating. Furthermore, research and epidemiological work more often than not require the use of non-fresh samples, like FFPE tissue and frozen tissue. For that reason, HPV DNA tests in this context have to be optimized to deal with the difficulties of working with genetic material from paraffin-embedded and frozen specimens. The definition of a HPV positive case in this context might therefore be standardized as: An HPV DNA positive case detected by the most sensitive available technology with optimal specificity (blue bubble in Figure 2.2).

Thus, what constitutes an HPV positive case should indeed vary by cancer type and **context**. However, within one context, the definition of 'HPV positivity' should be **standardized** and **homogenous**. Homogeneity within a definite context will yield exacting definitions of 'HPV positivity', more powerful meta-analytical conclusions, and increased clinical effectiveness of HPV as a screening and triage tool.

However, this definitional precision can only be achieved with technologies and principles rigorous and unique enough to generate this level of clarity. This approach to defining 'HPV positivity' therefore demands further assessment of HPV indicators and their implications,

and the acknowledgement and exploration of non-HPV indicators. A shift in focus to biomarkers related to HPV and the extent of its activity (e.g. upregulated proteins, methylation markers) is necessary to realize this clear-cut contextual standardization. These related biomarkers have the potential to give far more accurate and useful definitions of 'HPV positivity' for the multiplicity of contexts that demand one.

2.3.5 Re-defining 'HPV Positivity': The Role of Biomarkers

Biomarkers represent the way forward for defining 'HPV positivity' given methodological heterogeneity, cost effectiveness, the varying carcinogenesis of different kinds of HPV-related cancers, as well as the diverging demands of both the research and clinical arenas. Biomarkers provide a cost-effective mechanism for standardizing HPV-based triage in cervical screening and distinguishing clinically significant from transient HPV infections in head and neck, ano-genital, and penile cancers. They consequently have the potential to achieve the heterogeneous goals of clinical and research-based HPV testing whilst simultaneously homogenizing the definitions for 'HPV positivity' in these different contexts.

2.3.5.1 p16/ki-67

The p16 protein is a tumour suppressor protein that in humans is encoded by the CDKN2A gene, colloquially known as the p16 gene. The protein plays an important role in cell cycle regulation by decelerating cell progression from the G1 phase to the S phase in mitosis[72]. It is a well-documented protein in the carcinogenic pathways of HNSCC, penile carcinoma, and cervical cancer. Most importantly, up-regulation of p16 is a known by-product of HPV's integration into the host human genome. It serves as an indicator of an active as opposed to a transient HPV infection[72].

ki-67 is another nuclear protein that is associated with cellular proliferation. The protein is present during all active phases of the cell cycle, including G1, S, G2, and mitosis, but is absent in quiescent cells. Dual staining of both p16 and ki-67 is commonplace in cervical cancer research[74,75], but ki-67's diagnostic role in other HPV-related cancers is not yet very well explored.

In the cervical case, p16 and the dual p16/ki-67 stain have proven promising biomarkers in the risk-based triage of cervical screening, though they have not yet become standard in screening programs. A study nested in a large Italian trial of HPV-based primary screening evaluated the performance of p16 for triage of HPV-positive women[76,77]. p16-positive women were at high enough risk of immediate colposcopy referral, while p16-negative women had low enough risk that they were judged in the Italian context to not need repeat testing for at least two years[76,77].

In another large study within the Kaiser Permanente Northern California (KPNC) Health System, the p16/ki-67 dual stain was evaluated for triage of HPV positive women[74]. Compared to cytology, the dual stain had both higher sensitivity and higher specificity with lower colposcopy referral rates. HPV-positive women testing positive for the dual stain were at high enough risk for immediate referral to colposcopy, while dual stain negative women could be retested at extended intervals, similar to the observations in the Italian study of p16 alone[75]. Similar results were found for triage of HPV-positive, cytology-negative women. That dual stain slides may eventually be evaluated in an automated manner may lead to further improvements of its performance[75,78,79].

In new unpublished data from Trinity College Dublin in Ireland, the p16/ki67 dual stain seems to be the most risk-sensitive triage test being trialed in an extensive set of biomarkers[80]. The dual stain's emerging ability to distinguish between categories of dysplasia is promising for building different definitions for 'HPV positivity' that strictly define screening referral structures.

p16 also has enormous potential in determining the definition of 'HPV positivity' in the case of HNCs. HPV-related HNCs, mainly arising in the oropharynx, show much better survival outcomes than their HPV-unrelated counterparts[18]. p16 up-regulation alone shows better survival in oropharyngeal cancers[81–85]. In fact, p16 immunohistochemistry has long been considered a surrogate biomarker for 'HPV positivity' and has been used by itself in the clinic as proof of HPV infection. However, it is now well-established that p16 over- and under-expression is not always related to HPV's integration and/or replication of

oncoproteins, and can simply be a consequence of normal cellular processes[86]. This is further complicated by the fact that hybrid HNC cases induced both by smoking/alcohol and HPV may not overexpress p16.

For this reason, the pairing of HPV DNA positivity with p16 staining has been proposed as a mechanism for identifying clinically relevant HPV infections in HNC[63,87]. This pairing confirms HPV's presence in the tissue *and* signals that the virus is integrated, replicating, and active. In the clinical context, identification of the true population of cancers caused by HPV has critical survival implications. By eliminating those cases with transient HPV infections or those that are smoking/alcohol-related, it will be possible to optimize the already-documented better survival in HPV-related cases. Treatment types that work best and minimize morbidity for HPV-related patients will then be identified and exploited. Though survival comparisons are not as well documented in ano-genital cancers, p16's clinical relevance is similarly important in HPV-related penile cancers[62,88].

2.3.5.2 Host Methylation Markers

Host methylation markers represent another promising method of detailing the stratification and implications of 'HPV positivity', especially in the cervical case, though they are not yet used in clinical practice.

Increased methylation of host genes has been observed in women with pre-cancer and cancer compared with those with acute HPV infection[75]. Several of these genes are candidates for triage of HPV-positive women in the clinical cervical context. CADM1, MAL, and miR-124-2 have been evaluated in cervical cancer screening studies in the Netherlands, demonstrating similar performance as cytology for triage of HPV-positive women[79,89]. New unpublished research regarding these three genes is also emerging from Ireland[80]. The increased post-test risk for pre-cancer in methylation positive and HR HPV positive women is a promising characteristic underlying the positive triage capacity of these markers[75].

It should also be noted that molecular testing for methylation markers offers some advantages over cytology triage. Molecular tests are not subjective, may offer higher

throughput, and can be conducted using a variety of specimen types, including self-sampling specimens[75]. The list of potential host methylation markers for triage of HPV-positive women keeps growing however, and the list of useful markers and optimal combinations of markers needs consideration. The same consideration is required for determining the best sample type, assay platform, and appropriate applications of the markers (e.g. triage of all HPV-positive women or triage of HPV positive, HPV16/18 negative women)[75].

2.3.5.3 HPV Methylation Markers and Other Markers

In the cervical case, HPV methylation markers may serve to further standardize what constitutes an 'HPV positive' case during screening. HPV genome-wide methylation studies of carcinogenic HPV types have demonstrated an increase of viral methylation associated with pre-cancer and cancer compared with HPV infection[75]. Across a number of carcinogenic types, a characteristic pattern has been observed with increased methylation particularly in the E2, L2, and L1 regions[75,90]. Individual CpG sites from these regions have shown discrimination between HPV infection and pre-cancer, suggesting that HPV methylation testing could serve as triage marker for HPV-positive women[90]. Furthermore, a combined methylation assay for HPV16, HPV18, HPV31, and three host genes has shown promise for the triage of HPV-positive women[90].

This said, further studies are needed, particularly in comparison to established markers, to assess the value of HPV methylation for cervical screening triage[75]. If successful, an HPV methylation assay that covers the majority of carcinogenic types would yield combined HPV test results, HPV genotyping, and HPV methylation results in a single assay that would provide risk stratification for placing women in all but the highest risk groups[75]. A methylation-based triage assay would not require the extensive cost and personnel that cytology currently does and could provide the possibility of using self-collected samples[91].

Expression of viral oncogenes may be another refining marker for defining 'HPV positivity' in cervical screening and other non-cervical HPV-related disease. E6 and E7 expression is much higher in cervical pre-cancers compared to transient HPV infections[90]. The same is true of both oncogenes in developing HNCs[92,93]. However, what contribution E6 and E7 could make

to triage that would add to cytology-based assessment and the potential use of p16/ki-67 or methylation markers is uncertain.

HPV-related HNCs also show a significant correlation to the down-regulation of EGFR, which in turn, is a positive prognosticator[94,95]. EGFR, though not as valued as p16 in the head and neck context, might represent a further indication of a clinically active HPV infection, or rather, a triage marker of a purely HPV-related HNC that would react expectedly well to standard treatments.

The potential for biomarkers to refine and standardize the definitions of 'HPV positivity' within each required context (Figure 2.2) is blatant. There are an innumerable number of biomarkers that could ultimately form the basis of less invasive, and more cost-efficient cervical screening, along with the treatment triage of other HPV-related diseases like HNCs. The sheer mass of biomarkers available however is, for the moment, a detriment to their clinical validation. The current scope of biomarkers requires further investigation to determine what will likely be panels of markers optimized to define 'HPV positivity' in cervical, non-cervical, research, and clinical contexts.

2.4 Re-defining 'HPV Positivity': A Case Study in HNC

The suggested principles for re-defining 'HPV positivity' in Section 2.3.4 are clear, relying on stratification by cancer site, and the goals and demands of a particular context (Figure 2.2). To contextualize these concepts using the potential exacting power of biomarkers, the coming discussion illustrates the application of the systematic structure for defining 'HPV positivity' posited in Section 2.3.4 using a case study of the foremost emerging HPV-related disease: HNC.

In the head and neck space, two main areas of focus have emerged: Epidemiology, and clinical testing, as showcased in Figure 2.2. In the epidemiological context, the goal of HPV testing is to determine the general presence of HPV in a regional, national, or global population of HNCs. Whether or not a present HPV infection is the causal agent in the associated cancer is not important to this type of analysis. In fact, it is preferable to identify

as many HPV infections as possible to ascertain the population-level prevalence and incidence of the virus in these cancers. Thus, DNA-based technologies with the highest sensitivity, paired with a reasonably high specificity would be optimal for this context. Most epidemiological work in the HPV head and neck space is also conducted using FFPE or frozen samples. Technologies used must therefore be selected and/or developed on this basis. Thus, 'HPV positivity' in the head and neck epidemiological context could be standardized and defined as any FFPE or frozen case identified as HPV DNA positive by the most sensitive available detecting technology (Figure 2.2).

This logic for standardizing the definition of 'HPV positivity' in the epidemiological context of HNCs is already emerging. In-situ hybridization, dot-blot hybridization, pure PCR, and real-time PCR have all traditionally been used to detect HPV DNA at the population level. The WHO has recently made the first step towards eliminating this technological heterogeneity in the epidemiological arena in the form of the HPV AHEAD study[96,97]. This study is the first world-wide meta-analysis of the epidemiology of HPV in HNCs that has developed a standardized protocol for processing and testing FFPE tissue to detect HPV DNA. The technology being used is one of the most sensitive ever developed combining a Multiplex PCR[15] and laser-based Luminex technology[98,99]. The goal of the study is to assess, without methodological or definitional variation, the true epidemiological characteristics of HPV DNA around the world.

In the clinical and diagnostic context, clinicians require a different but equally homogenized definition of 'HPV positivity'. Unlike in the epidemiological sphere, the clinic does not overtly require the identification of as many HPV infections as possible. Instead, it demands the specification of HPV infections that are clinically significant – causal carcinogens in the associated HNC. The reason this delineation is crucial is given the opportunity that HPV-related SCCs represent for maximizing survival[18], and also the major role that the currently available prophylactic vaccinations[100,101] and any future HNC HPV-related screening tools may have in preventing these cancers to begin with. Additional methods are therefore needed to delineate those cases with carcinogenic HPV infections from those that simply contain transient infections.

Indeed, the need for HPV testing is already well-established in clinic. The most recent College of American Pathologists (CAP) recommendations of HPV testing in HNC of 2017 advise testing for HR HPV in newly diagnosed OPSCC, with some additional recommendations for HPV-related biomarkers, p16 in particular[42]. The exhaustive list of recommendations are summarized in Table 2.1.

Table 2.1 2017 CAP recommendations for HPV testing in HNC[102].

	Guideline Statements	Strength of Recommendation
1	Pathologists should perform HR HPV testing on all patients with newly diagnosed OPSCC, including all histologic subtypes.	Strong recommendation
2	For oropharyngeal tissue specimens that are non-cytological, pathologists should perform HR HPV testing by surrogate marker p16 immunohistochemistry. Additional HPV-specific testing may be done at the discretion of the pathologist and/or treating clinician, or in the context of a clinical trial.	Recommendation
3	Pathologists should not routinely perform HR HPV testing on patients with non-SCC tumours of the oropharynx.	Expert Consensus Opinion
4	Pathologists should not routinely perform HR HPV testing on patients with non-oropharyngeal primary tumours of the head and neck.	Recommendation
5	Pathologists should routinely perform HR HPV testing on patients with metastatic SCC of unknown primary in a cervical upper or mid jugular chain lymph node.	Recommendation
6	For tissue specimens that are non-cytological from patients presenting with metastatic SCC of unknown primary in a cervical upper or mid jugular chain lymph nodes, pathologists should perform p16 immunohistochemistry.	Expert Consensus Opinion
7	Pathologists should perform HR HPV testing on head and neck fine needle aspiration SCC samples from all patients with known SCC not previously tested for HR HPV with suspected OPSCC or with metastatic SCC of unknown primary.	Expert Consensus Opinion
8	Pathologists should report p16 immunohistochemistry positivity as a surrogate biomarker for HR HPV in tissue specimens that are non cytological when there is at least 70% nuclear and cytoplasmic expression with at least moderate to strong intensity.	Expert Consensus Opinion
9	Pathologists should not routinely perform LR HPV testing on patients with HNCs.	Expert Consensus Opinion
10	Pathologists should not repeat HPV testing on patients with locally recurrent, regionally recurrent, or persistent tumour if primary tumour HR HPV status has already been established.	Expert Consensus Opinion
11	Pathologists should not routinely perform HR-HPV testing on patients with distant metastases if primary tumour HR HPV status has been established.	Expert Consensus Opinion
12	Pathologists should report primary OPSCC that test positive for HR HPV or its surrogate marker p16 as HPV-positive and/or p16 positive.	Expert Consensus Opinion
13	Pathologists should not provide a tumour grade or differentiation status for HPV positive/p16 positive OPSCC	Expert Consensus Opinion
14	Pathologists should not alter HR HPV testing strategy based on patient smoking history.	Expert Consensus Opinion

What the CAP guidelines do not detail is how an HPV positive case should be defined. There are no strict suggestions regarding the type of technology that should be used, nor do they specifically recommend the use of one particular HPV indicator, DNA or otherwise. DNA is currently the most widely used HPV test in the clinic, with in-situ hybridization being the most common technology for HNC, though p16 has historically been the indicator of an active HPV infection upon which any possible treatment decisions are made. Neither of these tests alone however effectively distinguish between carcinogenic and transient HPV infections. HPV alone, especially DNA, identifies the spectrum HPV infections, from LR transient infections to HR carcinogenic infections with no way to determine other than by genotype if an infection is causal in the associated HNC. p16 alone is evidence of the downstream effects of a potential HPV infection, but the marker can be upregulated in HNC with no HPV infection.

One suggested method to overcome this is that of combining HPV DNA detection with p16 staining[63,87] (Figure 2.2). Indeed, HPV DNA positivity is associated with better survival itself[3,17,103], and p16 is an independently good prognosticator of survival in HNSCC. However, the best discrimination between survival groups in the literature is found when p16 and HPV DNA status are paired together. Better survival is observed in HPV DNA and p16 positive cases, more so than with either indicator separately[17,104]. This strongly intimates that it is cases of HNSCC that are definitively caused by HPV, rather than those that harbor transient infections, that are uniquely responsive to treatment and crucial to delineate.

There are inferential suggestions that HPV mRNA may identify the same cases as the pairing of HPV DNA and p16 in HNC (Figure 2.2). This is based on the assumption that the presence of HPV mRNA indicates the same integrated and active infection as p16 does in the company of HPV DNA positivity. However, mRNA testing is still for the most part used in research[43,56,105] and is currently much less cost effective than HPV DNA and p16 staining paired together. In-situ hybridization and the E*6 assay[43] are also the best methods available for mRNA testing in tissue specimens at present, and the associated labour is intensive.

In either scenario, the HNC clinical context requires the most sensitive available technology for detecting HPV. It is well-established that the mere presence of HPV in cells is enough to evidence carcinogenesis, and that high viral load is not indicative of higher carcinogenic potential or activity[54,55]. The detection of small copy numbers of HPV as either DNA or mRNA is therefore crucial. However, especially in the case of HPV DNA, high sensitivity might yield the detection of a large number of transient infections. The pairing of a highly sensitive HPV DNA test with p16 thus represents a good opportunity to not only distinguish carcinogenic infections in HNC from transient ones, but also to ensure that as few as possible HPV-related tumours are missed. This will prove important in the context of potential future screening in sampled tissue from the region and in that informing clinical decisions regarding treatment. The alternative would be to employ a less sensitive HPV DNA-detecting technology paired with p16 that might miss even a minimal number of patients with optimal survival opportunity or the need to be monitored more closely in the case of screening. Jeopardizing patient well-being in the clinic in this way would be unacceptable.

In-situ hybridization is currently used in the clinic to detect HPV DNA, but it is by no means the most sensitive currently available technology. Other assays have been developed in the laboratory with much higher sensitivities, including genome sequencing, the INNOLiPA®, the Linear Array®, and others making use of laser-based platforms, xMAP® technology[98], and the Luminex Analyser[99]. These can all be optimized for FFPE and frozen tissue. They are not yet validated for clinical use as they continue to be employed only for emerging epidemiological data in the research context. However, their potential value, or the value of other similarly sensitive technologies for fresh screening or diagnostic samples is extremely promising.

Thus, in the head and neck clinical context, ‘HPV positivity’ could be homogenously defined as any case positive for HPV DNA using the most sensitive available technology that is also positive for p16 based on diagnostic staining (Figure 2.2). This could evolve in future given current mRNA research (Figure 2.2), and another homogenized definition may be necessary for ‘HPV positivity’ in context of future HNC screening tools.

2.5 The Future of Defining 'HPV Positivity'

'HPV positivity' has never been a simple binary test using a single target and technological platform. The foregoing review of the various indicators available for detecting the virus and the multiplicity of technologies on the market to test for them, each with different sensitivities, specificities, and result presentation methods (e.g. binary, quantitative), makes this very clear. Despite this, the literature has tended to generalize results of HPV-related studies under the assumption that all technologies, principles, and therefore definitions of 'HPV positivity' are somehow equivalent and have the same broader significance.

On this basis, it is evident that one sweeping definition of 'HPV positivity' across all HPV-related cancer types, and in all research and clinical contexts, is not sufficient, useful, or valuable. Where the cervical clinical context requires a definition of 'HPV positivity' optimal for the triage of women in particular age groups using fresh samples, the head and neck epidemiological context demands one that yields the most representative assessment of the virus' presence in a given population using archival tissue. A singular definition will never suit both of these contexts, and it is unlikely that one technology will ever be optimized to satisfy each of their needs.

Thus, there should indeed be many heterogeneous definitions for 'HPV positivity.' However, each defined context requiring such a definition must have its own homogenized way of determining what constitutes an HPV positive case. These definitions should be generated using a standardized systematic structure, like that suggested in Section 2.3.4 (Figure 2.2). Given that clinical work and research findings are synergistic, these exacting definitions are essential to maximizing clinical outcomes.

However, it is not possible to definitively refine the definitions of 'HPV positivity' in all contexts presently as the study of non-cervical HPV-related cancers and the role of HPV and its associated markers in cervical screening are still evolving. These definitions will likely develop and change in years to come. Entirely new fields requiring their own definition will also materialize with the growing need for non-cervical screening tools, especially in the case of HNC. What is unequivocal though is the need for a systematic structure defining

'HPV positivity' in discerning contexts. Only then will the greatest number of lives be saved in a field of oncology that already represents a unique survival optimizing opportunity.

References

1. Center for Disease Control. STD Facts - Human papillomavirus (HPV). *Center for Disease Control* (2017). Available at: https://www.cdc.gov/std/hpv/stdfact-hpv.htm. (Accessed: 2nd February 2017)
2. Van Doorslaer, K. & Burk, R. D. Evolution of Human Papillomavirus Carcinogenicity. *Adv. Virus Res.* **77**, 41–62 (2010).
3. Woods, R. Human Papillomavirus-related Oropharyngeal Squamous Cell Carcinoma – Prevalence and Incidence in a Defined Population and Analysis of the Expression of Specific Cellular Biomarkers. (2016).
4. zur Hausen, H. Human papillomaviruses and their possible role in squamous cell carcinomas. *Curr. Top. Microbiol. Immunol.* **78**, 1–30 (1977).
5. Walboomers, J. M. M. *et al.* Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J. Pathol.* **189**, 12–19 (1999).
6. Liyanage, S. S. *et al.* The Aetiological Role of Human Papillomavirus in Oesophageal Squamous Cell Carcinoma: A Meta-Analysis. *PLoS One* **8**, e69238 (2013).
7. Aldabagh, B., Angeles, J. G. C., Cardones, A. R. & Arron, S. T. Cutaneous squamous cell carcinoma and human papillomavirus: is there an association? *Dermatol. Surg.* **39**, 1–23 (2013).
8. Syrjänen, K., Syrjänen, S., Lamberg, M., Pyrhönen, S. & Nuutinen, J. Morphological and immunohistochemical evidence suggesting human papillomavirus (HPV) involvement in oral squamous cell carcinogenesis. *Int. J. Oral Surg.* **12**, 418–24 (1983).
9. Löning, T. *et al.* Analysis of oral papillomas, leukoplakias, and invasive carcinomas for human papillomavirus type related DNA. *J. Invest. Dermatol.* **84**, 417–20 (1985).
10. O'Leary, J. J. *et al.* Cervical screening: A new way forward (tests of risk and tests of disease). *HRB Open Res.* **1**, 3 (2018).
11. Marur, S. & Forastiere, A. A. Head and Neck Squamous Cell Carcinoma: Update on Epidemiology, Diagnosis, and Treatment. *Mayo Clin. Proc.* **91**, 386–396 (2016).
12. Kleter, B. *et al.* Novel Short-Fragment PCR Assay for Highly Sensitive Broad-Spectrum

Detection of Anogenital Human Papillomaviruses. *Am. J. Pathol.* **153**, 1731–1739 (1998).

13. Kleter, B. *et al.* Development and clinical evaluation of a highly sensitive PCR-reverse hybridization line probe assay for detection and identification of anogenital human papillomavirus. *J. Clin. Microbiol.* **37**, 2508–17 (1999).
14. FlashGel™ DNA System by Lonza Inc. (2017). Available at: http://www.lonza.com/products-services/bio-research/electrophoresis-of-nucleic-acids-and-proteins/nucleic-acid-electrophoresis/fast-electrophoresis-flashgel-system-for-dna-rna-and-recovery/flashgel-dna-system.aspx. (Accessed: 22nd August 2017)
15. Gheit, T. *et al.* Development of a Sensitive and Specific Assay Combining Multiplex PCR and DNA Microarray Primer Extension To Detect High-Risk Mucosal Human Papillomavirus Types. *J. Clin. Microbiol.* **44**, 2025–2031 (2006).
16. Bishop, J. A. *et al.* Detection of Transcriptionally Active High-risk HPV in Patients With Head and Neck Squamous Cell Carcinoma as Visualized by a Novel E6/E7 mRNA In Situ Hybridization Method. *Am. J. Surg. Pathol.* **36**, 1874–1882 (2012).
17. Woods, R. *et al.* Role of human papillomavirus in oropharyngeal squamous cell carcinoma: A review. *World J. Clin. cases* **2**, 172–93 (2014).
18. Ang, K. K. *et al.* Human papillomavirus and survival of patients with oropharyngeal cancer. *N. Engl. J. Med.* **363**, 24–35 (2010).
19. Abreu, A. L. P., Souza, R. P., Gimenes, F. & Consolaro, M. E. L. A review of methods for detect human Papillomavirus infection. *Virol. J.* **9**, 262 (2012).
20. Qiagen. digene HC2 HPV DNA Test. *Qiagen Ltd* (2019). Available at: https://www.qiagen.com/ie/products/diagnostics-and-clinical-research/sexual-reproductive-health/cervical-cancer-screening/digene-hc2-hpv-dna-test/#orderinginformation. (Accessed: 3rd June 2019)
21. Hologic. Cervista HPV HR Assay. *2019* Available at: https://www.hologic.com/package-inserts/diagnostic-products/cervista-hpv-hr-assay. (Accessed: 3rd June 2019)
22. Pan American Health Organization. *INTEGRATING HPV TESTING IN CERVICAL CANCER SCREENING PROGRAMS A MANUAL FOR PROGRAM MANAGERS*. (2016).
23. Rahman, M. *et al.* High prevalence of intermediate-risk human papillomavirus infection in uterine cervices of kenyan women infected with human

immunodeficiency virus. *J. Med. Virol.* **83**, 1988–1996 (2011).

24. Brandstetter, T. *et al.* A polymer-based DNA biochip platform for human papilloma virus genotyping. *J. Virol. Methods* **163**, 40–48 (2010).
25. Greiner Bio-One. PapilloCheck®. *Greiner Bio-One* (2019). Available at: https://shop.gbo.com/en/row/products/diagnostics/hpvdiagnostics/papillocheck/. (Accessed: 3rd June 2019)
26. Roche Diagnostics. cobas® 4800 System. *Roche Diagnostics* (2019). Available at: https://diagnostics.roche.com/global/en/products/systems/cobas-4800-system.html. (Accessed: 3rd June 2019)
27. Heideman, D. A. M. *et al.* Clinical Validation of the cobas 4800 HPV Test for Cervical Screening Purposes. *J. Clin. Microbiol.* **49**, 3983–3985 (2011).
28. Mateos, M. L., Chacón de Antonio, J., Rodríguez-Domínguez, M., Sanz, I. & Rubio, M. D. Evaluación de un sistema de PCR a tiempo real (cobas 4800) para la detección separada de los genotipos 16 y 18 y otros genotipos de alto riesgo del virus del papiloma humano en la prevención del cáncer cervical. *Enferm. Infecc. Microbiol. Clin.* **29**, 411–414 (2011).
29. Roche Diagnostics, FDA & Government of the United States. *COBAS Specifications - FDA*. (2015).
30. Abbott Molecular. RealTime High Risk HPV Assay. *Abbott Molecular* (2019). Available at: https://www.molecular.abbott/int/en/products/infectious-disease/realtime-high-risk-hpv. (Accessed: 3rd June 2019)
31. FDA. BD Onclarity HPV Assay. *FDA* (2019). Available at: https://www.fda.gov/medical-devices/recently-approved-devices/bd-onclarity-hpv-assay-p160037. (Accessed: 3rd June 2019)
32. Cepheid. Xpert HPV. *Cepheid* (2019). Available at: http://www.cepheid.com/en/cepheid-solutions/clinical-ivd-tests/sexual-health/xpert-hpv. (Accessed: 3rd June 2019)
33. Smith, L. M. *et al.* Fluorescence detection in automated DNA sequence analysis. *Nature* **321**, 674–679 (1986).
34. Lee, L. G. *et al.* DNA sequencing with dye-labeled terminators and T7 DNA polymerase: effect of dyes and dNTPs on incorporation of dye-terminators and probability analysis of termination fragments. *Nucleic Acids Res.* **20**, 2471–83 (1992).

35. Novais, R. C. & Thorstenson, Y. R. The evolution of Pyrosequencing® for microbiology: From genes to genomes. *J. Microbiol. Methods* **86**, 1–7 (2011).
36. Genomica. CLART® HPV. *Genomica* (2019). Available at: http://genomica.es/en/in_vitro_diagnostics_products_clart_hpv2.cfm. (Accessed: 3rd June 2019)
37. Fujirebio. INNO-LiPA® HPV Genotyping Extra II. *Fujirebio* (2019). Available at: https://www.fujirebio-europe.com/products-services/product-browser/inno-lipar-hpv-genotyping-extra-ii-20t. (Accessed: 3rd June 2019)
38. Roche Diagnostics. LINEAR ARRAY® HPV Genotyping. *Roche Diagnostics* (2019). Available at: https://diagnostics.roche.com/global/en/products/params/linear-array-hpv-genotyping.html. (Accessed: 4th June 2019)
39. Genomica. Clinical Arrays. *Genomica* (2019). Available at: http://genomica.es/en/in_vitro_diagnostics_products_clart.cfm. (Accessed: 4th June 2019)
40. Wentzensen, N. & von Knebel Doeberitz, M. Biomarkers in cervical cancer screening. *Dis. Markers* **23**, 315–30 (2007).
41. Pretect.no. Pretect HPV-Proofer. *Pretect.no* (2019). Available at: https://www.pretect.no/pretecthpvproofer. (Accessed: 4th June 2019)
42. Hologic. APTIMA HPV Assay. *Hologic* (2019). Available at: https://www.hologic.com/sites/default/files/package-insert/AW-14517-001_003_01.pdf.
43. Halec, G. *et al.* Biological activity of probable/possible high-risk human papillomavirus types in cervical cancer. *Int. J. Cancer* **132**, 63–71 (2013).
44. Kajitani, N., Satsuka, A., Kawate, A. & Sakai, H. Productive Lifecycle of Human Papillomaviruses that Depends Upon Squamous Epithelial Differentiation. *Front. Microbiol.* **3**, 152 (2012).
45. Yoshida, T. *et al.* Quantitative real-time polymerase chain reaction analysis of the type distribution, viral load, and physical status of human papillomavirus in liquid-based cytology samples from cervical lesions. *Int. J. Gynecol. Cancer* **18**, 121–127 (2008).
46. Matovina, M., Sabol, I., Grubišić, G., Gašperov, N. M. & Grce, M. Identification of human papillomavirus type 16 integration sites in high-grade precancerous cervical

lesions. *Gynecol. Oncol.* **113**, 120–127 (2009).

47. Gao, G. *et al.* Mate pair sequencing of oropharyngeal squamous cell carcinomas reveals that HPV integration occurs much less frequently than in cervical cancer. *J. Clin. Virol.* **59**, 195–200 (2014).
48. Olthof, N. C. *et al.* Comprehensive Analysis of HPV16 Integration in OPSCC Reveals No Significant Impact of Physical Status on Viral Oncogene and Virally Disrupted Human Gene Expression. *PLoS One* **9**, e88718 (2014).
49. Tamalet, C. *et al.* Testing for Human Papillomavirus and Measurement of Viral Load of HPV 16 and 18 in Self-Collected Vaginal Swabs of Women Who Do Not Undergo Cervical Cytological Screening in Southern France. *J. Med. Virol. J . Med . Virol* **82**, 1431–1437 (2010).
50. Constandinou-Williams, C. *et al.* Is human papillomavirus viral load a clinically useful predictive marker? A longitudinal study. *Cancer Epidemiol. Biomarkers Prev.* **19**, 832–7 (2010).
51. Lowe, B., O'Neil, D., Loeffert, D. & Nazarenko, I. Distribution of Human papillomavirus load in clinical specimens. *J. Virol. Methods* **173**, 150–152 (2011).
52. Hesselink, A. T. *et al.* Comparison of GP5+/6+-PCR and SPF10-Line Blot Assays for Detection of High-Risk Human Papillomavirus in Samples from Women with Normal Cytology Results Who Develop Grade 3 Cervical Intraepithelial Neoplasia. *J. Clin. Microbiol.* **46**, 3215–3221 (2008).
53. Xi, L. F. *et al.* Human Papillomavirus Type 18 DNA Load and 2-Year Cumulative Diagnoses of Cervical Intraepithelial Neoplasia Grades 2-3. *JNCI J. Natl. Cancer Inst.* **101**, 153–161 (2009).
54. Chaturvedi, A. K. *et al.* Human Papillomavirus and Rising Oropharyngeal Cancer Incidence in the United States. *J. Clin. Oncol.* **29**, 4294–4301 (2011).
55. Holzinger, D. *et al.* Viral RNA Patterns and High Viral Load Reliably Define Oropharynx Carcinomas with Active HPV16 Involvement. *Cancer Res.* **72**, 4993–5003 (2012).
56. Gheit, T. *et al.* Role of mucosal high-risk human papillomavirus types in head and neck cancers in central India. *Int. J. Cancer* **141**, 143–151 (2017).
57. Näsman, A. *et al.* Incidence of human papillomavirus (HPV) positive tonsillar carcinoma in Stockholm, Sweden: An epidemic of viral-induced carcinoma? *Int. J. Cancer* **125**, 362–366 (2009).

58. Buttmann-Schweiger, N., Deleré, Y., Klug, S. J. & Kraywinkel, K. Cancer incidence in Germany attributable to human papillomavirus in 2013. *BMC Cancer* **17**, 682 (2017).
59. Wittekindt, C. *et al.* Increasing Incidence rates of Oropharyngeal Squamous Cell Carcinoma in Germany and Significance of Disease Burden Attributed to Human Papillomavirus. *Cancer Prev. Res.* (2019). doi:10.1158/1940-6207.CAPR-19-0098
60. Gillison, M. L. *et al.* Eurogin Roadmap: Comparative epidemiology of HPV infection and associated cancers of the head and neck and cervix. *Int. J. Cancer* **134**, 497–507 (2014).
61. Mourad, M. *et al.* Epidemiological Trends of Head and Neck Cancer in the United States: A SEER Population Study. *J. Oral Maxillofac. Surg.* **75**, 2562–2572 (2017).
62. Chaturvedi, A. K. Beyond Cervical Cancer: Burden of Other HPV-Related Cancers Among Men and Women. *J. Adolesc. Heal.* **46**, S20–S26 (2010).
63. Salazar, C. R. *et al.* Combined P16 and human papillomavirus testing predicts head and neck cancer survival. *Int. J. Cancer* **135**, 2404–2412 (2014).
64. Reimers, N. *et al.* Combined analysis of HPV-DNA, p16 and EGFR expression to predict prognosis in oropharyngeal cancer. *Int. J. Cancer* **120**, 1731–1738 (2007).
65. Mirghani, H. & Blanchard, P. Treatment de-escalation for HPV-driven oropharyngeal cancer: Where do we stand? *Clin. Transl. Radiat. Oncol.* **8**, 4–11 (2018).
66. Wierzbicka, M., Szyfter, K., Milecki, P., Składowski, K. & Ramlau, R. The rationale for HPV-related oropharyngeal cancer de-escalation treatment strategies. *Contemp. Oncol. (Poznan, Poland)* **19**, 313–22 (2015).
67. Brisson, R. J. *et al.* De-escalation in HPV-negative locally advanced head and neck squamous cell cancer (LA-HNSCC) in patients after induction chemotherapy: A retrospective case series. *J. Clin. Oncol.* **36**, e18090–e18090 (2018).
68. Mannweiler, S., Sygulla, S., Winter, E. & Regauer, S. Two major pathways of penile carcinogenesis: HPV-induced penile cancers overexpress p16ink4a, HPV-negative cancers associated with dermatoses express p53, but lack p16ink4a overexpression. *J. Am. Acad. Dermatol.* **69**, 73–81 (2013).
69. Stelzer, M. K. *et al.* A mouse model for human anal cancer. *Cancer Prev. Res. (Phila).* **3**, 1534–41 (2010).
70. Ueda, Y. *et al.* Two distinct pathways to development of squamous cell carcinoma of the vulva. *J. Skin Cancer* **2011**, 951250 (2011).

71. Carchman, E. H., Matkowskyj, K. A., Meske, L. & Lambert, P. F. Dysregulation of Autophagy Contributes to Anal Carcinogenesis. *PLoS One* **11**, e0164273 (2016).
72. Rautava, J. & Syrjänen, S. Biology of human papillomavirus infections in head and neck carcinogenesis. *Head Neck Pathol.* **6 Suppl 1**, S3-15 (2012).
73. Stransky, N. *et al.* The Mutational Landscape of Head and Neck Squamous Cell Carcinoma. *Science (80-.).* **333**, 1157–1160 (2011).
74. Wentzensen, N. *et al.* p16/Ki-67 Dual Stain Cytology for Detection of Cervical Precancer in HPV-Positive Women. *J. Natl. Cancer Inst.* **107**, djv257 (2015).
75. Wentzensen, N., Schiffman, M., Palmer, T. & Arbyn, M. Triage of HPV positive women in cervical cancer screening. *J. Clin. Virol.* **76**, S49–S55 (2016).
76. Carozzi, F. *et al.* Risk of high-grade cervical intraepithelial neoplasia during follow-up in HPV-positive women according to baseline p16-INK4A results: a prospective analysis of a nested substudy of the NTCC randomised controlled trial. *Lancet Oncol.* **14**, 168–176 (2013).
77. Carozzi, F. *et al.* Use of p16-INK4A overexpression to increase the specificity of human papillomavirus testing: a nested substudy of the NTCC randomised controlled trial. *Lancet Oncol.* **9**, 937–945 (2008).
78. Grabe, N. *et al.* A virtual microscopy system to scan, evaluate and archive biomarker enhanced cervical cytology slides. *Cell. Oncol.* **32**, 109–19 (2010).
79. Bierkens, M. *et al. CADM1* and *MAL* promoter methylation levels in hrHPV-positive cervical scrapes increase proportional to degree and duration of underlying cervical disease. *Int. J. Cancer* **133**, 1293–1299 (2013).
80. White, C. & Reynolds, S. *Assessing a Panel of Biomarkers for the Most Risk Sensitive Triage in Cervical Screening*. (2019).
81. Park, K. *et al.* p16 immunohistochemistry alone is a better prognosticator in tonsil cancer than human papillomavirus in situ hybridization with or without p16 immunohistochemistry. *Acta Otolaryngol.* **133**, 297–304 (2013).
82. Stephen, J. K. *et al.* Significance of p16 in site-specific HPV positive and HPV negative HNSCC. *Cancer Clin. Oncol.* **2**, 51–61 (2012).
83. Ernoux-Neufcoeur, P. *et al.* Combined analysis of HPV DNA, p16, p21 and p53 to predict prognosis in patients with stage IV hypopharyngeal carcinoma. *J. Cancer Res. Clin. Oncol.* **137**, 173–81 (2011).

84. Lundberg, M. *et al.* Association of BMI-1 and p16 as prognostic factors for head and neck carcinomas. *Acta Otolaryngol.* **136**, 501–5 (2016).
85. Wang, H., Sun, R., Lin, H. & Hu, W. P16 ^{INK4A} as a surrogate biomarker for human papillomavirus-associated oropharyngeal carcinoma: Consideration of some aspects. *Cancer Sci.* **104**, 1553–1559 (2013).
86. O'Regan, E. M. *et al.* p16INK4A genetic and epigenetic profiles differ in relation to age and site in head and neck squamous cell carcinomas. *Hum. Pathol.* **39**, 452–458 (2008).
87. Albers, A. E., Qian, X., Kaufmann, A. M. & Coordes, A. Meta analysis: HPV and p16 pattern determines survival in patients with HNSCC and identifies potential new biologic subtype. *Sci. Rep.* **7**, 16715 (2017).
88. Taylor, S., Bunge, E., Bakker, M. & Castellsagué, X. The incidence, clearance and persistence of non-cervical human papillomavirus infections: a systematic review of the literature. *BMC Infect. Dis.* **16**, 293 (2016).
89. De Strooper, L. M. A. *et al. CADM1* , *MAL* and *miR124-2* methylation analysis in cervical scrapes to detect cervical and endometrial cancer. *J. Clin. Pathol.* **67**, 1067–1071 (2014).
90. Ebisch, R. M. *et al.* Triage of high-risk HPV positive women in cervical cancer screening. *Expert Rev. Anticancer Ther.* **16**, 1073–1085 (2016).
91. Arbyn, M. *et al.* Accuracy of human papillomavirus testing on self-collected versus clinician-collected samples: a meta-analysis. *Lancet Oncol.* **15**, 172–183 (2014).
92. Strati, K., Pitot, H. C. & Lambert, P. F. Identification of biomarkers that distinguish human papillomavirus (HPV)-positive versus HPV-negative head and neck cancers in a mouse model. *Proc. Natl. Acad. Sci.* **103**, 14152–14157 (2006).
93. Psyrri, A., Sasaki, C., Vassilakopoulou, M., Dimitriadis, G. & Rampias, T. Future directions in research, treatment and prevention of HPV-related squamous cell carcinoma of the head and neck. *Head Neck Pathol.* **6 Suppl 1**, S121-8 (2012).
94. Kumar, B. *et al.* EGFR, p16, HPV Titer, Bcl-xL and p53, sex, and smoking as indicators of response to therapy and survival in oropharyngeal cancer. *J. Clin. Oncol.* **26**, 3128–37 (2008).
95. Langer, C. J. Exploring biomarkers in head and neck cancer. *Cancer* **118**, 3882–3892 (2012).

96. HPV-AHEAD - Role of human papillomavirus infection and other co-factors in the aetiology of head and neck cancer in Europe and India. (2017). Available at: http://hpv-ahead.iarc.fr/. (Accessed: 22nd August 2017)
97. International Agency for Research on Cancer & World Health Organization. HPV-AHEAD Study. *International Agency for Research on Cancer* (2019). Available at: http://hpv-ahead.iarc.fr/about/index.php. (Accessed: 27th April 2019)
98. xMAP® Technology for Research & Applied Markets. (2017). Available at: https://www.luminexcorp.com/research/our-technology/xmap-technology/. (Accessed: 22nd August 2017)
99. Luminex® 100/200™ for Clinical Diagnostic Use. Available at: https://www.luminexcorp.com/eu/clinical/instruments/luminex-100200/. (Accessed: 9th April 2018)
100. Wang, C. *et al.* Targeting Head and Neck Cancer by Vaccination. *Front. Immunol.* **9**, 830 (2018).
101. Guo, T., Eisele, D. W. & Fakhry, C. The potential impact of prophylactic human papillomavirus vaccination on oropharyngeal cancer. *Cancer* **122**, 2313–23 (2016).
102. College of American Pathologists. Human Papillomavirus Testing in Head and Neck Carcinomas. *College of American Pathologists* (2017). Available at: https://www.cap.org/protocols-and-guidelines/cap-guidelines/current-cap-guidelines/human-papillomavirus-testing-in-head-and-neck-carcinomas. (Accessed: 4th June 2019)
103. Syrjänen, S. Human papillomavirus (HPV) in head and neck cancer. *J. Clin. Virol.* **32**, 59–66 (2005).
104. Ndiaye, C. *et al.* HPV DNA, E6/E7 mRNA, and p16INK4a detection in head and neck cancers: a systematic review and meta-analysis. *Lancet Oncol.* **15**, 1319–1331 (2014).
105. Cerasuolo, A. *et al.* Comparative analysis of HPV16 gene expression profiles in cervical and in oropharyngeal squamous cell carcinoma. *Oncotarget* **8**, 34070–34081 (2017).

CHAPTER 3

MATERIALS AND METHODS

3 CHAPTER 3: MATERIALS AND METHODS

3.1 Study Design

The ECHO study was a nationwide retrospective study of HPV infection in oropharyngeal, laryngeal, and oral cavity SCC cases diagnosed between 1994 and 2013 in Ireland. It was the largest all-encompassing project of its kind to be conducted in Ireland.

3.2 Clinical Database

It was through the NCRI's database that the ECHO study identified relevant specimens to create its own databank of cases.

3.2.1 The NCRI

The NCRI was founded in 1991 to collect data on all cancer cases in the Republic of Ireland, and since 1994, it has recorded information on all tumours diagnosed in the country. The NCRI utilizes active registration methods, and trained tumour registration officers based in hospitals across the country to identify newly diagnosed cancers from a variety of sources including pathology laboratories, hospital administration systems, hospital in-patient episodes, and radiotherapy and chemotherapy records.

Following standard protocols based on international guidelines, the NCRI abstracts information and records it for each incident tumour. This information includes personal details (e.g. name, date of diagnosis, cancer site, morphology, clinical and pathological stage, grade) and clinical/treatment related information (e.g. hospitals of diagnosis and treatment, managing clinicians, treatment received within the first year post-diagnosis, GP). The NCRI also records whether or not the tumour has been histologically verified, together with the date of the histology, pathology laboratory, pathologist, and histology laboratory number for the specimen. The Registry then follows-up through linkage with national death certificate files, and details of any relevant date and cause of death are recorded.

3.2.2 Evaluation of the NCRI as a Source

The Registry is the only source to have kept a consistent record of cancer cases in Ireland for more than 20 years. The Registry itself estimates a completeness of registration of 97%.

Independent evaluation has found that the NCRI has a lower bound of completeness of case ascertainment of 94.3% for all invasive cancers excluding non-melanoma skin cancers[1].

3.3 Clinical Data Criteria

Very strict clinical data criteria were used to determine case eligibility for the ECHO study. Assessment of eligibility was determined on the basis of narrow search parameters entered into the NCRI database, patient consent, the examination of pathology reports, and the histopathological analysis of newly generated H+E slides. The coming discussion details precisely what kind of cases were included and excluded in the study.

3.3.1 Inclusion Criteria

Patients newly diagnosed with a primary, invasive, HNC between 1994 and 2013 were identified in the NCRI database. They were categorized into three groups modeled on the most recent classification recommended by IARC[2]: Oropharyngeal cancers (ICD10 C01.0, C02.4, C05.1, C05.2, C09.0-C09.9, C10.0-C10.9, C14.2); oral cavity cancers (ICD10 C02.0, C02.1, C02.2, C02.3, C02.8, C02.9, C03.0-C03.9, C04.0-C04.9, C05.0, C05.8, C05.9, C06.0, C06.1, C06.2, C06.9, C14.0, C014.8); and laryngeal cancers (ICD10 C32.0-C32.9).

Within this population, patients diagnosed with SCC whose tumours were histologically verified and/or who underwent surgery, were identified and included in the study.

3.3.2 Exclusion Criteria

Cancers of the salivary glands, nasopharynx, and hypopharynx were eliminated given that these did not fall within the HPV-vulnerable region of the head and neck. The study also excluded patients with no or insufficient tissue for all steps of the analysis from the primary oropharyngeal, laryngeal, and oral cavity sites.

Only pure SCC cases were included in the study to eliminate any possible confounding by SCC types or other tumour types that are not generally implicated in HPV carcinogenesis. A non-exhaustive list of the pathologies diverging from pure SCC that were excluded follows: Mucoepidermoid SCC, carotenoid SCC, basaloid SCC, adenoid cystic SCC, and spindle cell

SCC. Other tumour types that were excluded given their blatant non-SCC pathology were: Adenocarcinoma, verrucous carcinoma, acinic cell carcinoma, melanoma, epidermoid carcinoma, sarcomachordoma, and rhabdomyosarcoma.

Similarly, cases were excluded if they were only 'suspicious' for SCC diagnosis. This strict criteria ensured no skew in the data towards HPV-negativity and adhered to the literature's determination that SCCs comprise 95% of the relevant population[3].

Cases for which only lymph nodes were available as tissue, either with or without invasive SCC, were also excluded. It was important and necessary to only include samples of primary tumour specimen available to avoid distortion of results should metastatic tumour have different characteristics to primary tumours and suggest inaccurate carcinogenic pathways.

Specimens fixed in Bouins, or specimens that only had frozen tissue available were excluded given that this kind of tissue was not suitable for the formalin-fixed paraffin-embedded (FFPE)-based nature of the study and the technologies it employed.

3.3.3 Ethical Approval and Consent

Ethical approval for the use of archival tissue specimens was obtained from all relevant local hospital ethics committees. This consisted of 11 committees in total representing the 14 potential participating hospital sites, summarized in Table 3.1 below.

Table 3.1 A comprehensive list of hospitals willing to participate in the ECHO study.

Hospitals were represented collectively by 11 different ethics committees, from which ethical approval for all required aspects of the study was obtained.

Beaumont University Hospital, Dublin	**Sligo General Hospital, Sligo**
Cork University Hospital, Cork	**South Infirmary Victoria University Hospital, Cork**
Cork Dental Hospital, Cork	**St. James University Hospital, Dublin**
Dublin Dental Hospital, Dublin	**St. Vincent's University Hospital, Dublin**
Kerry General Hospital, Tralee	**The Royal Victoria Eye and Ear Hospital, Dublin**
Mater Misericordiae University Hospital, Dublin	**University Hospital Limerick, Limerick**
Midlands Regional Hospital, Tullamore	**Waterford Regional Hospital, Waterford**

The study was also conducted in compliance with the HSE National Consent Policy of May 2013[4] and was consistently benchmarked against best practice models internationally.

In terms of patient and pathologist consent, details of the pathology laboratories at which the specimens were analysed were extracted from the NCRI database by collaborators in the NCRI itself. Pathologists from the relevant hospital sites possessing specimens from eligible patients were contacted for their agreement to collaborate with the study. Specimens were sought for patients who were alive and for those who were dead at the time the study population was selected.

For those who were still alive, consent for access to their specimens was sought from the patient directly by collaborators in the NCRI. Considerable care was taken to ensure that patients were not contacted inappropriately. Collaborators in the NCRI carried out two steps to achieve this. The first was to access the data held by the NCRI (which included in-patient admission from Hospital Inpatient Episode Statistics (HIPE)), and the second involved establishing liaison with treating clinicians.

For the first step, careful checks were made of the data held by the NCRI to identify and exclude patients who had had a recent recurrence or a recent hospital admission. In addition, checks for recent deaths were completed using the General Record Offices and RIP.ie, methods standardly used by the Registry. Those patients who had died were moved to the “deceased patients” list.

In the second step, letters were sent to clinicians inquiring after identified patients and requesting confirmation that they could be included in the study. The treating clinicians were thus given the opportunity to further exclude any patients from amongst those believed to be still alive should they have reason to be believe the patients should not be contacted for any reason. To minimize the burden on clinicians, if no response was received by the specified date in the letter, the NCRI collaborators went ahead and contacted all the patients on the sent list. In accordance with the data protection requirements regarding secondary use of data, the patient details sent to the treating clinicians were provided solely for the purposes of the study and clinicians were informed of this and instructed to destroy these after review.

Patients still alive and eligible after this screening process were contacted by letter by the NCRI collaborators and asked to consent to provide access to their tumour specimens. Non-responders were followed-up with two reminder letters. The consent rate for alive patients in the study was 44.56% with 595 patients of the 1335 eligible alive patients identified in the NCRI database.

Originally, for those patients who had died at or before the time the study population was selected and checked, there were no specific consent issues, and specimens were therefore allowed to be accessed with the agreement of the pathologist concerned. However, with the introduction of new General Data Protection Regulation (GDPR) guidelines in 2018 that were implemented near the end of the study, the NCRI only allowed the Researcher to have access to data regarding samples that were actually tested. Access to information regarding those cases tested for patients who were deceased therefore ultimately required new and further documentation and applications to the NCRI.

3.4 Identification, Retrieval, Inclusion, and Storage of Clinical Specimens (FFPE Tissue)

3.4.1 Summary of Specimen Identification, Retrieval, and Inclusion

The identification, retrieval, and inclusion of FFPE tissue specimens, referred to in shorthand as 'blocks', that were ultimately included in the ECHO study was a multi-step process. At each stage, cases had to be eliminated from the study population due to continuous eligibility assessment based on criteria in Sections 3.3.1, 3.3.2, and 3.3.3, and due to accessibility limitations.

The first step of the process was the scanning of the NCRI database for incident tumour specimens, the narrowing of this population based on eligibility criteria in Sections 3.3.1 and 3.3.2, and the elimination of unconsented patients based on Section 3.3.3. This step was carried out by NCRI collaborators. The second, third, and fourth steps were carried out by the Researcher. The second step was the review of pathology reports for cases in this newly generated database and the further elimination of cases based on the same eligibility criteria. The third step was the retrieval of cases from relevant storage sites and the elimination of those cases that were inaccessible or missing. The fourth step was the histopathological analysis of newly generated H+E slides and the further elimination of cases on the basis of the same eligibility criteria. This multi-step process, including each step at which cases had to be eliminated due to continuous eligibility assessment, is summarized in Figure 3.1 below.

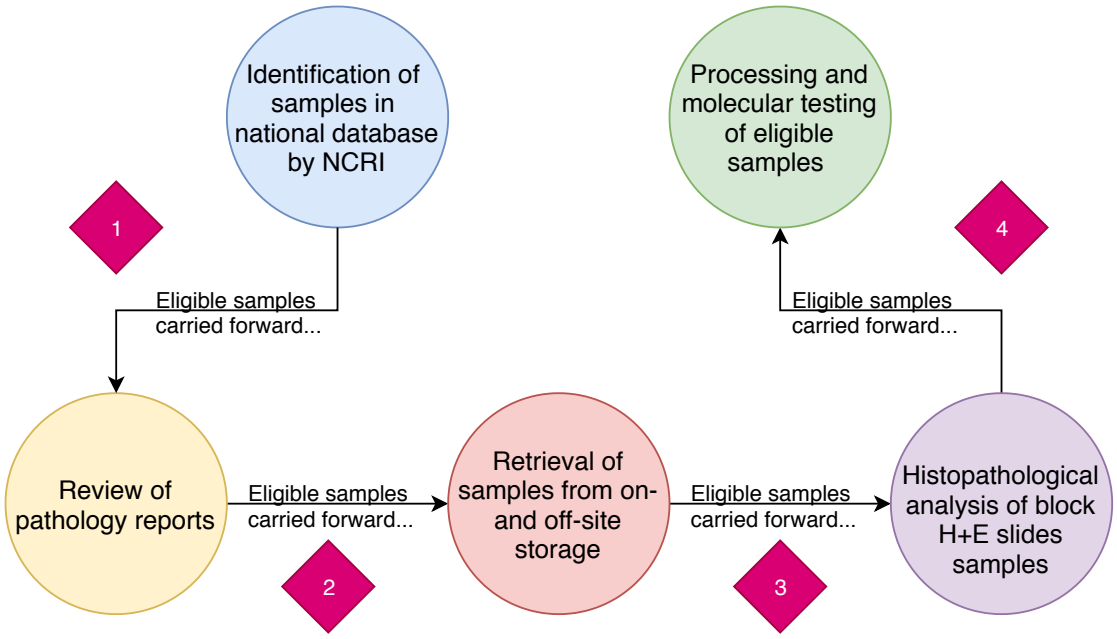

Figure 3.1 Summary of the steps involved in the process of identification, retrieval, and inclusion of FFPE blocks in the ECHO study.

Pink numbered diamonds indicate an eligibility assessment using criteria in Sections 3.3.1, 3.3.2, and 3.3.3 after which cases were eliminated. These pink diamonds will be referred to throughout the coming Chapters as "step 1", "step 2", "step 3", and "step 4". The blue bubble in the top left-hand corner is the starting point for the process. Incident tumours in the national database of cancers were identified here by collaborators in the NCRI. Step 1 is the elimination of cases from this database on the basis of eligibility criteria and consent. The pathology reports for those samples identified for the study were then reviewed (yellow bubble). Step 2 is the elimination of cases during pathology report review. FFPE blocks were then retrieved from appropriate on- and off-site storage facilities (red bubble). Step 3 is the elimination of cases during this retrieval due to inaccessibility or missing cases. Retrieved samples were then processed, generating H+E slides that were histopathologically evaluated by a pathology review board (purple bubble). Step 4 is the elimination of cases during histopathological analysis of H+E slides. All remaining cases were brought forward for molecular testing for the ECHO study (green bubble).

3.4.2 Step 1: Initial Identification of Cases in the NCRI Database

Step 1 (Figure 3.1), the initial scan of the NCRI database by NCRI collaborators, yielded the discovery of 5792 incident tumours potentially viable for use in the ECHO study. Once eligibility criteria in Sections 3.3.1 and 3.3.2 were applied to the database, 4166 cases remained. This number represented 2831 dead patients and 1335 alive patients. All dead patients were eligible for the study, and 595 (44.56%) of alive patients consented to participate. Thus, the 4166 cases were further reduced to 3426 after patient consent

procedures described in Section 3.3.3 were completed. Once this consented database was generated by the NCRI collaborators, specimen numbers were assigned a random reference study number which was used to identify the relevant block throughout all steps in the study, and to provide a link to a file of anonymised socio-demographic, clinical, treatment, and outcome data held securely in the NCRI. The Researcher therefore only ever had access to the reference study number.

3.4.3 Step 2: Review of Pathology Reports

Once step 1 (Figure 3.1) had been carried out by collaborators in the NCRI, the Researcher contacted the representing pathologists from the 14 potential hospital sites to review case pathology reports (step 2 in Figure 3.1). It was possible to establish protocols with 8 different hospital sites on the basis of this contact. The Researcher made a sincere effort to target hospitals with large head and neck treatment facilities, and to draw from a diverse range of hospitals around the country. It should also be noted that the Royal Victoria Eye and Ear University Hospital specimens were accessible through and in St. James' University Hospital storage and were thus considered merged as a site with the latter. Any reference to St. James' University Hospital therefore includes cases from both of these hospitals. The 8 (merged 7) hospitals that ultimately participated in the study are summarized below in Table 3.2.

Table 3.2 An exhaustive list of the hospitals included in the ECHO study and for which the Researcher was able to contact and establish pathology report review protocols.

Ultimately, The Royal Victoria Eye and Ear Hospital samples were accessible through St. James' University Hospital and these were from then on considered a merged site.

Beaumont University Hospital, Dublin
Cork University Hospital, Cork
St. James' University Hospital, Dublin
St. Vincent's University Hospital, Dublin
Kerry General Hospital, Tralee
Mater Misericodordiae University Hospital, Dublin
The Royal Victoria Eye and Ear Hospital, Dublin
University Hospital Limerick, Limerick

Report review for these hospitals required months of logistical preparation as each site had its own desired protocols. These included providing the Researcher with log-in details to electronic systems in many cases. Table 3.3 is a summary of the protocols established with each involved hospital for pathology report review, and the parties that undertook organization of these protocols and the actual review of the reports. It should be stressed that no matter how reports were accessed, either electronically or on printed paper, the Researcher was blinded in all cases from any patient data during the review of pathology reports, with the only reference to patients accessible being the anonymized study number.

Table 3.3 Summary of protocols established to review pathology reports, and the parties involved in both the organization of report review and the execution of report review itself.

Hospital	Description of Protocol Established	Parties Involved in Pathology Report Review
Cork University Hospital	Single meeting on hospital site to organize report review, followed by set-up of unique electronic log-in for report access, followed by the Researcher's review of all reports on-site.	**Organization**: Researcher, On-site Pathologist, Chief Medical Scientist **Report Review**: Researcher
Kerry General University Hospital	Phone meetings followed by review of pathology reports by on-site Pathologist themselves.	**Organization**: Researcher, On-site Pathologist, Chief Medical Scientist **Report Review**: On-site Pathologist
St. James' University Hospital	Several on-site and phone meetings to organize report review, followed by set-up of non-unique electronic log-in for report access, followed by the Researcher's review of all reports on-site.	**Organization**: Researcher, On-site Pathologist, Chief Medical Scientist **Report Review**: Researcher
Beaumont University Hospital	Several on-site and phone meetings to organize report review, followed by set-up of non-unique electronic log-in for report access, followed by the Researcher's review of all reports on-site.	**Organization**: Researcher, On-site Pathologist, Chief Medical Scientist **Report Review**: Researcher
St. Vincent's University Hospital	Several on-site and phone meetings to organize report review, followed by set-up of non-unique electronic log-in for report access, followed by the Researcher's review of all reports on-site.	**Organization**: Researcher, On-site Pathologist, Chief Medical Scientist **Report Review**: Researcher
University Hospital Limerick	Phone meetings and e-mail correspondence to organize report review, followed by print-off and compilation of necessary reports by on-site Pathologist, followed by the Researcher's review of all reports on-site.	**Organization**: Researcher, On-site Pathologist, Chief Medical Scientist **Report Review**: Researcher
Mater Misericordiae University Hospital	Several on-site and phone meetings to organize report review, followed by set-up of non-unique electronic log-in for report access, followed by the Researcher's review of all reports on-site.	**Organization**: Researcher, On-site Pathologist, Chief Medical Scientist **Report Review**: Researcher

In these 7 hospital sites, the Researcher was able to gain access to a total of 2527 pathology reports which she reviewed (with the exclusion of those in Kerry General University Hospital as indicated in Table 3.3). Given that most pathologists preferred for the Researcher to travel to the site and make use of their respective physical and electronic pathology report files, the pathologists made sure that information accessed by the Researcher was only

associated with the relevant laboratory/histology number for the specimen. This was essential to ensure patient anonymity.

The 2527 pathology reports were analyzed in-depth to confirm as accurately as possible, that the primary site of the tumour conformed to the study requirements as described in Sections 3.3.1 and 3.3.2; to confirm the relevant tumour's pure SCC pathology; to ensure that the tumour was invasive in pathological nature; and to select the precise blocks associated with the case that contained tumour tissue fitting all of the above criteria. Ultimately, this process was carried out to avoid processing a large number of blocks for every case simply to find the singular block with the most tumour. Not only did this minimize the number of blocks withdrawn and carried cross-country, but it also minimized the waste of a large amount of tissue. For the most part, reviewing reports narrowed down the number of blocks to only those potentially relevant, especially for cases with many (sometimes over 20) blocks.

Some case numbers retrieved pathology reports in hospital systems that re-directed to another case number associated with the same patient that actually diagnosed the relevant tumour. The Researcher accepted this redirection after consulting with hospital-based pathologists to confirm that the patient was indeed the same for each case number given the anonymous nature of the review process.

It should also be noted that in the review of each pathology report, the Researcher prioritized obtaining FFPE biopsy and/or resection specimens for all patients in the study population. For patients whose ultimate treatment was non-surgical, the tissue samples used were those obtained at diagnosis. For those patients treated surgically, resected tumour samples were used if others obtained at diagnosis were unavailable.

3.4.4 Step 3: The Retrieval of FFPE Blocks

Once relevant blocks were identified through pathology report review, the Researcher organized FFPE block retrieval and transport protocols with the representing pathologists responsible for on-site FFPE tissue storage, and with the private companies responsible for

FFPE tissue storage off-site (Step 3 in Figure 3.1). Each hospital had its own protocol and individualized process for acquiring blocks.

Many of the sites made use of private off-site storage which generally charged for the pulling of a single block. This charge ranged from 5EU per block to 20EU per block. Given that this was an unforeseen expense, no budget remained in the study for this paid pathway of block acquisition. The Researcher therefore decided to negotiate with the off-site storage companies individually for every hospital to visit the private off-site warehouses herself. The Researcher agreed with each company and hospital that visiting the warehouses would be possible and would come at no cost to the project.

Table 3.4 below summarizes by hospital the protocols established for block retrieval and the parties involved in organizing and executing retrieval.

Table 3.4 Summary of protocols established for block retrieval and the parties involved in the organization of block retrieval and the execution of block retrieval itself.

Hospital	Description of Protocol Established	Parties Involved in Block Retrieval
Cork University Hospital	Telephone and in-person meetings with pathologist and medical scientist to organize retrieval, followed by Researcher visit to on-site storage to retrieve blocks. Discussions with off-site company unable to yield cost-free retrieval of other blocks.	**N/A**: It was not possible to establish a cost-free retrieval of blocks at this site.
Kerry General University Hospital	Retrieval of all blocks required carried out by on-site Pathologist themselves followed by Researcher in-person pick-up of blocks retrieved.	**Organization**: Researcher, On-site Pathologist, Chief Medical Scientist **Retrieval**: On-site Pathologist
St. James' University Hospital	Pathologist withdrew first half of blocks by internal request. The Researcher then organized the collection of the next half of blocks through a secondary application process to a BioBank within the hospital in line with newly introduced GDPR regulations. The Researcher then visited private off-site storage facilities, searched for, and collected the blocks.	**Organization**: Researcher, On-site Pathologist, Chief Medical Scientist, St. James' Private Hospital Storage Management, St. James' Biobank Ireland Trust **Retrieval**: On-site Pathologist, Researcher
Beaumont University Hospital	Researcher organized with on-site pathologist and chief medical scientist to withdraw on-site blocks from internal hospital storage. Lengthy telephone and in-person meetings with private off-site company, pathologists, and medical scientist to organize retrieval, followed by Researcher off-site storage visits to acquire blocks.	**Organization**: Researcher, On-site Pathologist, Chief Medical Scientist, Oasis Private Storage Management Group, Beaumont University Hospital Internal Storage Management **Retrieval**: Researcher
St. Vincent's University Hospital	Researcher organized with on-site pathologist and chief medical scientist to withdraw on-site blocks from internal hospital storage.	**Organization**: Researcher, On-site Pathologist, Chief Medical Scientist, St. Vincent's Internal Hospital Storage Management **Retrieval**: Researcher
University Hospital Limerick	Researcher organized with on-site pathologist and chief medical scientist to withdraw on-site blocks from internal hospital storage. Lengthy telephone and in-person meetings with private off-site company, pathologists, and medical scientist to organize retrieval, followed by Researcher off-site storage visits to acquire blocks.	**Organization**: Researcher, On-site Pathologist, Chief Medical Scientist, DSM Private Data Management and Storage Ltd, University Hospital Limerick's Internal Hospital Storage Management **Retrieval**: Researcher
Mater Misericordiae University Hospital	Researcher organized with on-site pathologist and chief medical scientist to withdraw on-site blocks from internal hospital storage.	**Organization**: Researcher, On-site Pathologist, Chief Medical Scientist, Mater Misericordiae, University Hospital Internal Hospital Storage Management **Retrieval**: Researcher

The Researcher retrieved all blocks for the ECHO study except in the case of the 66 blocks from Kerry General University Hospital and those from Cork University Hospital given that it

was not possible to establish a cost-free retrieval of blocks at this site (Table 3.4). The on-site Pathologist for St. James' University Hospital also retrieved the first half of the blocks for this site by internal request (Table 3.4).

It was possible to acquire a total of 1115 blocks from the on- and off-site storage sites for the participating hospitals. Once these blocks were acquired, the Researcher transported the blocks to the Coombe Women and Infant's University Hospital (CWIUH) in secure, waterproof boxes. CWIUH was the study center for the ECHO project, and the site for all block processing. No couriers were used at the request of the pathologists. The specimen blocks were stored securely in closed FFPE-specific boxes within the Molecular Pathology laboratory at the CWIUH, which was only accessible by laboratory staff with fob security access. The study database documenting these cases was also established using an encrypted, password-protected Microsoft Access file and was kept in the fob-secured laboratory office.

3.4.5 Step 4: The Generation and Histopathological Review of H+E Slides

Once the 1115 blocks were transported to the CWIUH, they were processed as described later in Sections 3.6.2 and 3.6.3. This processing included the generation of two new H+E slides, one before and after molecular sections were cut. Thus, the Researcher newly generated a total of 2230 H+E slides for the study. For anticipated p16 immunohistochemistry in future studies, another slide was cut for each case, yielding another 1115 slides. In total, the Researcher generated 3345 slides.

The 2230 H+E slides were histopathologically reviewed by a board of six pathologists, with 20% of them being reviewed by two of these pathologists to ensure consistency in diagnosis. Blocks were again evaluated on the basis of their eligibility according to Sections 3.3.1 and 3.3.2. Only those deemed eligible were brought forward for molecular testing. After the review of these slides for the 1115 FFPE blocks retrieved, 861 were ultimately included in the study.

3.4.6 Overview of Acquisition of FFPE Blocks and the Cases Eliminated During Each Step of the Process

Cases were eliminated at each stage of the FFPE block acquisition process due to eligibility concerns and inaccessibility. A summary of the blocks remaining eligible at each step of the process outlined in Figure 3.1 is showcased in Table 3.5 below. It should be noted that given that NCRI collaborators generated the study database during Step 1, no information on the hospitals from which cases were eliminated during this step was available. Information was only available for the 2527 cases for which pathology reports were accessible and reviewed.

Table 3.5 Number of cases remaining in the study after each of steps 1, 2, 3, and 4 involved in sample identification and retrieval summarized in Figure 3.1.

Hospital | Step	After step 1: Number of cases for which pathology reports were accessed/reviewed	After step 2: Number of cases eligible after pathology report review	After step 3: Number of cases found during block retrieval	After step 4: Number of samples eligible after H+E slide review
Cork University Hospital	232	220	0	0
Kerry General University Hospital	66	65	65	40
St. James' University Hospital	1172	1079	357	269
Beaumont University Hospital	313	290	235	156
St. Vincent's University Hospital	136	118	60	52
University Hospital Limerick	315	277	268	237
Mater Misericordiae University Hospital	293	248	130	107
Total	2527	2297	1115	861

An in-depth analysis of the procedures, parties, lengths of time taken to acquire FFPE blocks, and reasons for case elimination is the subject of Chapter 4.

3.5 Overview, Principles, and Application of Laboratory Methodologies

3.5.1 Overview of Laboratory Methodologies

Once brought to the CWIUH, FFPE blocks were subjected to a variety of laboratory techniques and processes. These included sectioning, slide-making, slide-staining, extraction, polymerase chain reaction (PCR), and genotyping.

The project commenced with an initial pilot study comparing two different HPV DNA-detecting technologies and sterility protocols. The first attempt of the pilot study used an "original sterility protocol" which is described later in Section 3.6.2. The first attempt made use of both the Multiplex PCR Luminex Genotyping and SPF10 PCR Gel Electrophoresis technologies described later in Section 3.7. Evidence of contamination appeared using this "original sterility protocol" with the Multiplex PCR Luminex Genotyping technology. No such evidence appeared using the "original sterility protocol" with the SPF10 PCR Gel Electrophoresis. Thus, a second attempt of the pilot study was conducted using the "IARC sterility protocol" described later in Section 3.6.2 paired with the Multiplex PCR Luminex Genotyping technology alone. The comparisons of these technologies and sterility protocols are the subject of the analysis of Chapter 5.

After this pilot study was conducted, the full ECHO study was carried out with the most valid sterility protocol and technology determined by the pilot study: The "IARC sterility protocol" using the Multiplex PCR Luminex Genotyping technology for HPV DNA detection.

For ease of understanding, the schematic Figure 3.2 below is an exhaustive summary of the techniques that were applied, in what order they were applied, and to how many samples they were applied.

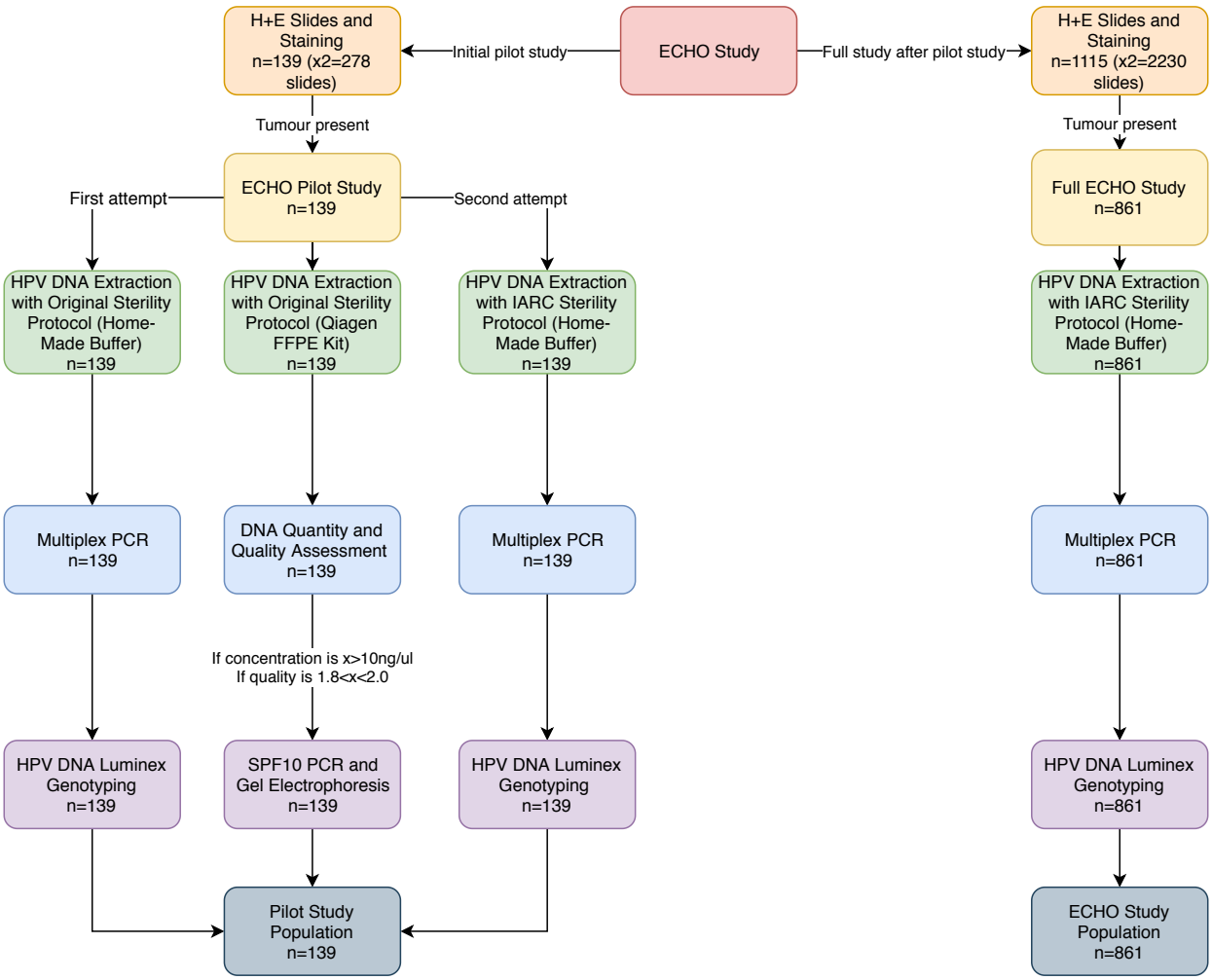

Figure 3.2 Schematic diagram representing the laboratory techniques applied to cases throughout the ECHO study.

An initial pilot study was conducted on 139 cases to validate two types of HPV DNA-detecting technologies and sterility protocols (left-hand side of diagram). The first attempt of the pilot study made use of an "original sterility protocol" outlined in Section 3.6.2. The sections cut using this protocol were tested using both Multiplex PCR Luminex Genotyping and SPF10 PCR Gel Electrophoresis. The extraction procedures for these two different technologies were different, the Multiplex PCR making use of a home-made digestion buffer described in Section 3.7.3.2, and the SPF10 PCR Gel Electrophoresis using the QIAGEN QIAamp® DNA FFPE Tissue Kit (QIAGEN Ltd, Manchester, UK)[5] method outlined in Section 3.7.2.1. Contamination issues were detected in results using this sterility protocol with the Multiplex PCR Luminex Genotyping method alone. The second attempt (left-hand side of diagram) of the pilot study therefore refers to the same 139 cases being re-sectioned with a new sterility protocol established by IARC ("IARC Sterility Protocol") and re-tested using the Multiplex PCR Luminex Genotyping. No contamination issues were detected here and results were used for pilot study analysis and determination of the best methodology for the overall ECHO study. The full ECHO study (right-hand side of diagram) was thus conducted after the pilot study, including 1115 cases sectioned with the "IARC sterility protocol", of which 861 had relevant tumour tissue that was tested using the Multiplex PCR Luminex Genotyping technology.

It should be noted that all procedures were carried out in the CWIUH. However, though cases subjected to Multiplex PCR Luminex Genotyping (Figure 3.2) were sectioned and extracted by the Researcher in the CWIUH, testing using this technology was carried out in the Infections and Cancer Biology Laboratory in IARC, Lyon, France. The samples were sent by DHL in batches. The Researcher attended the laboratory in IARC in the summer of 2018 to be trained on the technology and to test two batches of the samples (n=330) sent in that time-period. The rest of the samples were tested by the laboratory in IARC without the Researcher present.

3.6 Sectioning of FFPE Blocks, and Creation and Review of H+E Slides

All cases in the ECHO study and the initial pilot study were subjected to sectioning using a microtome (Figure 3.2) by the Researcher in the CWIUH. The coming sections describe the principles of microtomy and the precise protocols that were applied to the FFPE blocks in both the initial pilot study and the ECHO study as a whole.

3.6.1 Principles of Microtomy

Microtomes are instruments for cutting extremely thin sections of material for examination under a microscope. In many cases, they are used to cut thin sections of um thickness from rectangular FFPE blocks. Sections are used to create slides of varying types including H+Es to diagnose tissue type and immunohistochemistry slides to stain for the presence of biomarkers including proteins.

3.6.2 Sectioning of FFPE Blocks

The ECHO study and its initial pilot study made use of a fully manual microtome. The microtome model was: Leica® RM2135 (Leica Biosystems, Wetzlar, Germany)[6]. This microtome however stopped functioning after the pilot study and 600 cases cut for the study as a whole. A replacement was used for the rest of the sections cut for the whole study. The replacement microtome was a semi-automated one which allowed for either manual cutting of each section or electronic cutting of each section using only a button to stop and start the process. The model used for the rest of the 1115 cases cut for the whole study was: pfm Rotary 3006 EM (pfmmedical, Köln, Germany)[7].

The sectioning protocol used in the ECHO study and its initial pilot study was that of the international HPV-AHEAD Study Consortium[8,9], as directed by the Infections and Cancer Biology Laboratory in IARC, Lyon, France. The sectioning allowed for the cutting of two H+E slides (S1 and S9 in Figure 3.3), 1 p16 slide (S2 in Figure 2.3) for future use, three sections for HPV mRNA detection (T1 in Figure 3.3) for future use, and three sections for HPV DNA Multiplex PCR Luminex Genotyping (T2 in Figure 3.3). It should be noted that for the purposes of the first attempt of the initial pilot study (Figure 3.2), T1 sections in Figure 3.3 were actually those cut for SPF10 PCR Gel Electrophoresis, and T2 sections in Figure 3.3 were cut for the first attempt of the Multiplex PCR Luminex Genotyping. For the second attempt of the initial pilot study (Figure 3.2), only T1 sections were cut for Multiplex PCR Luminex Genotyping alone.

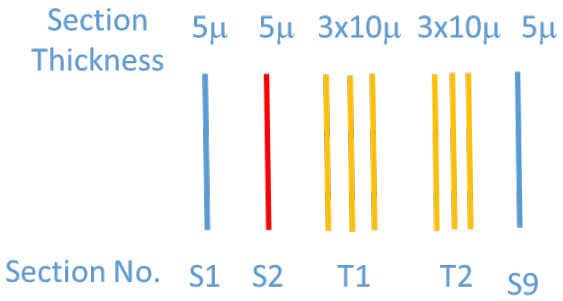

Figure 3.3 Sectioning protocol for each FFPE block in the ECHO study as defined by the HPV-AHEAD[8,9] study directed by IARC, Lyon, France.

Sections S1 and S9 are H+E slides, one cut at the beginning and end of processing. S2 is a p16 slide cut for future immunohistochemistry. These three sections were all cut at 5um as pictured. For the initial pilot study and first attempt, T1 sections were cut for SPF10 PCR Gel Electrophoresis. T2 sections were cut for Multiplex PCR Luminex Genotyping. For the second attempt of the pilot study with altered sterility protocols, only T1 sections were cut for Multiplex PCR Luminex Genotyping. For the ECHO study as a whole, T1 sections were cut for future HPV mRNA testing, and T2 sections were cut for HPV DNA Multiplex PCR Luminex Genotyping.

The functional and sterility protocols for the cutting of all sections in Figure 3.3 is summarized in Figure 3.4. The "IARC Sterility Protocol" mentioned previously (Figure 3.2) is that described exactly in Figure 3.4. It makes use of DNA ZAP™ (ThermoFisher™ Scientific, MA, USA)[10] and 70% ethanol to clean the microtome. This was the sterility protocol used in the second attempt of Multiplex PCR Luminex Genotyping in the initial pilot study, and for

the full ECHO study (Figure 3.2). The first attempt of Multiplex PCR Luminex Genotyping and the SPF10 PCR Gel Electrophoresis in the initial pilot study (Figure 3.2) instead used Industrial Methylated Spirits (IMS) and Parapel™ (ThermoFisher™ Scientific, MA, USA)[11], a paraffin wax melting solution, to clean the microtome. This is the previously referred to "Original Sterility Protocol" (Figure 3.2).

1. Place FFPE blocks ready for sectioning on a cooling plate to be chilled.
2. Prepare all Leica® Bond Plus slides (Leica Biosystems, Wetzlar, Germany)[12] and DNase/RNase-free 1.5ml Micro tubes (Sarstedt, Wexford, Ireland)[11] with study ID numbers.
3. Wearing a new set of disposable gloves, remove any blade left in the microtome and clean the microtome completely with DNA ZAP™ (ThermoFisher™ Scientific, MA, USA)[10] and 70% Ethanol using disposable tissue paper.
4. Insert a new microtome blade.
5. Mount a chilled FFPE specimen block on to the microtome.
6. Trim the block by flattening and removing any rough surface at 10um.
7. Clean the microtome thoroughly with DNA ZAP™ (ThermoFisher™ Scientific, MA, USA)[10] and 70% Ethanol using disposable tissue paper to minimize cross contamination between specimens by debris created in step 6.
8. Cut two consecutive 5um sections onto separate Leica® Bond Plus slides (Leica Biosystems, Wetzlar, Germany)[12], one for H+E staining (S1 in Figure 3.3), and the other for future p16 immunohistochemistry (S2 in Figure 3.3).
9. Cut three consecutive 10um sections for future mRNA extraction (T1 in Figure 3.3).
10. Collect the sections into a sterile DNase/RNase-Free 1.5ml Micro tube (Sarstedt, Wexford, Ireland)[11] using sterile tweezers or a disposable pair of tweezers.
11. Cut three consecutive 10um sections for DNA extraction depicted (T2 in Figure 3.3).
12. Collect the sections into a sterile DNase/RNase-Free 1.5ml Micro tube (Sarstedt, Wexford, Ireland)[11] using sterile tweezers or a disposable pair of tweezers.
13. Cut one 5um section onto a Leica® Bond Plus slide (Leica Biosystems, Wetzlar, Germany)[12] for H+E staining (S9 in Figure 3.3).
14. Return to step 3 above and process a new FFPE specimen block, until 10 specimens have been processed. Then, move to step 15.
15. Process a blank block negative control, beginning by wearing a new disposable glove set, cleaning the microtome and surface and bench with DNA ZAP™ (ThermoFisher™ Scientific, MA, USA)[10] and 70% ethanol using disposable tissue paper, and inserting a new microtome blade.
16. Mount the blank paraffin block and cut three consecutive 10um sections for a future mRNA control, and another 10um section for a DNA control, collecting each set of 3 sections into separate DNase/RNase-Free 1.5ml Micro tubes (Sarstedt, Wexford, Ireland)[11] using a sterile pair of tweezers or a disposable pair of tweezers.
17. Repeat steps 15 through 17 using a block of SiHA or HeLA cells as a positive control.
18. Start a new set of 10 samples with step 1.

Figure 3.4 Detailed sectioning and sterility protocols for the ECHO study with reference to sections outlined in Figure 3.3.

For the second attempt of the initial pilot and the ECHO study as a whole, this precise protocol was used and named the "IARC Sterility Protocol". The first attempt of the initial pilot study for the ECHO study (Figure 3.2) used this protocol as well but with changes in the sterilization products used. Anywhere in this protocol where DNA ZAP™ (ThermoFisher™ Scientific, MA, USA)[10] or 70% ethanol are mentioned, IMS and Parapel™ (ThermoFisher™ Scientific, MA, USA)[11] were used instead. This was called the "Original Sterility Protocol."

3.6.3 Creation and Review of H+E Slides

The sectioning protocol described in section 3.6.2 includes sections cut for H&E and p16 slides for future staining. A more detailed sub-protocol for the fabrication of these slides immediately after sections were cut is detailed below in Figure 3.5. For both H+E and p16 slides for future staining, Leica® Bond Plus charged slides (Leica Biosystems, Wetzlar, Germany)[12] were used.

1. Before sectioning any blocks, prepare cold (room temperature) and hot water (37°C) baths and a hot plate (40°C) (Bibby Scientific Ltd, Staffordshire, UK)[21].
2. Before sectioning any blocks, prepare stable slide racks.
3. Before sectioning any blocks, pre-heat oven (ThermoFisher™ Scientific, Waltham, MA, USA)[21] to 70°C.
4. After cutting section S1 in Figure 3.3, place the section into the cold water bath with tweezers.
5. Scoop up the flat section using the prepared Leica® Bond Plus slide (Leica Biosystems, Wetzlar, Germany)[12].
6. Drop the section gently into the hot water bath (Bibby Scientific Ltd, Staffordshire, UK)[21] from the slide to remove any folds in the paraffin.
7. Scoop up the flat, unfolded, section again with the slide and place the slide on the hot plate (Bibby Scientific Ltd, Staffordshire, UK)[21].
8. Label the slide in pencil.
9. Repeat steps 3 to 8 for section S2 in Figure 3.3.
10. Repeat steps 3 to 8 for section S9 in Figure 3.3.
11. Once a maximum number of slides able to fit in the oven are on the hot plate, place the slides into stable slide racks.
12. Bake the slides at 70°C for 30 minutes in the oven (ThermoFisher™ Scientific, Waltham, MA, USA)[21].
13. Stain slides S1 and S9 as H+Es and store S2 slides for p16 for future studies.

Figure 3.5 Step-by-step protocol for the fabrication of H+E slides and p16 slides for future staining.

This protocol is standard in hospital histopathology departments and is drawn from the procedure used in St. James' University Hospital and the CWIUH in Dublin. This protocol was an integral part of the sectioning protocol in Figure 3.4, as slides were fabricated, but not stained, immediately after sections S1, S2, and S9 (Figure 3.3) were cut. Leica® Bond Plus slides (Leica Biosystems, Wetzlar, Germany)[12] were used for S1, S2, and S9 (Figure 3.3). S1 and S9 were stained as H+Es, and S2 was kept for future p16 staining.

Slides generated with sections S1 and S9 in Figure 3.3 using the protocol described in Figure 3.5 were stained using the Ventana HE 600 System (Roche, Basel, Switzerland)[13]. This is an automated machine that offers individual slide staining using racks of 20 slides at a time. The system allows for the incubation of sections followed by the simultaneous staining, dehydrating, and coverslipping of slide trays.

Once sections S1 and S9 (Figure 3.3) were mounted on to Leica® Bond Plus slides (Leica Biosystems, Wetzlar, Germany)[12], they were placed on a Ventana staining slide tray and loaded onto the system. The system first incubated the slides at 70°C or above for a further 30 minutes to that applied in Figure 3.5 to ensure adherence of the tissue section. The slides were then washed using the Ventana HE600 washing solution (Roche, Basel, Switzerland)[13] and stained with Ventana HE600 Haematoxylin (Roche, Basel, Switzerland)[13]. The Haematoxylin was then blued with the Ventana HE600 Bluing agent (Roche, Basel, Switzerland)[13]. Ventana HE600 Eoisin (Roche, Basel, Switzerland)[13] was applied followed by the Ventana HE600 Organic Solution (Roche, Basel, Switzerland)[13]. The Ventana Eoisin (Roche, Basel, Switzerland)[13] was an acidic dye that had an affinity for basic tissue elements (cytoplasm, most connective tissue fibers), and stained these various shades of pink, red, and orange. Following this, differentiation was attained using the Ventana HE600 Acid Differentiator (Roche, Basel, Switzerland)[13]. A sequence of Ventana HE600 Transfer Fluid (Roche, Basel, Switzerland)[13] and accompanying Cleaning Solution (Roche, Basel, Switzerland)[13] was then applied before the slides were coverslipped.

The system coversliped the slides with the use of the Ventana HE600 Coverslipper Activator (Roche, Basel, Switzerland)[13] and pre-glued Ventana HE600 Glass Coverslips (Roche, Basel, Switzerland)[13]. Once slides were stained, they were stored in 100-slide boxes in the fob-accessed Molecular Pathology Laboratory of the CWIUH.

The slides were then reviewed by a member of a six-person Pathology Review Board composed of pathologists from four different hospitals. The pathologists assessed the slides to confirm the presence of relevant tumour according to the eligibility criteria in Sections 3.3.1 and 3.3.2. A subset of 20% of cases were reviewed a second time by another member of the Pathology Review Board to confirm diagnosis.

If diagnosis of pure SCC was not confirmed for both H+E slides (S1 and S9 in Figure 3.3) for the selected FFPE block, the Researcher returned to the database to select another FFPE block identified as containing tumour by its associated pathology report. Protocols were repeated and both H+E slides (S1 and S9 in Figure 3.3) were again reviewed. This was

repeated until all FFPE specimen blocks highlighted in the pathology reports as containing relevant tumour were examined. If none contained appropriate tissue according to review of the H+E slides, the case was excluded from the study.

For the ECHO study as a whole, 1115 cases with two H+E slides each were histopathologically analyzed. Those generated for the 139 pilot study cases were also reviewed. With two H+Es per case, 2230 H+E slides were analyzed for the ECHO study as a whole along with an additional 278 for the pilot study. For the ECHO study as a whole however, this is a conservative estimate as several blocks had to be cut for multiple cases to find appropriate tissue. 861 cases remained in the ECHO study as a whole after H+E review.

3.7 HPV DNA Detection

T1 and T2 sections (Figure 3.3) were subjected to HPV DNA detection following sectioning. For the initial pilot study, two HPV DNA-detecting technologies were used: SPF10 PCR Gel Electrophoresis and the Multiplex PCR Luminex Genotyping technology. Only Multiplex PCR Luminex Genotyping was used for the ECHO study as a whole. The coming section outlines the principles of PCR forming the basis for both technologies, and proceeds to describe each DNA-detecting technique in detail including the principles of the Luminex xMAP® Technology (Luminex Corp, TX, USA)[14,15].

3.7.1 Principles of PCR

PCR is a powerful biochemical method for amplifying a pre-selected segment of DNA. It is a distinct process, different from both cloning and propagation within the host cell.

PCRs are composed of a single thermal cycle repeated many times to denature DNA, segregate double-stranded DNA, and synthesize a new strand based on the template of those single strands. The first step of a PCR thermal cycle entails the heating of extracted DNA to sever the bonds between the two strands forming double-stranded DNA to produce single strands. These single strands form a template for the desired sequence that is to be amplified. Following this, a sequence-specific oligonucleotide primer pair is annealed to each DNA single strand template. Third, a thermostable *Taq* DNA polymerase enzyme

directs the synthesis of a complimentary DNA strand to the DNA template formed by the initial single strand using deoxynucleotide substrates. These nucleotides are added to the 3' end of the primer, thus generating an extended region of double stranded DNA for each template, thereby doubling the DNA in each cycle.

The thermal cycle is then repeated multiple times to significantly amplify the target sequence. Results can be measured by various methods including gel electrophoresis and laser-based detection.

3.7.2 SPF10 PCR and Gel Electrophoresis

For the initial pilot study (Figure 3.2) on 139 cases, T1 sections in Figure 3.3 were subjected to HPV DNA detection through SPF10 PCR and Gel Electrophoresis. This was a binary detection process whereby the presence or absence without genotype of HPV DNA was determined. This process was conducted by the Researcher in the CWIUH and included DNA extraction, quantification and purity analysis of extracted DNA using a spectrophotometer, SPF10 PCR, and gel electrophoresis.

3.7.2.1 DNA Extraction

Based on the optimization of DNA extraction and cost-effectiveness analysis[16], the QIAGEN QIAamp® DNA FFPE Tissue Kit (QIAGEN Ltd, Manchester, UK)[5] was used to extract DNA from FFPE tissue sections T1 (Figure 3.3) in accordance with the manufacturer's instructions (QIAGEN Ltd, Manchester, UK)[5]. All associated positive (HeLa or SiHa) and negative (empty paraffin or placenta-containing paraffin blocks) controls that were cut alongside each set of 10 samples were also extracted.

The extraction was performed in a ventilated and sterilized hood: ESCO Airstream® Class II BSC Hood (ESCO, Barnsley, UK)[13]. Extraction commenced with the deparaffinisation of the sections by adding 1ml of xylene to each sample. Samples were then vortexed vigorously for 10 seconds and centrifuged at full speed for 2 minutes at room temperature. Supernatant was removed, and 1ml of ethanol (96-100%) was added to the remaining pellets for each sample. All were again vortexed and centrifuged for 2 minutes at full speed at room

temperature. Supernatant was again removed. All samples, with their tube lids open, were left to incubate at room temperature for 10 minutes, or, until all residual ethanol had evaporated.

All sample pellets were then resuspended in 180ul Buffer ATL from the kit. 20ul of proteinase K (20 mg/ml) from the kit was added to each sample. All were then vortexed. Samples were then incubated overnight at 56°C to allow for the best possible lysis. Samples were then incubated for 1 hour at 90°C and briefly centrifuged.

200ul Buffer AL was added to the samples, followed by vortexing, the addition of 200ul of ethanol (96-100%), further vortexing, and brief centrifugation. The remaining lysate was transferred to a QIAamp MinElute column that itself was placed in a 2ml collection tube. All columns were centrifuged at 8000 rpm for 1 minute. The 2ml collection tube containing the flow-through was disposed of and replaced with a clean collection tube.

500ul of Buffer AW1 was added to each of the columns, which were then centrifuged at 8000 rpm for 1 minute. With a replacement 2ml collection tube, this process was repeated with the addition of 500ul of Buffer AW2. Replacing the 2ml collection tube for the last time, all columns were centrifuged at 14000 rpm for 3 minutes to dry the membrane of the column completely. The collection tube was discarded and replaced with a clean 1.5ml DNase/RNase-free 1.5ml Micro tube (Sarstedt, Wexford, Ireland)[11].

50ul of Buffer ATE was then added to the column, and all samples were incubated at room temperature with their column lids open for 15 minutes to allow for maximal DNA yield. Lids were closed and all samples were then centrifuged at 14000 rpm 1 minute. Columns were disposed of, lids of the sample tubes were closed, and all samples were stored at -80°C.

3.7.2.2 Quantification and Purity Analysis of Extracted DNA using a Spectrophotometer

For these 139 cases and their controls, DNA concentration and quality was assessed using the Thermo Scientific NanoDrop 2000c Spectrophotometer (ThermoFisher™ Scientific, Waltham, MA, USA)[17].

Each sample was retained between two fiberoptic cables using surface tension. The machine was first calibrated by pipetting 1ul of DNase/RNase-free water (ThermoFisher™ Scientific, MA, USA)[18] onto the optical surface and closing the lever arm before zeroing on the Nanodrop software. The Nanodrop dock was then wiped with a clean, dry, lint-free tissue, and 1ul of the first sample was pipetted onto the optical surface for assessment. Between each sample, the optical surfaces were again cleaned with new clean, dry, lint-free tissue. The software assessed nucleic acid concentration, absorbance at 260nm and 280nm, and the spectra of each sample. Purity of DNA was assessed by absorbance ratio 260/80 and the spectra of each sample.

Only samples with a defined single-peak spectra, DNA concentration exceeding 10ng/ml, and a purity of 260/80 between 1.8-2.0 were carried forward for SPF10 PCR and Gel Electrophoresis. These cut-offs represented strict standards for DNA quantity and quality and have been utilized as a standard for further PCR[16].

3.7.2.3 SPF10 PCR

HPV DNA was amplified by the SPF10 short PCR fragment HPV primer set (Life Technologies Inc, Grand Island, NY, USA)[11]. This set has the potential to amplify at least 54 HPV types, and together, the primers amplify a 65-base pair fragment from the conserved L1 open reading frame in the viral capsid region of the HPV genome[19,20].

Each PCR reaction consisted of a mastermix mixture of 300nM of each of the six SPF10 primers (Table 3.6), 200uM deoxynucleoside triphosphates (Sigma Aldrich/Merck, Darmstadt, Germany)[21], 10x PCR buffer at 1x concentration (containing 10mM Tris-HCl pH 8.3) (ThermoFisher™ Scientific, MA, USA)[11], 50mM KCl (ThermoFisher™ Scientific, MA, USA)[11], 3mM $MgCl_2$ (ThermoFisher™ Scientific, MA, USA)[11], 1U of AmpliTaq Gold Taq polymerase (Life Technologies Inc, Grand Island, NY, USA)[22], and UltraPure™ DNAase/RNase free water (ThermoFisher™ Scientific, MA, USA)[18] .

Table 3.6 Sequences of HPV SPF10 primers used in SPF10 PCR. HPV DNA from all extracted samples was amplified using a master mix containing 300nM of each of the six of these primers.

Primer Name	Primer Sequence (5' to 3' and 11-45 bases)
SPF10 1A	GCICAGGGICACAATAATGG
SPF10 1B	GCICAGGGICATAACAATGG
SPF10 1C	GCICAGGGICATAATAATGG
SPF10 1D	GCICAAGGICATAATAATGG
SPF10 2B	GTIGTATCIACAACAGTAACAAA
SPF10 2D	GTIGTATCIACTACAGTAACAAA

The 17ul of the mastermix for each reaction was placed in a separate well of a marked Microamp 96-well PCR plate (ThermoFisher™ Scientific, MA, USA)[23], followed by the 3ul of each sample, yielding a total reaction volume of 20ul. The plate was then sealed with an optical seal and centrifuged briefly. The PCR was run on the 7500 real-time PCR System (ThermoFisher™ Scientific, MA, USA)[24], although it was not a real-time reaction. The plate was inserted into the 7500 real-time PCR System (ThermoFisher™ Scientific, MA, USA)[24] and the samples were labeled as per the plate markings.

The PCR was initiated by a five-minute denaturation and enzyme activation step at 95°C. There were two PCR stages, the first of which contained one step, and the second of which contained two steps. The first step of stage one was 30 seconds at 95°C for denaturation. The first step of stage two was 45 seconds at 52°C for annealing, followed by step two for 45 seconds at 72°C for extension. The PCR was completed after 40 cycles of the thermacycling. The corresponding negative and positive controls that were sectioned at the same time as a particular set of samples were processed through the PCR alongside those samples. These controls were used to monitor the performance of the PCR method in each experiment. The amplicons were then analysed by gel electrophoresis.

3.7.2.4 Gel Electrophoresis

Amplicons of SPF10 PCR were then subjected to Gel Electrophoresis to determine HPV status. This was done using 2.2% agarose with 16-wells and the Lonza FlashGel™ System (Lonza Inc, Basel, Switzerland)[25]. Together these visualized a 65bp PCR product by

comparison to a 50bp DNA ladder. This system allowed for fast separation of DNA fragments greater than 10bp. For each PCR run, clinical specimens were run alongside a negative template control of DNase-free water (ThermoFisher™ Scientific, MA, USA)[18] along with the negative tissue block and positive HeLA/SiHA tissue block controls.

Wells of the gels were first flooded with deionized water. 1ul of gel loading dye, Blue 6X (New England Biolabs Inc, Ipswhich, MA, USA)[26], was mixed by pipetting up and down with 4ul of each sample's PCR product in a new clean plate. The dye is a premixed tracking dye with loading buffer for agarose gel electrophoresis. It contains SDS to obtain sharper DNA bands, EDTA to chelate magnesium for stopping enzymatic reaction if present, and bromophenol blue which is a universal tracking dye for electrophoresis.

The homogenous mixture of sample PCR product with loading dye was loaded by pipetting into the sample wells of the 2.2% agarose gel. Each row in the 16-well gel contained 12 PCR products, 1 positive control (SiHA/HeLA block), 1 negative control (blank or placenta block), 1 PCR negative control (DNase-free water (ThermoFisher™ Scientific, MA, USA)[18]), and a DNA ladder (New England Biolabs Inc, Ipswhich, MA, USA)[26]. The ladder consisted of fragments ranging from 50 base pairs to 10 kilobases, with increased intensity at 1kb.

Once full, the 2.2% agarose FlashGel™ Casette (Lonza Inc, Basel, Switzerland)[25] was inserted in to the FlashGel™ Dock (Lonza Inc, Basel, Switzerland)[25] and the lid was closed. The Dock was attached to the power supply and a constant voltage of 180V was applied until the DNA separated enough to obtain as specific an amplicon length reading as possible. The accompanying FlashGel™ Camera (Lonza Inc, Basel, Switzerland)[25] was used to record the results. A case positive for HPV was determined by a band at 65bp level in relation to the DNA ladder on the recorded image. An example of wells captured by camera using the system is exhibited in Figure 3.6 below.

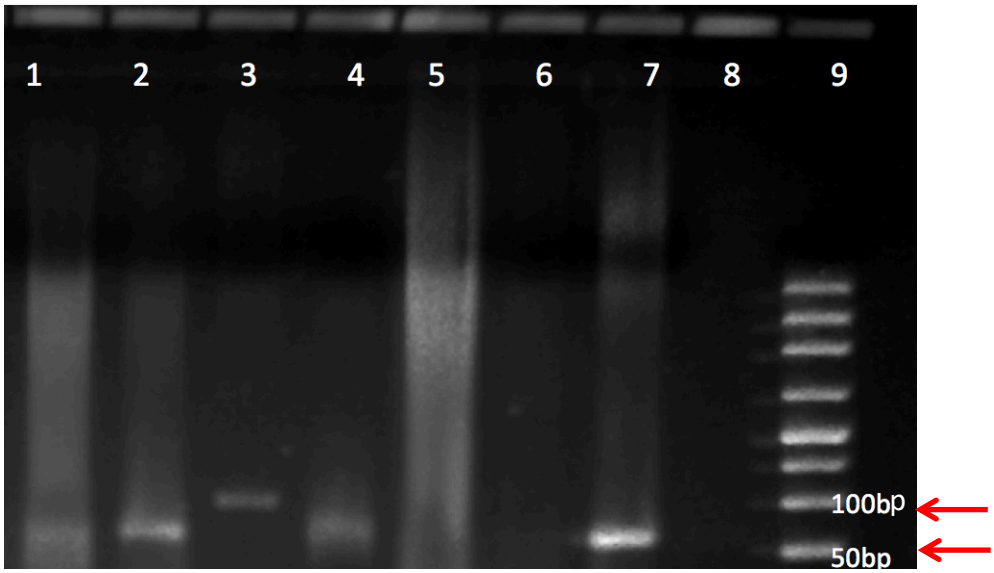

Figure 3.6 Image of Gel Electrophoresis of SPF10 Amplified DNA.

Lane 9: DNA ladder labelled at the 100bp and 50bp bands. Lane 8: PCR negative control (DNase/RNase free water). Lane 7: Positive control (SiHA). Lane 6: Blank block negative control. Lanes 1, 2, and 4: Positives indicated by a band at 65bp. Lanes 3 and 5: Negatives with no band at 65bp.

3.7.3 Multiplex PCR and Luminex Genotyping

For both the initial pilot study and the ECHO study as a whole (Figure 3.2), T2 sections in Figure 3.3 were subjected to another HPV DNA genotyping procedure. This involved DNA extraction, a Multiplex PCR, oligonucleotide probe coupling to fluorescent beads, a hybridization of PCR product to probes, and a fluorescence intensity measurement by the Luminex 200 Analyser (Luminex Corp, TX, USA)[27]. Thus, HPV genotype was detected by a type-specific multiplex genotyping (TS-MPG) assay, which combined Multiplex PCR and bead-based Luminex Technology (Luminex Corporation, TX, USA)[27].

This method was developed and optimized in the IARC[28,29] with the help of other collaborators[30,31]. It is the standardized method utilized for the HPV-AHEAD Study[8,9], the worldwide meta-analysis of HPV in HNC being conducted in the IARC and for which the present study will provide the Irish data.

The coming section describes the principles of Luminex xMAP® Technology (Luminex Corp, TX, USA)[14] and goes onto outline the precise methods used to extract DNA, implement the

Multiplex PCR, couple probes to fluorescent beads, hybridize PCR product to probes, and analyze HPV genotypes using the Luminex.

3.7.3.1 Principles of Luminex xMAP® Technology

The Luminex System (Luminex Corp, TX, USA)[27] is a flexible analyser based on the principles of flow cytometry[14,15,27]. The system enables the multiplex measuring (simultaneous measuring of multiple analytes) up to 100 analytes in a single microplate well using small sample volumes. The bioassay offers results in many formats including gene expression, transcription factor profiling, and immunoassay profiling.

The Luminex System (Luminex Corp, TX, USA)[27] is a combination of three central xMAP® technologies. The first is xMAP® microspheres, a family of 100 fluorescently dyed 5.6 micron-sized polystyrene microspheres that act as both the identifier and the solid surface to build the assay[15,27]. The second is a flow cytometry-based instrument, the Luminex Analyser (Luminex Corp, TX, USA)[27], which integrates key xMAP® detection components such a lasers, optics, advanced fluidics, and high-speed digital signal processors[15,27]. The third component is the assays that are designed around the microspheres to function with the system's integrated software.

The 5.6 micron polystyrene microspheres that xMAP® is based on are internally dyed with red and infrared fluorophores. Using different ratios of the two dyes for different batches of microspheres, up to 100 different microsphere sets can be created[14]. Thus, each bead is unique with a spectral signature determined be a red and infrared dye mixture. The bead is filled with a specific known ratio of the two dyes. As each microsphere carries a unique signature, the xMAP® detection system can identify to which set it belongs, allowing for the multiplexing of up to 100 tests in a single reaction volume[14,15,27].

In a typical molecule-based assay (e.g. DNA), these beads will be coupled to specific and known oligonucleotide probes, with one pre-determined coloured bead being associated with one probe. These probes, coupled to the bead, will then be hybridized to most commonly, a biotinlyated PCR product. If a particular probe, associated with its coupled colour, hybridizes, then the complimentary molecule exists in the sample.

To detect this hybridization by way of the associated fluorescent bead, the Luminex Analyser (Luminex Corp, TX, USA)[27] combines two lasers, fluidics, and real-time digital signal processing to distinguish the up to 100 different sets of colour-coded beads, each representing a different probe and assay.

With respect to fluidics, the reader detects individual beads by flow cytometry. The fluidics system aligns the beads into single file as they enter a stream of sheath fluid and then enter a flow cell. Once the beads are in a single file within the flow cell, each bead is individually interrogated for bead colour (e.g. the coupled probe hybridized to its associated molecule) and assay signal strength (e.g. streptavidin-R-phycoerythrin (Strep-PE) present as a result of biotinylation, representing fluorescence intensity).

There are two lasers in the system. The first is a 532nm green laser used to excite the Strep-PE dye of the assay. It determines the magnitude of the Strep-PE-derived signal, which is in direct proportion to the amount of bound analyte. Thus, a semi-quantitative measurement of the amount of HPV present in each sample is obtained. The second is a 635nm solid state laser used to excite the dyes inside the beads to determine their colour and thus their associated probe/molecule. It is also used for doublet discrimination by light scatter.

Finally, the system has four detectors. These detectors are used to measure the fluorescence of the assay, to determine what bead colour is being detected, and to discriminate between single and aggregate beads.

3.7.3.2 DNA Extraction

All samples were extracted in the CWIUH by the Researcher. To prepare samples for analysis using the Luminex, T2 (Figure 3.3) sections were extracted for HPV DNA using a home-made digestion buffer developed by the Infections and Cancer Biology Laboratory, IARC, Lyon, France for the HPV-AHEAD Study[9]. This homemade digestion buffer was prepared under sterile conditions in a ventilated hood (ESCO, Barnsley, UK)[13]. For each sample, 242.7ul of 10 mM Trizma® hydrochloride solution pH 7.4 diluted from 1M (Sigma Aldrich/Merck, Darmstadt, Germany)[32], 6.3ul of proteinase K (20mg/ml) (QIAGEN Ltd, Manchester, UK)[33],

and 1ul of Tween 20 (Sigma Aldrich/Merck, Darmstadt, Germany)[34] were needed. The Trizma® was diluted appropriately using DNAse/RNase free water (ThermoFisher™ Scientific, MA, USA)[18].

For each batch of 100 samples and 20 associated controls, appropriate volumes of each of Trizma®, proteinase K, and Tween 20 were made into a stock solution. 250ul of the digestion buffer was then added directly into each 1.5ml Micro tube (Sarstedt, Wexford, Ireland)[11] containing the T2 sections for each specimen. The samples were then incubated for 2 hours at 56°C under a 750rpm agitation using a GeneChip® Hybridization Oven 640 (ThermoFisher™ Scientific, Waltham, MA, USA)[21]. To inactivate the proteinase K, each tube was heated for 10 minutes on a Stuart® SBH130D Heat Block (Stuart®, Staffordshire, UK)[35] at 95°C. They were then centrifuged for 2 minutes at 13000rpm using the Eppendorf 5418R (Eppendorf, Hamburg, Germany)[36]. They were finally chilled on ice and stored in a -20°C freezer (ThermoFisher™ Scientific, Waltham, MA, USA)[11], for further use.

3.7.3.3 Multiplex PCR

Samples were then shipped by DHL to the Infections and Cancer Biology Laboratory, IARC, Lyon, France, for the rest of the HPV DNA detection procedures, including the Multiplex PCR. The Researcher traveled to the IARC to be trained in the Multiplex PCR and Luminex technology using a batch of samples sent (n=330), but the rest of the testing occurred without the Researcher present.

10ul of extracted DNA for each sample was subjected to Multiplex PCR described previously[28]. This PCR used HPV type-specific primers targeting the E7 region of 19 HR or probably HR HPV genotypes (16, 18, 26, 31, 33, 35, 39, 45, 51, 53, 56, 58, 59, 66, 68a, 68b, 70, 73, 82) and two LR HPV genotypes (6, 11). Detection limits ranged from 10 to 1,000 copies of the viral genome per reaction. Two primers for amplification of the β-globin gene were also included to control for the quality of the template DNA. The slight modification of the protocol described previously[29,37] for the amplification of shorter (~100bp) fragments for ten HPV genotypes (16, 18, 31, 33, 35, 52, 56, 66, 6, 11) and 117bp for β-globin were applied.

The primers used to amplify the HPV DNA were originally developed by IARC[38] with some updates later made for greater sensitivity[39]. These primers are summarized in Table 3.7 below. All reverse primers were biotinylated to allow later detection with a Luminex 200 Analyser (Luminex Corp, TX, USA)[27].

Table 3.7 Sequences of forward and reverse HPV-type specific primers and sizes of the PCR-amplified fragments[a] as developed and reported by IARC for Multiplex PCR.

All reverse primers were biotinylated to allow later detection with a Luminex Analyser (Luminex Corp, TX, USA)[27]. Associated GenBank Sequences that were used as references to develop the primers are also shown.

HPV Genotype	Primer Sequence[b]	PCR Fragment Size (bp)	GenBank Sequence
16	F. 5′-TTATGAGCAATTAAATGACAGCTCAG-3′ R. 5′-TGAGAACAGATGGGGCACACAAT-3′	212	K02718
18	F. 5′-GACCTTCTATGTCACGAGCAATTA-3′ R. 5′-TGCACACCACGGACACACAAAG-3′	236	X05015
26	F. 5′-CGAAATTGACCTACGCTGCTACG-3′ R. 5′-TGGCACACCAAGGACACGTCTTC-3′	239	X74472
31	F. 5′-AGCAATTACCCGACAGCTCAGAT-3′ R. 5′-GTAGAACAGTTGGGGCACACGA-3′	210	J04353
33	F. 5′-ACTGACCTAYACTGCTATGAGCAA-3′ R. 5′-TGTGCACAGSTAGGGCACACAAT-3′	229	M12732
35	F. 5′-CAACTGACCTATACTGTTATGAGC-3′ R. 5′-TGTGAACAGCCGGGGCACACTA-3′	234	M74117
39	5′-GGTTTGCAGTTGCACACCACGG-3′ R. 5′-GACACTGTGTCGCCTGTTTGTTTA-3′	357	M62849
45	F. 5′-GACCTGTTGTGTTACGAGCAATTA-3′ R. 5′-TGCACACCACGGACACACAAAG-3′	236	X74479
51	F. 5′-GCTACGAGCAATTTGACAGCTCAG-3′ R. 5′-ATCGCCGTTGCTAGTTGTTCGCA-3′	242	NC_001533
52	F. 5′-ACTGACCTAYACTGCTATGAGCAA-3′ R. 5′-CAGCCGGGGCACACAACTTGTAA-3′	229	NC_001592
53	F. 5′-ACCTGCAATGCCATGAGCAATTGAA-3′ R. 5′-TTATCGCCTTGTTGCGCAGAGG-3′	253	X74482
56	F. 5′-ACCTACARTGCAATGAGCAATTGG-3′ R. 5′-TGATGCGCAGAGTGGGCACGTTA-3′	244	NC_001594
58	F. 5′-GCTATGAGCAATTATGTGACAGCT-3′ R. 5′-TGTGCACAGSTAGGGCACACAAT-3′	219	NC_001443
59	F. 5′-ACCTTGTGTGCTACGAGCAATTAC-3′ R. 5′-GCTGCACACAAAGGACACACAAA-3′	243	NC_001635
66	F. 5′-ACCTACARTGCAATGAGCAATTGG-3′ R. 5′-TGATGCGCAGAGTGGGCACGTTA-3′	244	NC_001695
68	F. 5′-TTGTATGTCACGAGCAATTAGGAG-3′ R. 5′-GATTACTGGGTTTCCGTTGCACAC-3′	258	Y14591

70	F. 5′-CACGAGCAATTAGAAGATTCAGACA-3′ R. 5′-TTCCCGATGCACACCAGGGACA-3′	237	U21941
73	F. 5′-CTTACATGTTACGAGTCATTGGAC-3′ R. 5′-GTTTCTGGAACAGTTGGGGCAC-3′	221	X94165
82	F. 5′-GCTACGAGCAATTTGACAGCTCAG-3′ R. 5′-CATTGCCGATGTTAGTTGGTCGCA-3′	240	AB027021

[a]Due to the homology in the E7 gene of different HPV types, the following primers had identical sequences: HPV18R/HPV45R, HPV33/HPV52F, HPV33R/HPV58R, HPV39F/HPV68F, HPV51F/HPV82F, HPV56F/HPV66F, and HPV56R/HPV66R.

[b]F: forward sequence; R: reverse sequence.

Oligonucleotides were synthesized by Eurofins MWG Operon (Eurofins Genomics, Ebersberg, Germany)[39] and mixed to obtain a 10x solution containing 2uM of each primer. The total number of primers used in the mix was 31 including the two primers for the amplification of β-globin (GenBank Accession number AY260740). Some primers were used for more than one HPV genotype due to the high similarities between the E7-region genes (Table 3.7).

The PCR was performed with the QIAGEN Multiplex PCR Kit (QIAGEN Ltd, Manchester, UK)[39] according to the instructions of the manufacturer. Each PCR reaction was composed of 25ul 2x QIAGEN Multiplex PCR Master Mix (giving a final concentration of 3mM $MgCl_2$) (QIAGEN Ltd, Manchester, UK)[39], 5ul 10x biotinylated primer mix (containing each primer in Table 3.7 at 2uM), 15ul DNase/RNase-free water (ThermoFisher™ Scientific, MA, USA)[18], and 5ul of extracted DNA. The final reaction volume was therefore 50ul.

The PCR was initiated by a 15-minute activation step at 95°C. This was followed by 45 cycles of 30 seconds at 94°C for denaturation, 90 seconds at 63°C for annealing, and 90 seconds at 72°C for extension. Final extension was 10 minutes at 72°C. The PCR products ranged in size from 210bp to 350bp. The PCR products were then hybridized to oligonucleotide probes coupled to fluorescent beads for detection by the Luminex 200 Analyzer (Luminex Corp, TX, USA)[27] (Section 3.7.3.5).

3.7.3.4 Coupling of oligonucleotide probes to fluorescent beads

Before the hybridization assay proceeded, 5'-Amino-modifier C_{12}-linked oligonucleotide probes were coupled to distinctly coloured sets of carboxylated seroMAP beads (Luminex Corp, TX, USA)[27] by a carbodiimide-based coupling procedure as described previously[31]. This was necessary so that any hybridization would result in a bead's fluorescence, then detectable by a Luminex 200 Analyser (Luminex Corp, TX, USA)[27].

For each combination of probe and bead set, 2.5 million carboxylated beads (Luminex Corp, TX, USA)[27] were suspended in 25ul of 0.1M 2-(N-morpholino)-ethanesulfonic acid, pH 4.5 (MES). Probe oligonucleotides (400 pmol) and 200ug of N-(3-dimethylaminopropyl)-N-ethylcarbodiimide (EDC) were added and thoroughly mixed with the beads. Incubation was performed in the dark under agitation for 30 minutes and was interrupted by a thorough mixing step after 15 minutes.

The addition of EDC and incubation steps were repeated, and the coupled beads were finally washed once with 0.5ml of 0.2g/L Tween 20 (Sigma Aldrich/Merck, Darmstadt, Germany)[34] and once with 0.5ml of 1.0g/L sodium dodecyl sulfate before being stored in 100ul of TE buffer at 4°C in the dark.

3.7.3.5 Hybridization Assay and Luminex 200 Fluorescence Analysis

The beads coupled to oligonucleotide probes in Section 3.7.3.4 were then hybridized to the biotinylated PCR products generated in Section 3.7.3.3 as described previously[30,31].

10ul of each reaction mixture was transferred to 96-well plates (Costar, Wiesdbaden, Germany)[31] containing 33ul of tetramethylammonium chloride (TMAC) hybridization solution (0.15M TMAC, 75mM Tris-HCl, 6mM EDTA, 1.5 g/liter Sarkosyl, pH 8.0) and a mixture of 2,000 probe-coupled beads of each set (Section 3.7.3.4). TE buffer (7.0ul) was added, followed by gentle mixing with a 12-channel pipette (Biohit PLC, Helsinki, Finland)[39]. The mixture was heated to 95°C for 10 minutes in a laboratory oven (Bachofer, Reutlingen, Germany)[31], immediately placed on ice for 2 minutes and then transferred to a thermomixer (Eppendorf, Hamburg, Germany)[40]. Hybridization was performed at 41°C for 30 minutes under agitation.

Using a 12-channel pipette (Brand, Roskilde, Denmark), the samples were transferred to a 96-well wash plate (Millipore, Bedford, MA, USA)[41] preequilibrated with blocking buffer (phosphate-buffered saline, 1mg/ml casein). Subsequently, the beads were washed on with 100ul of blocking buffer on a vacuum wash station (Millipore, Bedford, MA, USA)[41]. On a horizontal shaker at room temperature, beads were resuspended for 20 minutes in 75ul of Strep-PE (Molecular Probes/Life Technologies, OR, USA)[11] diluted 1:1,600 in 2.0M TMAC, 75mM Tris-HCl, 6mM EDTA, 1.5 g/L Sarkosyl, 1.0g/L casein, pH 8.0. Beads were then washed three times with 100ul blocking buffer and finally resuspended in 100ul blocking buffer for 5 minutes on a shaker.

Beads were then analyzed by a Luminex 200 Analyser (Luminex Corp, TX, USA)[27] reporting on internal bead colour and Strep-PE reporter fluorescence. The results were expressed as the median fluorescence intensity (MFI) of at least 100 beads per bead set.

For each probe, the MFI values obtained when no PCR product was added to the hybridization mixture were considered the background values. The cutoff for HPV positivity in each case was therefore computed by adding 5 MFI to 1.1*median background value. For all probes, this cutoff value was above the mean background plus 3 times the standard deviation.

3.8 Statistical Analysis and Sample Size

3.8.1 Statistical Analysis

Statistical analyses were conducted using IBM® SPSS® Statistics Version 25, XLSTAT 2019.1.3, Joinpoint Regression 4.7.0, and Microsoft Excel Version 16.25.

In Chapter 4, comparisons of means in any scenario were conducted using T-tests and One-Way ANOVAs where data was normally distributed, and Mann-Whitney U tests and Kruskal-Wallis tests where data was not normally distributed. Tabular tests for independence of variables including hospitals and reasons for sample attrition were conducted using Chi-square tests and Fisher's exact tests when expected values were lower than 5.

In Chapter 5, prevalence and genotype distribution of HPV DNA by the SPF10 PCR Gel Electrophoresis and Multiplex PCR Luminex were calculated using simple proportions. Chi-square association and Fisher's exact tests were used to compare any difference in prevalence, the latter being using when expected and/or observed values were less than 5. Concordance between results by the different methods was determined by proportion and Kappa values.

Throughout Chapters 6, 7, and 8, patient and tumour characteristics for which between 0% and 10% of cases were missing data were analyzed with missing cases excluded. Variables with more than 10% of cases missing data were analyzed with an additional "unknown/missing" category as is conventional in the epidemiological literature to account for non-random missing data. For treatment-related analyses, only patients who received treatment (which was limited to treatment within the first year of diagnosis) were analyzed.

In Chapter 6, socio-demographic and tumour characteristics of patients for whom samples were obtained and successfully analyzed were compared to characteristics of patients whose samples were not obtained and/or who did not have samples available using aggregated data using Chi-square tests. This provided information on the representativeness of the HPV-tested patients, and the generalizability of the results to the population level. Correlations between patient and tumour characteristics were assessed with Chi-square, Fisher's exact, T-, Mann-Whitney, ANOVA, and Kruskal-Wallis tests where appropriate. Raw incidence was represented by moving average graphically, whereby the incidence (number of cases) for a single year was the average of the previous year, the year itself, and the year following (e.g. 1995 = (1994+1995+1996)/3). Average annual percentage change for raw incidence was calculated using Joinpoint Regression 4.7.0. Survival by patient and tumour characteristics was also assessed. Disease-specific survival was measured in months from the date of diagnosis to censoring (either the date of latest follow-up or the date of death by other cause) or death from disease, whichever occurred first. Overall survival was measured in the same way except "death from disease" was replaced by "death by any cause". Survival, both overall and cancer-specific, was assessed using Kaplan-Meier analysis, with a cross marking censored data, by log-rank statistic, and

was further evaluated with the use of univariate and multivariate cox proportional hazards models, including calculation of hazard ratios (HR). All variables significantly predictive at the univariate level were included in the initial multivariate models. Backward regression, whereby the least significant predictor was taken out, the model was run again, the least significant predictor was taken out, the model was run again, and so on, was applied until all variables in the model were significant or until removal of another variable rendered the model as a whole insignificant.

In Chapter 7, overall HPV prevalence and type-specific prevalence were determined. Only HR HPV types were then carried forward for further analysis from this point onwards due to the extremely small number of LR cases and their unlikely involvement in carcinogenesis. Overall and type-specific HR prevalence were compared between the three core subsites in the population (oropharynx, oral cavity, and larynx) using Chi-square tests for independence and Fisher's exact tests in the case that expected values fell below a value of 5. If necessary, HPV genotype categories were combined to allow the generation of meaningful statistics. Raw incidence was represented by moving average graphically, whereby the incidence (number of cases) for a single year was the average of the previous year, the year itself, and the year following (e.g. 1995 = (1994+1995+1996)/3). Average annual percentage change for raw incidence was calculated using Joinpoint Regression 4.7.0. The association between patient and tumour characteristics and HR HPV status was carried out using Chi-square, Fisher's exact, T-, Mann-Whitney, and Krusal-Wallis tests where appropriate. Univariate and multivariate logistic regression methods, including odds ratio (OR) calculation, were used to identify patient- and tumour-related factors significantly associated with HR HPV positivity. For multivariate logistic regression, all univariately significant variables were included in the initial models and backward regression, as described above, was then carried out.

In Chapter 8, patterns of treatment received by HR HPV-positive and HPV-negative patients were compared. To compare observed survival in HR HPV-positive and HPV-negative patients, survival analysis employed the Kaplan-Meier method with different strata again assessed for significance using log-rank tests. The same method was employed when comparing survival by treatment type for HPV positive and negative cases individually. Variation in treatment administered to HPV positive and negative cases was assessed using

Chi-square and Fisher's exact tests where appropriate. Univariate and multivariate Cox proportional hazards models were then used to determine the impact of HR HPV status on risk of death, adjusting for other significant confounders (e.g. patient, tumour and clinical/treatment related). Analyses were repeated for deaths from all causes and cancer-specific deaths.

The threshold for statistical significance was set at $p \leq 0.05$ and all tests were two-sided. All confidence intervals (CI) were 95% CIs.

3.8.2 Sample size

4166 cases of oropharyngeal, oral cavity, and laryngeal SCC were identified in the NCRI database as eligible for the ECHO study. Based on an estimate from systematic review for HPV prevalence of 25.6%[42], a sample size of 681 cases would be required to estimate prevalence in the population at the 95% confidence level and a 3% margin of error. The 861 cases included in the ECHO study fulfill this requirement.

References

1. O'Brien, K., Comber, H. & Sharp, L. Completeness of case ascertainment at the Irish National Cancer Registry. *Ir. J. Med. Sci.* **183**, 219–224 (2014).
2. World Health Organization. ICD-10 Code Descriptions for All Disease Types. *World Health Organization* (2016). Available at: https://icd.who.int/browse10/2016/en. (Accessed: 27th May 2019)
3. Ferlay, J. *et al.* Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer* **127**, 2893–2917 (2010).
4. Hughes, A. & Lawless, M. *National Consent Policy May 2019 V.1.2*. (2013).
5. QIAamp® FFPE DNA Tissue Kit. (2017). Available at: https://www.qiagen.com/us/shop/sample-technologies/dna/genomic-dna/qiaamp-dna-ffpe-tissue-kit/#orderinginformation. (Accessed: 22nd August 2017)
6. Leica Biosystems. Leica Biosystems. *Leica Biosystems* (2019). Available at: https://www.leicabiosystems.com/. (Accessed: 28th May 2019)
7. Pfmmedical. pfmmedical Rotary 2006 EM. *pfmmedical* (2019).

8. HPV-AHEAD - Role of human papillomavirus infection and other co-factors in the aetiology of head and neck cancer in Europe and India. (2017). Available at: http://hpv-ahead.iarc.fr/. (Accessed: 22nd August 2017)
9. International Agency for Research on Cancer & World Health Organization. HPV-AHEAD Study. *International Agency for Research on Cancer* (2019). Available at: http://hpv-ahead.iarc.fr/about/index.php. (Accessed: 27th April 2019)
10. ThermoFisher Scientific. DNAZap PCR DNA Degradation Solutions. *ThermoFisher Scientific* (2019). Available at: https://www.thermofisher.com/order/catalog/product/AM9890?SID=srch-srp-AM9890. (Accessed: 28th May 2019)
11. Life Technologies. Life Technologies: ThermoFisher Scientific. *Life Technologies* (2019). Available at: https://www.thermofisher.com/ie/en/home/brands/life-technologies.html. (Accessed: 27th May 2019)
12. Leica Biosystems. Leica Microsystems BOND Plus Slides. *Leica Biosystems* (2019). Available at: https://www.leicabiosystems.com/ihc-ish-fish/bond-ancillary-ihc-ish-reagents/consumables/products/leica-microsystems-plus-slides/. (Accessed: 27th May 2019)
13. Roche Diagnostics. VENTANA HE 600 System. *Roche Diagnostics* (2019). Available at: https://diagnostics.roche.com/global/en/products/systems/ventana-he-600-system.html. (Accessed: 28th May 2019)
14. xMAP® Technology for Research & Applied Markets. (2017). Available at: https://www.luminexcorp.com/research/our-technology/xmap-technology/. (Accessed: 22nd August 2017)
15. Affymetrix Panomics. Luminex Assays: How it Works. *Affymetrix Panomics* 1 (2014). Available at: http://cdn.panomics.com/products/luninex-assays/technical-overview/how-it-works. (Accessed: 15th November 2017)
16. Woods, R. Human Papillomavirus-related Oropharyngeal Squamous Cell Carcinoma – Prevalence and Incidence in a Defined Population and Analysis of the Expression of Specific Cellular Biomarkers. (2016).
17. ThermoFisher Scientific. NanoDrop™ 2000/2000c Spectrophotometers. *ThermoFisher Scientific* (2019). Available at: https://www.thermofisher.com/order/catalog/product/ND-2000. (Accessed: 27th

May 2019)

18. Thermofisher Scientific. UltraPure DNase/RNase-Free Distilled Water. *ThermoFisher Scientific* (2019). Available at: https://www.thermofisher.com/order/catalog/en/US/adirect/lt?cmd=catDisplayStyle&catKey=101&filterDispName=UltraPure%26trade%3B+DNase%2FRNase-Free+Distilled+Water&filterType=1&OP=filter&filter=ft_1201%2Ff_154501*&_bcs_=H4sIAAAAAAAAANM1VDWwCCjKTylNLinWVjUy0. (Accessed: 28th May 2019)
19. Kleter, B. *et al.* Novel Short-Fragment PCR Assay for Highly Sensitive Broad-Spectrum Detection of Anogenital Human Papillomaviruses. *Am. J. Pathol.* **153**, 1731–1739 (1998).
20. Kleter, B. *et al.* Development and clinical evaluation of a highly sensitive PCR-reverse hybridization line probe assay for detection and identification of anogenital human papillomavirus. *J. Clin. Microbiol.* **37**, 2508–17 (1999).
21. Cole-Parmer Scientific Experts. Bibby Scientific. *Bibby Scientific* (2019). Available at: http://www.bibby-scientific.com/. (Accessed: 28th May 2019)
22. ThermoFisher Scientific/Life Technologies. AmpliTaq Gold DNA Polymerases. *Life Technologies* (2019). Available at: https://www.thermofisher.com/ie/en/home/life-science/pcr/pcr-enzymes-master-mixes/taq-dna-polymerase-enzymes/amplitaq-360-dna-polymerase.html. (Accessed: 27th May 2019)
23. ThermoFisher Scientific. MicroAmp Optical 96-Well Reaction Plate. *ThermoFisher Scientific* (2019). Available at: https://www.thermofisher.com/order/catalog/product/N8010560. (Accessed: 29th May 2019)
24. ThermoFisher Scientific. Applied Biosystems® 7500 fast and 7500 Real-Time PCR System. *ThermoFisher Scientific/Life Technologies* (2019). Available at: https://www.thermofisher.com/ie/en/home/life-science/pcr/real-time-pcr/real-time-pcr-instruments/7500-fast-real-time-pcr-system.html. (Accessed: 27th May 2019)
25. FlashGel™ DNA System by Lonza Inc. (2017). Available at: http://www.lonza.com/products-services/bio-research/electrophoresis-of-nucleic-acids-and-proteins/nucleic-acid-electrophoresis/fast-electrophoresis-flashgel-system-for-dna-rna-and-recovery/flashgel-dna-system.aspx. (Accessed: 22nd August 2017)
26. New England BioLabs. 1 kb DNA Ladder. *New England BioLabs* (2019). Available at:

https://international.neb.com/products/n3232-1-kb-dna-ladder#Product Information. (Accessed: 27th May 2019)

27. Luminex® 100/200™ for Clinical Diagnostic Use. Available at: https://www.luminexcorp.com/eu/clinical/instruments/luminex-100200/. (Accessed: 9th April 2018)
28. Gheit, T. *et al.* Development of a Sensitive and Specific Assay Combining Multiplex PCR and DNA Microarray Primer Extension To Detect High-Risk Mucosal Human Papillomavirus Types. *J. Clin. Microbiol.* **44**, 2025–2031 (2006).
29. Gheit, T. *et al.* Role of mucosal high-risk human papillomavirus types in head and neck cancers in central India. *Int. J. Cancer* **141**, 143–151 (2017).
30. Schmitt, M. *et al.* Abundance of multiple high-risk human papillomavirus (HPV) infections found in cervical cells analyzed by use of an ultrasensitive HPV genotyping assay. *J. Clin. Microbiol.* **48**, 143–9 (2010).
31. Schmitt, M. *et al.* Bead-based multiplex genotyping of human papillomaviruses. *J. Clin. Microbiol.* **44**, 504–12 (2006).
32. Aldrich/Merck, S. Trizma hydrochloride solution. *Merck* (2019). Available at: https://www.sigmaaldrich.com/catalog/product/sigma/t3038?lang=en&region=IE.
33. Qiagen. Proteinase K. *Qiagen Ltd* (2019). Available at: https://www.qiagen.com/ie/products/discovery-and-translational-research/lab-essentials/enzymes/qiagen-proteinase-k/#orderinginformation. (Accessed: 28th May 2019)
34. Aldrich/Merck, S. Tween 20. *Merck* (2019). Available at: https://www.sigmaaldrich.com/catalog/product/sigma/p9416?lang=en&region=IE.
35. Stuart. Stuart Block Heater Digital SBH130D. *Stuart* (2019). Available at: http://www.stuart-equipment.com/product.asp?dsl=17. (Accessed: 28th May 2019)
36. Aldrich/Merck, S. Eppendorf Centrifuge 5418R. *Merck* (2019).
37. Mena, M. *et al.* Development and validation of a protocol for optimizing the use of paraffin blocks in molecular epidemiological studies: The example from the HPV-AHEAD study. *PLoS One* **12**, e0184520 (2017).
38. Gheit, T. *et al.* Development of a sensitive and specific assay combining multiplex PCR and DNA microarray primer extension to detect high-risk mucosal human papillomavirus types. *J. Clin. Microbiol.* **44**, 2025–31 (2006).

39. QIAGEN. QIAGEN Multiplex PCR Kit. *QIAGEN* (2019). Available at: https://www.qiagen.com/ie/products/discovery-and-translational-research/pcr-qpcr/pcr-enzymes-and-kits/end-point-pcr/qiagen-multiplex-pcr-kit/#resources. (Accessed: 27th May 2019)
40. Eppendorf UK. Eppendorf – Laboratory equipment, supplies, and services. *Eppendorf* (2019). Available at: https://www.eppendorf.com/UK-en/. (Accessed: 27th May 2019)
41. Merck. Millipore/Merck Science Products. *Merck* (2019). Available at: http://www.merckmillipore.com/IE/en?ReferrerURL=https%3A%2F%2Fwww.google.com%2F. (Accessed: 27th May 2019)
42. Kreimer, A. R., Clifford, G. M., Boyle, P. & Franceschi, S. Human Papillomavirus Types in Head and Neck Squamous Cell Carcinomas Worldwide: A Systematic Review. *Cancer Epidemiol. Biomarkers Prev.* **14**, 467–475 (2005).

Chapter 4

THE ORGANIZATION AND EXECUTION OF SAMPLE ACQUISITION: REFLECTIONS ON THE EFFICIENCY AND IMPACT OF CLINICAL RESEARCH IN IRELAND

4 CHAPTER 4: THE ORGANIZATION AND EXECUTION OF SAMPLE ACQUISITION: REFLECTIONS ON THE EFFICIENCY AND IMPACT OF CLINICAL RESEARCH IN IRELAND

4.1 Introduction

Ireland boasts some of the world's most renowned research institutions. From 1996 to 2017, the country's research bodies have published almost 60,000 citable journal articles in the field of Medicine which have an average of approximately 23 citations each[1]. The country's contribution to Medical advances is favorably disproportionate to its population size.

The acquisition of data and patient samples for medical studies is a foundational necessity for Ireland's research impact. Some methods of data collection form an integral part of patients being treated in hospitals, including clinical trials, whilst some require no patient samples at all. Others involve more complex retrieval methods through archives of frozen or FFPE samples. The latter relies on the continuous, long-term recording and archiving of patient samples by hospitals, clinics, and biobanks. All of these face financial, organizational, and governance challenges[2–5].

The archiving, storage, and biobanking of patient samples in Ireland is organized in a hierarchical manner, starting at the level of the Department of Public Health, and descending to the Health Service Executive (HSE), and public and private healthcare facilities including hospitals and clinics[6,7]. Each level utilizes both internal and private storage facilities for patient material. In addition to this, there are many standalone biobanks focused on the centralized recording and storage of particular types of patient samples. One of the most recently created is a national biobank of diagnosed brain tumours[8].

In the case of FFPE tissue, almost every hospital across the country has its own individual procedure, database, and location for the archiving of blocks. These are homogenous within a single hospital for all patient types, agglomerated within one or two databases accessible electronically through software unique to the hospital site or through paper filing systems.

Up until 2018, has been very little national synchronization between hospital databases, on- and private off-site storage, and protocols for requesting and gaining access to FFPE tissue. This extends to each hospital and BioBanks' unique requirements for the ethical approval for using FFPE blocks in the research context. The only part of the archival process that is universal to all hospitals is that of the laboratory or histology number associated with the particular sample and/or patient. This is the identifier that is sourced by registries at the national level to pool personal data. The NCRI[9] acts as this registry for all histology numbers associated with cancers diagnosed across the country in all hospitals.

The heterogeneity in the archival records, and storage of and access to FFPE blocks has significant implications for the cyclical relationship between the Irish clinic, clinical research, and the impact of research on patients[4]. The greater the number and diversity of steps and parties involved in gaining access to patient data across different hospital sites, the longer vital research is postponed as a result of repetitive, excess documentation and organizational exchanges. The longer this organization process, the less time and fewer samples researchers have to take full advantage of the acquired funding to conduct studies. The larger the extent of dissociation between each party involved in research, including registries, hospitals, clinics, and storage managements, the more likely samples will need to be eliminated from the study. This is mostly due to mistaken, mismatched, or differentially formatted patient identifiers and details in different biobanking, registry, and hospital databases, and the greater potential for missing cases in the multiplicity of storage sites involved. National level research therefore becomes more inefficient and potentially less impactful for patients over time.

To ensure the highest quality, greatest output, and largest impact of research, especially that of the retrospective variety, the procedures and protocols for FFPE block acquisition must be assessed and standardized. No studies currently exist in the literature examining the fine detail of exactly how heterogenous or complex processes of acquisition are in the Irish research system. There are also no present analyses identifying what components and stages of the acquisition process create the longest time delays or have the largest influence on the attrition of samples from studies.

The FFPE sample acquisition for the ECHO study provides a unique opportunity to identify both the efficiencies and the shortcomings of conducting nationwide cancer research in Ireland. The study's large sample size and collaboration with parties across the research spectrum including the NCRI, 14 different hospital sites, BioBanks, and private storage companies makes it representative of the current complexities of sample acquisition. Thus, in the coming sections, the process of acquiring FFPE blocks, from the identification of relevant samples in the NCRI database to the transport of samples to the site of research (Coombe Women and Infants University Hospital, Dublin 8) will be analyzed, and its significance extrapolated.

4.2 Aims

- To establish the number of procedures and parties necessary to acquire pathology reports and retrieve sample FFPE blocks from hospital sites.
- To determine the lengths of time (days) taken to organize and execute the review of pathology reports and the retrieval of FFPE blocks from hospital sites.
- To pinpoint the steps of the pathology report review and FFPE block retrieval process contributing most to the attrition of cases from the study.
- To identify the reasons for which attrition occurred at each step of the pathology report review and FFPE block retrieval process.

4.3 Materials and Methods

4.3.1 Study population

Sections 3.3 through 3.4 provide a comprehensive overview of the procedures used to identify and retrieve the samples needed to carry out the ECHO study. The coming investigation is an assessment of each stage of this process. The primary study population for this analysis was therefore the 5792 incident tumours diagnosed in Ireland between 1994 and 2013 identified in the NCRI database. However, after a series of eliminations based on eligibility, the Researcher was able to access pathology reports for 2527 cases. The study population for more detailed analysis of attrition of cases was therefore 2527.

4.3.2 Outline of the pathology report review and FFPE block retrieval process

Using the database of the identified 5792 cases generated by the NCRI, the Researcher began the process of narrowing the study population on the basis of the eligibility and ethics criteria outlined in Section 3.3. When a final eligible and consented database of patients was compiled, the Researcher contacted the representing pathologists for the 14 hospital sites collaborating with the ECHO study. These hospitals are summarized in Table 2.1.

On a hospital-by-hospital basis, the Researcher and the pathologists then organized the review of pathology reports associated with all identified cases. Pathology report review, whether electronic or paper-based, carried out by the Researcher was anonymized using the study number assigned by the NCRI to the ECHO study. The review process was necessary to confirm as accurately as possible, that the primary site of the tumour conformed to the study requirements as described in Section 3.3; to confirm the relevant tumour's SCC pathology; to ensure that the tumour was invasive in pathological nature; and to select the precise blocks that contained tumour tissue fitting all of the above criteria. Ultimately, this process was intended to avoid processing a large number of blocks for every case simply to find the singular block with the most tumour. Not only did this minimize the number of blocks withdrawn and carried cross-country, but it also minimized the waste of a large amount of tissue. For the most part, reviewing reports narrowed down the number of blocks to only those potentially relevant, especially for cases with many (sometimes over 20) blocks.

Some case numbers retrieved pathology reports in hospital systems that re-directed to another case number associated with the same patient that actually diagnosed the relevant tumour. The Researcher accepted this redirection after consulting with hospital-based pathologists to confirm that the patient was indeed the same for each case number given the anonymous nature of the review process.

Once relevant blocks were identified through pathology report review, retrieval and transport protocol was organized with the representing pathologists responsible for on-site FFPE tissue storage and the private companies responsible for FFPE tissue storage off-site. Each hospital had its own protocol and individualized process for acquiring blocks. Those

hospitals utilizing private off-site storage generally charged for the pulling of a single block. This charge ranged from 5EU per block to 20EU per block. Given that this was an unforeseen expense, no budget remained for the normal pathways of block acquisition. The Researcher therefore decided to negotiate with the off-site storage companies and pathologists for every hospital to visit the warehouses herself. The Researcher agreed with each company and hospital that visiting the warehouses herself would be possible and would come at no cost to the project. In some instances, companies did not agree to have the Researcher withdraw blocks herself free of charge.

Once blocks were collected at on- and off-site storage sites, the Researcher transported the blocks to the Coombe Women and Infant's University Hospital (CWIUH) in secure, waterproof boxes. No couriers were used at the request of the pathologists.

Blocks were then sectioned and processed in accordance with the protocols described in Section 3.6. Two H+E slides for each block were generated and reviewed by a board of pathologists using the protocol in Section 3.6.2. The pathologists provided full diagnosis for all of the cases and the Researcher eliminated any cases that did not adhere to the eligibility criteria outlined in Section 3.3. Cases still deemed eligible were then carried forward for molecular testing.

The foregoing summary establishes a clear number of stages involved in identifying blocks eligible for the ECHO study. These stages varied by hospital, with each site requiring the Researcher to follow different procedures and to interact with different parties to review pathology reports and acquire blocks. Furthermore, at each juncture, cases were reassessed on the basis of eligibility criteria and were thus vulnerable to elimination. The stages involved in acquiring and reviewing pathology reports and FFPE blocks are summarized in Figure 4.1. Between each stage is a numbered pink diamond indicating an eligibility assessment on the basis of accessibility and eligibility criteria in Section 3.3 after which cases were eliminated. In Figure 4.1, step 1 (pink diamond 1) is the evaluation and subsequent attrition of incident tumours identified in the NCRI database based on the eligibility criteria. Step 2 (pink diamond 2) is the same evaluation and attrition but following pathology report review. Step 3 (pink diamond 3) is the attrition of cases during block

retrieval itself. Step 4 (pink diamond 4) is the final eligibility evaluation and attrition following the histopathological analysis of H+E slides by the pathology review board. These pink diamonds will be referred to throughout the analysis as “step 1”, “step 2”, “step 3”, and “step 4”.

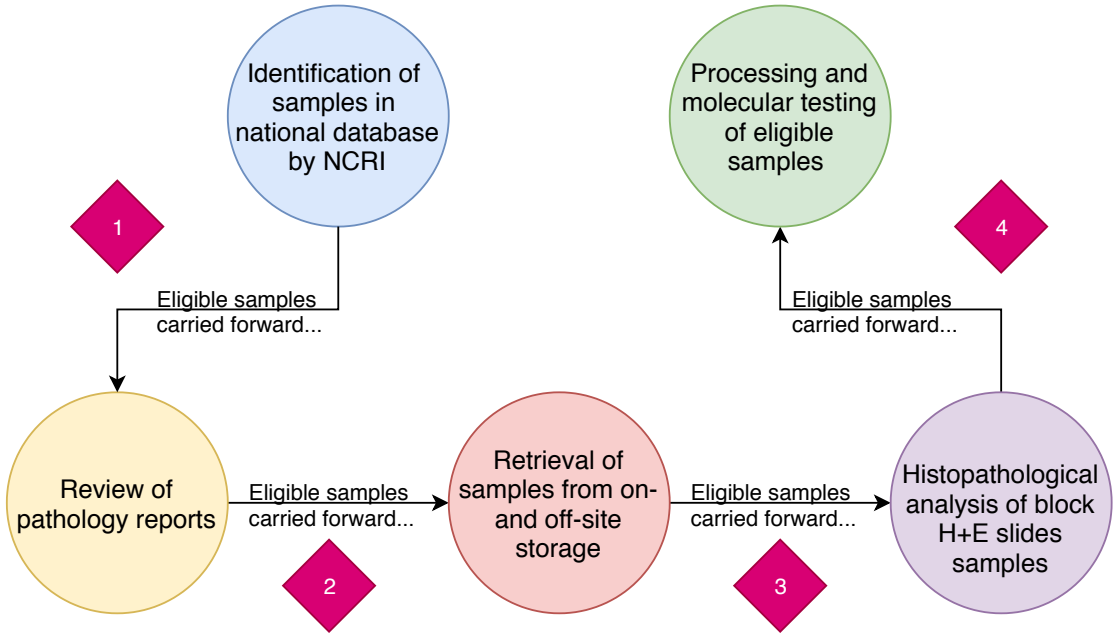

Figure 4.1 Summary of the stages in the acquisition of FFPE blocks for the ECHO study.

Pink numbered diamonds indicate an eligibility assessment using criteria in Section 3.3 after which cases were eliminated. These will be referred to throughout the analysis as “step 1”, “step 2”, “step 3”, and “step 4”. The blue bubble in the top left-hand corner is the starting point for the process. Incident tumours in the national database of cancers were identified here by the NCRI. Step 1 is the elimination of cases from this database on the basis of eligibility criteria and consent. The pathology reports for those samples identified for the study were then reviewed (yellow bubble). Step 2 is the elimination of cases during pathology report review. FFPE samples were then retrieved from appropriate on- and off-site storage facilities (red bubble). Step 3 is the elimination of cases during this retrieval due to inaccessibility or missing cases. Retrieved samples were then processed, generating H+E slides that were histopathologically evaluated by a pathology review board (purple bubble). Step 4 is the elimination of cases during histopathological analysis of H+E slides. All remaining cases were brought forward for molecular testing for the ECHO study (green bubble).

4.3.3 Statistical Analysis

Statistics were generated using IBM SPSS Statistics Version 25, XLSTAT 2019.1.3, and Microsoft Excel Version 16.25. T-tests and one-way ANOVAs were carried out to compare means where relevant. Where normality was not respected, Mann-Whitney and Kruskal-

Wallis tests were employed instead. Chi-square and Fisher's exact tests were carried out to assess associations where necessary, the latter being used when observed or expected values were less than 5. All significance tests were two-tailed and carried out at the 95% level.

4.4 The number and type of procedures and parties necessary to complete the review of pathology reports and retrieve FFPE blocks

As outlined in Section 3.4, each hospital involved in the ECHO study had its own protocols for the review of case pathology reports and for the retrieval of required blocks associated with each case. Table 4.1 below shows the generalized definitions and codes for the different procedure and party types needed to acquire all necessary pathology reports and block retrieval. A procedure type is defined as a general sequence of actions a hospital required the Researcher and others to carry out in order to access reports and blocks. A party type is considered a category of professional, individual, or a company.

Table 4.1 Assigned letter codes for the procedure and party types required to review pathology reports and retrieve FFPE blocks for the ECHO study.

Procedure types are defined as general sequences of actions a hospital required the Researcher and others to carry out in order to access reports and blocks. A party type is considered a category of professional, individual, or a company.

Assigned Code	Description of Procedure/Party Type
A	Single meeting on hospital site to organize report review, followed by set-up of unique electronic log-in for report access, followed by the Researcher's review of all reports on-site.
B	Phone meetings followed by review of pathology reports by on-site Pathologist themselves.
C	Several on-site and phone meetings to organize report review, followed by set-up of non-unique electronic log-in for report access, followed by the Researcher's review of all reports on-site.
D	Phone meetings and e-mail correspondence to organize report review, followed by print-off and compilation of necessary reports by on-site Pathologist, followed by the Researcher's review of all reports on-site.
E	Retrieval of all blocks required carried out by on-site Pathologist themselves followed by Researcher in-person pick-up of blocks retrieved.
F	Pathologist withdrew first half of blocks by internal request. The Researcher then organized the collection of the next half of blocks through a secondary application process to a BioBank within the hospital in line with newly introduced GDPR regulations. The Researcher then visited private off-site storage facilities, searched for, and collected the blocks.
G	Researcher organized with on-site pathologist and/or chief medical scientist to withdraw on-site blocks from internal hospital storage.
H	Lengthy telephone and in-person meetings with private off-site company, pathologists, and medical scientist to organize retrieval, followed by Researcher off-site storage visits to acquire blocks.
I	Telephone and in-person meetings with pathologist and medical scientist to organize retrieval, followed by Researcher visit to on-site storage to retrieve blocks. Discussions with off-site company unable to yield cost-free retrieval of other blocks.
J	Researcher
K	On-site Pathologist
L	Chief Medical Scientist
M	St. James' Private Hospital Storage Management
N	St. James' Biobank Ireland Trust
O	Oasis Private Storage Management Group
P	St. Vincent's Internal Hospital Storage Management
Q	DSM Private Data Management and Storage Ltd
R	Mater Misericordiae University Hospital Internal Hospital Storage Management
S	University Hospital Limerick's Internal Hospital Storage Management
T	Beaumont University Hospital Internal Storage Management

Ultimately, the Researcher was able to establish connections with 8 of the 14 different hospital sites originally ethically inducted into the study. It should be noted that all relevant

samples from the Royal Eye and Ear University Hospital were referred for clinical pathology review to St. James' University Hospital. All Royal Eye and Ear cases were therefore within the St. James' database. Thus, any reference to St. James' University Hospital in this analysis also includes all identified Royal Eye and Ear University Hospital samples. The total number of hospital sites considered in the analysis is therefore 7.

Using the coded procedure and party types in Table 4.1, Table 4.2 summarizes the report review and block retrieval process for each hospital involved in the ECHO study.

Table 4.2 Compilation and sum of procedure and party types required to carry out pathology report review and block retrieval for each hospital involved in the ECHO study.

Each coded letter represents a generalized procedure or party type, each of which are detailed in Table 4.1. Row 9 of this Table is the sum of the number of generalized procedure and party types involved at each stage of the process. Column 6 is the sum of procedure and party types involved at each stage for every individual hospital. Though procedure and party types are generalized by coded letters in Table 4.1, each was unique to every hospital site. Thus, the cumulative number of distinct procedures and parties involved for all hospitals is summed in the last row of column 6.

Hospital | Procedures and Parties	Procedure types involved in report review	Procedure types involved in block retrieval	Party types involved throughout report review and block retrieval	Summary of procedure and party types required	*Sum of party/procedure types for each hospital*
Cork University Hospital	A	I	J, K, L	A, I, J, K, L	*5*
Kerry General University Hospital	B	E	J, K, L	B, E, J, K, L	*5*
St. James' University Hospital	C	F	J, K, L, M, N	C, F, J, K, L, M, N	*7*
Beaumont University Hospital	C	G, H	J, K, L, O, T	C, G, H, J, K, L, O, T	*8*
St. Vincent's University Hospital	C	G	J, K, L, P	C, G, J, K, L, P	*6*
University Hospital Limerick	D	G, H	J, K, L, Q, S	D, G, H, J, K, L, Q, S	*8*
Mater Misericordiae University Hospital	C	G	J, K, L, R	C, G, J, K, L, R	*6*
Sum of procedure/party types for all hospitals at each stage	*(A-D) 4*	*(E-I) 5*	*(J-T) 11*	*(A-T) 20*	*45*

The mean number of generalized procedure and party types required to review reports and retrieve blocks per hospital was 6.43 (CI: 5.34, 7.52) (n=7) based on the sums in Column 6 of Table 4.2. The Table also shows that to obtain cases from all 7 involved hospitals, 4 procedure types for report acquisition, 5 procedure types for block retrieval, and 11

different party types were necessary. A total of 20 generalized procedure and party types were required to review and retrieve FFPE blocks for all hospitals. That said, though each procedure and party type were similar for each hospital, there were distinct organizational processes for each site. Thus, Table 4.2 reflects at the base of Column 6 that there was actually a total of 45 unique procedures and parties required to review reports and retrieve FFPE blocks for the ECHO study.

4.5 Time taken to organize and execute pathology report review and FFPE block retrieval

For each hospital, the organization of pathology report review, the organization of block retrieval, the execution of report review, and the execution of block retrieval took varying amounts of time. Table 4.3 summarizes the time in days taken for each activity along with associated means and standard errors.

Table 4.3 Lengths of time (days) taken to organize pathology report review, execute pathology report review, organize FFPE block retrieval, and execute FFPE block retrieval for each hospital involved in the ECHO study. *For the length of time (days) to organize and review pathology reports, n=7. Given that no agreement could be reached for cost-free retrieval of FFPE blocks by the Researcher for Cork University Hospital, n=6 for length of time (days) to organize and execute FFPE block retrieval.*

Hospital | Length of time (days)	Length of time (days) taken to organize report review (n=7)	Length of time (days) taken to review reports (n=7)	Length of time (days) taken to organize block retrieval (n=6)	Length of time (days) taken to retrieve blocks (n=6)
Cork University Hospital	61	2	*	*
Kerry General University Hospital	30	2	61	1
St. James' University Hospital	60	15	184	4
Beaumont University Hospital	122	4	212	4
St. Vincent's University Hospital	62	2	30	1
University Hospital Limerick	183	2	244	3
Mater Misericordiae University Hospital	423	3	31	2
Mean Time (days)	134.43 (CI: 16.99, 251.90)	4.29 (CI: 0.192, 8.39)	127 (CI: 33.98, 220.02)	2.5 (CI: 1.18, 3.82)

*Indicates that block retrieval was not possible due to inability to agree on cost-free retrieval of blocks by the Researcher.

To assess the difference between length of time in days for organization and execution in general, a new Table was created combining Columns 2 and 4 followed by Columns 3 and 5 of Table 4.3. In other words, Table 4.4 represents the comparison between the total time taken to organize both report review and block retrieval to the time it took to execute both the review of reports and retrieval of blocks.

Table 4.4 Lengths of time (days) taken to organize both pathology report review and FFPE bock retrieval, and to execute both pathology report review and FFPE block retrieval.

Column 2 is a combined sum of columns 2 and 4 of Table 4.3. Column 3 is a combined sum of columns 3 and 5 of Table 4.3. Given that no agreement could be reached in time for cost-free retrieval of FFPE blocks by the Researcher for Cork University Hospital, n=6.

Hospital | Length of time (days)	Length of time (days) taken to organize both report review and block retrieval (n=6)	Length of time (days) taken to execute both report review and block retrieval (n=6)
Kerry General University Hospital	30 + 61 = 91	2 + 1 = 3
St. James' University Hospital	60 + 184 = 244	15 + 4 = 19
Beaumont University Hospital	122 + 212 = 334	4 + 4 = 8
St. Vincent's University Hospital	62 + 30 = 92	2 + 1 = 3
University Hospital Limerick (UHL)	183 + 244 = 427	2 + 3 = 5
Mater Misericordiae University Hospital	423 + 31 = 454	3 + 2 = 5
Mean Time (days)	**273.67 (CI: 121.04, 426.30)**	**7.17 (CI: 1.35, 12.99)**

T-tests, Mann-Whitney tests, and a Kruskal-Wallis test were carried out on all means generated in Table 4.3 between lengths of time (days) taken to organize report review, execute report review, organize FFPE block retrieval, and execute FFPE block retrieval. A Mann-Whitney test was then conducted using the combined means calculated in Table 4.4 for length of time (days) taken for organization of report review and FFPE block retrieval, and length of time (days) taken to execute report review and FFPE block retrieval. The results of these tests are summarized in Table 4.5.

Table 4.5 Statistical comparisons between mean lengths of time (days) taken to organize report review, execute report review, organize FFPE block retrieval, and execute FFPE block retrieval.

These are the means calculated and shown in Table 4.3. Comparison of all four means is reported in the fifth row of the Table. Drawing from Table 4.4, a comparison of the combined mean lengths of time (days) to organize both report review and FFPE block retrieval, and to execute report review and FFPE block retrieval is also shown in the last row of the Table.

Mean length of time (days) associated with particular activity	Mean length of time (days) to review reports (n=7)	Mean length of time (days) to organize block retrieval (n=6)	Mean length of time (days) to retrieve blocks (n=6)
Mean length of time (days) to organize report review (n=7)	Mann-Whitney p=0.002*	Mann-Whitney p=0.923	Mann-Whitney p=0.003*
Mean length of time (days) to review reports (n=7)		Mann-Whitney p=0.003*	Mann-Whitney P=0.674
Mean length of time (days) to organize block retrieval (n=6)			T-test p=0.011*
Comparison of all 4 time (days) means	Kruskal-Wallis p=0.00029*		
Comparison of combined organization time (days) and execution time (days) means	Mann-Whitney p=0.005*		

*Denotes a significant finding at the 95% level.

4.6 Extent and causes of sample attrition throughout the sample acquisition process

Each stage of the sample identification and retrieval process summarized in Figure 4.1 resulted in the attrition of cases from the study as a result of ineligibility based on either inaccessibility and consent, or the criteria described in Section 3.3. The numbered pink diamonds in Figure 4.1 indicate the points in the process at which cases were eliminated and are referred to as “step 1”, “step 2”, “step 3”, and “step 4”. Step 1 (pink diamond 1) is the evaluation of incident tumours identified in the NCRI database based on eligibility criteria. Step 2 (pink diamond 2) is the same evaluation but during pathology report review. Step 3 (pink diamond 3) is the attrition of cases during block retrieval itself. Step 4 (pink diamond 4) is the final eligibility evaluation during the histopathological analysis of H+E slides.

The coming analysis is first a summary of the number and extent of sample attrition at each step of the sample acquisition process, followed by a more detailed analysis of the causes and origins of sample attrition during each step for each involved hospital.

4.6.1 Summary of sample attrition during each step of the sample acquisition process

The number of cases remaining eligible after each step of the review and retrieval process showcased in Figure 4.1 are shown in Table 4.6 below.

Table 4.6 Number of cases remaining in the study after each step of steps 1, 2, 3, and 4 involved in sample identification and retrieval summarized in Figure 4.1.

Hospital | Step	After step 1: Number of cases for which pathology reports were accessed/reviewed	After step 2: Number of cases eligible after pathology report review	After step 3: Number of cases found during block retrieval	After step 4: Number of samples eligible after H+E slide review
Cork University Hospital	232	220	0	0
Kerry General University Hospital	66	65	65	40
St. James' University Hospital	1172	1079	357	269
Beaumont University Hospital	313	290	235	156
St. Vincent's University Hospital	136	118	60	52
University Hospital Limerick	315	277	268	237
Mater Misericordiae University Hospital	293	248	130	107
Total	2527	2297	1115	861

The Researcher then calculated the number of cases eliminated during each stage of the sample acquisition and retrieval process. This was simply done by subtracting each column in Table 4.6 from the column preceding it. For example, the number of cases eliminated for

the Mater Misericordiae University Hospital during step 2 is 293-248, which is 45 cases. Given that no information was available on the hospital sites from which cases were eliminated during step 1, this analysis was only possible for step 2 onwards. The results of this data processing are shown in Table 4.7.

Table 4.7 Summary of the number of cases eliminated during each of steps 2, 3, and 4 involved in sample acquisition and retrieval broken down by hospital site (n=1666).

Hospital | Step	Step 2	Step 3	Step 4	Total
Cork University Hospital	12	220	0	232
Kerry General University Hospital	1	0	25	26
St. James' University Hospital	93	722	88	903
Beaumont University Hospital	23	55	79	157
St. Vincent's University Hospital	18	58	8	84
University Hospital Limerick	38	9	31	78
Mater Misericordiae University Hospital	45	118	23	186
Total	230	1182	254	1666

To conduct valid statistical analysis on Table 4.7, it was necessary to combine two of the smaller hospital sites, Kerry General University Hospital and Cork University Hospital. The resulting tabulation is below in Table 4.8 with the relevant association test.

Table 4.8 Summary of the number of cases eliminated during each of steps 2, 3, and 4 in the sample acquisition process with combined categories from Table 4.7.

Cork University Hospital and Kerry General University Hospital cases were combined and are represented as 'Cork and Kerry University Hospitals' in order to facilitate the most valid statistical association test, the result of which is showcased at the base of the table (n=1666).

Hospital | Step	Step 2	Step 3	Step 4	Total
Cork and Kerry University Hospitals	13	220	25	258
St. James' University Hospital	93	722	88	903
Beaumont University Hospital	23	55	79	157
St. Vincent's University Hospital	18	58	8	84
University Hospital Limerick	38	9	31	78
Mater Misericordiae University Hospital	45	118	23	186
Total	230	1182	254	1666
Association Test	Chi square p<0.0001*			

*Denotes a significant finding at the 95% level.

4.6.2 Causes of attrition during step 1

In step 1, the evaluation of incident tumours identified in the NCRI database, the originally identified 5792 incident tumours were whittled down to 4166 cases on the basis of sample eligibility (71.93% of the 5792) and then further reduced to 3426 cases (59.15% of the 5792) after patient consent procedures were completed. Given ethical considerations, it was not possible to determine the hospitals from which ineligible samples came from until patients were actually consented. Of those 3426 consented patients, the Researcher was able to access pathology reports for 2527 from 7 different hospital sites. Thus, 43.6% of samples (of the original 5792) remained eligible and accessible for the study after step 1.

4.6.3 Summary of the causes of attrition during steps 2, 3, and 4

Samples eliminated during steps 2, 3, and 4 of the sample acquisition process were disqualified for a number of generalizable reasons. The Researcher therefore broke down the cases eliminated during steps 2, 3, and 4 by categories of 'reasons for ineligibility' on the basis of notes recorded throughout the process. These 'reasons for ineligibility' are

summarized and detailed in Table 4.9. The details associated with each reason give an exhaustive list of every documented cause falling within the category.

Table 4.9 Categories of reasons for which cases were rendered ineligible during steps 2, 3, and 4 of sample acquisition for the ECHO study with further details regarding each classification.

Reason for ineligibility	Further details regarding reason/cause
Recurrence	Tissue was found for recurrences of the primary tumour after treatment but no associated reports detailing the primary tumour or incident were found.
Report not printed	The report was not printed for review, a mistake that was only identified after retrieval took place.
No blocks available for case	Case had no FFPE blocks associated. It may have had frozen or fresh sections available, or only slides cut for H+E or immunohistochemistry.
Only lymph nodes available	The only tumour tissue available was that in lymph nodes. No primary tumour tissue existed according to reports.
Not found	The case number registered in the NCRI database identified at the beginning of the study was not found or associated with any reports in the relevant hospital database.
No malignancy evident	The Researcher deemed any report or H+E that that cited any of the below descriptions as no malignancy evident: No evidence of malignancy, benign tumour, no tumour identified, no sign of tumour, no tumour seen, tumour previously taken out but no new tumour seen or associated with retrievable relevant reports, previous malignancy but no SCC seen here or found in all retrievable reports, no reference to or detail about malignancy anywhere, no tumour evident, and no confirmation of malignancy or tumour type.
Unconfirmed but highly suspicious for SCC	Any case whose reports or H+E slides were only suspicious of SCC diagnosis in the appropriate physiological sites was classified here. A detailed summary of the variation in these description follows: Grander analysis suggests SCC but blocks are only described as in-situ carcinoma, only highly suspicious of SCC, not clear that case is SCC, no SCC confirmed - advised only for further biopsy and no further reports available, marked dysplasia with possible early micro-invasion, highly suspicious but not diagnostic of SCC in biopsy, strongly suspicious of micro-invasion, not certain SCC diagnosis, only suspicious of invasive tumour, metastatic carcinoma but unsure if SCC, oral mucosa with high grade squamous dysplasia, carcinoma in-situ and suspicious for invasion but unconfirmed, at least carcinoma in-situ and may be SCC, suspicious for early SCC but no confirmation, biopsy with features suspicious of early micro-invasive SCC but not certain, severe dysplasia suspicious of potential SCC, squamous epithelial dysplasia with suspicion for SCC.
Tissue or tumour in another part of body	Any case report or H+E that was associated with a part of the body that did not fall within the criteria established in section 3.3 was classified here. The following is an exhaustive list of the descriptions that fell within this category of exclusion: Placenta and cord showing inflammation, endometrial tumour, non-secretory endometrial tissue with no appearance of malignancy, cervix-uterus-fallopian tube with no malignancies, thyroid tumour, thyroid tissue, prostatic biopsy, benign prostatic tumour, bone marrow biopsy, bone marrow sample, breast biopsy, basal cell carcinoma on skin, skin with intradermal melanocytic nevus without atypical features, minor salivary origin, external skin lesion, benign intradermal melanocytic nevus of the skin from right buttock, lipoma from sternum and right thigh, low-grade salivary tumour, acinic cell carcinoma from salivary gland, esophagus sample, esophageal mucosa, nasopharynx tissue, liver metastatic poorly differentiated SCC, renal biopsy, colon tumour, rectal biopsy with no evidence of malignancy, normal duodenal mucosa, intestinal tumour, and neuroendocrine tumour specimen.

Impure SCC variant or other tumour type originating in appropriate sites	Cases were classified here on the basis of reports or H+Es if the primary tumour was associated with the appropriate ICD10 code but was not pure SCC, or was simply another tumour type entirely. What follows is an exhaustive list of these divergent SCC or other tumour types: Melanoma, spindle cell melanoma, malignant melanoma, morphology of high-grade neoplasia but uncertain, melanoma, carcinoma, early invasive carcinoma, carcinoma cells, favouring poorly differentiated carcinoma, poorly differentiated carcinoma, suspicious of carcinoma in-situ, favours high-grade carcinoma, anaplastic carcinoma and carcinoma, undifferentiated carcinoma, areas of intraepithelial carcinoma and some foci micro-invasive carcinoma, spindle cell carcinoma, verrucous carcinoma, invasive well-differentiated verrucous carcinoma, adenoid cystic carcinoma, polymorphous low-grade adenocarcinoma with adenoid cystic carcinoma, acinic cell carcinoma, muco-epidermoid carcinoma, metastatic epidermoid carcinoma, sarcomachordoma, rhabdomyosarcoma, sarcomatoid SCC, basaloid SCC, adenocarcinoma, low-grade adenocarcinoma, invasive moderately differentiated adenocarcinoma, polymorphous adenocarcinoma, mild actinic keratosis and solar elastosis, verrucous hyperkeratotic stratified squamous epithelium, atypical carcinoid tumour, and tubular adenoma with mild dysplasia.
No or not enough tissue	Cases for which reports and H+Es revealed no tissue at all, no tumour at all, or not enough tumour to conduct meaningful molecular analysis were eliminated and classified in this category. This included cases for which the first H+E slide had tissue, but the second did not, suggesting that so little tumour existed in the block that there was no guarantee tumour was present in the cut molecular samples.
Samples inaccessible	Cases eliminated for this reason were simply not physically accessible by the Researcher. In some hospitals, agreements were not able to be reached to access samples stored in private storage without a major cost incurred. In others, boxes containing older samples from the 1990s were either not able to be located in storage or their whereabouts were entirely unknown.
Missing samples	Any case falling within this category was searched for in all storage sites but had either been withdrawn already for other research or was missing from its supposed position in storage boxes with no label written to indicate the reason for its removal or whereabouts.
Protocol change	During the study, new GDPR regulations came into effect that changed the paperwork associated with withdrawing samples from some hospitals. This caused major time delays in withdrawal and ultimately prevented the retrieval of these cases.

With these ‘reasons for ineligibility’ or ‘causes’ detailed, each remaining step of the process was then analyzed by cause of attrition and hospital site.

4.6.4 Causes and origins of attrition during step 2

For those cases eliminated during step 2 (n=230), the review of pathology reports, Figure 4.2 shows the percentage of cases rendered ineligible by cause.

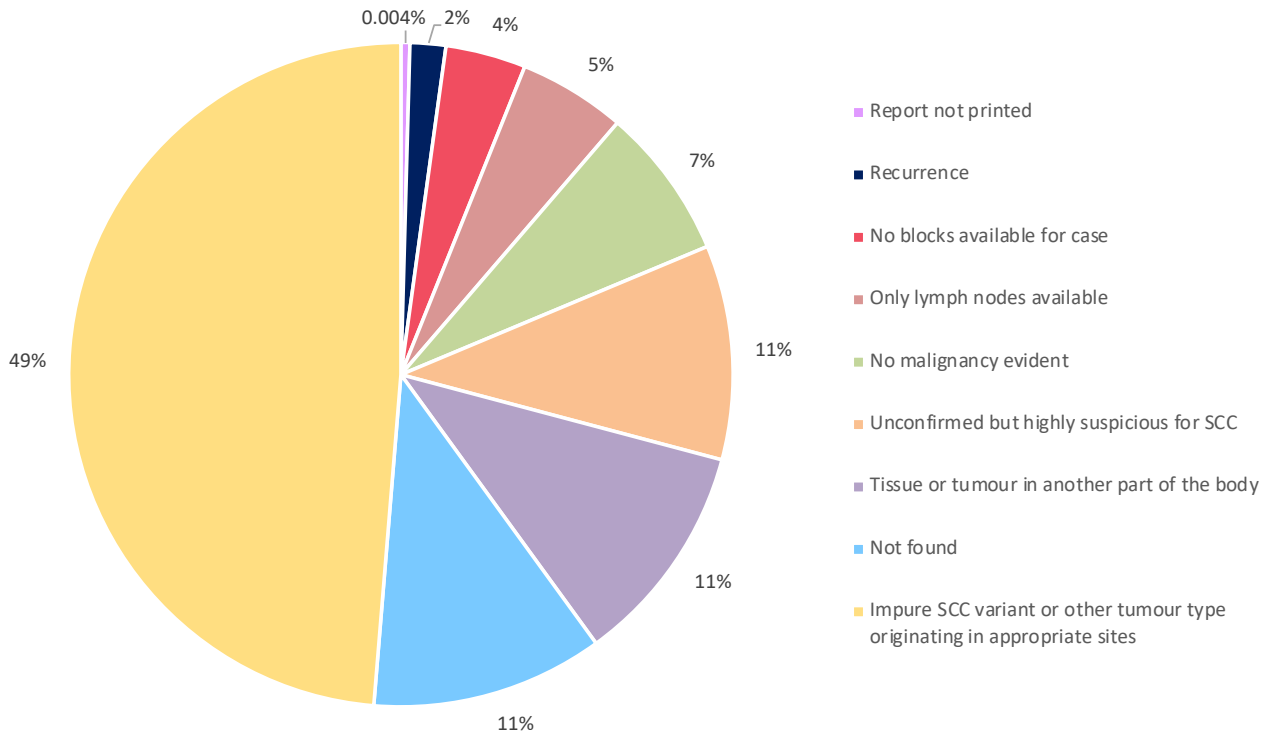

Figure 4.2 Percentage of cases rendered ineligible by cause during step 2, pathology report review.

Details of these causes can be found in Table 4.9 (n=230).

The percentage of cases eliminated during step 2 from each hospital site is shown in Figure 4.3.

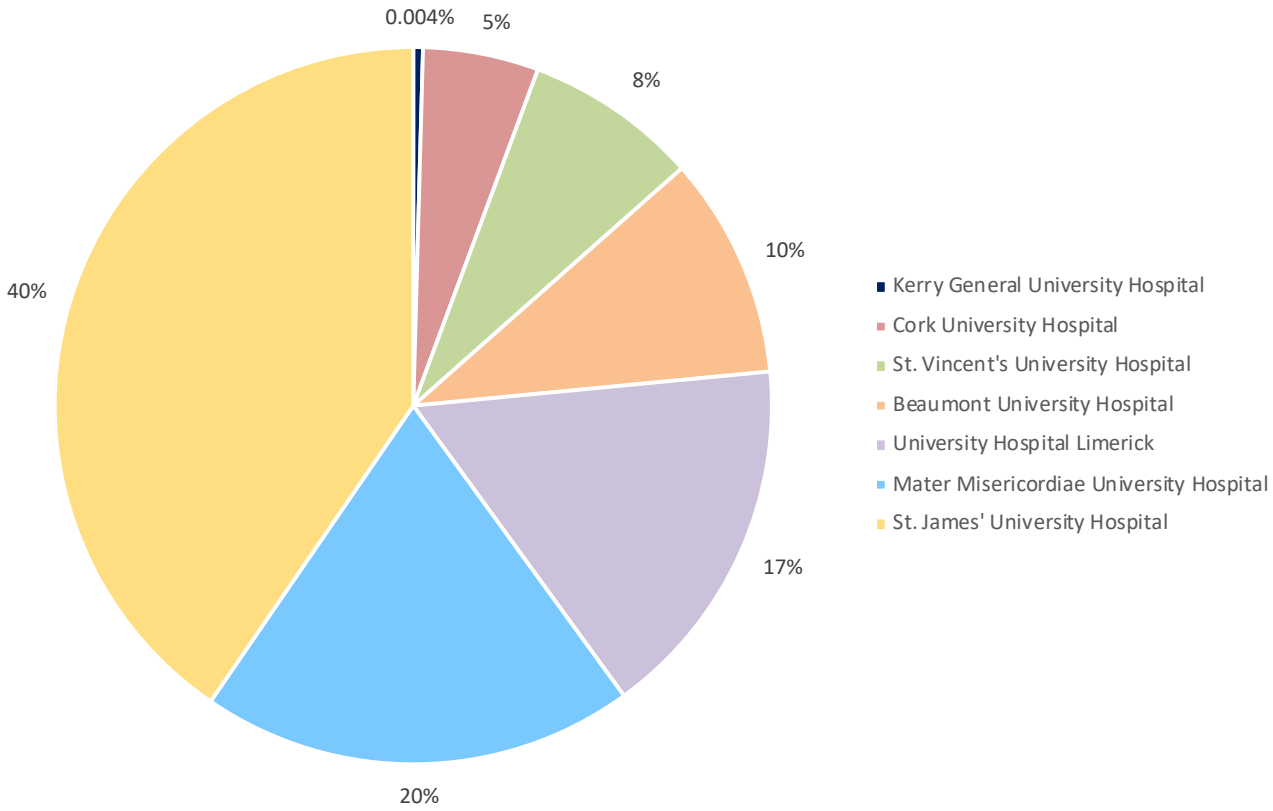

Figure 4.3 Percentage of cases rendered ineligible by hospital site during step 2, pathology report review.
Details of these causes can be found in Table 4.9 (n=230).

Table 4.10 below shows the raw direct comparison between the hospitals from which and the causes for which cases were eliminated during step 2.

Table 4.10 Summary of cases eliminated during step 2, review of pathology reports, by hospital site and cause of elimination.

Details of these causes can be found in Table 4.9 (n=230).

Hospital | Cause	Tissue or tumour in another part of body	Unconfirmed but highly suspicious for SCC	Impure SCC variant or other tumour type	Not found	No malignancy evident	Only lymph nodes available	No blocks available	Recurrence or Printing failure	Total
Beaumont University Hospital	2	4	7	7	2	1	0	0	23
St. James' University Hospital	7	5	45	14	15	3	4	0	93
Limerick University Hospital	0	14	14	0	8	2	0	0	38
Mater Misericordiae University Hospital	11	4	15	2	0	6	5	2	45
Kerry General University Hospital	0	0	1	0	0	0	0	0	1
St. Vincent's University Hospital	4	4	4	3	0	0	0	3	18
Cork University Hospital	1	5	6	0	0	0	0	0	12
Total	25	36	92	26	25	12	9	5	230

The only way to meaningfully assess this data was to combine various causes that, across all hospitals, were not as prevalent as others. ‘Tissue or tumour in another part of body’ and ‘unconfirmed but highly suspicious for SCC’ were combined due to thematic similarity and sample size. ‘Not found’, ‘no malignancy evident’, ‘only lymph nodes available’, and ‘recurrence or printing failure’, were combined due to the very small number of samples eliminated for all of these reasons. Similarly, hospital sites that contributed very few eliminated cases at this stage were combined. These were Kerry General, St. Vincent’s, and Cork. An association test was performed on the resulting tabulation in Table 4.11.

Table 4.11 Summary of cases eliminated during step 2, pathology report review, by hospital site and cause of elimination with combined categories from Table 4.10 and the relevant association test.

For the purposes of conducting valid statistical tests, 'Tissue or tumour in another part of body' and 'unconfirmed but highly suspicious for SCC' cases were combined and are represented as the 'Suspicious for SCC and other part of body' column. 'Not found', 'no malignancy evident', 'only lymph nodes available', and 'recurrence or printing failure' were combined and are represented as 'Not found/no malignancy/lymph/no blocks/recurrence/print'. Kerry General University Hospital, St. Vincent's University Hospital, and Cork University Hospital were also combined and are represented as 'Kerry, St. Vincent's, Cork University Hospitals' (n=230).

Hospital | Cause	Suspicious for SCC and other part of body	Impure SCC variant or other tumour type	Not found/no malignancy/lymph/no blocks/recurrence/print	Total
Beaumont University Hospital	6	7	10	23
St. James' University Hospital	12	45	36	93
Limerick University Hospital	14	14	10	38
Mater Misericordiae University Hospital	15	15	15	45
Kerry, St. Vincent's, Cork University Hospitals	14	11	6	31
Total	61	92	77	230
Association Test	Chi-square p=0.012*			

*Denotes a significant finding at the 95% level.

4.6.5 Causes and origins of attrition during step 3

For those eliminated during step 3 (n=1182), block retrieval, Figure 4.4 shows the percentage of cases rendered ineligible by cause.

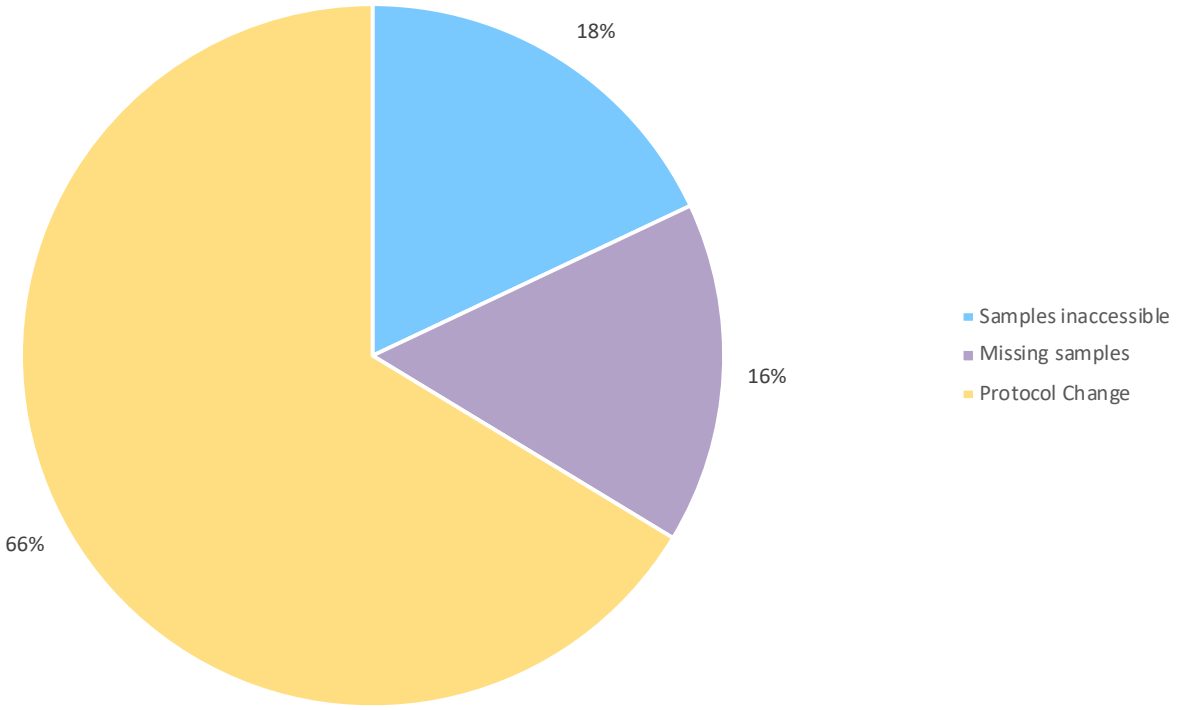

Figure 4.4 Percentage of cases rendered ineligible by cause during step 3, block retrieval.

Details of these causes can be found in Table 4.9 (n=1182).

For those eliminated during step 3 (n=1182), block retrieval, Figure 4.5 shows the percentage of cases rendered ineligible by hospital.

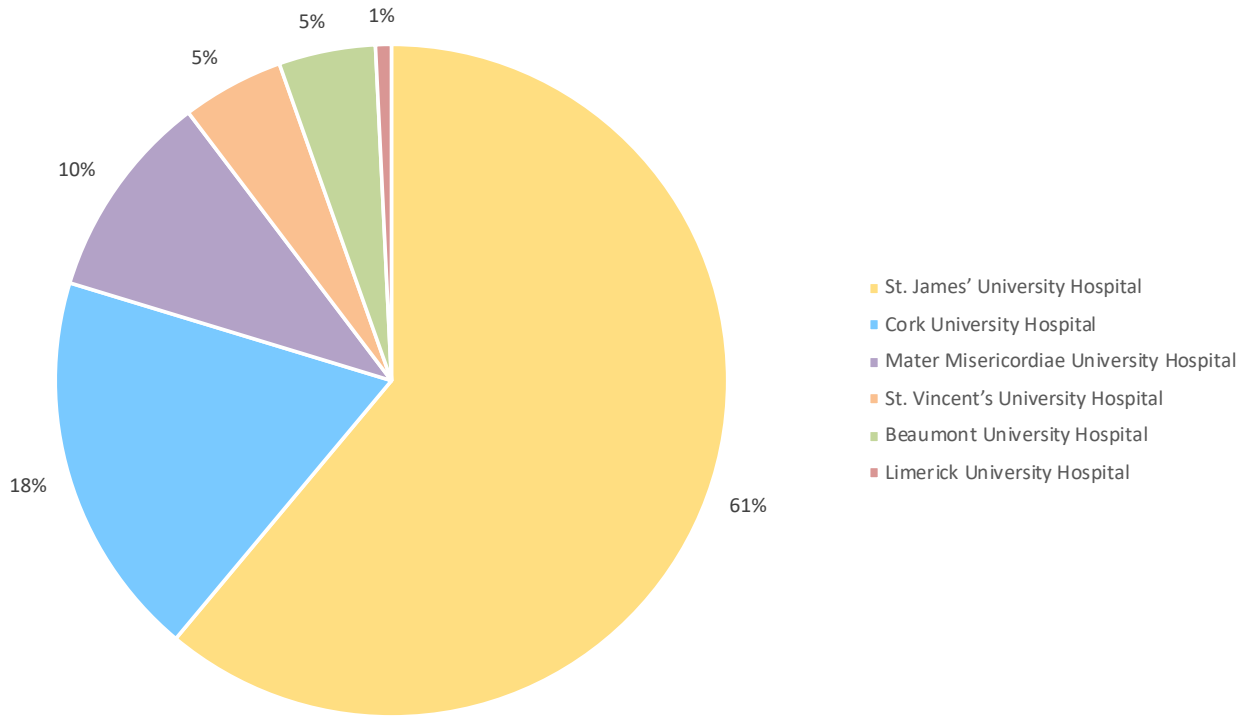

Figure 4.5 Cases rendered ineligible by hospital site during step 3, block retrieval.

Details of these causes can be found in Table 4.9 (n=1182).

Table 4.12 below shows the raw direct comparison between the hospitals from which and the causes for which cases were eliminated during step 3.

Table 4.12 Summary of cases eliminated during step 3, block retrieval, by hospital site and cause of elimination.

Details of these causes can be found in Table 4.9 (n=1182).

Hospital | Cause	Samples inaccessible	Missing samples	Protocol Change	Total
St. James' University Hospital	0	84	638	722
Cork University Hospital	220	0	0	220
Mater Misericordiae University Hospital	108	10	0	118
St. Vincent's University Hospital	50	8	0	58
Beaumont University Hospital	15	40	0	55
Limerick University Hospital	0	9	0	9
Total	393	151	638	1182

To conduct meaningful statistical analysis on this data, it was necessary to combine cases eliminated due to 'samples inaccessible' and 'protocol change' causes, along with Cork University Hospital and Limerick University Hospital cases. The results of this tabulation and the associated statistical test are shown in Table 4.13.

Table 4.13 Summary of cases eliminated during step 3, block retrieval, by hospital site and cause of elimination with combined categories from Table 4.12 and the relevant association test.

'Samples inaccessible' and 'protocol change' were combined and are represented by 'Samples inaccessible/Protocol change'. Cork University Hospital and Limerick University Hospital were also combined and are represented by 'Cork and Limerick University Hospitals' (n=1182).

Hospital | Cause	Samples inaccessible/Protocol change	Missing samples	Total
St. James' University Hospital	638	84	722
Cork and Limerick University Hospitals	220	9	229
Mater Misericordiae University Hospital	108	10	118
St. Vincent's University Hospital	50	8	58
Beaumont University Hospital	15	40	55
Total	1031	151	1182
Association Test	Chi square p<0.0001*		

*Denotes a significant finding at the 95% level.

4.6.6 Causes and origins of attrition during step 4

For cases eliminated during step 4 (n=254), histopathological analysis of H+E slides, Figure 4.6 shows the percentage of cases rendered ineligible by cause.

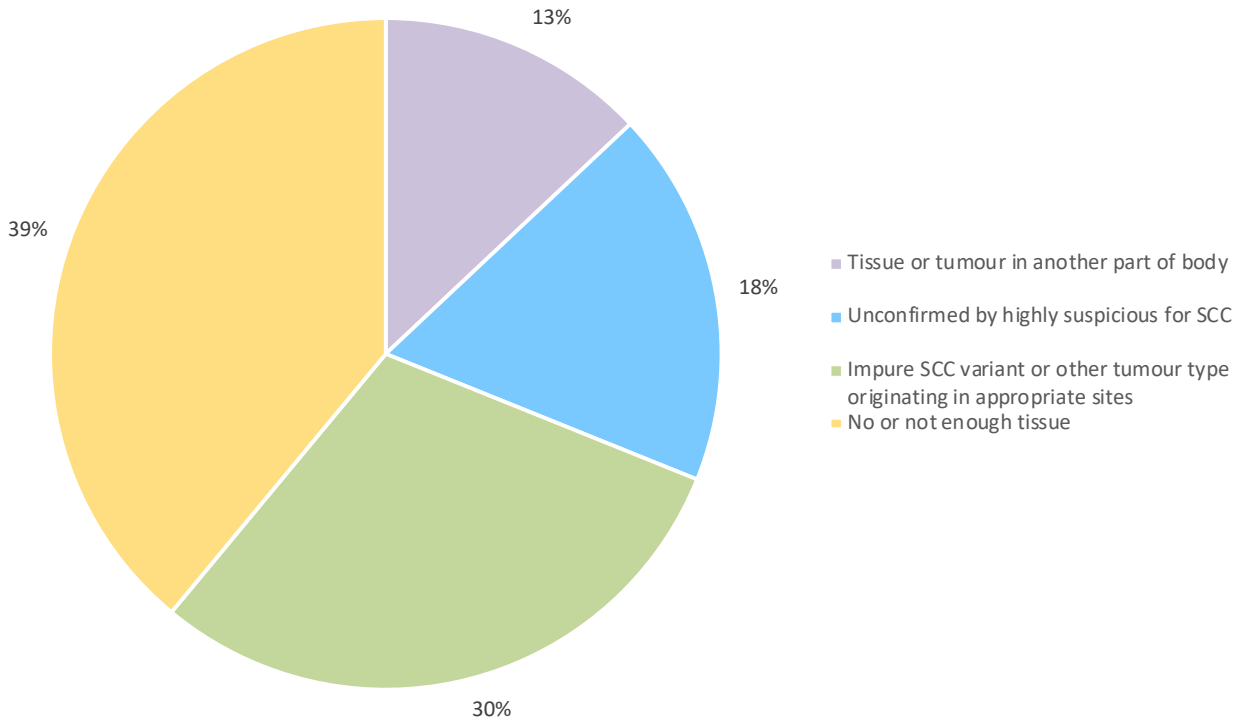

Figure 4.6 Percentage of cases rendered ineligible by cause during step 4, histopathological analysis of H+E slides.

Details of these causes can be found in Table 4.9 (n=254).

For those eliminated during step 4, histopathological analysis using H+E slides, Figure 4.4 shows the number and proportion of cases rendered ineligible by hospital site.

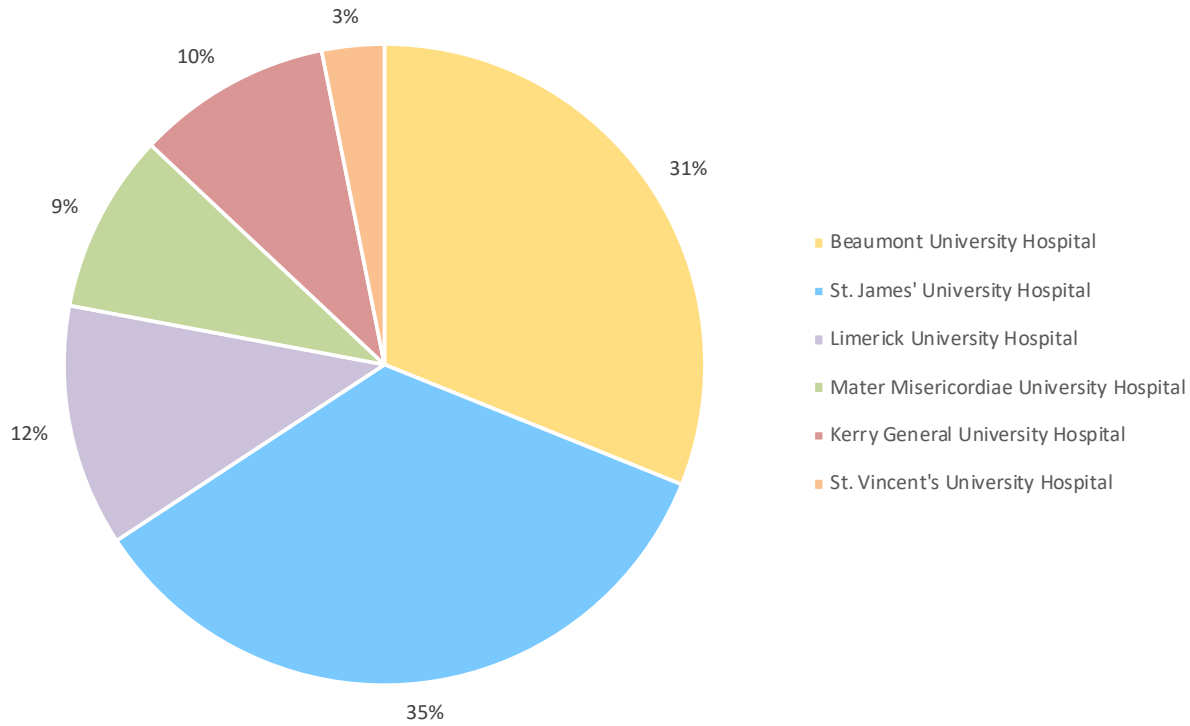

Figure 4.7 Percentage of cases rendered ineligible by hospital site during step 4, histopathological analysis of H+E slides.

Details of these causes can be found in Table 4.9 (n=254).

Table 4.14 below shows the direct comparison between the hospitals from which and the causes for which cases were eliminated during step 4.

Table 4.14 Summary of cases eliminated during step 4, histopathological review of H+E slides, by hospital site and cause of elimination.

Details of these causes can be found in Table 4.9 (n=254).

Hospital | Cause	Tissue or tumour in another part of body	Unconfirmed but highly suspicious for SCC	Impure SCC variant or other tumour type	No or not enough tissue	Total
Beaumont University Hospital	25	3	23	28	79
St. James' University Hospital	2	18	37	31	88
Limerick University Hospital	2	11	6	12	31
Mater Misericordiae University Hospital	0	8	5	10	23
Kerry General University Hospital	0	6	4	15	25
St. Vincent's University Hospital	4	0	1	3	8
Total	33	46	76	99	254

As evident from Table 4.14, Kerry General, the Mater, and St. Vincent's contributed the least number of eliminated cases at this stage. To facilitate the best possible and reliable statistical assessment, cases from the Mater, Kerry General, and St. Vincent's were combined. 'Unconfirmed but highly suspicious for SCC' and 'Tissue or tumour in another part of body' were also combined due to thematic similarity and sample size. The resulting tabulation and relevant association test are shown in Table 4.15.

Table 4.15 Summary of cases eliminated during step 4, histopathological analysis of H+E slides, by hospital site and cause of elimination with combined categories from Table 4.14 and the relevant association test. *'Unconfirmed but highly suspicious for SCC' and 'Tissue or tumour in another part of body' were combined and are represented by 'Suspicious for SCC and other part of body.' Kerry General University Hospital, the Mater Misericordiae University Hospital, and St. Vincent's University Hospital cases were combined and are represented by 'Mater, Vincent's, and Kerry University Hospitals' (n=254).*

Hospital | Cause	Suspicious for SCC and other part of body	Impure SCC variant or other tumour type	No or not enough tissue	Total
Beaumont University Hospital	28	23	28	79
St. James' University Hospital	20	37	31	88
Limerick University Hospital	13	6	12	31
Mater, Vincent's, Kerry University Hospitals	18	10	28	56
Total	79	76	99	254
Association Test	Chi square p=0.028*			

*Denotes a significant finding at the 95% level.

4.7 Discussion

The aims of the foregoing analysis were four-fold. The first aim was to establish the number of procedures and parties necessary to acquire pathology reports and retrieve sample FFPE blocks from all hospitals for the ECHO study. The second was to determine the lengths of time (days) taken to organize and execute the review of pathology reports and retrieval of FFPE blocks from these same sites. The third and fourth were to pinpoint the steps of the report review and FFPE block retrieval process contributing most to the attrition of cases from the study, and to identify the reasons for the attrition of cases at each step of the

process. The significance of generating results for these aims is rooted in the constructive evaluation of the way in which Irish clinical research is conducted.

The study established a total of 20 different procedure and party types necessary to organize and execute the review of pathology reports and the retrieval of blocks (Table 4.2). These procedure and party types were generalized across all 7 hospital sites. Based on these generalized descriptions, each hospital required an average of 6.43 procedure and party types to acquire requested blocks. However, each hospital had its own unique version of the coded procedure, and its own hospital-based party. In other words, if a Chief Medical Scientist was involved in the process, a unique Chief Medical Scientist was involved at each site. Thus, a grand total of 45 unique procedures and parties were involved in the final acquisition of sample blocks (Table 4.2).

This said, Table 4.2 identifies 11 different party types required to facilitate sample retrieval compared to 4 different procedure types to review pathology reports, and 5 different procedure types to retrieve blocks. The much smaller number of generalized procedure types suggests that the way in which each individual hospital coordinates sample acquisition, though still varied, is actually relatively similar. It just so happens that because each hospital manages pathology reports and associated FFPE blocks separately, rather than through a centralized system, a large number of party types are required to carry out the same tasks in each site.

The enormous number of organizational steps required to acquire samples is clearly reflected in the data generated regarding the length of time (days) it took to organize and execute the review of pathology reports and the retrieval of FFPE blocks. The means exhibited in Table 4.3 and statistically compared in Table 4.5 showcase the significant difference between the length of time (days) it took to: Organize pathology report review and execute pathology report review; organize block retrieval and execute block retrieval; organize pathology report review and execute block retrieval; and organize block retrieval and execute pathology report review. The four-way Kruskal-Wallis comparison further establishes the significant divergence in the lengths of time (days) for each of these processes. Most summative however is the analysis of the combined means in Table 4.4.

This unequivocally reflects the fact that the organizing of report review and block retrieval was the major contributor to delaying the acquisition of blocks. An average of 273.67 (CI: 121.04, 426.30) days to organize the review of reports or block retrieval speaks for itself in comparison to an average of 7.17 (CI: 1.35, 12.99) days to actually carry out the review and retrieve blocks.

These results highlight the fact and impact of the decentralized manner in which clinical research is conducted in Ireland, especially for studies that require the collaboration of multiple registries, biobanks, and hospitals. In comparison to other projects, the number of parties and procedures required in this study is enormous[10–13]. Indeed, the ECHO study is one of the largest of its kind regarding HPV and HNSCC, but some even bigger than the present study have had fewer organizational steps to overcome. As a consequence, the organizational process for studies of this scale in Ireland is extremely lengthy and far longer than the actual execution of the protocols decided upon.

Table 4.2 emphasizes that this is mainly due to the fact that each hospital site involved required its own procedures and roster of parties to ultimately retrieve blocks from varying storage sites across the country. However, that these procedures are relatively similar is promising. Though it primarily indicates inefficiency because the same procedures have to be repeated many times to access samples, it also suggests that the centralization of managing patient samples from one national database is not only necessary but possible. Hospitals in Ireland have individually established similar ways to access samples. They may therefore be well prepared and amenable to using these same procedures but sourced from a singular and integrated database for all hospitals. Such a move would change little for the manner in which hospitals operate currently, but would simply require the homogenization of both database software and patient sample storage across the country. The latter is often centralized from one storage location in other nations[10], something which is feasible in Ireland given its smaller and concentrated geographic size.

With respect to the attrition of cases, it is significant that 56.40% of incident tumours had to be eliminated in step 1 due to ineligibility and lack of access to pathology reports before a single pathology report was even reviewed. Indeed, 28.08% of this was a result of a lack of

eligibility of samples in the original search of the NCRI database. This narrowing down of the population is typical and necessary in most studies. The very particular specifications of samples needed for the study, outlined in Section 3.3 actually make this 28.08% a reasonable number of cases to exclude. A further 12.78% of this 56.40% was eliminated due a lack of alive patient consent. Though the consent rate for the study was just above 44% for alive patients, with all deceased patients not requiring specific active consent, this makes it clear that ultimately, lack of consent played a relatively small role in attrition during step 1.

More disconcerting is the remaining 15.54% of the 56.40% of all potential cases eliminated during step 1 that were removed simply because pathology reports could not be accessed due to an inability to establish procedural protocols with 6 of the 14 participating hospitals sites in a timely manner. In other words, not only were more than the established average of 237.67 days involved in trying to organize FFPE block retrieval for these remaining 6 hospitals, but no protocols could actually be established before the end of the study. This is by no fault of any individual site, but simply another consequence of the decentralized manner in which patients and samples are managed. Should a single pathway for patient management, consent, access have been available, an even greater diversity of patients and hospitals may have been easily incorporated into the study. Thus, though the study did establish as representative a population as possible, contacting hospitals across the country in an effort to avoid sole focus on Dublin centers, an integrated data management system would have improved the efficiency of its creation.

Sample attrition from step 2 onwards is summarized in Table 4.7. Along with the Table's associated significance test, it unequivocally showcases that hospitals were affected by step number differently, with some seeing greater or less attrition than others at each step. Examining the Table more closely reveals that Limerick had more cases than expected eliminated during step 2, pathology report review, the same being true of Beaumont during step 4, the histopathological analysis of H+E slides. What is even more evident however is that the grand majority of cases eliminated from St. James', St. Vincent's, the Mater, and Cork were removed during step 3, the retrieval of blocks. Overall, the inability to retrieve blocks resulted in 1182 of the 1666 cases eliminated, thus making step 3 the largest contributor to attrition overall.

A closer examination of step 2, pathology report review, in Figure 4.2 highlights that despite the elimination of seemingly ineligible samples from the 5792 incident tumours identified in the NCRI database, more cases had to be disqualified using the same eligibility criteria. This may be a reflection of the laborious and sometimes manual way in which data is recorded (or transferred) from hospital records to the national database. It may also be an artefact of the more general way in which the pathology of tumours are recorded, simply being denoted as "Squamous" for the purposes of database searches. The 49% of cases eliminated during step 2 due to 'impure SCC variant or other tumour type originating in the appropriate anatomical site' justifies this conclusion. The same is true of the additional 11% eliminated that were only suspicious and unconfirmed for SCC. The exhaustive lists detailing the eliminated highly suspicious cases, and the impure SCC variants and other tumour types in Table 4.9 is also a testament to the wide range of case types the database included even after step 1.

That said, cumulatively, 31% of the cases eliminated after pathology report review were removed because: The histological identification numbers provided in the NCRI-generated database were not found in the hospital databases; the case referred to a tumour in an irrelevant part of the body; no malignancy was reported; or the case was a recurrence and not a primary tumour. Almost one third of cases eliminated during step 2 were thus ineligible due to a disconnect between the recording of patient data into the national database, and the hospital databases from which information was drawn. Those cases for which no blocks were available, and the one case for which the pathology report was not printed made a minimal contribution to attrition in step 2. These were therefore not disqualified due to any database faults.

Figure 4.3 shows that the majority of cases eliminated during step 2 were from St. James' University Hospital, Mater Misericordiae University Hospital, and University Hospital Limerick. 40% of the cases disqualified came from St. James', an attrition twice as high as that of the Mater. Based on the number of cases from each hospital for which reports were reviewed (Table 4.6), step 2 saw slight disproportionate elimination of cases from the three aforementioned hospitals.

This said, the significance of the statistical test conducted on Table 4.11 shows a clear relationship between hospital site and the cause of elimination in step 2. The specific breakdown of the numbers in Table 4.10 would suggest that St. James' cases were disproportionately eliminated due to impure SCC variants or other tumour types being detected. The same is true of the Mater and cases drawn from other parts of the body. Limerick saw a larger proportion of cases disqualified for only being suspicious of SCC. This indicates that either the scan of the NCRI database during step 1 did not apply eligibility criteria in the same way across all hospitals, or that the information registered in the database was done so differently for each hospital. The latter is more likely as the scan of the database was unilateral, of a standardized nature, electronic, and conducted at one point in time.

Figures 4.4 and 4.5, along with Tables 4.12 and 4.13 and the associated significant statistical test, showcase that protocol changes and inaccessible samples were the main reasons why the 1182 of 1666 disqualified cases were eliminated during step 3. These originated overwhelmingly from St. James' Hospital as a result of regulations put in place after the arrival of new GDPR guidelines which did not prohibit access by any means, but required a re-application for access that caused months of time delays (Table 4.9). The remaining eliminated cases during step 3 were removed slightly disproportionately from the Mater and Cork due to samples being inaccessible (Table 4.9).

It is evidently impossible to predict the introduction of new regulations regarding data protection and patient and sample management. However, that the consequences of these guidelines would disproportionately affect some hospitals and their management of samples is unexpected. St. James' was the largest hospital site involved in the study given its status as the center for referrals and treatment of head and neck cancer patients. It was still very well represented in the study population despite step 3 attrition, but the new protocols established changed the management of cases completely and due to time delays, prevented even more cases from the important site being included from this site.

Where these new guidelines presented a challenge in one hospital, they did not in another where accessibility of samples was instead a roadblock. In the Mater and Cork, inaccessibility was simply due to an inability to establish a cost-free procedure for retrieving blocks from off-site private storage companies. This is by no means the responsibility of the hospitals themselves, but instead just another artefact of the diversity of ways in which patients and samples are currently managed.

Finally, another 254 cases were eliminated during step 4, the histopathological analysis of H+E slides. This is a large number of cases to disqualify during the last step of the retrieval process given that the samples had already been subjected to two eligibility screening stages during steps 1 and 2. In fact, Figure 4.6 reveals that 61% of the eliminated cases were disqualified because they did not adhere to the strict eligibility criteria of the study (e.g. suspicious for SCC, tumours in another part of the body, and impure SCC or other tumour type). The vast majority of this 61% were eliminated due to the diagnosis of an impure SCC variant or other tumour type in the appropriate region, or they were only suspicious for SCC. This is justifiable for impure SCC variants and suspicious cases if pathology reports simply reported the general nature of the tumour without specifying detailed tumour sub-pathology. Different tumour types (e.g. adenocarcinoma, acinic cell carcinoma) and the remaining 13% of this 61% originating in an irrelevant part of the body are less explicable. Their differential diagnosis is not easily confused with SCC and would have been expected to be documented specifically in the pathology reports (step 2) and thus in the NCRI database (step 1).

The continued appearance of cases originating in irrelevant parts of the body and other tumour types could indicate a lack of synchronization between the hospital databases and the labelling of FFPE blocks in storage. It is also possible that in some hospitals, one identification number is associated with more than one patient. Often, a single additional letter in the identification number may distinguish between these patients. This alphabetical nuance is however sometimes absent from the labels of FFPE blocks in storage. Thus, blocks are confused, stored in the incorrect bins, or simply withdrawn by mistake.

The persistence of cases only highly suspicious of SCC after steps 1 and 2 is indicative of disconnect between pathology reports and the real-time embedding of tissue in FFPE blocks. As outlined in Section 3.4.3, reports were reviewed specifically to identify blocks that had relevant tumour tissue. That blocks pinpointed in reports as having ample confirmed SCC proved only suspicious of SCC suggests that either reports documented blocks incorrectly or blocks were labelled mistakenly. It is possible that pure SCC was present in other blocks associated with the case, but due to this epistolary discord, blocks containing surrounding less defined tumour were drawn instead.

It should be noted however that overall, as shown in Figure 4.6, the many cases disqualified during step 4 (39%) either had no tissue left in the associated blocks, or had insufficient tumour tissue to reliably conduct molecular analysis. This is simply a product of former use of the tissue for research or clinical needs, or a result of the very small size of available biopsies. It is difficult to assess the size and block-by-block availability of ample tumour specimen during steps 1, 2, and 3, so this attrition is almost unavoidable.

The three hospitals most heavily represented in case elimination during step 4 were Beaumont University Hospital, St. James' University Hospital, and University Hospital Limerick. However, Beaumont and Kerry disproportionately contributed to attrition in this step. The former saw 79 cases disqualified with 235 cases processed, with the latter having 25 cases disqualified when only 65 were processed. This is in comparison to for instance, 88 cases disqualified in St. James' and 357 cases processed. The significant association test and data in Table 4.15 highlights this disparate contribution. Beaumont appears to have suffered more from attrition due to cases appearing from irrelevant parts of the body. Both James' and Beaumont saw more elimination due to impure SCC or other tumour types in the appropriate sites. Limerick showed disproportionate elimination due to only suspicious SCC cases whilst almost half of those cases eliminated from Kerry were due to a lack of tissue in blocks.

The cumulative analysis of attrition during steps 2, 3, and 4 definitively shows that the refining of the eligible study population for nation-wide Irish investigations varies by hospital. Each site, likely as a result of their own unique pathology report and FFPE bock

filing and storing procedures analyzed above, experiences differential sample attrition for statistically diverging reasons. Some have more flexible agreements with off-site private storage companies while others do not. Some do not make use of private off-site storage at all but instead have less precise translations between pathology reports and the labelling of associated FFPE blocks. The combinations of strengths and limitations for each is diverse.

What can therefore be extrapolated is that the communication and synchronization between decentralized registries, storage companies, biobanks, and hospitals is not evenly distributed. The innumerable number of procedures that must be fulfilled and parties that must generously dedicate their time to making sure procedures are followed ultimately creates a sincere distance both temporally and in terms of impact between all of these clinical and research bodies, and patients. It should be noted that this is not just an Irish issue. Many database analyses from across Europe and North America have noted similar heterogeneity in documentation and access to cases and information, whether it be for observational, epidemiological, or treatment/drug-related data[14–16].

The need for a homogenized database and biobanking system for the nation is therefore evident. Several studies have noted that in the era of big data, the larger the extent of centralization in the research process, the greater the efficiency, quality, and impact of the research on patients[17,18]. Since the ultimate goal of clinical research is a realized improvement in patient outcomes, the faster meaningful results can be generated, and the greater the quality of these results, the more research serves its real purpose.

This is especially true in the case of HPV-related diseases, specifically those of the head and neck. Based on the current literature, HNSCC is at least in part a preventable and imminently treatable affliction[19–26]. For instance, the better survival of HPV-related HNSCC patients has raised the question of whether or not such severe treatment is necessary[12,27–30]. The possibility of de-escalation of treatment in HPV-related HNSCC has been discussed at length for many years, the most significant result of which has been a change in the staging of HNSCC tumours on the basis of viral origins[31,32]. This said, de-escalation has still not been introduced into the clinic due to continued uncertainties[27,33–36]. The rate at which studies analyzing de-escalation is not yet enough to provide evidence for its implementation or

conversely, its lack of legitimacy. The longer a firm conclusion takes to be drawn, the greater the chance that more patients will be subjected to potentially unnecessary and debilitating treatments. The same is true of the uncertain role the HPV vaccination in HNSCC prevention[23,24].

It should be emphasized that the need for a centralized and efficient patient and sample management system is not simply for the benefit of research's long-term effects on patients. First and foremost, the centralization and efficiency of health information technology has a direct impact on the quality and success of patient care[37]. This is especially true in Ireland where several hospitals may be responsible for managing a single patient's care. This was even evident in pathology reports reviewed for the ECHO study where cases for some patient could not be found or accessed because they were registered to one hospital but were actually being managed by another hospital. As a consequence, this was understandably not readily obvious to any external registries, including the NCRI. If health information and associated patient material and samples can be managed from one point of access, integrated within hospitals as a national standard, patients can be more efficiently and effectively treated. Referrals between hospitals would be processed more quickly, and the location of associated patient material would be less indeterminable with singular tracking and management.

It should be noted that efforts to create such an integrated biobanking system are already underway. A new International Organization for Standardization (ISO) guideline, "Requirements for Biobanking" (ISO 20387), was adopted in Ireland in 2018. Furthermore, the new GDPR Health Research Regulations implemented in 2018 attempt to standardize the management of patient information and data for research purposes. Importantly, BIOBANK: Ireland Trust is the country's most noted push towards establishing an Irish BioBank Network[6]. Their aims are to ensure that hospital biobanks collect samples using the same procedures and database, with a view towards online access to samples and data for potential projects. The establishment of a national BioBanking network would enhance or complement other aims of national cancer strategy including better information technology systems, communications, and a national research database. Most importantly, it would help connect hospitals and focus on patient care rather than on competing institutions.

Collaborating with other countries whose national biobanking is already in order would also be made far more efficient.

The BIOBANK: Ireland Trust project already has functioning arms in St. James' University Hospital and University Hospital Cork with goals, limited by funding, to expand further. That the main wing of this BioBank is in St. James' is particularly significant in the context of this study. The present analysis determined 1182 of the 1666 total cases eliminated from the ECHO study were disqualified simply due to an inability to physically access the cases. St. James' disproportionately contributed 722 cases to this 1182 due to time delays as a result of new guidelines implemented after the introduction of GDPR through the BioBank itself. Thus, though poorly timed in the framework of this study, the standardization of access protocols through the BioBank across the nation will ultimately centralize all sample access permissions to all storage sites, whether private or internally managed by the hospitals. Not only will this avoid the inaccessibility of cases in St. James', but also those emanating from the other hospitals represented in Table 4.12.

The BIOBANK: Ireland Trust is the first step in a series for the implementation of several proposals outlining the best strategies for homogenizing patient and healthcare information databases to increase the quality of patient care and associated and varying types of clinical research[38–40]. There would need to be further assessment of the Irish context, the number of databases and facilities that exist for patient information and samples, and how they are managed to determine which exactly which strategy is best. That said, this analysis unequivocally points to three conclusions that make this centralization an achievable goal:

1. Hospitals already use similar databases and procedures for managing patients and patient material and the integration of this process would not unduly disrupt systems already in place, but simply improve them.
2. Singular patient and case identification numbers are already a facet of how hospitals, registries, biobanks, and research bodies collaborate and would therefore not need to be newly instated as in the case of other nations.
3. The disconnect between clinical, registry, and research bodies in Ireland only exists due to the large number of procedures and parties involved in their management, something that if

standardized and made less complex, could prevent miscommunications or mistaken documentation of patients and patient material.

Ireland already garners more funding and publishes more impactful research at greater rates than would be expected on average from a country its size. The impact on patients in many realms is visible in real-time. Nonetheless, this research is mainly focused on small-scale clinical studies conducted by research units attached to individual hospitals. Indeed, collaborations and communications between hospitals on this scale are numerous and successful. For instance, Ireland boasts the most highly cited groups studying rare metabolic conditions due to enormous patient-focused output generated in the Mater Misericordiae Hospital and Temple Street Hospital[40].

For larger scale projects like the ECHO study however, requiring data and samples from a huge number of hospitals and existing biobanks, such collaboration evidently becomes much more difficult to coordinate and execute efficiently. This is extensively evidenced by the foregoing analysis which determined: The enormous number of procedures and party types necessary to acquire samples from 7 hospital sites; the significantly larger time taken (days) to organize pathology report review and block retrieval compared to that required for the execution of report review and block retrieval; the diverse reasons for which samples must be eliminated due to ineligibility despite multiple screening stages to avoid continuous attrition; and the unevenly distributed impact of the causes of sample attrition for each involved hospital site at every step of the sample acquisition process. Thus, to avoid attrition, lengthy organizational procedures, the potential compromise of research's impact, and miscommunications between clinical and research bodies, a centralized management of all Irish patients and samples is urgently needed. Only then will Irish patients see the real-time results of efficient research that they contribute to, and experience the highest possible quality care.

References

1. Scimago Journal and Country Rank. SJR - Citable Medicine Documents Ranking by Country. (2017). Available at: https://www.scimagojr.com/countryrank.php?area=2700&order=cd&ord=desc&min=0&min_type=itp. (Accessed: 22nd January 2019)
2. Manson, N. C. The ethics of biobanking: Assessing the *right to control* problem for broad consent. *Bioethics* (2019). doi:10.1111/bioe.12550
3. Rao, A. *et al.* Critical Financial Challenges for Biobanking: Report of a National Cancer Institute Study. *Biopreserv. Biobank.* **17**, 129–138 (2019).
4. Patil, S. *et al.* Cancer oriented biobanks: a comprehensive review. *Oncol. Rev.* **12**, 357 (2018).
5. Langhof, H., Schwietering, J. & Strech, D. Practice evaluation of biobank ethics and governance: current needs and future perspectives. *J. Med. Genet.* **56**, 176–185 (2019).
6. Flanagan, C. & Gaffney, E. Cancer Research in Dublin Ireland | Biobank Ireland Trust. *Biobank Trust Ireland* (2019). Available at: http://www.biobankireland.com/. (Accessed: 30th April 2019)
7. Mee, B. *et al.* Development and progress of Ireland's biobank network: Ethical, legal, and social implications (ELSI), standardized documentation, sample and data release, and international perspective. *Biopreserv. Biobank.* **11**, 3–11 (2013).
8. thejournal.ie. Irish cancer research gets a boost with its first ever brain tumour biobank. *thejournal.ie* (2017). Available at: https://www.thejournal.ie/ireland-tumour-biobank-3410402-May2017/. (Accessed: 30th April 2019)
9. National Cancer Registry Ireland. National Cancer Registry Ireland. *National Cancer Registry* (2019). Available at: https://www.ncri.ie/. (Accessed: 30th April 2019)
10. Schache, A. G. *et al.* HPV-Related Oropharynx Cancer in the United Kingdom: An Evolution in the Understanding of Disease Etiology. *Cancer Res.* **76**, 6598–6606 (2016).
11. Castellsagué, X. *et al.* HPV Involvement in Head and Neck Cancers: Comprehensive Assessment of Biomarkers in 3680 Patients. *J. Natl. Cancer Inst.* **108**, (2016).
12. Anantharaman, D. *et al.* Combined effects of smoking and HPV16 in oropharyngeal cancer. *Int. J. Epidemiol.* dyw069- (2016). doi:10.1093/ije/dyw069

13. Chaturvedi, A. K. *et al.* Human Papillomavirus and Rising Oropharyngeal Cancer Incidence in the United States. *J. Clin. Oncol.* **29**, 4294–4301 (2011).
14. Madigan, D. *et al.* Evaluating the Impact of Database Heterogeneity on Observational Study Results. *Am. J. Epidemiol.* **178**, 645–651 (2013).
15. Crossfield, S. S. R. & Clamp, S. E. Centralised Electronic Health Records Research Across Health Organisation Types. in 394–406 (2014). doi:10.1007/978-3-662-44485-6_27
16. Pacurariu, A. *et al.* Electronic healthcare databases in Europe: descriptive analysis of characteristics and potential for use in medicines regulation. *BMJ Open* **8**, e023090 (2018).
17. Ehrenstein, V., Nielsen, H., Pedersen, A. B., Johnsen, S. P. & Pedersen, L. Clinical epidemiology in the era of big data: new opportunities, familiar challenges. *Clin. Epidemiol.* **9**, 245–250 (2017).
18. Mooney, S. J., Westreich, D. J. & El-Sayed, A. M. Commentary: Epidemiology in the era of big data. *Epidemiology* **26**, 390–4 (2015).
19. Albers, A. E., Qian, X., Kaufmann, A. M. & Coordes, A. Meta analysis: HPV and p16 pattern determines survival in patients with HNSCC and identifies potential new biologic subtype. *Sci. Rep.* **7**, 16715 (2017).
20. Albers, A. E., Qian, X., Kaufmann, A. M. & Coordes, A. Meta analysis: HPV and p16 pattern determines survival in patients with HNSCC and identifies potential new biologic subtype. *Sci. Rep.* **7**, 16715 (2017).
21. Tong, F. *et al.* Prevalence and Prognostic Significance of HPV in Laryngeal Squamous Cell Carcinoma in Northeast China. *Cell. Physiol. Biochem.* **49**, 206–216 (2018).
22. Sedghizadeh, P. P. *et al.* Is p16-positive oropharyngeal squamous cell carcinoma associated with favorable prognosis? A systematic review and meta-analysis. *Oral Oncol.* **54**, 15–27 (2016).
23. Wang, C. *et al.* Targeting Head and Neck Cancer by Vaccination. *Front. Immunol.* **9**, 830 (2018).
24. Guo, T., Eisele, D. W. & Fakhry, C. The potential impact of prophylactic human papillomavirus vaccination on oropharyngeal cancer. *Cancer* **122**, 2313–23 (2016).
25. Brotherton, J. M. L., Giuliano, A. R., Markowitz, L. E., Dunne, E. F. & Ogilvie, G. S. Monitoring the impact of HPV vaccine in males-Considerations and challenges.

Papillomavirus Res. (Amsterdam, Netherlands) **2**, 106–111 (2016).

26. Pinto, L. A. *et al.* Quadrivalent Human Papillomavirus (HPV) Vaccine Induces HPV-Specific Antibodies in the Oral Cavity: Results From the Mid-Adult Male Vaccine Trial. *J. Infect. Dis.* **214**, 1276–83 (2016).
27. Wierzbicka, M., Szyfter, K., Milecki, P., Składowski, K. & Ramlau, R. The rationale for HPV-related oropharyngeal cancer de-escalation treatment strategies. *Contemp. Oncol. (Poznan, Poland)* **19**, 313–22 (2015).
28. Woods, R. Human Papillomavirus-related Oropharyngeal Squamous Cell Carcinoma – Prevalence and Incidence in a Defined Population and Analysis of the Expression of Specific Cellular Biomarkers. (2016).
29. Woods, R. *et al.* Role of human papillomavirus in oropharyngeal squamous cell carcinoma: A review. *World J. Clin. cases* **2**, 172–93 (2014).
30. Ang, K. K. *et al.* Human papillomavirus and survival of patients with oropharyngeal cancer. *N. Engl. J. Med.* **363**, 24–35 (2010).
31. American Joint Committee on Cancer & American Cancer Society. *AJCC Cancer Staging Manual*. (Springer Publishing, 2016).
32. American Joint Committee on Cancer & American Cancer Society. Head and Neck Sites. in *AJCC Cancer Staging Manual* 24–59 (Lippincott-Raven Publishers, 1997).
33. Brisson, R. J. *et al.* De-escalation in HPV-negative locally advanced head and neck squamous cell cancer (LA-HNSCC) in patients after induction chemotherapy: A retrospective case series. *J. Clin. Oncol.* **36**, e18090–e18090 (2018).
34. Mirghani, H. & Blanchard, P. Treatment de-escalation for HPV-driven oropharyngeal cancer: Where do we stand? *Clin. Transl. Radiat. Oncol.* **8**, 4–11 (2018).
35. Villaflor, V. M. *et al.* Response-adapted volume de-escalation (RAVD) in locally advanced head and neck cancer. *Ann. Oncol.* **27**, 908–913 (2016).
36. Chen, A. M. *et al.* Reduced-dose radiotherapy for human papillomavirus-associated squamous-cell carcinoma of the oropharynx: a single-arm, phase 2 study. *Lancet Oncol.* **18**, 803–811 (2017).
37. Chaudhry, B. *et al.* Systematic Review: Impact of Health Information Technology on Quality, Efficiency, and Costs of Medical Care. *Ann. Intern. Med.* **144**, 742 (2006).
38. Gini, R. *et al.* Data Extraction and Management in Networks of Observational Health Care Databases for Scientific Research: A Comparison of EU-ADR, OMOP, Mini-

Sentinel and MATRICE Strategies. *EGEMS (Washington, DC)* **4**, 1189 (2016).

39. Quantin, C. *et al.* Medical record: systematic centralization versus secure on demand aggregation. *BMC Med. Inform. Decis. Mak.* **11**, 18 (2011).

40. Quantin, C. *et al.* Centralised versus decentralised management of patients' medical records. *Stud. Health Technol. Inform.* **150**, 700–4 (2009).

Chapter 5

VALIDATING HPV DNA DETECTING METHODOLOGIES

5 CHAPTER 5: VALIDATING HPV DNA DETECTING METHODOLOGIES

5.1 Introduction

HPV positivity is ill-defined in the current literature. There are not only a number of indicators for HPV positivity (e.g. DNA, mRNA, viral load, viral integration), but there is also a multiplicity of different scientific principles used for detecting each indicator with even more numerous technological platforms available to apply them. These technologies have varying specificities, sensitivities, and methods for quantifying HPV status (e.g. binary, continuous). They are also optimized for differing sample types (e.g. FFPE, fresh, frozen). As a consequence, there is exponential diversity in the literature as to the definition of a HPV positive sample, and the validity and reliability of various technologies.

For instance, to detect HPV DNA alone, there are three central approaches: nucleic acid hybridization, signal amplification, and nucleic acid amplification[1,2]. Signal amplification methods are used in several technologies including the Digene HPV platform with hybrid capture 2 and the Cervista HPV HR assay[3]. Nucleic acid amplification is PCR-based, with studies using different PCR types from a selection of pure PCR, PCR-FRFLP, real-time PCR, and Abbott real-time PCR. This PCR product is then subjected to various tests of different sensitivities and specificities including gel electrophoresis, microarray analysis, and laser-based xMAP® Luminex detection. Each of these technologies quantifies HPV's presence differently. For example, gel electrophoresis uses a binary visual determination and xMAP® Luminex is semi-quantitative. The diversity of available platforms for HPV DNA is similar for those used to detect HPV mRNA, viral load, and viral integration.

It should be noted that experimental disparities are not simply differences in technologies, principles, and classifications used, but also discrepancies in preparatory techniques, cleaning and sterilizing protocols, and extraction methods. In the case of FFPE tissue used for molecular testing, it is standard to use a new sterile blade for each case cut[4,5]. However, heterogeneity in other very basic procedures to prepare tissue for complex tests is rampant across the literature, especially in the case of protocols for extracting DNA and mRNA. This is particularly true with respect to sterilization of the microtome used for cutting. Some studies employ combinations of chemicals to melt away paraffin and disinfect including

industrial methylated salts (IMS) and ethanol[6–9]. Others are extremely particular about adding solutions that specifically disintegrate DNA and mRNA, including DNA and/or mRNA Zap[4,5,10,11]. All of these approaches have been validated, though the addition of products like DNA Zap is sanctioned by IARC, Lyon, France[5]. Furthermore, extraction techniques are by no means homogenous. In the case of DNA, the QIAGEN QIAamp DNA FFPE Tissue Kit, one of several xylene-driven methods, is commonly employed[9,12–14]. That said, many studies also use and compare a multiplicity of unique digestion buffers and protocols[4,5,8,10,11].

This variation, which spans the breadth of the kind of HPV indicator used to the sterility protocols employed, has severe consequences for the validity, reliability, and comparability of results published in the HPV field. This is especially true for emerging HPV-related cancers including those of the head and neck given that HPV detection technologies have traditionally been optimized for liquid-based cervical samples.

For example, though it is difficult to quantify the impact this heterogeneity has on the results of HPV-detection studies, the enormous variation of HPV prevalence in head and neck cancers in geographically and culturally similar nations is indicative. Across the data from developed countries, prevalence of HPV DNA has been cited as low as 18.5% to as high as 90%[15–19]. Most studies used to establish this range employed different indicators, combinations of indicators, and varying platforms to define a HPV positive case. Whether or not this large prevalence range is due to genuine variation in the penetration of HPV in each population or simply due to experimental disparities is unclear.

It is evident that approaches to HPV detection require standardization not only for research itself, but for the eventual effect that it will have on clinical practice. In the cervical case, standardization and its impact are already well-established. HPV is successfully being used as a triage mechanism in the clinic in several countries, with more expecting to introduce it in the coming years[3,20–23]. No such homogeneity exists in the head and neck case despite the fact that HPV's use as a triage and treatment-decision mechanism for head and neck cancers is becoming more and more apparent[1,24–31]. With survival significantly better in HPV-related head and neck cases, there is a real opportunity to de-escalate harsh treatment[28,32–35]. It is also important to accurately determine whether or not the current

HPV vaccine and its included genotypes could be deployed to prevent this subset of head and neck cancers all together[36–43].

Thus, validating HPV-detecting technologies for the head and neck is essential if HPV testing is to eventually be used in public health for prevention of head and neck cancers as well as in the clinic as a diagnostic, prognostic, and treatment-directing tool. Ensuring that results will be similar using the same basic science and molecular protocols is not only tantamount to assuring that results from the lab are valid no matter the technology being used, but also to guaranteeing that public health schemes will be as cost-effective as possible and that patients will be appropriately treated to maximize their survival and minimize morbidity and long-term side effects.

To date however, there have been no major comparative studies between different HPV-detecting technologies for head and neck samples to assess their correspondence and determine which is best suited for potential use in epidemiological or clinical studies. The pilot study for the ECHO project afforded the unique opportunity to do exactly this: to compare HPV detection methods developed to find the same HPV indicator (DNA) in head and neck cancers on the basis of diverging principles.

5.2 Aims

- To compare HPV DNA detection between SPF10 PCR Gel Electrophoresis, and Multiplex PCR paired with xMAP® Luminex laser-based technology.
- To assess the effect of altering sterility protocols on the validity of results.
- To determine which method is best suited for the epidemiological endeavors of the ECHO study as a whole.

5.3 Materials and Methods

5.3.1 Study Population and Overview

The pilot study was conducted using 139 FFPE HNSCC cases diagnosed in Ireland between 1994 to 2013. Cases were selected as detailed in Section 3.3. All 139 cases were subjected to both SPF10 PCR and gel electrophoresis, and Multiplex PCR Luminex technologies.

Upon first trial ("first attempt" in Figure 3.2) using the "Original Sterility Protocol" described in Section 3.6.2, evidence of contamination appeared in the Multiplex PCR Luminex results which were not apparent using the SPF10 PCR gel electrophoresis. Specifically, the results generated in parallel to the SPF10 PCR gel electrophoresis by the Multiplex PCR Luminex showed negative controls appearing as positive. SPF10 PCR gel electrophoresis controls showed no signs of contamination. This discrepancy occurred despite the fact that sections of FFPE tissue generated for each technology were cut at exactly the same time and sequentially one after the other, including positive and negative controls. On the basis of this comparison, corrective action was taken.

Adjustments were made in cleaning and sterilizing protocols, and a "second attempt" (Figure 3.2) of the pilot was carried out and all cases were repeated with Multiplex PCR Luminex alone given that this was clearly the more sensitive technology. This new protocol was the "IARC sterility protocol" described in Section 3.6.2. The new Multiplex PCR Luminex controls showed no abnormalities in control blocks with these alterations.

Though prevalence was determined for both sets of Multiplex PCR Luminex results, the contaminated "first attempt" Multiplex results using the "Original Sterility Protocol" were not fit for comparison. Only the Multiplex PCR "second attempt" results using the "IARC Sterility Protocol" and the SPF10 PCR results were analyzed. Throughout the analysis, the results generated by the Multiplex PCR Luminex before sterilization adjustments are referred to as "Multiplex PCR Luminex PRE". Those obtained after adjustments are referred to as "Multiplex PCR Luminex POST".

5.3.2 Sectioning of FFPE tissue

The FFPE blocks for the pilot study were cut in accordance with the procedure described in Section 3.6.2. Three 10um sections for each sample were cut for Multiplex PCR Luminex testing immediately followed by another three 10um sections for SPF10 PCR. A new sterile blade was used for each case. Positive and negative controls were cut every 10 cases.

Cleaning and sterilizing protocol differed in the first instance from that described in Section 3.6.2. Instead of DNA Zap and 70% ethanol, the microtome was cleaned with Richard Allen™ Parapel™ (an anti-wax solution) and IMS. This was the "Original Sterility Protocol". This sterilization protocol is used widely across the literature, including in other HPV-related HNSCC studies as summarized in Section 5.1. Results obtained with this sterilization protocol are those named Multiplex PCR Luminex PRE.

After contamination become apparent in these results, the sterilization protocol was changed to that exactly described in Section 3.6.2, another literature-wide procedure sanctioned and used by the Infections and Cancer Biology Laboratory in IARC, Lyon, France. This was the "IARC Sterility Protocol". The same 139 samples were cut in the same way (three 10um per sample) for the Multiplex PCR Luminex alone. This was to ensure no further cross-contamination continued to appear using the more sensitive technology. Results obtained with the new sterilization protocol are those named Multiplex PCR Luminex POST.

5.3.3 Extraction of DNA

Those sections used for SPF10 PCR were processed and extracted for DNA using the QIAGEN QIAamp DNA FFPE Tissue Kit. Quantification and purity analysis of the extracted DNA was then done using the Thermo Scientific NanoDrop 2000x Spectrophotometer. This is detailed extensively in Section 3.7.

DNA extraction for the Multiplex PCR Luminex PRE and POST methods were carried out using a home-made digestion buffer developed by the Infections and Cancer Biology Laboratory, IARC, Lyon, France. It is a protocol used for the world-wide HPV-AHEAD Study Consortium and has thus been validated and sanctioned at the international level. This extraction protocol is detailed in Section 3.7.3.2.

5.3.4 HPV DNA Detection and Genotyping

Sections cut for the 139 cases extracted with the QIAGEN QIAamp FFPE Tissue kit were tested for HPV DNA using SPF10 PCR, with gel electrophoresis of the PCR products visually showcasing the binary results (either HPV positive or negative with no genotype detail). This process is summarized in detail in Section 3.7.2.

Sections cut for the same 139 cases extracted using the home-made digestion buffer were then tested for HPV DNA using the Multiplex PCR Luminex technology based in the Infections and Cancer Biology Laboratory, IARC, Lyon, France. The Multiplex PCR applied to the extracted DNA is detailed in 3.7.3. The preparation of this PCR product and analysis by the laser-based Luminex platform is then explained in the same section. This testing protocol was precisely the same for Multiplex PCR Luminex PRE and POST results.

5.3.5 Statistical Analysis

Statistics were generated using IBM SPSS Statistics Version 25, XLSTAT 2019.1.3, and Microsoft Excel Version 16.25. Prevalence of HPV DNA by the SPF10 PCR Gel Electrophoresis and Multiplex PCR Luminex was calculated using simple proportions. Z-tests, Chi-square association, and Fisher's exact tests were used to compare any difference in prevalence detected, the latter being using when expected and/or observed values were less than 5. Concordance between results by the different methods was determined by proportion and Kappa values. All significance tests were two-tailed and carried out at the 95% level.

5.4 Prevalence Comparison

In total, three prevalence statistics were generated using the 139 cases: Multiplex PCR Luminex PRE, Multiplex PCR Luminex POST, and SPF10 PCR Gel Electrophoresis. Prevalence results are summarized in Table 5.1 below.

Table 5.1 HPV DNA prevalence statistics generated by SPF10 PCR Gel Electrophoresis, Multiplex PCR Luminex PRE, and Multiplex PCR Luminex POST (n=139).

Method Employed	SPF10 PCR Gel Electrophoresis	Multiplex PCR Luminex PRE	Multiplex PCR Luminex POST
HPV DNA Positive Cases	55	111	56
HPV DNA Negative Cases	84	28	83
Total Cases	139	139	139
HPV DNA Prevalence	39.57%	79.86%	40.29%

Multiplex PCR Luminex PRE results were not carried forward from this stage on for further analysis for the reasons detailed in Section 5.3.1.

There was no significant association detected between HPV DNA status and method employed (either SPF10 PCR or Multiplex PCR Luminex POST) (Chi-square: p=0.903)., nor was there a difference between prevalences detected (Z-test: p=0.904). The prevalence and complimentary absence of HPV DNA in Table 5.1 is graphically represented in Figure 5.1.

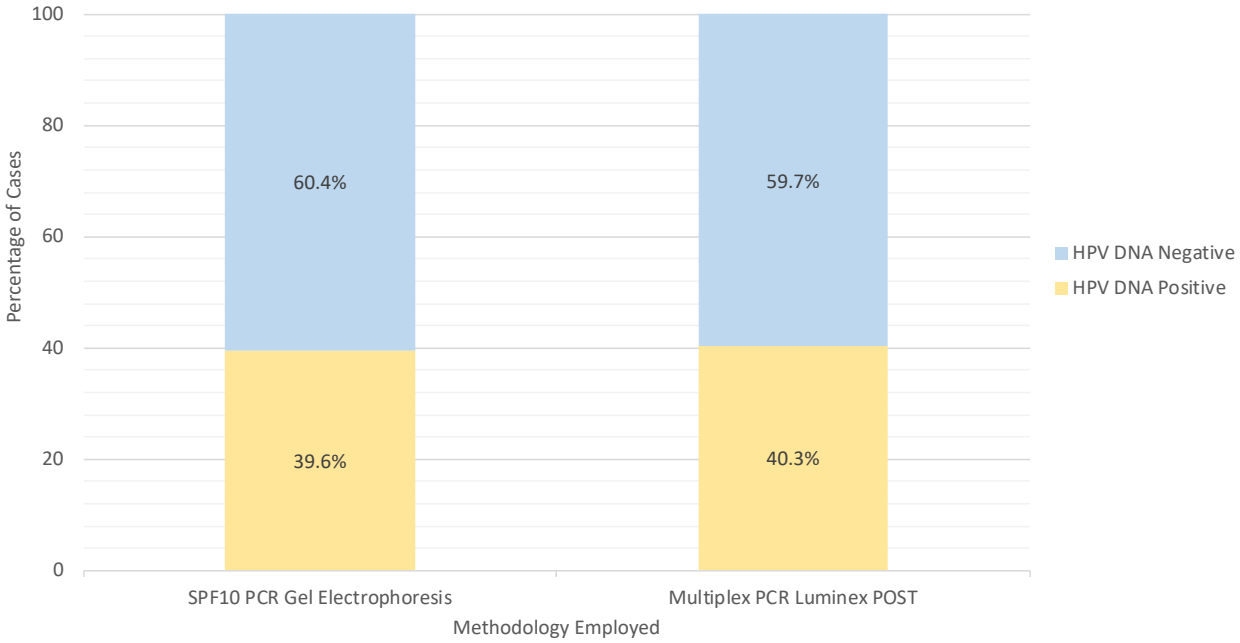

Figure 5.1 Percentage of cases deemed HPV DNA positive and HPV DNA negative by SPF10 PCR Gel Electrophoresis and Multiplex Luminex POST (n=139).

5.5 Concordance Between SPF10 PCR Gel Electrophoresis and Multiplex PCR Luminex POST Methods

The concordant and disconcordant cases with respect to HPV DNA status for SPF10 PCR Gel Electrophoresis and Multiplex PCR Luminex POST methods are individually shown in Table 5.2. The table also includes associated concordance and Kappa statistics.

Table 5.2 Tabular comparison of HPV positive and negative cases as determined by SPF10 PCR Gel Electrophoresis and Multiplex PCR Luminex POST methods.

Concordance and Kappa coefficients are also shown (n=139).

Method Employed		**Multiplex PCR Luminex POST**	
SPF10 PCR Gel	***HPV STATUS***	*HPV POSITIVE*	*HPV NEGATIVE*
Electrophoresis	*HPV POSITIVE*	30	25
	HPV NEGATIVE	26	58
Concordance	63.31% (88/139)		
Kappa Coefficient	0.235		

The HPV DNA status concordance analyzed in Table 5.2 is presented graphically in Figure 5.2.

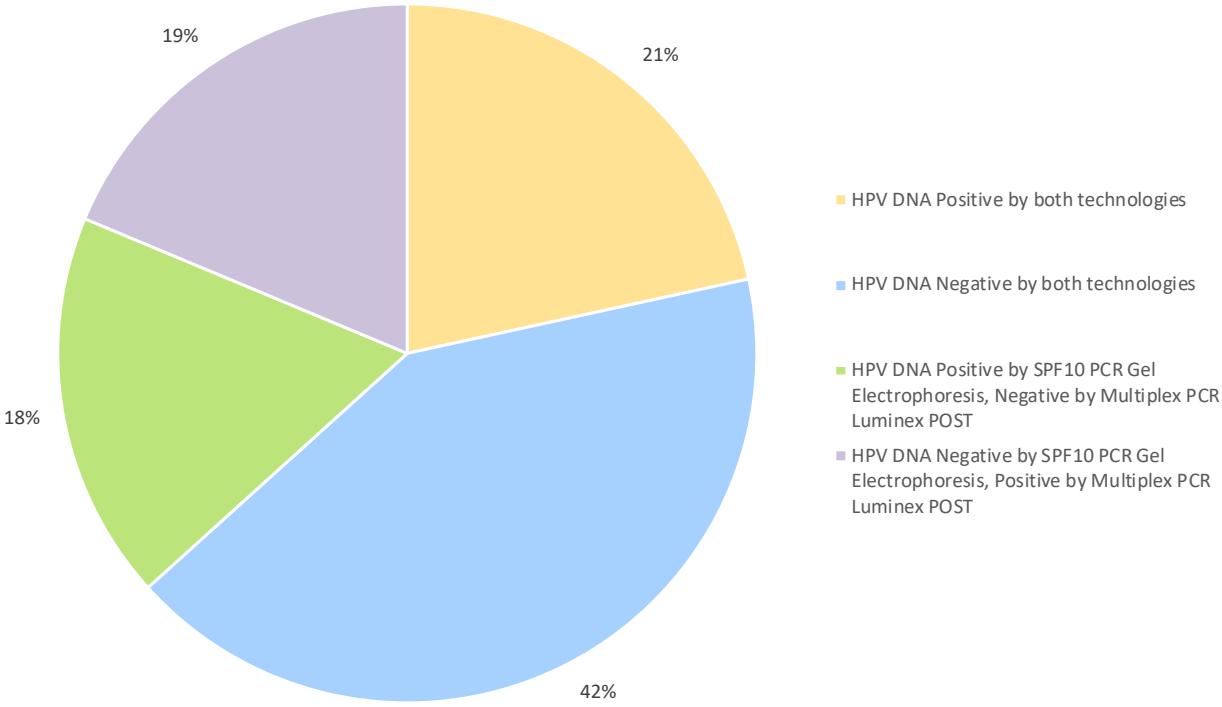

Figure 5.2 HPV DNA status concordance between SPF10 PCR Gel Electrophoresis and Multiplex PCR Luminex POST (n=139).

5.6 Discussion

The aim of this pilot study was to compare HPV DNA detection between SPF10 PCR Gel Electrophoresis, and Multiplex PCR paired with xMAP® Luminex laser-based technology. On this basis, the study aimed to determine the effect of altering sterility protocols on the validity of results and thus identify which method is best suited for the epidemiological endeavors of the ECHO study as a whole.

To begin, in developed countries, HPV DNA is detected in 18.5% to 90% of HNSCC[44]. HPV DNA prevalence, as determined by SPF10 PCR Gel Electrophoresis, Multiplex PCR Luminex

PRE, and Multiplex PCR Luminex POST were in agreement with this range. However, it is clear that the prevalence generated by the Multiplex PCR Luminex PRE was significantly higher than its counterparts. Prevalence was almost 80% using this method compared to the prevalence of approximately 40% detected by the SPF10 PCR Gel Electrophoresis and Multiplex PCR Luminex POST.

Contamination in the Multiplex PCR Luminex PRE using the "Original Sterility Protocol" is evidently the explanation for this large difference. Negative controls were detected as HPV DNA positive. It should be emphasized that the SPF10 PCR Gel Electrophoresis sections were cut sequentially at the same time as those for the Multiplex PCR Luminex PRE. This suggests that similar contamination evidence should have been evident using the SPF10 PCR Gel Electrophoresis. Given that the SPF10 PCR amplifies 54 different HPV genotypes[9,45–47] of which the Multiplex PCR Luminex detects a subset[10,11], it is unlikely that the scope of the Multiplex's genotype detection played a role in these discrepancies. Furthermore, because all controls were as expected using SPF10 PCR Gel Electrophoresis, the results established an unequivocal disparity in the sensitivities of the two technologies. The original cleaning protocol (Richard Allen™ Parapel™ and IMS) clearly did not adequately sterilize instruments for the Multiplex PCR Luminex technology, resulting in evident contamination picked up by the highly sensitive platform.

The higher sensitivity of the Multiplex PCR Luminex technology is noted by those who developed it[11,48]. It is one of the most sensitive HPV DNA detection technologies in existence and its lowest limit of detection is 10 to 20 copies less than that of SPF10 PCR[10,49]. Indeed, a difference of 20 copies may be enough to cease the amplification of enough detectable HPV DNA during PCR. Even more probable is that the diverging ways in which both technologies present a positive result affects their sensitivities. Specifically, the Multiplex system's ability to detect copy-number by the individual fluorescence of colour-coded beads as opposed to the end-point detection of the SPF10 PCR by gel electrophoresis is likely to blame. The SPF10 PCR's reliance on gel electrophoresis for the visualization of ample PCR product may render evidence of even slight contamination or positivity difficult to perceive.

The initial microtome sterilization protocol for the SPF10 PCR Gel Electrophoresis and the Multiplex PCR Luminex PRE (Richard Allen™ Parapel™ and IMS) is thus put into question. The use of an anti-wax solution paired with IMS is one that has been used in the literature for many years, specifically for SPF10 PCR Gel Electrophoresis. No studies published have reported problems with contamination evidence in the case of this technology. This is the first time that this protocol has been used with the more sensitive Multiplex PCR Luminex, suggesting that this cleaning method is at the very least not valid when using the more sensitive technology and at most not valid in general.

This said, the lack of contamination evidence with adherence to the standard cleaning protocol used for the SPF10 PCR Gel Electrophoresis makes these results at least comparable to those of the Multiplex PCR Luminex POST, the latter using the standard sterilization method that the IARC uses for the technology.

First, that both the SPF10 PCR and Multiplex PCR Luminex POST methods determined almost identical HPV DNA prevalence in the population, something statistically evidenced in insignificant Chi-square statistics ($p=0.903$) and graphically in Figure 5.1, is promising. However, the similar prevalence detected by the Multiplex Luminex PCR POST and the SPF10 PCR Gel Electrophoresis makes the low concordance between results surprising. This is reflected in the 'poor' Kappa coefficient and agreement shown in Table 5.2 and Figure 5.2. 36.70% of cases were classified differently by the platforms signifying that sensitivity and specificity may vary not only by technology but also by sample.

For instance, 80.77% (21/26) of those cases positive by the Multiplex PCR Luminex POST and negative by the SPF10 PCR were very small biopsies with limited but present tumour tissue. HPV DNA copy number is likely to be influenced by the size of the available tumour. Thus, SPF10 PCR's reliance on the visualization of copy numbers that may fall outside its detection limits may explain this disparity. Those cases positive by SPF10 PCR and negative by Multiplex PCR Luminex POST are more difficult to reconcile with the sensitivities, specificities, and mechanisms of the technologies. This may be due however to a carryover of the now questionable sterilization protocol used for the SPF10 PCR.

The impact of these observations is significant. Given that both technologies detected similar HPV DNA prevalence, the comparability of results in the context of epidemiological studies may be justifiable. The detail of which particular cases are deemed positive does not appear to have a significant effect on the overarching estimate of the presence of the virus in the population. Nonetheless, the enormous range of HPV DNA prevalence in the literature may indicate that this agreement is not replicable using other technologies. Additionally, the evident lack of concordance between the technologies jeopardizes the validity of comparing them. This is especially true if sample characteristics including size of available tumour tissue have any bearing on the ability of a technology to detect HPV DNA accurately and precisely. This has even larger implications for other HPV-detecting methods. If the comparability of two platforms developed to detect the same HPV indicator is questionable, even more uncertainty arises when equating results between platforms detecting different indicators (e.g. mRNA, viral load, viral integration).

Comparability and reliability of methods of HPV detection is especially crucial for head and neck cancers if conclusions at the epidemiological level are to have implications in the public health and clinical contexts. Australia, Austria, Brazil, Canada, Germany, Israel, New Zealand, Norway and the United States already offer gender-neutral HPV vaccination program as a result of emerging data relating HPV DNA to HNSCCs and other ano-genital cancers[50]. Others are awaiting more data to do the same. In Ireland, health authorities have already recommended the school vaccination program's extension to include boys, something the Taoiseach committed to implementing in 2018[38]. Thus, it is of the utmost importance to generate reliable epidemiological data in Ireland to justifiably expand the implementation of the HPV vaccine containing the most relevant genotypes for non-cervical HPV-related cancers. The development of targeted head and neck screening tools may also look to HPV screening for an efficient and effective way to detect HPV-related HNSCC at early stages. Furthermore, the opportunity that the better survival of HPV-related HNSCCs represents for further increasing life expectancy and minimizing morbidity with de-escalation renders the reliable detection of HPV DNA or otherwise the difference between life and death for patients.

The type of technology needed to optimally detect HPV DNA in the epidemiological and clinical contexts may not be the same given variation in tissue types tested (e.g. FFPE, fresh, frozen). It is certain however that the most sensitive technology available for any of these tissues needs to be implemented for the best preventative and curative outcomes for patients. Missed cases of HPV DNA positivity in HNSCC due to lack of sensitivity would lead to misrepresentative epidemiological indicators and potentially squandered de-escalation opportunities for patients. The cleanliness and sterilization protocols for these technologies need to be extremely reliable as well if the high sensitivities are to be valid.

The Multiplex PCR Luminex POST results were generated in line with WHO protocols developed by the Infections and Cancer Biology Laboratory in IARC, Lyon, France[5,10,11]. This protocol was validated and optimized for the technology and its high sensitivity. Results using the sterilization technique with the Multiplex PCR Luminex have been reliably replicated across HNSCC populations from Europe, South Asia, and East Asia[10,51]. That the Multiplex PCR Luminex POST results were not suspicious for contamination using the highly sensitive platform suggests the success of the adjustment of the sterility protocol to that effectively used many times over by the WHO. Combined with its comparison to two other HPV DNA detection methodologies in this pilot study, this confirms that the Multiplex Luminex PCR, using the sanctioned DNA Zap- and ethanol-based sterility protocol ("IARC Sterility"), is valid for the epidemiological requirements and clinical implications of the ECHO study as a whole.

That is not to say that HPV DNA analysis alone represents the final link between the virus and an associated tumour. The identification of specifically carcinogenic virus in tumour tissue, rather than potentially transient ones additionally picked up by DNA, is extremely important in the clinical context. There is widespread evidence in research and in current practice to suggest that in the clinic, p16 immunohistochemistry status may be a solution to this heterogeneity[33,52–59]. The concurrent detection of p16 and HPV DNA is one strategy of identifying clinically significant HPV infections given that p16 indicates E7 mediated cellular transformation and is considered a marker for oncogenic HPV status[60,61]. HPV mRNA detection in HNSCC, evidence of an integrated and actively transcribing virus, has also been

shown to detect HPV in the majority of the same cases as those positive for both HPV DNA and p16 immunohistochemistry[10,24,62].

In the epidemiological context however, establishing the prevalence of the virus in general in the HNSCC population is best achieved using HPV DNA. This is especially in true in Ireland where this baseline epidemiological data does not yet exist. Thus, on the basis of the results generated in this pilot study, it is clear that the most valid and reliable technology to use for this purpose is the Multiplex PCR Luminex with the validated WHO sterilization protocol outlined in Section 3.6.2. Not only will this yield long-awaited epidemiological indicators for HPV in oropharyngeal, oral cavity, and laryngeal cancer in Ireland, but it will also contribute the data to the WHO's global meta-analysis on the same subject. Named the HPV-AHEAD Study[51], this meta-analysis is particularly important as all of the data will have been obtained using precisely the same protocols and technology, right down to the products used to sterilize implements. It will therefore be the first study of its kind to produce results from around the world that are undoubtedly comparable as a result of procedural standardization.

References

1. Venuti, A. & Paolini, F. HPV detection methods in head and neck cancer. *Head Neck Pathol.* **6 Suppl 1**, S63-74 (2012).
2. Abreu, A. L. P., Souza, R. P., Gimenes, F. & Consolaro, M. E. L. A review of methods for detect human Papillomavirus infection. *Virol. J.* **9**, 262 (2012).
3. Pan American Health Organization. *INTEGRATING HPV TESTING IN CERVICAL CANCER SCREENING PROGRAMS A MANUAL FOR PROGRAM MANAGERS*. (2016).
4. Lagheden, C. *et al.* Validation of a standardized extraction method for formalin-fixed paraffin-embedded tissue samples. *J. Clin. Virol.* **80**, 36–9 (2016).
5. Mena, M. *et al.* Development and validation of a protocol for optimizing the use of paraffin blocks in molecular epidemiological studies: The example from the HPV-AHEAD study. *PLoS One* **12**, e0184520 (2017).
6. Bønløkke, S. *et al.* Evidence of No Association Between Human Papillomavirus and Breast Cancer. *Front. Oncol.* **8**, 209 (2018).

7. Jain, M. R. *et al.* Quantitative Proteomic Analysis of Formalin Fixed Paraffin Embedded Oral HPV Lesions from HIV Patients. *Open Proteomics J.* **1**, 40–45 (2008).
8. Van Doorslaer, K., Chen, Z. & McBride, A. A. Detection and Genotyping of Human Papillomaviruses from Archival Formalin-Fixed Tissue Samples. *Curr. Protoc. Microbiol.* **43**, 14B.9.1-14B.9.20 (2016).
9. Woods, R. Human Papillomavirus-related Oropharyngeal Squamous Cell Carcinoma – Prevalence and Incidence in a Defined Population and Analysis of the Expression of Specific Cellular Biomarkers. (2016).
10. Gheit, T. *et al.* Role of mucosal high-risk human papillomavirus types in head and neck cancers in central India. *Int. J. Cancer* **141**, 143–151 (2017).
11. Gheit, T. *et al.* Development of a Sensitive and Specific Assay Combining Multiplex PCR and DNA Microarray Primer Extension To Detect High-Risk Mucosal Human Papillomavirus Types. *J. Clin. Microbiol.* **44**, 2025–2031 (2006).
12. QIAamp® FFPE DNA Tissue Kit. (2017). Available at: https://www.qiagen.com/us/shop/sample-technologies/dna/genomic-dna/qiaamp-dna-ffpe-tissue-kit/#orderinginformation. (Accessed: 22nd August 2017)
13. Lowe, B., O'Neil, D., Loeffert, D. & Nazarenko, I. Distribution of Human papillomavirus load in clinical specimens. *J. Virol. Methods* **173**, 150–152 (2011).
14. Daigrepont, J., Cameron, J. E., Wright, K. L., Cordell, K. G. & Rosebush, M. S. Detection of human papillomavirus DNA in formalin-fixed, paraffin-embedded squamous papillomas of the oral cavity. *J. Clin. Exp. Dent.* **10**, e979–e983 (2018).
15. D'Souza, G. *et al.* Six-month natural history of oralversus cervical human papillomavirus infection. *Int. J. Cancer* **121**, 143–150 (2007).
16. Kreimer, A. R., Clifford, G. M., Boyle, P. & Franceschi, S. Human Papillomavirus Types in Head and Neck Squamous Cell Carcinomas Worldwide: A Systematic Review. *Cancer Epidemiol. Biomarkers Prev.* **14**, 467–475 (2005).
17. Näsman, A. *et al.* Incidence of human papillomavirus (HPV) positive tonsillar carcinoma in Stockholm, Sweden: An epidemic of viral-induced carcinoma? *Int. J. Cancer* **125**, 362–366 (2009).
18. Hernandez, B. Y. *et al.* Human papillomavirus prevalence in invasive laryngeal cancer in the United States. *PLoS One* **9**, e115931 (2014).
19. Laskaris, S. *et al.* Prevalence of human papillomavirus infection in Greek patients with

squamous cell carcinoma of the larynx. *Anticancer Res.* **34**, 5749–53 (2014).

20. O'Leary, J. J. *et al.* Cervical screening: A new way forward (tests of risk and tests of disease). *HRB Open Res.* **1**, 3 (2018).
21. Heideman, D. A. M. *et al.* Clinical Validation of the cobas 4800 HPV Test for Cervical Screening Purposes. *J. Clin. Microbiol.* **49**, 3983–3985 (2011).
22. Wentzensen, N., Schiffman, M., Palmer, T. & Arbyn, M. Triage of HPV positive women in cervical cancer screening. *J. Clin. Virol.* **76**, S49–S55 (2016).
23. Ebisch, R. M. *et al.* Triage of high-risk HPV positive women in cervical cancer screening. *Expert Rev. Anticancer Ther.* **16**, 1073–1085 (2016).
24. Castellsagué, X. *et al.* HPV Involvement in Head and Neck Cancers: Comprehensive Assessment of Biomarkers in 3680 Patients. *J. Natl. Cancer Inst.* **108**, (2016).
25. Hübbers, C. U. & Akgül, B. HPV and cancer of the oral cavity. *Virulence* **6**, 244–8 (2015).
26. Sanders, E. A Comparison of Clinical Outcomes Between HPV Positive and HPV Negative Squamous Cell Carcinomas of the Oropharynx. *ORL. Head. Neck Nurs.* **34**, 11–4 (2016).
27. de Martel, C., Plummer, M., Vignat, J. & Franceschi, S. Worldwide burden of cancer attributable to HPV by site, country and HPV type. *Int. J. Cancer* **141**, 664–670 (2017).
28. Wierzbicka, M., Szyfter, K., Milecki, P., Składowski, K. & Ramlau, R. The rationale for HPV-related oropharyngeal cancer de-escalation treatment strategies. *Contemp. Oncol. (Poznan, Poland)* **19**, 313–22 (2015).
29. Chaturvedi, A. K. Beyond Cervical Cancer: Burden of Other HPV-Related Cancers Among Men and Women. *J. Adolesc. Heal.* **46**, S20–S26 (2010).
30. Chaturvedi, A. K. Epidemiology and Clinical Aspects of HPV in Head and Neck Cancers. *Head Neck Pathol.* **6**, 16–24 (2012).
31. Chaturvedi, A. K. *et al.* Worldwide Trends in Incidence Rates for Oral Cavity and Oropharyngeal Cancers. *J. Clin. Oncol.* **31**, 4550–4559 (2013).
32. Mirghani, H. & Blanchard, P. Treatment de-escalation for HPV-driven oropharyngeal cancer: Where do we stand? *Clin. Transl. Radiat. Oncol.* **8**, 4–11 (2018).
33. Chen, A. M. *et al.* Reduced-dose radiotherapy for human papillomavirus-associated squamous-cell carcinoma of the oropharynx: a single-arm, phase 2 study. *Lancet Oncol.* **18**, 803–811 (2017).

34. Psyrri, A., Sasaki, C., Vassilakopoulou, M., Dimitriadis, G. & Rampias, T. Future directions in research, treatment and prevention of HPV-related squamous cell carcinoma of the head and neck. *Head Neck Pathol.* **6 Suppl 1**, S121-8 (2012).
35. Eriksen, J. G. & Lassen, P. Human Papilloma Virus as a Biomarker for Personalized Head and Neck Cancer Radiotherapy. *Recent Results Cancer Res.* **198**, 143–161
36. D'Souza, G. & Dempsey, A. The role of HPV in head and neck cancer and review of the HPV vaccine. *Prev. Med. (Baltim).* **53**, S5–S11 (2011).
37. Palefsky, J. M. *et al.* HPV Vaccine against Anal HPV Infection and Anal Intraepithelial Neoplasia. *N. Engl. J. Med.* **365**, 1576–1585 (2011).
38. Health Information and Quality Authority (HIQA). PRESS RELEASE: HIQA advises changing to a more effective HPV vaccine and extending the vaccine to boys. *HIQA News Updates* (2018). Available at: https://www.hiqa.ie/hiqa-news-updates/hiqa-advises-changing-more-effective-hpv-vaccine-and-extending-vaccine-boys. (Accessed: 27th February 2019)
39. Gillison, M. L. Human papillomavirus-related diseases: oropharynx cancers and potential implications for adolescent HPV vaccination. *J. Adolesc. Health* **43**, S52-60 (2008).
40. Brotherton, J. M. L., Giuliano, A. R., Markowitz, L. E., Dunne, E. F. & Ogilvie, G. S. Monitoring the impact of HPV vaccine in males-Considerations and challenges. *Papillomavirus Res. (Amsterdam, Netherlands)* **2**, 106–111 (2016).
41. Guo, T., Eisele, D. W. & Fakhry, C. The potential impact of prophylactic human papillomavirus vaccination on oropharyngeal cancer. *Cancer* **122**, 2313–23 (2016).
42. Pinto, L. A. *et al.* Quadrivalent Human Papillomavirus (HPV) Vaccine Induces HPV-Specific Antibodies in the Oral Cavity: Results From the Mid-Adult Male Vaccine Trial. *J. Infect. Dis.* **214**, 1276–83 (2016).
43. Wang, C. *et al.* Targeting Head and Neck Cancer by Vaccination. *Front. Immunol.* **9**, 830 (2018).
44. Babiker, A. Y. *et al.* Screening for high risk human papilloma virus (HR-HPV) subtypes, among Sudanese patients with oral lesions. *Int. J. Clin. Exp. Med.* **6**, 275–81 (2013).
45. Kleter, B. *et al.* Novel Short-Fragment PCR Assay for Highly Sensitive Broad-Spectrum Detection of Anogenital Human Papillomaviruses. *Am. J. Pathol.* **153**, 1731–1739 (1998).

46. Kleter, B. *et al.* Development and clinical evaluation of a highly sensitive PCR-reverse hybridization line probe assay for detection and identification of anogenital human papillomavirus. *J. Clin. Microbiol.* **37**, 2508–17 (1999).
47. Safaeian, M. *et al.* Comparison of the SPF10-LiPA system to the Hybrid Capture 2 Assay for detection of carcinogenic human papillomavirus genotypes among 5,683 young women in Guanacaste, Costa Rica. *J. Clin. Microbiol.* **45**, 1447–54 (2007).
48. Schmitt, M. *et al.* Abundance of Multiple High-Risk Human Papillomavirus (HPV) Infections Found in Cervical Cells Analyzed by Use of an Ultrasensitive HPV Genotyping Assay. *J. Clin. Microbiol.* **48**, 143–149 (2010).
49. Micalessi, M. I., Boulet, G. A. & Bogers, J. A Real-Time PCR Approach Based on SPF10 Primers and the INNO-LiPA HPV Genotyping Extra Assay for the Detection and Typing of Human Papillomavirus. in *Methods in molecular biology (Clifton, N.J.)* **1249**, 27–35 (2015).
50. Kmietowicz, Z. Boys in England to get HPV vaccine from next year. *BMJ* **362**, k3237 (2018).
51. International Agency for Research on Cancer & World Health Organization. HPV-AHEAD Study. *International Agency for Research on Cancer* (2019). Available at: http://hpv-ahead.iarc.fr/about/index.php. (Accessed: 27th April 2019)
52. Albers, A. E., Qian, X., Kaufmann, A. M. & Coordes, A. Meta analysis: HPV and p16 pattern determines survival in patients with HNSCC and identifies potential new biologic subtype. *Sci. Rep.* **7**, 16715 (2017).
53. Mai, S. *et al.* Prognostic Relevance of HPV Infection and p16 Overexpression in Squamous Cell Anal Cancer. *Int. J. Radiat. Oncol.* **93**, 819–827 (2015).
54. Fonmarty, D. *et al.* Study of the concordance between p16 immunohistochemistry and HPV-PCR genotyping for the viral diagnosis of oropharyngeal squamous cell carcinoma. *Eur. Ann. Otorhinolaryngol. Head Neck Dis.* **132**, 135–9 (2015).
55. Sedghizadeh, P. P. *et al.* Is p16-positive oropharyngeal squamous cell carcinoma associated with favorable prognosis? A systematic review and meta-analysis. *Oral Oncol.* **54**, 15–27 (2016).
56. Albers, A. E., Qian, X., Kaufmann, A. M. & Coordes, A. Meta analysis: HPV and p16 pattern determines survival in patients with HNSCC and identifies potential new biologic subtype. *Sci. Rep.* **7**, 16715 (2017).

57. Brisson, R. J. *et al.* De-escalation in HPV-negative locally advanced head and neck squamous cell cancer (LA-HNSCC) in patients after induction chemotherapy: A retrospective case series. *J. Clin. Oncol.* **36**, e18090–e18090 (2018).

58. Marur, S. *et al.* E1308: Phase II Trial of Induction Chemotherapy Followed by Reduced-Dose Radiation and Weekly Cetuximab in Patients With HPV-Associated Resectable Squamous Cell Carcinoma of the Oropharynx— ECOG-ACRIN Cancer Research Group. *J. Clin. Oncol.* **35**, 490–497 (2017).

59. Tong, F. *et al.* Prevalence and Prognostic Significance of HPV in Laryngeal Squamous Cell Carcinoma in Northeast China. *Cell. Physiol. Biochem.* **49**, 206–216 (2018).

60. Wang, H., Sun, R., Lin, H. & Hu, W. P16 ^{INK4A} as a surrogate biomarker for human papillomavirus-associated oropharyngeal carcinoma: Consideration of some aspects. *Cancer Sci.* **104**, 1553–1559 (2013).

61. Stephen, J. K. *et al.* Significance of p16 in site-specific HPV positive and HPV negative HNSCC. *Cancer Clin. Oncol.* **2**, 51–61 (2012).

62. Ndiaye, C. *et al.* HPV DNA, E6/E7 mRNA, and p16INK4a detection in head and neck cancers: a systematic review and meta-analysis. *Lancet Oncol.* **15**, 1319–1331 (2014).

CHAPTER 6

CHARACTERISATION OF THE STUDY POPULATION

6 CHAPTER 6: CHARACTERISATION OF THE STUDY POPULATION

6.1 Introduction

The population for the ECHO study was created through the NCRI's database, as described in Sections 3.1, 3.2, and 3.3. Ethical approval for use of archival tissue specimens was obtained from all relevant local hospital ethics committees detailed in Section 3.3.3.

The population comprised of 861 cases of newly-diagnosed (between 1994 and 2013), primary, invasive, oral (oropharyngeal, oral cavity, and largyngeal) SCC retrieved from 7 different hospitals sites around Ireland including St. James' University Hospital (including the Royal Eye and Ear University Hospital), the Mater Misericordiae University Hospital, Beaumont University Hospital, St. Vincent's University Hospital, Kerry General University Hospital, and University Hospital Limerick. The ICD10 codes included in the study can be found in Section 3.3.1.

On the basis of these codes, their groupings recommended by the WHO, and those present in the study population, three generalized classifications of oropharyngeal, oral cavity, and laryngeal sub-site emerged. These are summarized in Table 6.1.

Table 6.1 Summary of ICD10 codes represented in the study population and the classification under which they were placed for the analysis.

ICD10 Codes	Classification
1.0, 2.4, 5.1, 5.2, 9.0-9.9, 10.0-10.9, 14.2	*Oropharynx*
2.0, 2.1, 2.2, 2.3, 2.8, 2.9, 3.0-3.9, 4.0-4.9, 5.0, 5.8, 5.9, 6.0, 6.1, 6.2, 6.9, 14.0, 14.8	*Oral Cavity*
32.0-32.9	*Larynx*

Clinical information regarding all cases included in the study (n=861) was provided by the NCRI using an anonymous study ID number linking the ECHO study database to the national database. Available patient and tumour characteristics in addition to sub-site information are presented in Table 6.2 below. Some adjustments to variables were made for the benefit of the analysis throughout the entire study and these are also detailed in the table.

Table 6.2 Variables made available by the NCRI for the population of the ECHO study and notes on any adjustments made for the purposes of the analysis.

Variable Code	Meaning	Variable Definition	Notes on Adjustments
SEX	Sex of patient	Sex of patient	N/A
AGE	Integer age at date of diagnosis	Integer age at date of diagnosis	Age was assessed both continuously and based on age younger than or equal to, and older than, age 50. Only continuous age was brought forward for multivariate analysis where relevant.
SMOKER_ID	Smoking status	Indication of current, ex-, or never-smoked behavior	N/A
GRADE	Grade of primary tumour	Poorly-, moderately, well-, or un-differentiated grade of tumour	Only 2 undifferentiated cases were detected in the population. These were excluded **after** distribution was determined for all grade statistics generated to avoid skew in results.
T5	T stage	T category of stage (5th edition for cases diagnosed up to 2013) derived from best available clinical or pathological T data	Due to low frequencies for sub-stages, these were combined to yield the following T stage categories: T1, T2, T3, T4.
N5	N Stage	N category of stage (5th edition for cases diagnosed up to 2013) derived from best available clinical or pathological N data	Due to low frequencies for sub-stages, these were combined to yield the following N stage categories: N0, N1, N2, N3. N2 and N3 were also combined due to extremely low numbers of N3 patients.
M5	M Stage	M category of stage (5th edition for cases diagnosed up to 2013) derived from best available clinical or pathological M data	N/A
TNM5	TNM Stage	TNM stage (5th edition for cases diagnosed up to 2013) derived from best available clinical or pathological data	Due to low frequencies for sub-stages of Stage IV, TNM stages were combined to yield the following categories: Stage I, II, III, IV.
COUNTY_RES	County of residence	County of residence of patient at time of diagnosis	Due to low frequencies for many counties, county was assessed based on both residence in counties with large urban centers

			(Dublin/Limerick/Cork) and residence in or outside Dublin.
DEPRIV_POBAL_2011	Socio-economic status/Social Deprivation Score	Pobal index of deprivation from 1 to 5 for 2011 patient's Electoral Division (ED) of residence at diagnosis re-expressed as quintiles of 2011 population	Social deprivation score was categorical on a scale of 1 to 5, with 5 being the most deprived. It was assessed both categorically and as a continuous variable.
MARITAL	Marital status	Indication of single, separated, widowed, or divorced status of patient	Divorced and separated individuals were grouped together due to similarity in status and low numbers of divorced patients.

Treatment administered and survival data was also provided by the NCRI, variables for which are summarized and detailed in Table 6.3 below.

Table 6.3 Variables relating to treatment type and survival made available by the NCRI for the population of the ECHO study, their meanings, and definitions.

Variable Code	Meaning	Definition
Chemo_1y	Chemotherapy	Binary indication of whether or not patient was treated with chemotherapy targeting the cancer within 1 year of diagnosis.
Radio_1y	Radiotherapy	Binary indication of whether or not patient was treated with radiotherapy targeting the cancer within 1 year of diagnosis.
Surg_1y	Surgery	Binary indication of whether or not patient was treated with surgery targeting the cancer within 1 year of diagnosis
VITAL_STAT	Overall survival	All-cause vital status of patient (0 alive or 1 dead) at common censoring date based mainly on death-certificate matching.
VITAL_CAN	Disease-specific (cancer-specific) survival	Cause-specific vital status (0 alive or died of other cause or different cancer or 1 died from the cancer of interest) at common censoring date.
SURVIVAL_MONTHS	Survival in months	Number of complete months from diagnosis of a specific tumour to common censoring date.

Before delving into analysis based on HPV status, it was necessary to establish the distribution of study population based on all available patient and tumour characteristics, compare these distributions to available national-level aggregate data for key characteristics, assess any correlations between variables, and evaluate survival. Together, these analyses revealed the kind of population being used for further analysis, and its comparability to the overall Irish oropharyngeal, oral cavity, and laryngeal SCC population as a whole. On this basis, the aims of this Chapter are delineated below.

6.2 Aims

- To determine the distribution of the present population (n=861) by patient and tumour characteristics.
- To assess the comparability of the study population to the Irish oropharyngeal, oral cavity, and laryngeal SCC population as a whole using key patient and tumour characteristics.
- To evaluate the correlation between patient and tumour characteristics.
- To determine predictors of overall and cancer-specific survival in this population.

6.3 Materials and Methods

6.3.1 Statistical Analysis

The analysis was carried out using IBM® SPSS® Statistics Version 25, XLSTAT 2019.1.3, Joinpoint Regression 4.7.0, and Microsoft Excel Version 16.25. The distribution of the population was assessed by simple proportions. Raw incidence was represented by moving average, whereby the incidence of cases in a single year was the average of the year prior, the year itself, and the year following (e.g. incidence in 1995=((1994+1995+1996)/3)). Average annual percentage change in cases diagnosed over the 1994 to 2013 period was assessed using Joinpoint Regression.

Variables for which more than 10% of data was missing were assessed as stated in Table 6.2 but with an additional “unknown” category representing missing data to detect any non-random patterns in missing data as is conventional in the epidemiological literature. For variables with 0% to 10% of cases with missing data, missing cases were eliminated for all

relevant analysis. TNM stage was populated by the NCRI itself using standard registry-based assumptions for unknown N and M stage, classifying these as N0 and M0, a method that has proven to accurately reflect cumulative TNM stage. Only patients who were treated within 12 months of diagnosis were included in any analysis involving treatment.

Association tests for discrete variables were performed using Chi-square and Fisher exact tests (where expected counts were less than 5) for independence. For continuous variables, T-tests and one-way ANOVAs or, if normality was violated, Mann-Whitney U and Kruskal-Wallis tests were performed. To assess the comparability of the sample to the national oropharyngeal, oral cavity, and laryngeal SCC population, Chi-square tests were performed comparing the study population to those patients eligible but not included in the study (e.g. the overall population minus the sample population).

The relationship between different variables and overall and cancer-specific survival was determined using Kaplan-Meier analysis by log-rank test. Univariate cox proportional hazard analysis, including HR calculation, was carried out to determine patient and tumour factors that were significantly predictive of survival. Multivariate cox proportional hazard models using varying combinations and sequences of variables was also employed to assess confounding variables and to further identify significant predictors of overall and cancer-specific survival. All significant variables by univariate models were included in the initial multivariate model. The least significant predictor was then taken out, and the model was run again. The least significant predictor was again taken out, and the model was run again. This continued until all variables remaining in the model proved significantly predictive of survival and risk of death, or until taking another variable out rendered the model as a whole insignificant.

It should also be noted that where age was assessed by univariate analysis in both the continuous form and by categorical variable (Age$\leq$50), it was only used in its continuous form in multivariate analysis. A significance level of $\leq$0.05 was used for all tests which were also all two-sided.

6.4 Distribution of Patient, Tumour, and Treatment Characteristics

6.4.1 Patient and Tumour Characteristics

To begin, the distributions and associated means and medians of the population by patient characteristics listed in Table 6.2 were assessed and results are presented in Table 6.4 below. Cases for which no data was available or missing is labelled as "unknown".

Table 6.4 Patient and tumour characteristics of the ECHO study population (n=861).

Distribution of the population within each characteristic along with means and medians are presented where appropriate.

Variable/Characteristic	Sub-set of Variable	Proportion/Mean/Median
Sex	*Male*	(661/861)=76.8%
	Female	(200/861)=23.2%
Age (Continuous)		Mean=63.30 (CI: 62.52, 64.08) Median=63.00
Age≤50	*≤50*	(121/861)=14.1%
	>50	(740/861)=86.9%
Smoking Status	*Current smoker*	(479/861)=55.6%
	Ex-smoker	(110/861)=12.8%
	Never-smoked	(156/861)=18.1%
	Unknown	(116/861)=13.5%
Sub-site	*Oropharynx*	(209/861)=24.3%
	Oral Cavity	(331/861)=38.4%
	Larynx	(321/861)=37.3%
Social Deprivation Score	*All known*	Mean=3.39(CI: 3.29, 3.49) Median=4.00
	1	(120/861)=13.9%
	2	(122/861)=14.2%
	3	(128/861)=14.9%
	4	(178/861)=20.7%
	5	(246/861)=28.6%
	Unknown	(67/861)=7.8%

County of Residence	*Urban (Dublin/Limerick/Cork)*	(513/861)=59.6%
	Rural (all other counties)	(348/861)=40.4%
	Dublin only	(380/861)=44.1%
	All other counties	(481/861)=55.9%
Marital Status	*Divorced/Separated*	(77/861)=8.9%
	Married	(434/861)=50.4%
	Single	(202/861)=23.5%
	Widowed	(117/861)=13.6%
	Unknown	(31/861)=3.6%
Grade	*Well-differentiated*	(88/861)=10.2%
	Moderately differentiated	(475/861)=55.2%
	Poorly-differentiated	(187/861)=21.7%
	Un-differentiated	(2/861)=0.2%
	Unknown	(109/861)=12.7%
T Stage	*T1*	(151/861)=17.5%
	T2	(234/861)=27.2%
	T3	(137/861)=15.9%
	T4	(210/861)=24.4%
	Unknown	(129/861)=15.0%
N Stage	*N0*	(341/861)=39.6%
	N1	(131/861)=15.2%
	N2	(233/861)=27.1%
	N3	(13/861)=1.5%
	Unknown	(143/861)=16.6%
M Stage	*M0*	(425/861)=49.4%
	M1	(46/861)=5.3%
	Unknown	(390/861)=45.3%
TNM Stage	*Stage I*	(119/861)=13.8%
	Stage II	(126/861)=14.6%
	Stage III	(133/861)=15.4%

	Stage IV	(376/861)=43.8%
	Unknown	(107/861)=12.4%

Figures 6.1 through 6.3 showcase the results in Table 6.4 for particularly pertinent variables including sub-site, smoking status, and TNM stage.

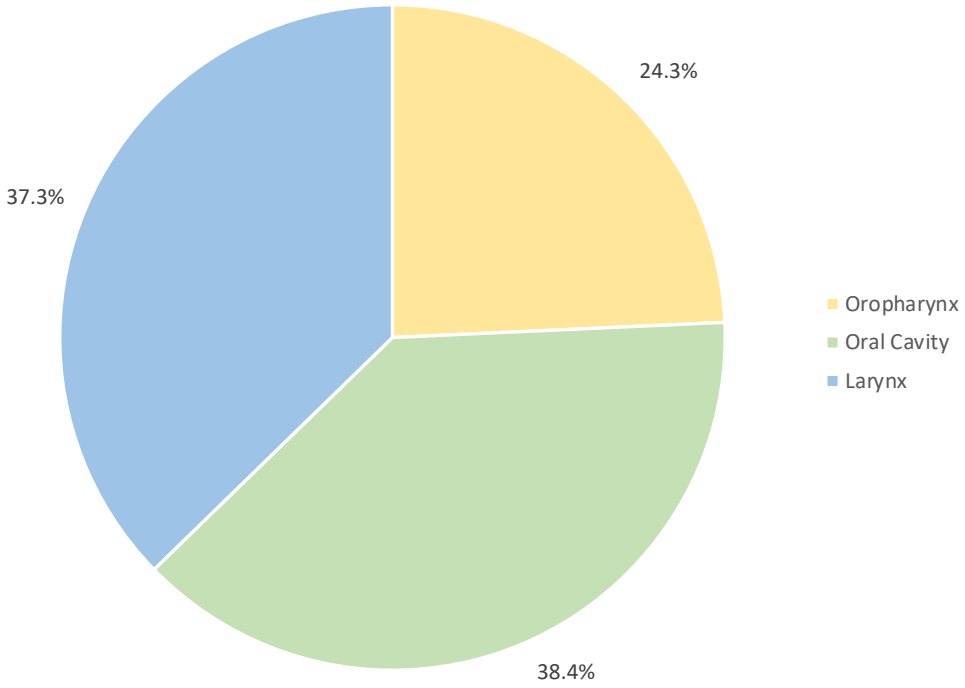

Figure 6.1 Distribution of the population by sub-site (n=861).

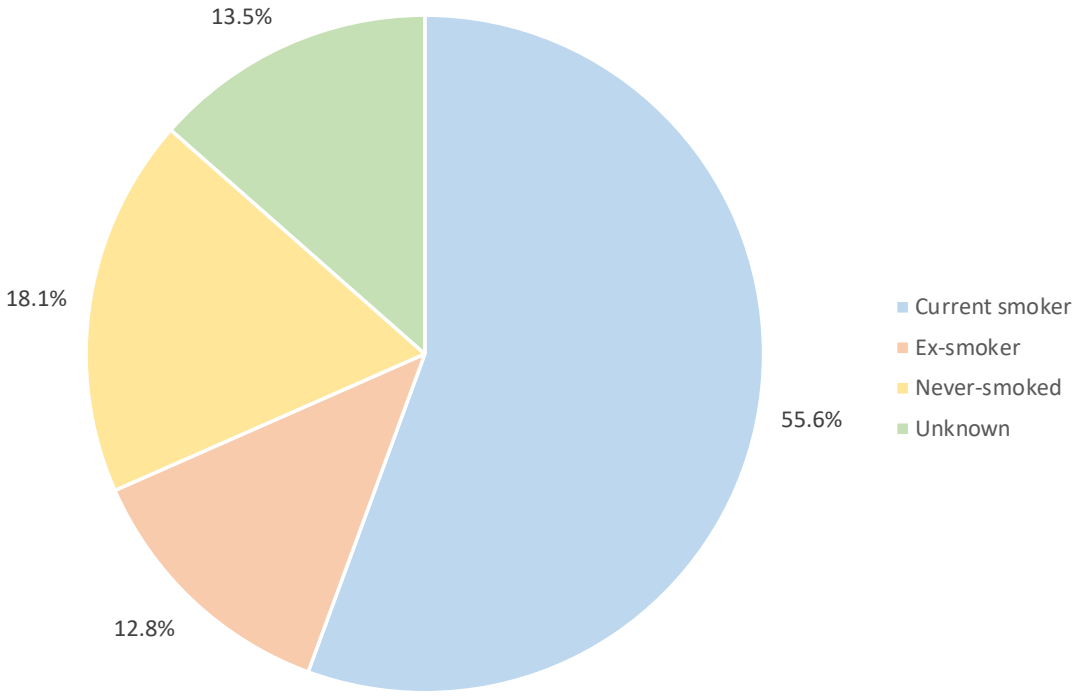

Figure 6.2 Distribution of the population by smoking status (n=861).

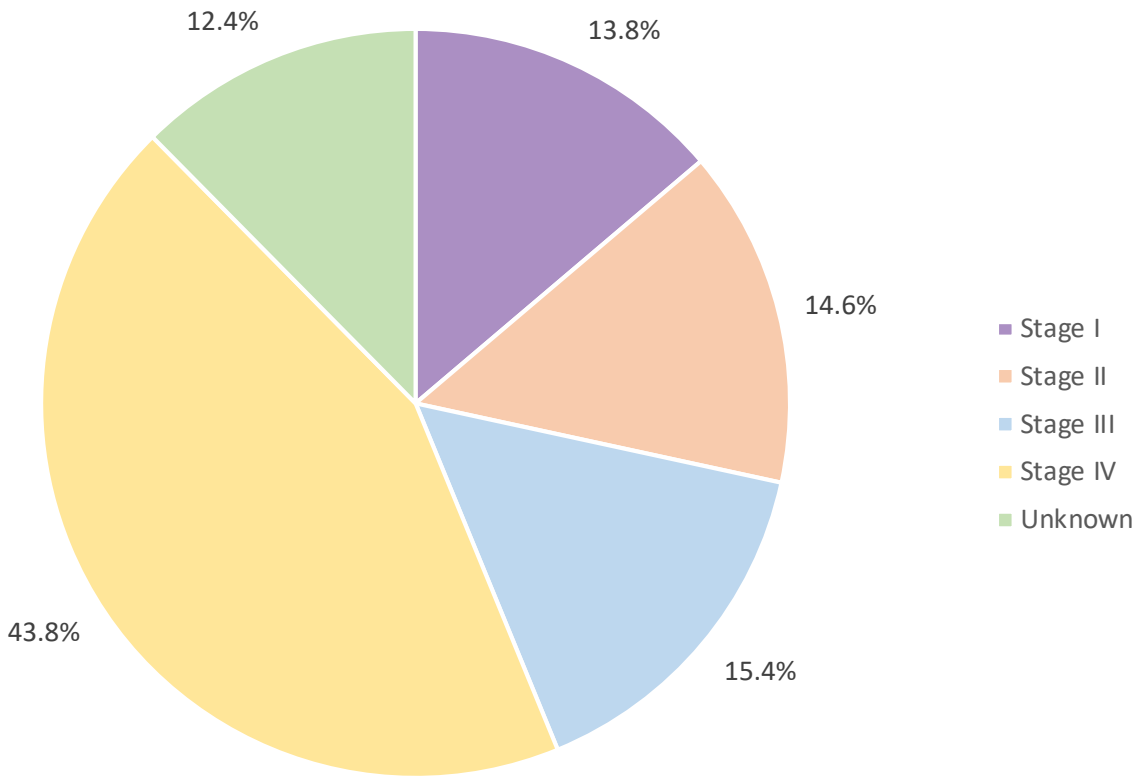

Figure 6.3 Distribution of the population by TNM stage (n=861).

6.4.2 Treatments Administered

The distribution of treatments administered within 12 months of diagnosis (n=758) were then analyzed. Results are presented in Table 6.5 below.

Table 6.5 Treatment administered to patients included in the ECHO study within 12 months of diagnosis (n=758).

Treatment Type	Proportion of Patients
Chemotherapy	(4/758)=0.5%
Radiotherapy	(238/758)=31.4%
Surgery	(133/758)=17.5%
Radiotherapy/Chemotherapy	(119/758)=15.7%
Surgery/Chemotherapy	(2/758)=0.3%
Surgery/Radiotherapy	(183/758)=24.1%
Surgery/Radiotherapy/Chemotherapy	(79/758)=10.4%

Figure 6.4 represents the distribution of treatment types administered to patients in the ECHO study population graphically.

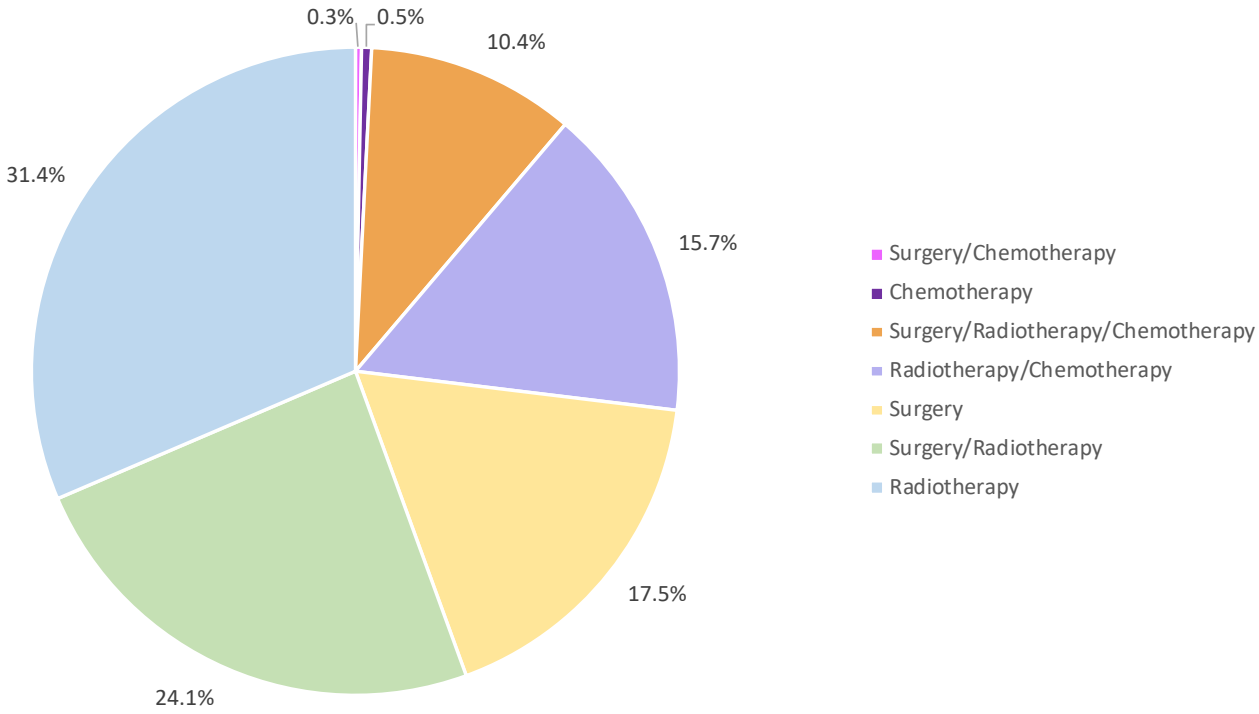

Figure 6.4 Proportion of patients treated using various combinations of radiotherapy, surgery, and chemotherapy (n=758).

6.4.3 Raw Incidence

The distribution of all 861 cases by year of diagnosis and individual sub-site is shown in Figure 6.5 below. The incidence in the Figure is represented by a moving or rolling average due to the discrete nature of the data and in some instances the small sample sizes as is standard in the statistical literature. This moving average is simply the average of three years represented by the central year (e.g. the number of cases diagnosed in 1995 is the average of those diagnosed in 1994, 1995, and 1996).

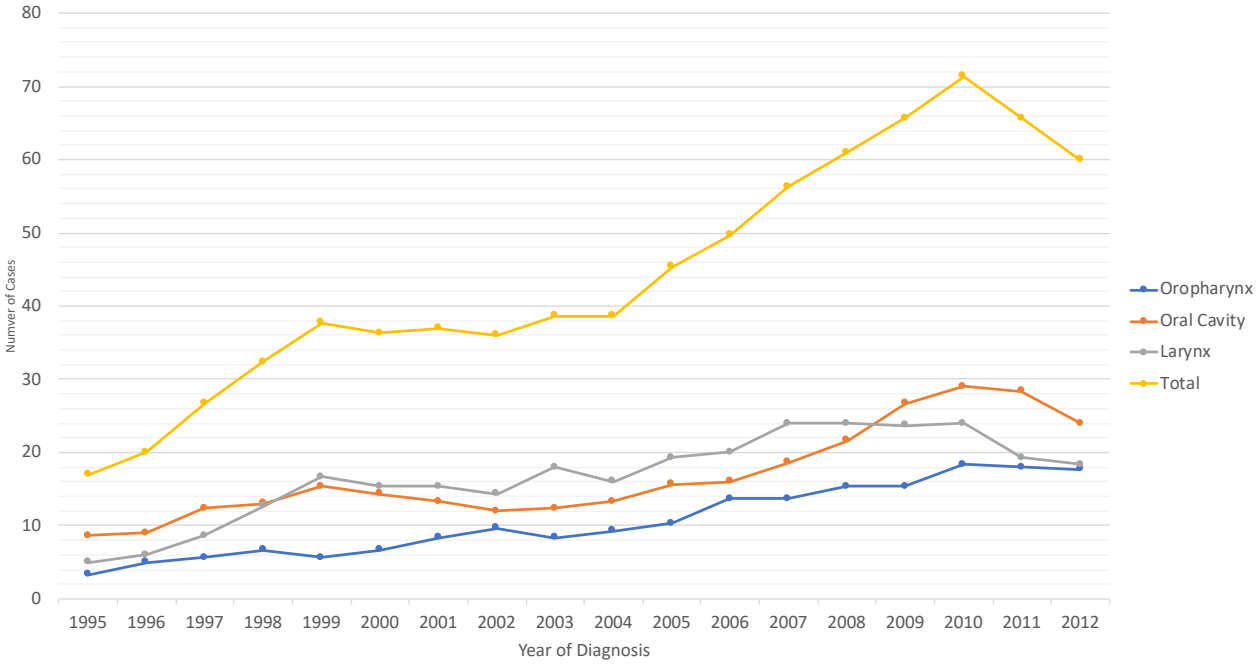

Figure 6.5 Raw incidence of total oropharyngeal, oral cavity, and laryngeal SCC diagnosed in Ireland between 1994 and 2013 represented by moving average further broken down by sub-site of origin.

For oropharynx n=209, oral cavity n=331, and larynx n=321. For the total population n=861.

The average annual percentage change for oropharyngeal cancers was 9.4 (CI: 5.6, 13.4)($p<0.0001$); for oral cavity cancers it was 6.6 (CI: 4.3, 8.9)($p<0.0001$); for laryngeal cancers it was 7.3 (CI: 3.6, 11.2)($p<0.0001$); and for all cases it was 4.9 (CI: -1.3, 11.4) ($p=0.100$).

6.5 Comparability of the Study Population to the National Oropharyngeal, Oral Cavity, and Laryngeal SCC Population

To assess comparability of the present study population (n=861) to the aggregated data, a Chi-square test of selection bias was conducted comparing the sample population to those samples eligible for the study but not included (e.g. the total available population minus the sample population).

Aggregated data was available from the NCRI for all patient and tumour characteristics included in the study. All variables were assessed as described in Tables 6.1, 6.2, and 6.3.

However, since aggregate data for age-group at diagnosis was only available in 5-year ranges with youngest patients classified <50 and oldest patients classified as 80+, it was assessed in this way. Year of diagnosis was similarly analyzed using 5- and 10-year (1994-2003 vs. 2004-2013) groups between 1994 and 2013. Treatment aggregate data was available in a form indicating whether or not patients had received surgery, radiotherapy, or chemotherapy in any capacity, regardless of combination with other treatment. Study population data was manipulated to reflect the same distribution of treatment by simply adding all cases having received each treatment in any form together. Unknown cases for all variables were included as a category to assess whether or not the dataset was disproportionately affected by missing data in comparison to cases not included. Only patients treated within 12 months since diagnosis were analyzed. Results of these two forms of representative analysis are summarized in Table 6.6.

Table 6.6 Representative nature and selection bias of sample population using Chi-square analyses comparing differences between the sample and those patients eligible but not included, by each patient and tumour characteristic.

Notes explaining any incomparability are also included.

Patient/Tumour Characteristic	Result	Notes
Year of Diagnosis (n=861)	*5-year groups*: Chi-square=34.528, 3 d.f., p<0.0001 *10-year groups*: Chi-square=24.202, 1 d.f., p<0.0001	Significance emanated from slight under-sampling between 1994 and 2003.
Age (n=861)	*5-year groups*: Chi-square = 10.032, 7 d.f., p=0.187	N/A
Sex (n=861)	Chi-square=1.599, 1. d.f., p=0.206	N/A
Sub-site (n=861)	Chi-square=22.1761, 2 d.f., p=0.000015	Significance emanated from over-sampling of the oropharyngeal site and under-sampling of the oral cavity site.
Grade (n=861)	Chi-square=3.677, 3 d.f., p=0.298	N/A
Smoking Status (n=861)	Chi-square=30.335, 3 d.f., p<0.0001	Significance emanated from slight under-sampling of ex-smokers and over-sampling of never-smokers.
T Stage (n=861)	Chi-square=13.552, 4 d.f., p=0.009	Significance emanated from slight under-sampling of Stage T1 and over-sampling of T3.
N Stage (n=861)	Chi-square=15.852, 4 d.f., p=0.003	Significance emanated from slight over-sampling in the N2 nodal

		category and slight under sampling in the N0 category.
M Stage (n=861)	Chi-square=0.380, 2 d.f., p=0.827	N/A
TNM Stage (n=861)	Chi-square=12.837, 4 d.f., p=0.012	Significance emanated from slight under-sampling of Stage I and over-sampling of Stage IV.
Social Deprivation Score (n=861)	Chi-square=4.642, 5 d.f., p=0.461	N/A
County of Residence (n=861)	*Dublin/Limerick/Cork vs all:* Chi-square=45.199, 1 d.f., p<0.0001 *Dublin vs all:* 16.426, 1 d.f., p<0.0001	Significance resulted from slight oversampling from both larger urban centers and Dublin alone, and thus under-sampling from non-urban non-Dublin counties.
Marital Status (n=861)	Chi-square=5.227, 4 d.f., p=0.265	N/A
Treatment (n=758)	Chi-square=12.434, 2 d.f., p=0.002	Significance resulted from slight over-sampling of patients who received chemotherapy in any capacity and under-sampling of patients who received surgery in any capacity.

It should be noted that results from Table 6.6 indicate that the study did not disproportionately suffer from missing patient data in the registry database. Any selection bias evident was due to over-sampling or under-sampling of known patient data.

6.6 Correlation Between Patient and Tumour Characteristics

Relevant association and continuous comparison tests were conducted to compare every available patient and tumour characteristic to the others. Social deprivation was evaluated in its original categorical form. Age was evaluated as a continuous variable. As previously highlighted, variables for which more than 10% of cases had missing data included a category called “unknown” to assess any non-random patterns in missing data. Given the extremely low numbers of patients treated using chemotherapy and surgery/chemotherapy, these 6 cases were excluded for treatment association and mean tests. The same was true of the 2 cases alone that were grade undifferentiated. Table 6.7 summarizes the results of all of the tests conducted.

Table 6.7 Correlation between patient and tumour characteristics.

Relevant association and continuous assessments were employed to compare each variable to the other. Urban/Rural refers to urban centers (Dublin/Limerick/Cork) vs all other counties. Treatment refers to treatment type/administered to patients. χ^2 refers to Chi-square, T refers to T-test, F refers to ANOVA, K-W refers to Kruskal-Wallis, M-W refers to Mann-Whitney, and Fisher's refers to Fisher's exact.

Variable	Age	Sub-site	Smoking Status	Grade	T Stage	N Stage	M Stage	TNM Stage	Urban /Rural	Dublin /Other	Social Deprivation	Marital Status	Treatment
Sex	T=0.194 859 d.f. P=0.846	**χ^2=30.455 2 d.f. P<0.0001**	**χ^2=31.274 3 d.f. P<0.0001**	χ^2=2.451 3 d.f. P=0.484	χ^2=4.613 4 d.f. P=0.329	χ^2=2.852 3 d.f. P=0.415	χ^2=1.760 2 d.f. P=0.415	χ^2=3.054 4 d.f. P=0.549	χ^2=0.633 1 d.f. P=0.426	χ^2=1.197 1 d.f. P=0.274	χ^2=5.206 4 d.f. P=0.267	**χ^2=33.758 3 d.f. P<0.0001**	χ^2=4.423 4 d.f. P=0.352
Age	X	**F=26.784 2 d.f. P<0.0001**	**K-W P<0.0001**	F=1.426 3 d.f. P=0.234	**K-W P=0.011**	**F=16.364 3 d.f. P<0.0001**	**F=4.428 2 d.f. P=0.012**	**F=8.113 4 d.f. P<0.0001**	T=-1.944 869 d.f. P=0.052	T=-1.660 859 d.f. P=0.097	**K-W P=0.026**	**K-W P<0.0001**	**K-W P<0.0001**
Sub-site	X	X	**χ^2=21.273 6 d.f. P=0.002**	**χ^2=38.795 6 d.f. P<0.0001**	**χ^2=24.800 8 d.f. P=0.002**	**χ^2=91.915 6 d.f. P<0.0001**	**χ^2=10.399 4 d.f. P=0.034**	**χ^2=56.965 8 d.f. P<0.0001**	χ^2=0.839 2 d.f. P=0.657	χ^2=0.583 2 d.f. P=0.747	χ^2=5.958 d.f. p=0.652	**χ^2=17.916 6 d.f. P=0.006**	**χ^2=136.788 8 d.f. p<0.0001**
Smoking Status	X	X	X	**χ^2=11.569 9 d.f. P=0.239**	**χ^2=32.666 12 d.f. P=0.001**	**χ^2=35.884 9 d.f. P<0.001**	χ^2=10.055 6 d.f. P=0.122	**χ^2=24.225 12 d.f. P=0.019**	χ^2=0.832 3 d.f. P=0.842	**χ^2=10.424 3 d.f. P=0.015**	**χ^2=27.745 12 d.f. P=0.006**	**χ^2=35.540 9 d.f. P<0.001**	**χ^2=43.504 12 d.f. P<0.001**
Grade	X	X	X	**X**	**χ^2=53.645 12 d.f. P<0.0001**	**χ^2=44.528 9 d.f. P<0.0001**	**Fisher's=15 .442 P=0.015**	**χ^2=52.844 12 d.f. P<0.0001**	χ^2=1.769 3 d.f. P=0.622	**χ^2=8.636 3 d.f. P=0.035**	χ^2=14.479 12 d.f. p=0.271	**χ^2=17.446 9 d.f. P=0.042**	χ^2=20.643 12 d.f. P=0.056
T Stage	X	X	X	X	X	**χ^2=262.523 12 d.f. P<0.0001**	**χ^2=47.943 8 d.f. P<0.0001**	**χ^2=1938.749 16 d.f. P<0.0001**	χ^2=1.199 4 d.f. P=0.878	**χ^2=14.268 4 d.f. P=0.006**	χ^2=18.044 16 d.f. P=0.321	**χ^2=32.006 12 d.f. P=0.001**	**χ^2=82.597 16 d.f. p<0.0001**
N Stage	X	X	X	X	X	X	**χ^2=113.511 6 d.f. P<0.0001**	**χ^2=847.112 12 d.f. P<0.0001**	χ^2=2.968 3 d.f. P=0.397	χ^2=1.917 3 d.f. P=0.590	χ^2=9.973]12 d.f. P=0.618	**χ^2=24.490 9 d.f. P=0.001**	**χ^2=84.620 12 d.f. P<0.0001**

M Stage	X	X	X	X	X	X	X	**χ^2=98.012 8 d.f. P<0.0001**	χ^2=2.342 2 d.f. P=0.310	χ^2=2.368 2 d.f. P=0.306	χ^2=12.010 8 d.f. P=0.151	**Fisher's=16.385 P=0.010**	**χ^2=24.256 8 d.f. P=0.002**
TNM Stage	X	X	X	X	X	X	X	X	χ^2=4.646 4 d.f. P=0.326	**χ^2=13.438 4 d.f. P=0.009**	χ^2=12.763 16 d.f. P=0.690	**χ^2=40.552 12 d.f. P<0.001**	**χ^2=106.972 16 d.f. P<0.0001**
Urban / Rural	X	X	X	X	X	X	X	X	X	**χ^2=461.428 1 d.f. P<0.0001**	**χ^2=43.144 4 d.f. P<0.0001**	χ^2=1.449 3 d.f. P=0.694	**χ^2=13.374 4 d.f. P=0.010**
Dublin / Other	X	X	X	X	X	X	X	X	X	X	**χ^2=45.554 4 d.f. P<0.0001**	χ^2=4.609 3 d.f. P=0.203	χ^2=5.186 4 d.f. P=0.269
Social Deprivation	X	X	X	X	X	X	X	X	X	X	X	χ^2=13.920 12 d.f. P=0.306	χ^2=15.208 16 d.f. P=0.509
Marital Status	X	X	X	X	X	X	X	X	X	X	X	X	χ^2=12.266 12 d.f. P=0.425

A closer look at those key relationships significant in Table 6.7 are elaborated upon in Table 6.8, detailing precisely where significance emanated from. Several Figures (6.6 through 6.10) were also generated to graphically highlight some of these relationships.

Table 6.8 Detailed explanation of key significant relationships presented in Table 6.7.

Key Significant Relationship	Detail of Relationship
Sex v Sub-site	25.9% of male cases were oropharyngeal compared to 19.0% of female cases. 55.0% of female cases occurred in the oral cavity, compared to 33.4% of male cases. Almost twice as many laryngeal cases occurred in men (40.7%) than in women (26.0%). Figure 6.6 showcases the relationship between these two variables.
Sex v Smoking Status	Twice as many women (22.5%) were ex-smokers than men (9.8%). More men were never-smokers than women (19.8% vs 12.5%).
Sex v Marital Status	26.6% of men were single compared to 17.0% of women. Over 80% of single people were men. Women were twice as often widowed as men (26.3% vs. 10.4%).
Age v Sub-site	Oropharyngeal cases presented at the earliest age (58.72, CI:57.35, 60.09), followed by oral cavity cancers (63.48, CI:62.15, 64.82), followed by laryngeal cancers (66.09, CI: 64.88, 67.30). A graphical representation of this is presented in Figure 6.9.
Age v Social Deprivation	Age at diagnosis dropped from 65.63 (CI:63.33, 67.94) by almost exactly a year for every increased unit in deprivation score between scores 1 and 4, with score 4's mean age being 61.47 (CI: 59.72, 63.22). Score five saw an age equivalent to that of score 2, as depicted in Figure 6.10.
Smoking Status v Sub-site	17.8% of oral cavity patients were ex-smokers compared to 11.5% of oropharyngeal patients and 8.4% of laryngeal patients. Figure 6.7 showcases the relationship between these two variables.
Smoking Status v Social Deprivation	Current smokers were half as likely as ex-smokers and never smokers to be in the lowest deprived group (score 1) (11.3% vs. 22.1% for ex-smokers vs. 19.5% for never smokers).
Smoking Status v Marital Status	Never smokers were less likely to be widowed, while current and never smokers were more likely than ex-smokers to be divorced/separated.

	64.5% of never smokers were married compared to between 47% and 54% for current and ex-smokers.
Sub-site v T Stage	10% of oropharyngeal cases presented at T1 compared to twice this rate amongst oral cavity and laryngeal cases.
Sub-site v N Stage	Laryngeal cases were twice as likely as oropharyngeal cases to present at N0. 49.8% of oropharyngeal cases presented at N2/3 compared to 27.5% of oral cavity cases and 15.9% of laryngeal cases.
Sub-site v TNM Stage	4.3% of oropharyngeal cases presented at Stage I in comparison to 15.1% of oral cavity cases and 18.7% of laryngeal cases. 61.2% of oropharyngeal cases presented at Stage IV in comparison to 45.0% of oral cavity cases and only 30.8% of laryngeal cases. Figure 6.8 showcases the relationship between these two variables.
Sub-site v Treatment	Oropharyngeal cases were treated almost twice as often with all three treatment modalities than laryngeal cases (14.3% compared to 8.7%), where oral cavity cancers were treated three times more often with surgery alone (30.9% compared to 9.5% in the oropharynx and 9.5% in the larynx), and laryngeal cancers were treated more than twice as often as oral cavity cancers with only radiotherapy (48.0% compared to 16.7%).
TNM Stage v Treatment	Stage I patients were more likely to be treated with surgery alone and Stage IV patients were treated with all three modalities (surgery/radiotherapy/chemotherapy).
TNM Stage v T Stage	All Stage I cases were T1 stage. The same was true of Stage II and T2. All T4 cases were Stage IV.
Social Deprivation v Geographic Location	Patients scoring either the lowest or highest deprivation scores were more likely to come from urban centers. 70% of score 1 patients resided in Dublin, where 60% of middle-level scores (2-4) came from outside Dublin.
Patterns in Missing Data	Missing T stage patients were older (Mean=66.96 (CI:64.84, 69.08)) than all known T stages (Mean between 62.00 and 63.00). Missing N stage patients were significantly older than other nodal stage patients. The same was true of missing M stage patients and missing TNM stage patients (Mean=68.79 (CI: 66.49, 71.09) compared to between 57.00 and 61.00 for known TNM stages). Missing grade, N stage, and M stage

patients emanated disproportionately from the larynx. Missing grade data was more likely to come from stage IV, and more missing TNM data came from well-differentiated grade. Most M stage missing data came from TNM Stage IV patients. For smoking, T, N, M, TNM, and grade patients, missing data from any of these variables was disproportionately missing in any of the others.

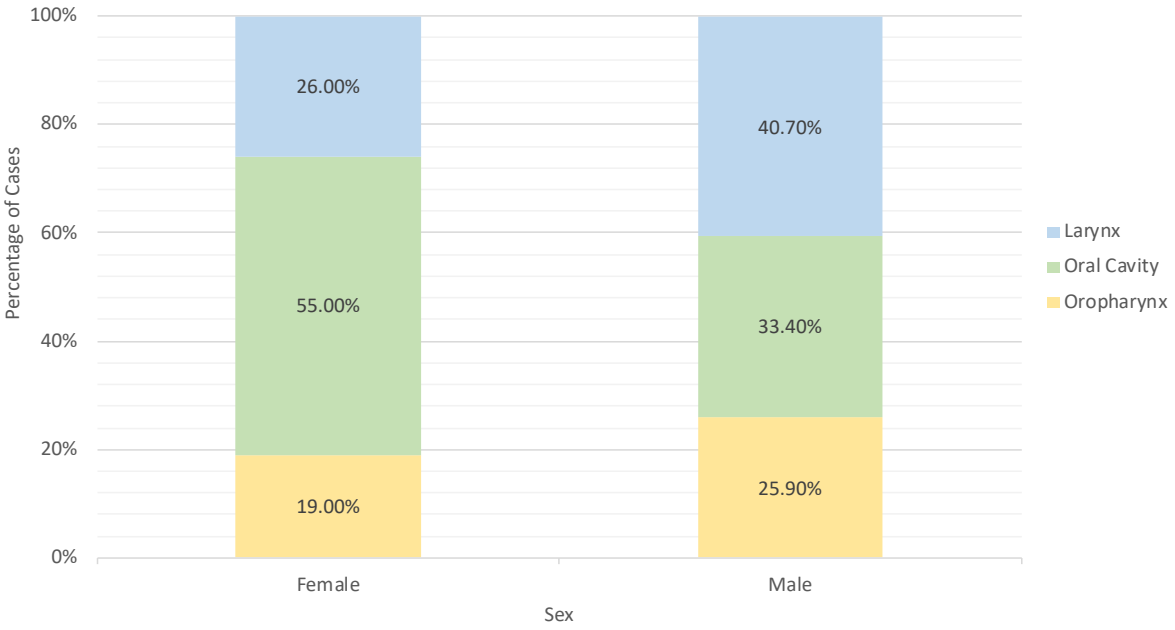

Figure 6.6 Percentage of male and female cases arising in the oropharyngeal, oral cavity, and laryngeal sub-sites.

For female n=200, and for male n=661. For the total population n=861.

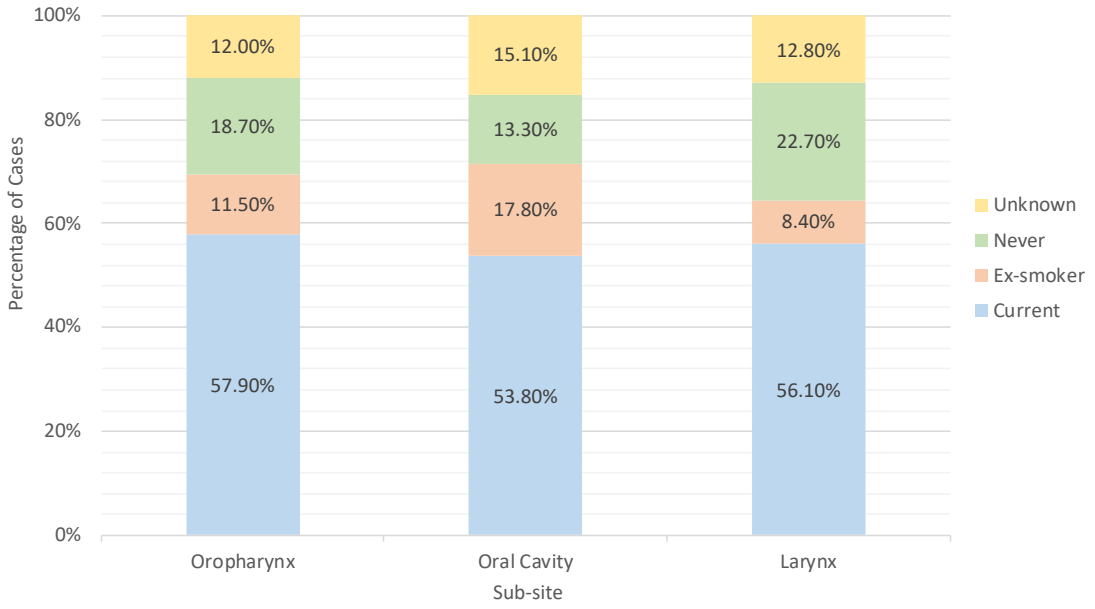

Figure 6.7 Percentage of oropharyngeal, oral cavity, and laryngeal cases that were current, ex, never, or unknown smokers.

For oropharynx n=209, for oral cavity n=331, and for larynx n=321. For the total population n=861.

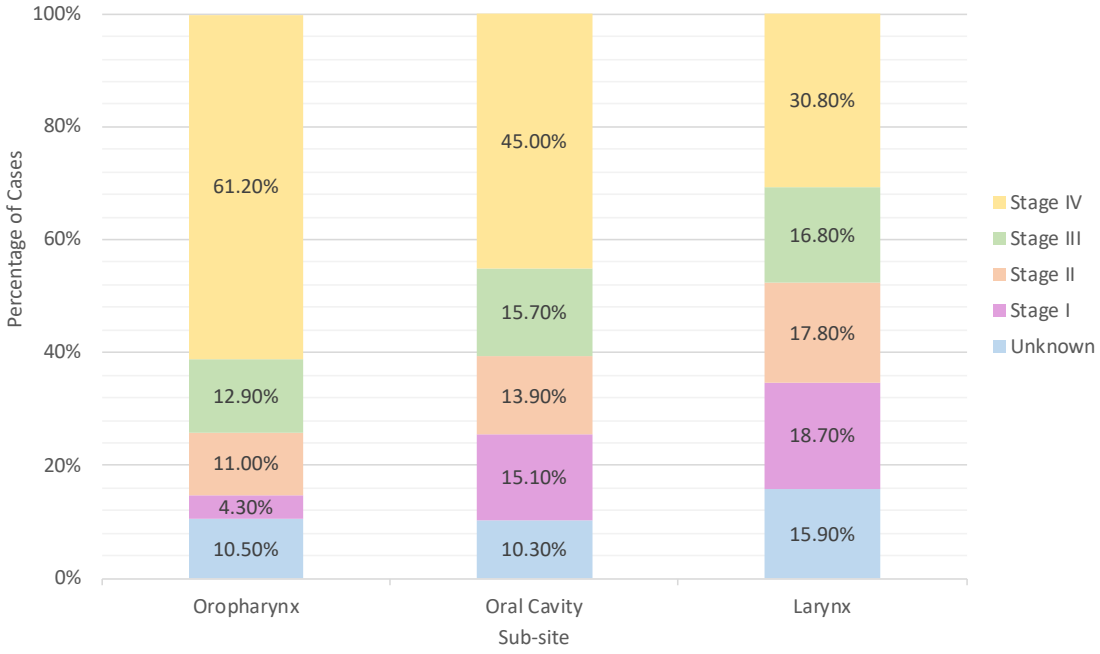

Figure 6.8 Percentage of oropharyngeal, oral cavity, and laryngeal cases arising at TNM Stages I, II, III, and IV.

For oropharynx n=209, for oral cavity n=331, and for larynx n=321. For the total population n=861.

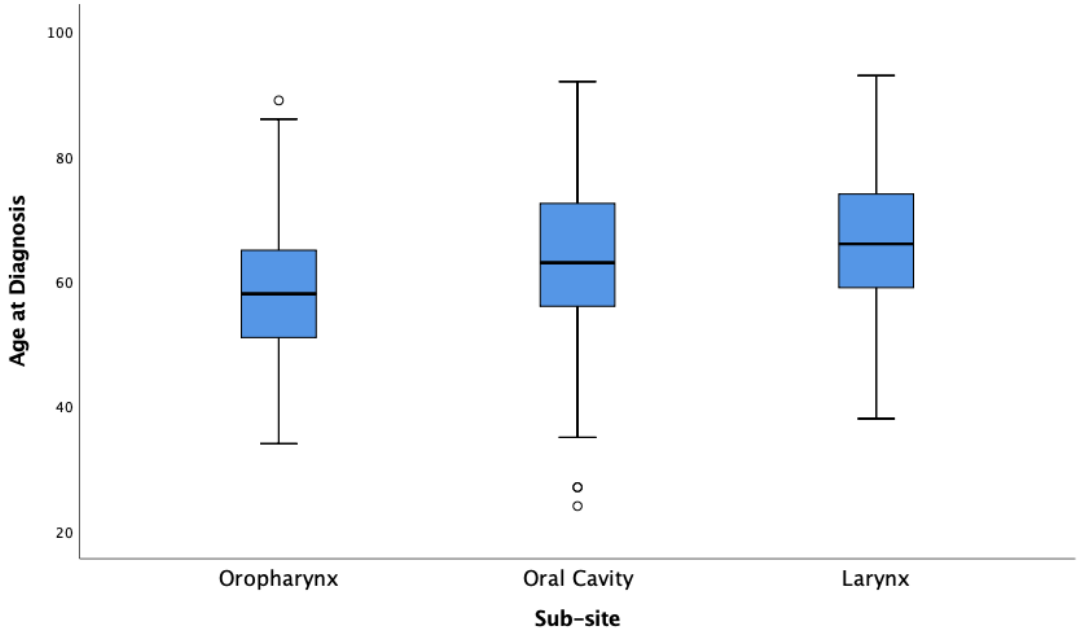

Figure 6.9 Age at diagnosis for oropharyngeal, oral cavity, and laryngeal cases.

For oropharynx n=209, for oral cavity n=331, and for larynx n=321. For the total population n=861.

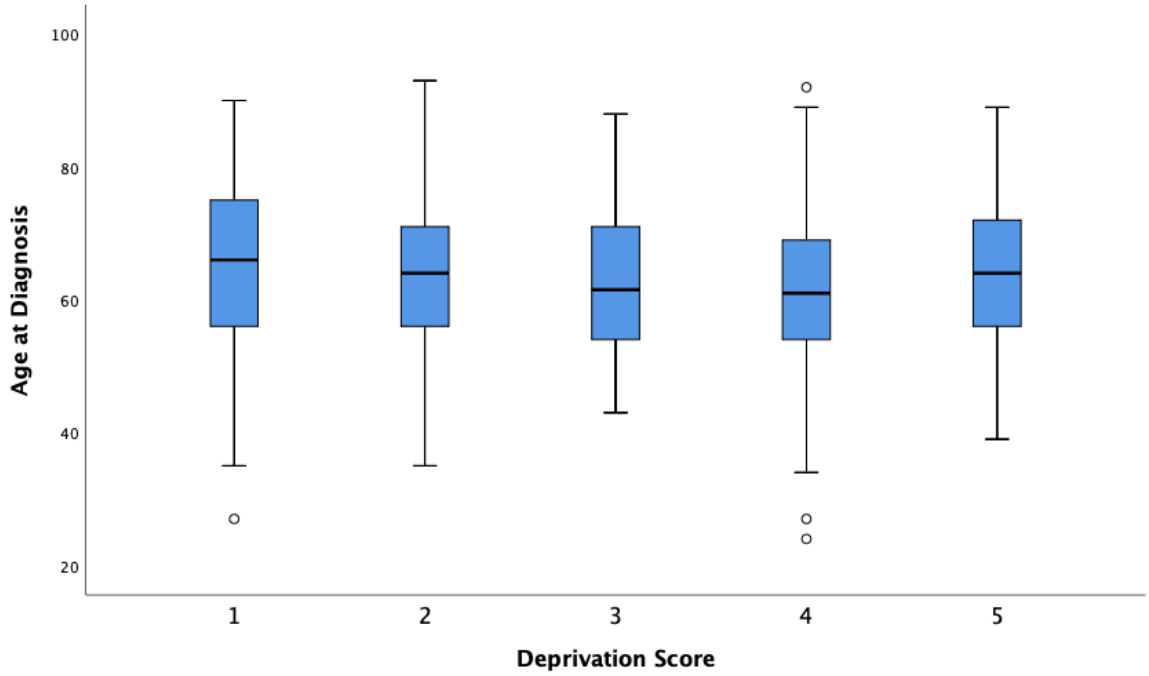

Figure 6.10 Age at diagnosis by social deprivation score.

For 1 n=120, for 2 n=122, for 3 n=128, for 4 n=178, and for 5 n=246. For the total population n=794.

6.7 Overall and Cancer-Specific Survival

Information regarding survival (months) for overall and cancer-specific survival was available for all 861 cases included in the study. Overall survival refers to the vital status of patients on 31/12/2016 and the corresponding date for cancer-specific survival was 31/12/2015. Patients still alive at these dates or for cancer-specific survival, died of other causes, were censored at these dates.

6.7.1 Overall and Cancer-Specific Survival by Characteristics

Kaplan-Meier analysis by log-rank test was used to assess differences in survival, both overall and cancer-specific, for all available patient and tumour characteristics. The significant results of these tests are presented in Figures 6.11 through 6.28. Overall survival results follow between Figures 6.11 and 6.20.

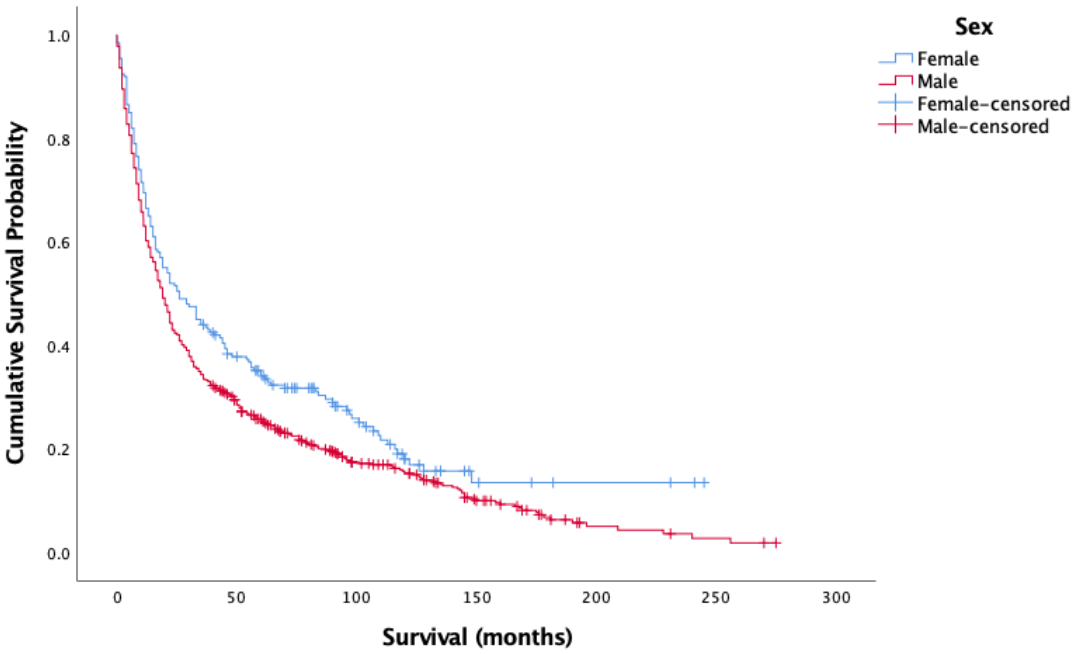

Figure 6.11 Kaplan-Meier analysis for overall survival by sex (n=861)

(Log-rank=6.727, 1 d.f., p=0.009)

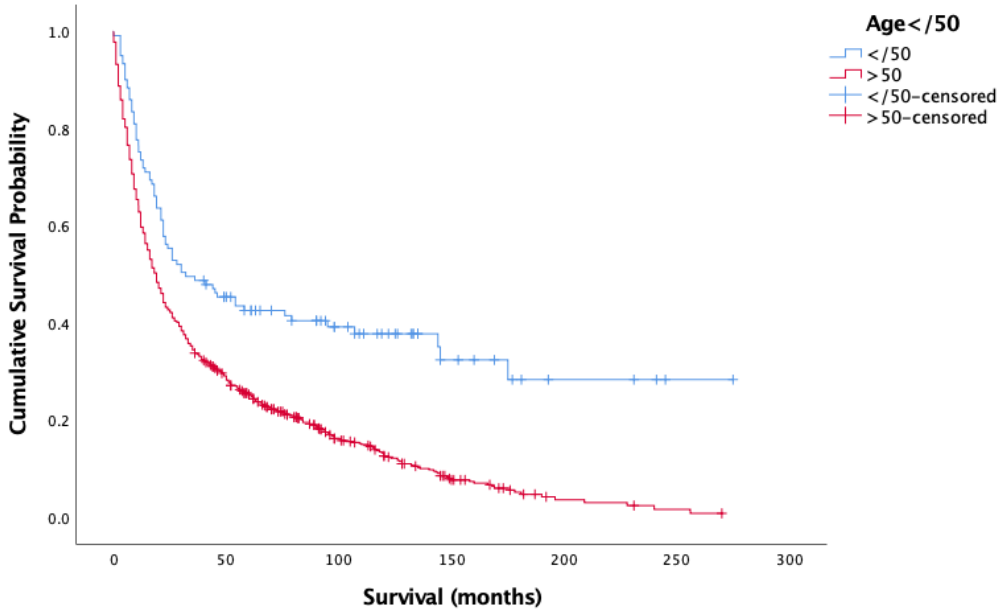

Figure 6.12 Kaplan-Meier analysis for overall survival by age younger than or equal to 50 (n=861)

(Log-rank=31.133, 1 d.f., p<0.0001)

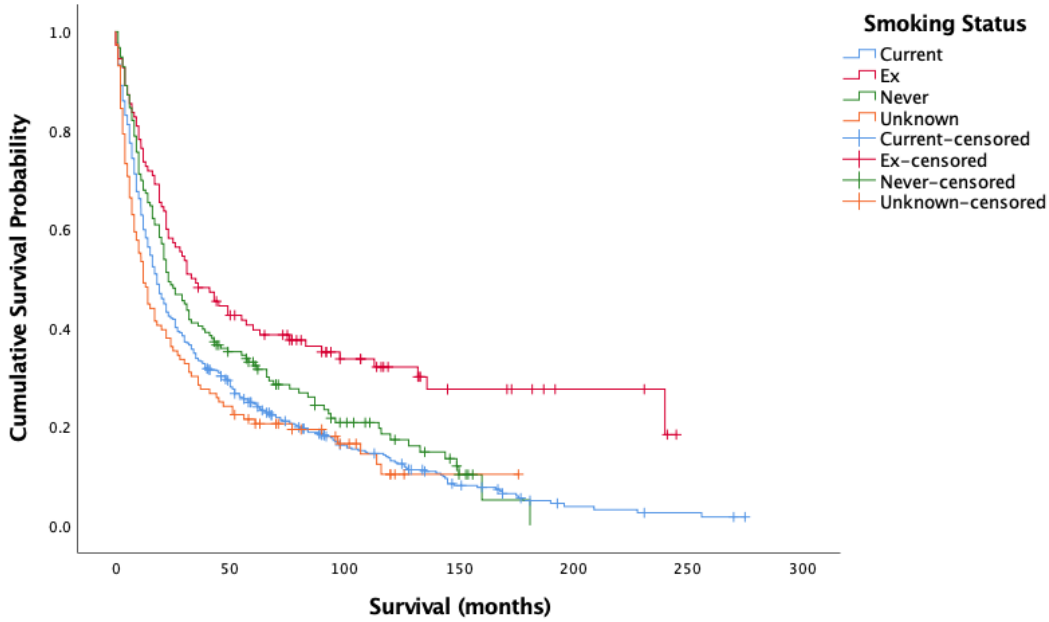

Figure 6.13 Kaplan-Meier analysis for overall survival by smoking status (n=861)

(Log-rank=26.856, 3 d.f., p<0.0001).

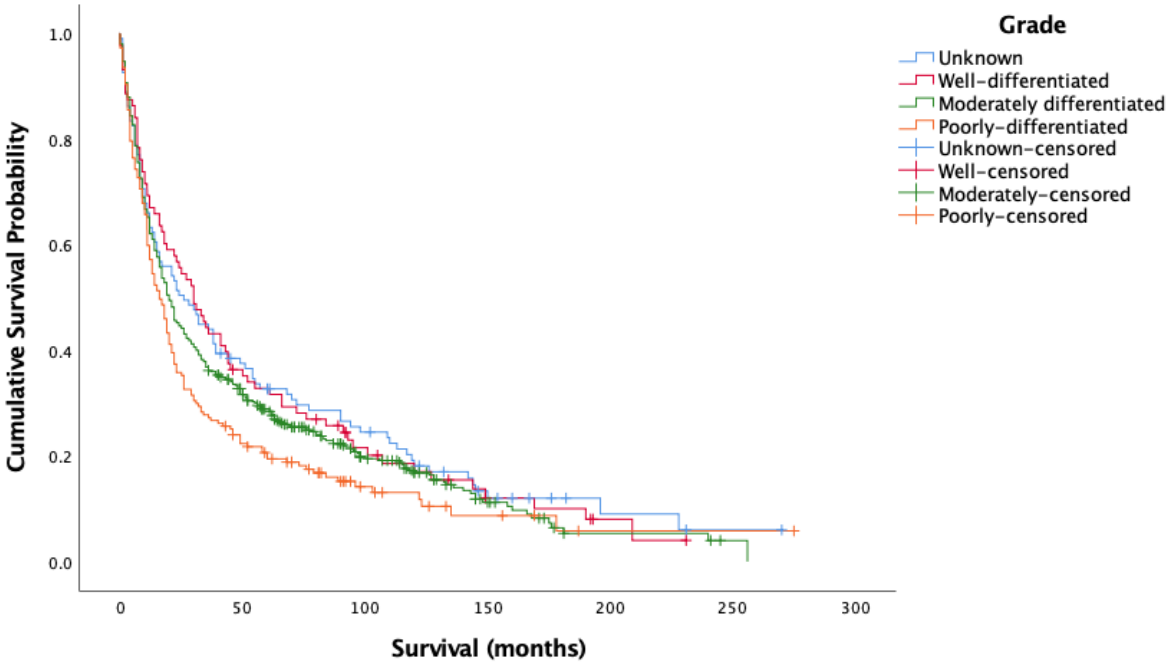

Figure 6.14 Kaplan-Meier analysis for overall survival by grade (n=861)

(Log-rank=7.900, 3 d.f., p=0.048)

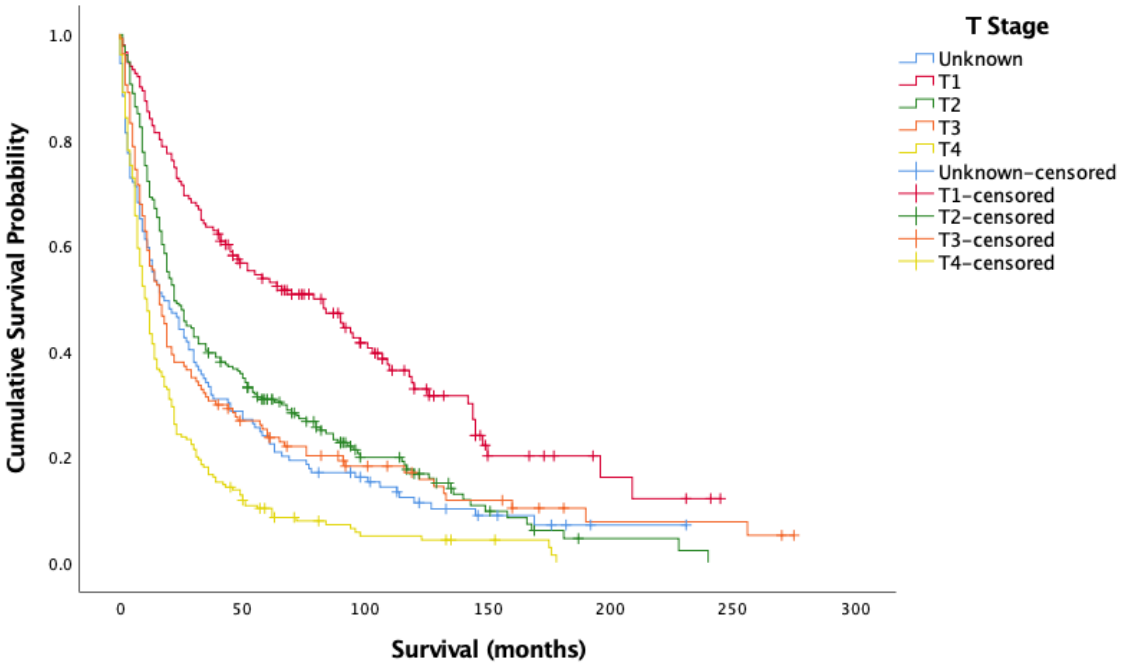

Figure 6.15 Kaplan-Meier analysis for overall survival by T stage (n=861)

(Log-rank=107.746, 4 d.f., p<0.0001)

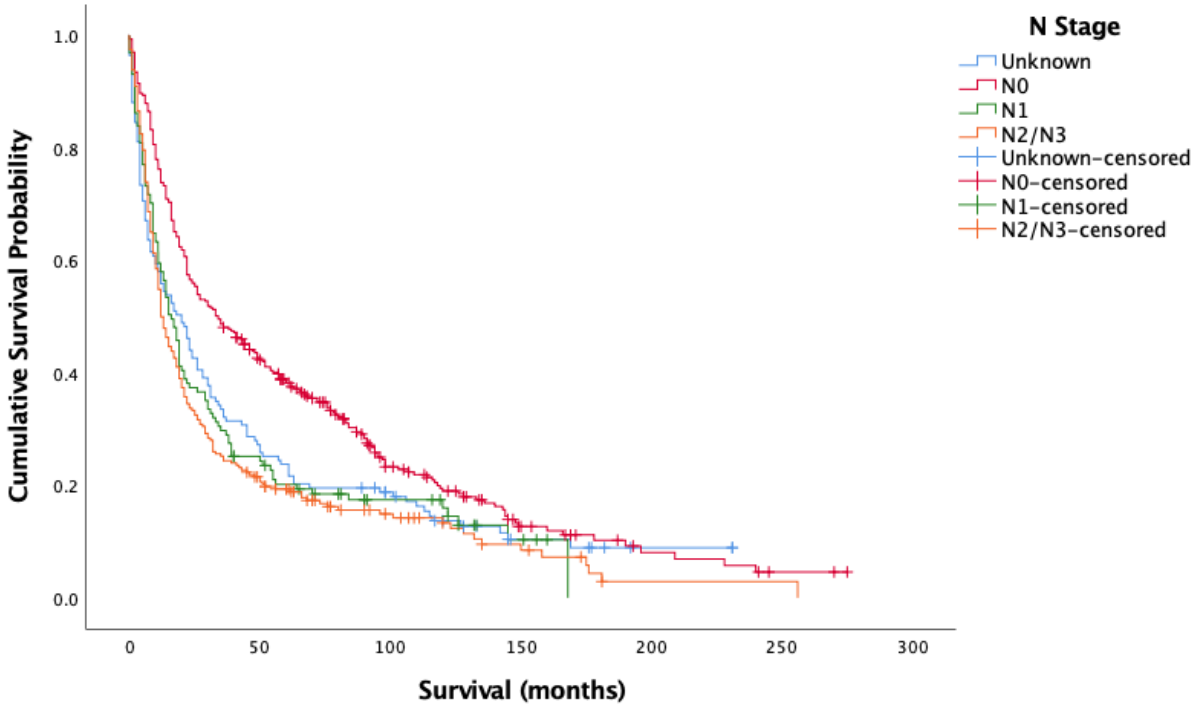

Figure 6.16 Kaplan-Meier analysis for overall survival by N stage (n=861)

(Log-rank=32.762, 3 d.f., p<0.0001)

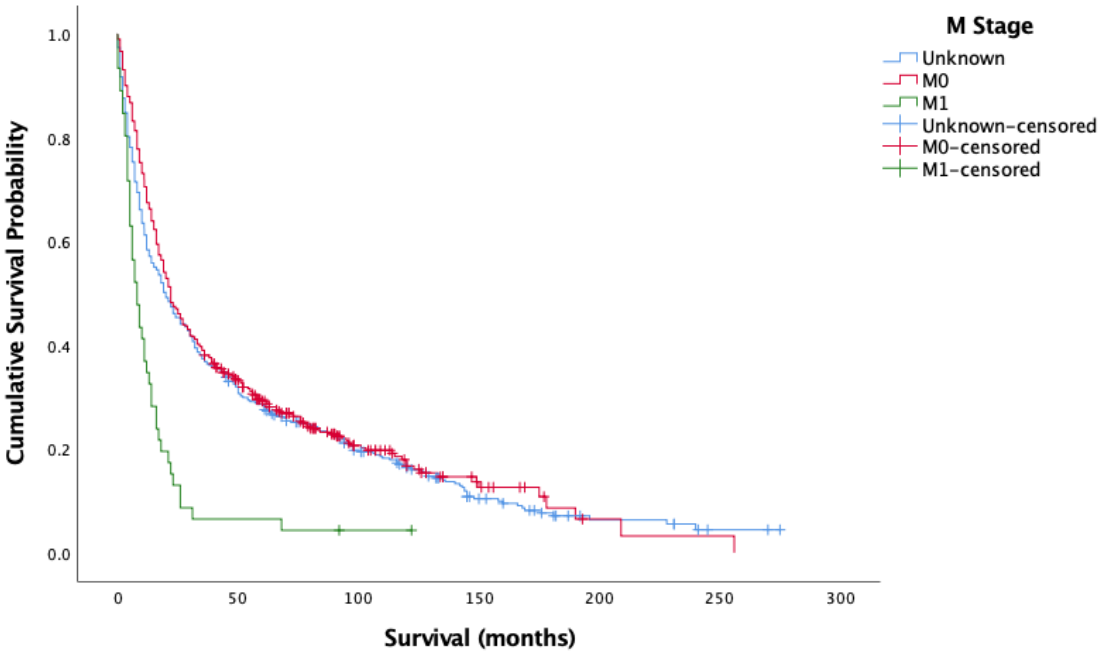

Figure 6.17 Kaplan-Meier analysis for overall survival by M stage (n=861)

(Log-rank=31.900, 2 d.f., p<0.0001)

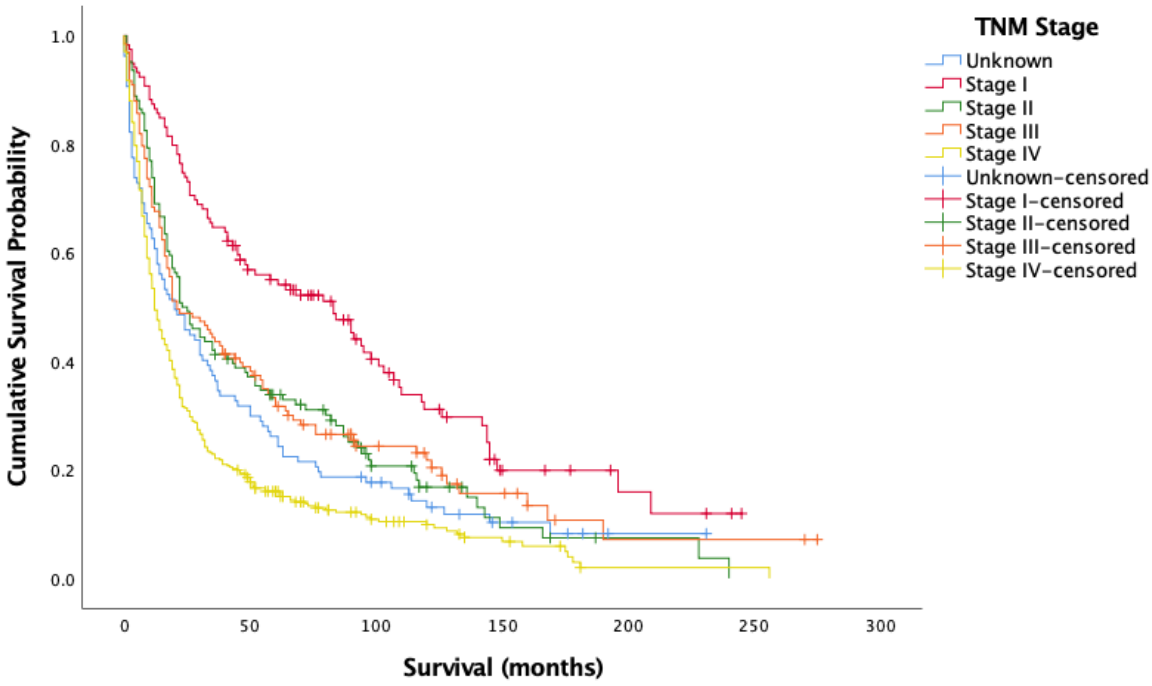

Figure 6.18 Kaplan-Meier analysis for overall survival by TNM stage (n=861)

(Log-rank=78.056, 4 d.f., p<0.0001)

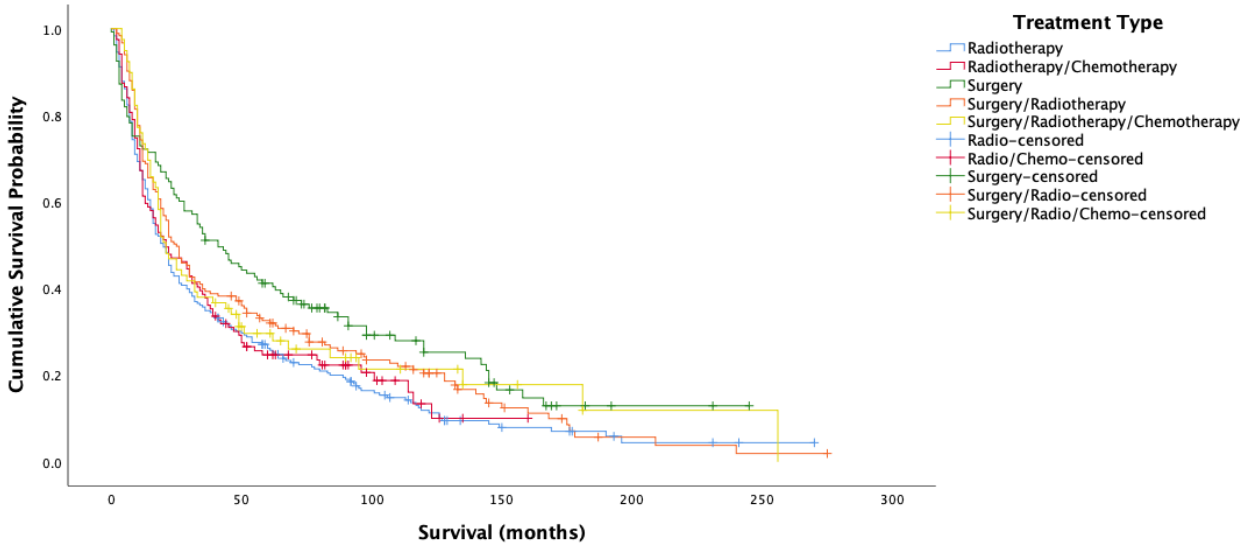

Figure 6.19 Kaplan-Meier analysis for overall survival by treatment type (n=752)

(Log-rank=12.402, 4 d.f., p=0.015)

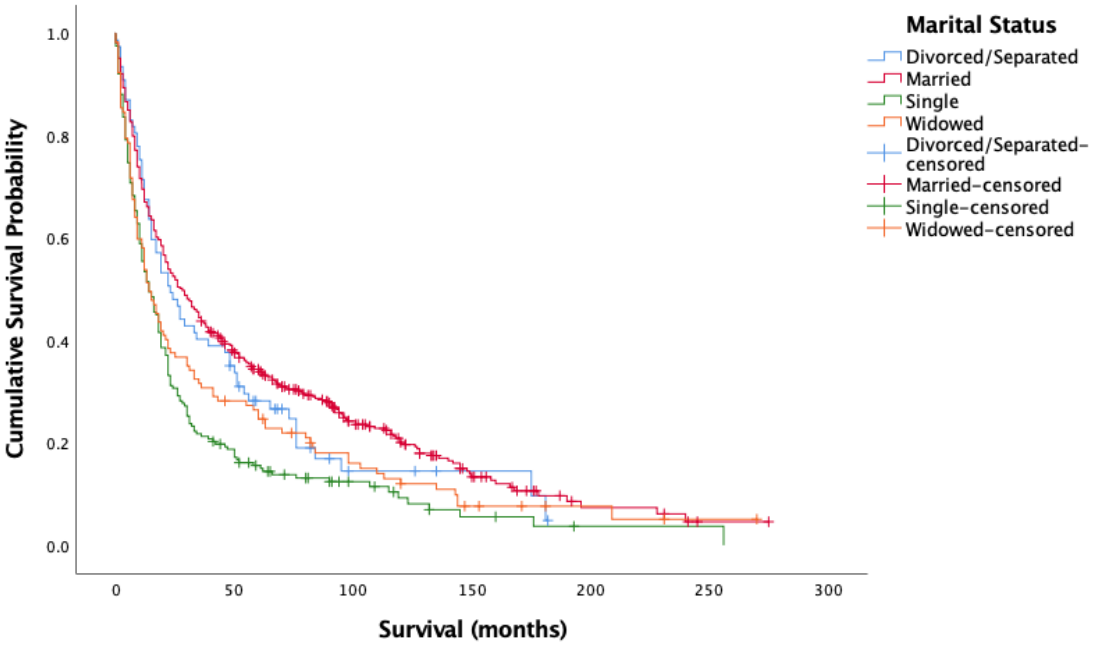

Figure 6.20 Kaplan-Meier analysis for overall survival by marital status (n=830)

(Log-rank=30.634, 3 d.f., p<0.0001)

Cancer-specific results follow between Figures 6.21 and 6.28.

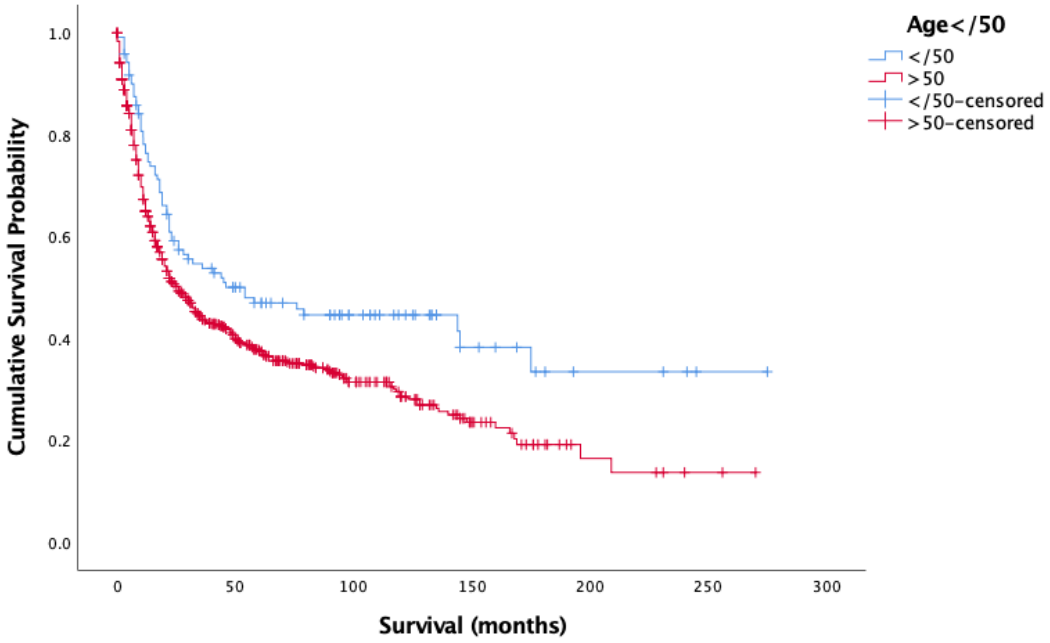

Figure 6.21 Kaplan-Meier analysis for cancer-specific survival by age younger than or equal to 50 (n=861)

(Log-rank=8.791, 1 d.f., p=0.003)

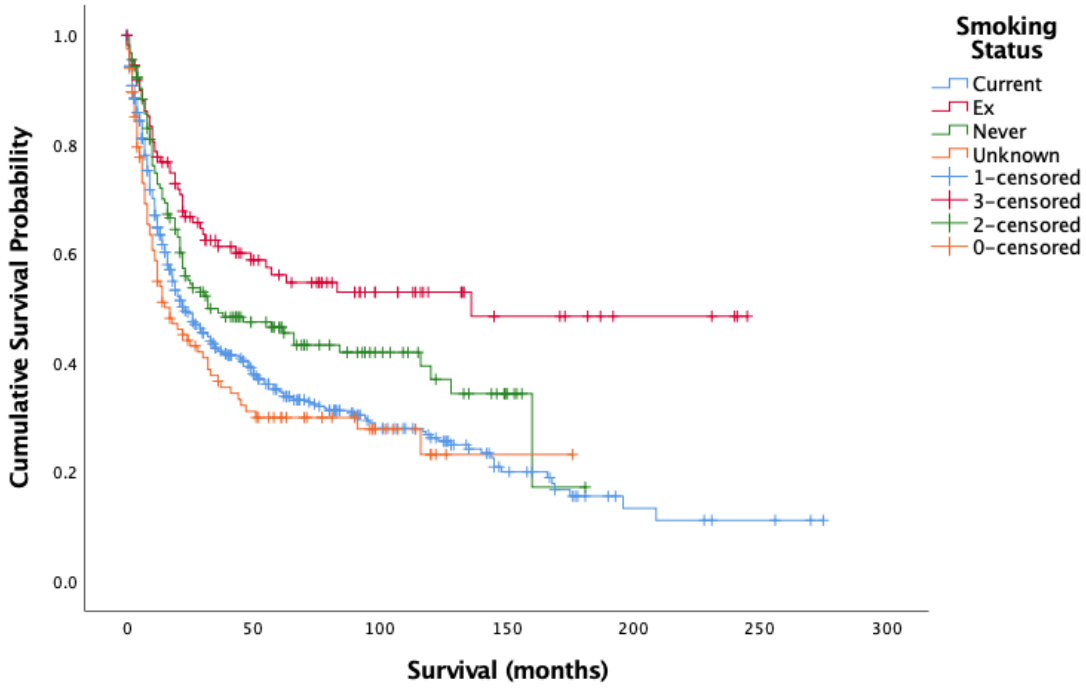

Figure 6.22 Kaplan-Meier analysis for cancer-specific survival by smoking status (n=861)

(Log-rank=28.074, 3 d.f., p<0.0001)

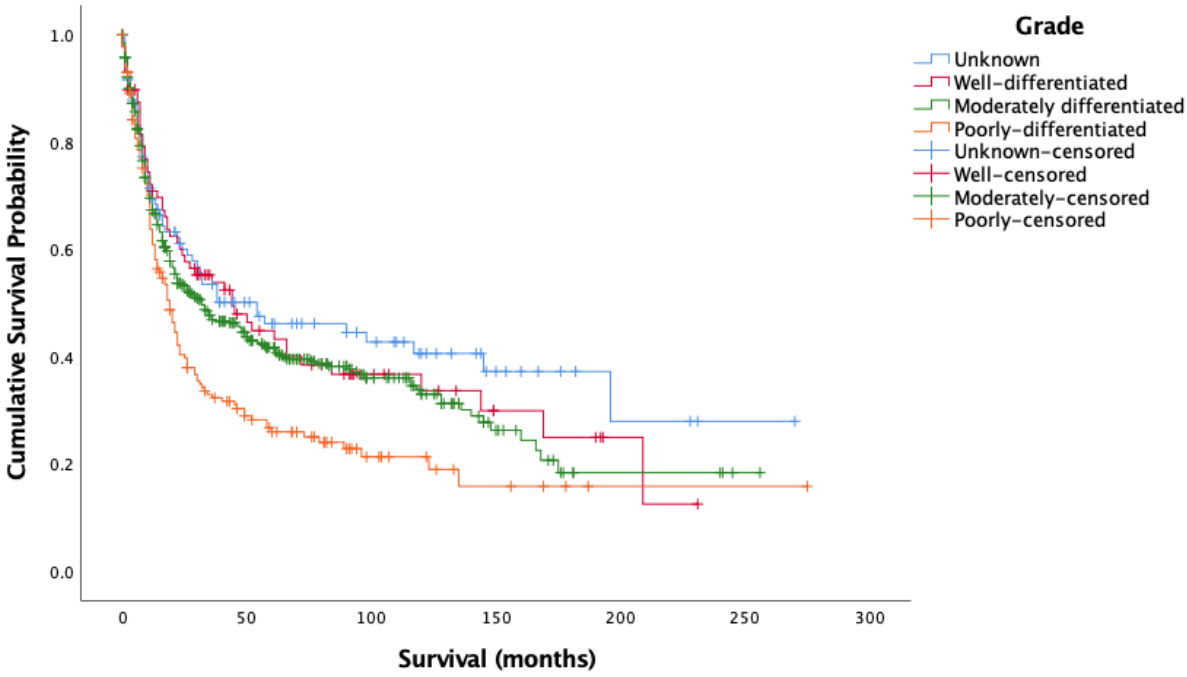

Figure 6.23 Kaplan-Meier analysis for cancer-specific survival by grade (n=861)

(Log-rank=15.310, 3 d.f., p=0.002)

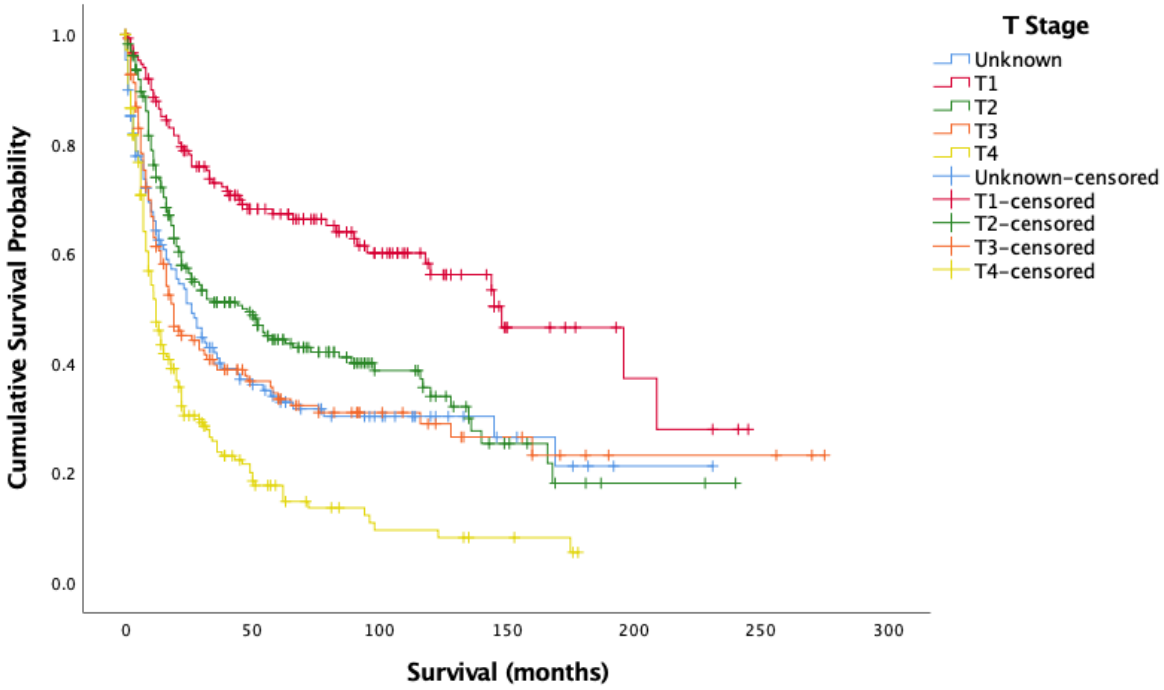

Figure 6.24 Kaplan-Meier analysis for cancer-specific survival by T stage (n=861)

(Log-rank=107.811, 4 d.f., p<0.0001)

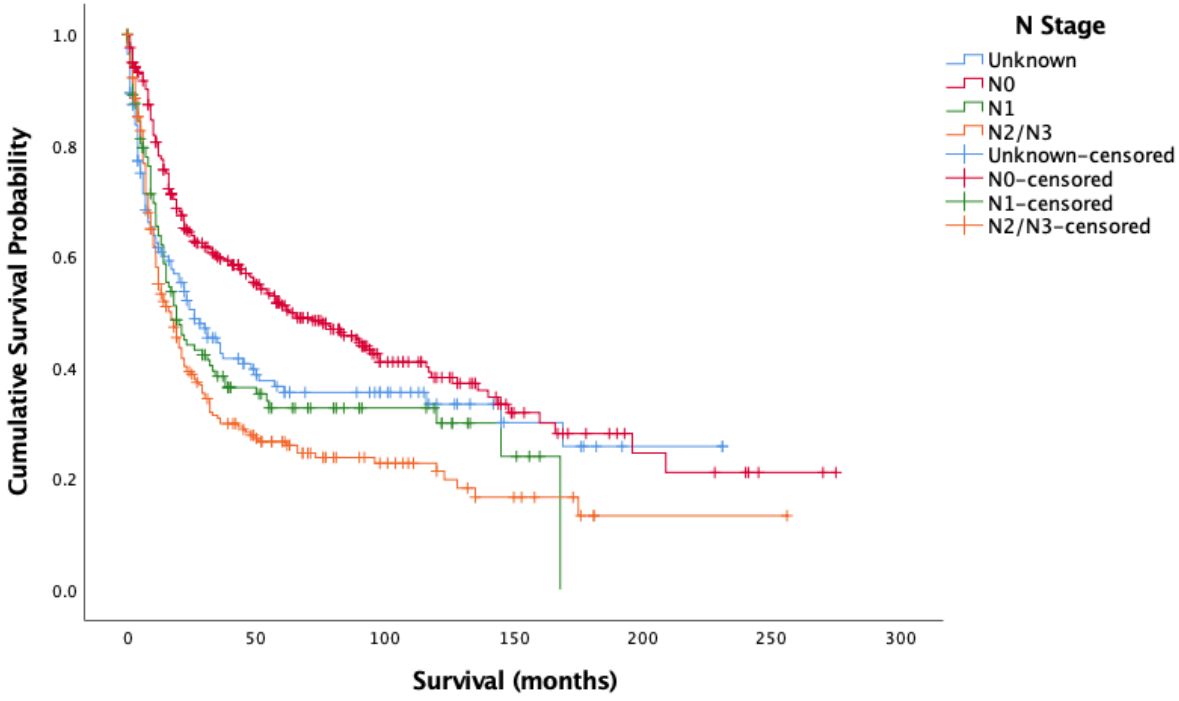

Figure 6.25 Kaplan-Meier analysis for cancer-specific survival by N stage (n=861)

(Log-rank=40.524, 3 d.f., p<0.0001)

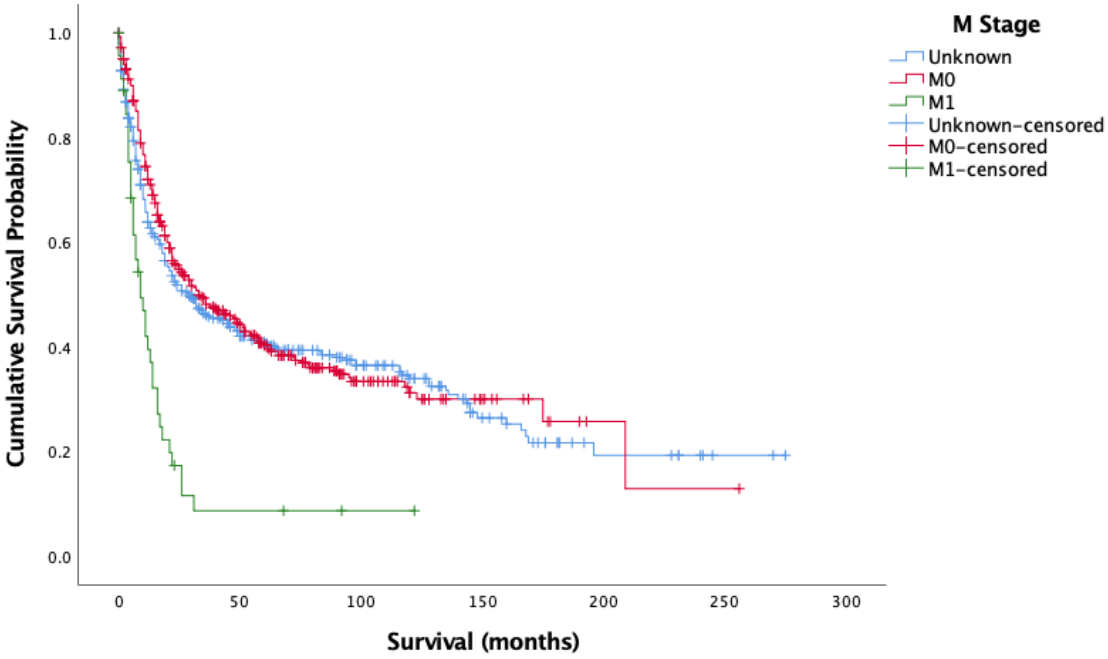

Figure 6.26 Kaplan-Meier analysis for cancer-specific survival by M stage (n=861)

(Log-rank=34.361, 2 d.f., p<0.0001)

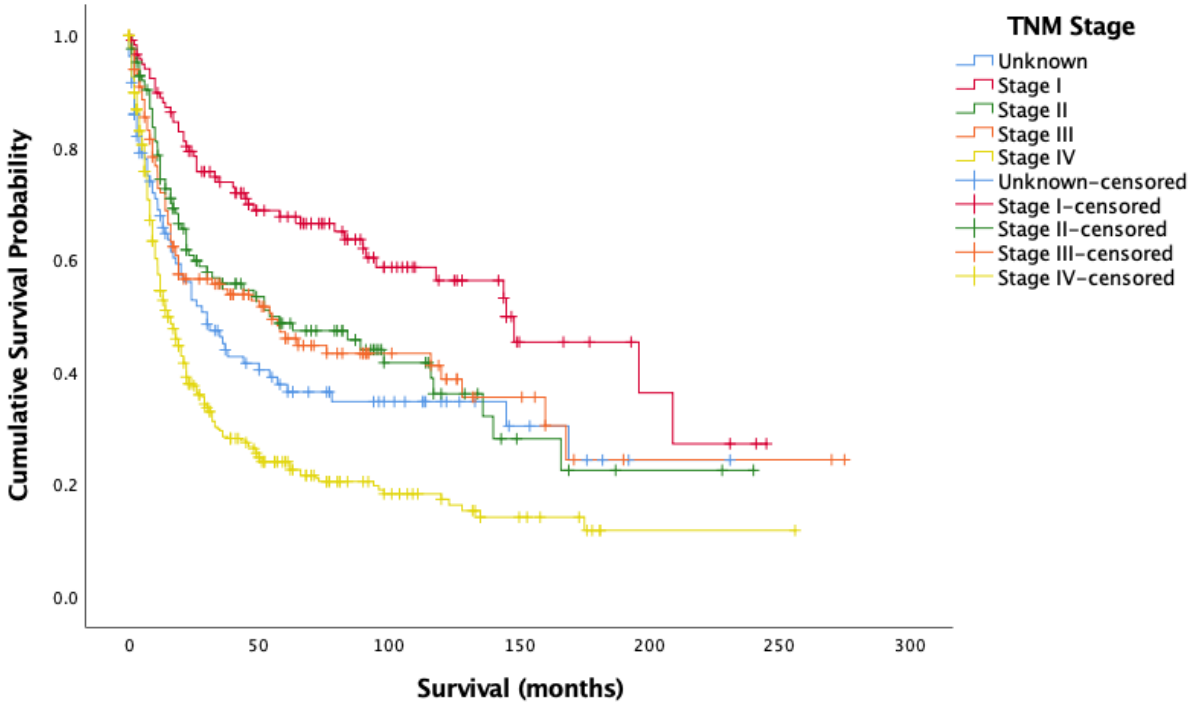

Figure 6.27 Kaplan-Meier analysis for cancer-specific survival by TNM stage (n=861)

(Log-rank=85.238, 4 d.f., p<0.0001)

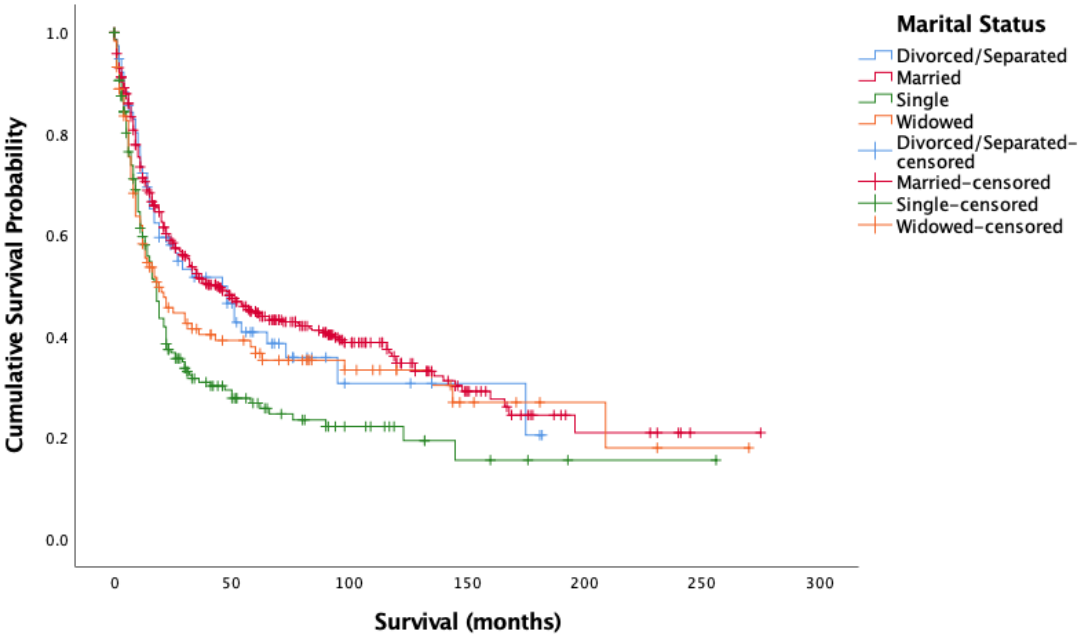

Figure 6.28 Kaplan-Meier analysis for cancer-specific survival by marital status (n=830)

(Log-rank=22.079, 3 d.f., p<0.0001)

6.8 Predictors of Survival

Univariate and multivariate cox proportional hazard analysis was conducted to assess significant predictors of survival in this population and to adjust for confounding factors between them. Table 6.9 showcases those variables significantly predictive of overall survival.

Table 6.9 Variables significantly predictive of overall survival by univariate cox proportional hazard models.

Variable/Factor	Statistic	Increased Risk of Death
Age (Continuous) (n=861)	HR=0.028 SE=0.003 P<0.0001	Older age
Age ≤50 (n=861)	HR=-0.659 SE=0.122 P<0.0001	>50
Sex (n=861)	HR=-0.232 SE=0.091 P=0.011	Male>Female
Smoking Status (n=861)	HR=-0.559, -0.189, 0.130 SE=0.125, 0.101, 0.112 P<0.0001, 0.0001, 0.006, 0.247	Current, Missing > Ex-smoker
T Stage (n=861)	HR=-0.436, -1.192, -0.638, -0.495 SE=0.118, 0.124, 0.101, 0.117 P<0.0001, 0.0001, 0.0001, 0.0001, 0.0001	T4>T3,T2, Missing>T1
N Stage (n=861)	HR=-0.161, -0.484, -0.097 SE=0.112, 0.091, 0.116 P<0.0001, 0.151, 0.0001, 0.404	N2/3, N1, Missing>N0
M Stage (n=861)	HR=-0.792, -0.872 SE=0.162, 0.161 P<0.0001, 0.0001, 0.0001	M1>M0, Missing

TNM (n=861)	HR=-0.294, -0.970, -0.468, -0.503 SE=0.117, 0.124, 0.112, 0.111 P<0.0001, 0.012, 0.0001, 0.0001, 0.0001	Stage IV>Missing>Stage III, Stage II>Stage I
Marital Status (n=830)	HR=-0.168, -0.291, 0.193 SE=0.159, 0.111, 0.123 P<0.0001, 0.291, 0.009, 0.116	Single > Married
Treatment Type (n=752)	HR=0.193, 0.152, -0.197, -0.004 SE=0.144, 0.163, 0.162, 0.150 P=0.017, 0.183, 0.351, 0.223, 0.977	Radiotherapy, Radiotherapy/Chemotherapy > Surgery

Due to evident and expected confounding between T, N, M, and TNM stages, TNM stage was used as a representative variable for each cancer stage in the multivariate model. Results of this model are presented in the Table below.

Table 6.10 Variables significantly predictive of overall survival by multivariate cox proportional hazard model. The initial model included all univariately significant variables except T, N, and M stage (n=727).

Variable/Factor	Statistic	Increased Risk of Death
Age	HR=0.030 SE=0.004 P<0.0001	Older age
Smoking Status	HR=-0.605, -0.218, -0.065 SE=0.139, 0.111, 0.136 P<0.0001, 0.001, 0.051, 0.632	Missing, Current, Never > Ex-smoker
TNM Stage	HR=-0.491, -0.931, -0.435, -0.469 SE=0.142, 0.134, 0.125, 0.120 P<0.0001, 0.001, 0.0001, 0.0001, 0.0001	Stage IV>Missing, Stage III, Stage II>Stage I

Marital Status	HR=0.098, 0.528, 0.241 SE=0.038, -0.025, 0.338 P<0.004, 0.0830, 0.0844, 0.016	Single > Married, Widowed, Divorced/Separated

Table 6.11 reveals those variables significantly predictive of cancer-specific survival by univariate analysis.

Table 6.11 Variables significantly predictive of cancer-specific survival by univariate cox proportional hazard models.

Variable/Factor	Statistic	Increased Risk of Death
Age (Continuous) **(n=861)**	HR=0.018 SE=0.004 P<0.0001	Older age
Age≤50 **(n=861)**	HR=-0.385 SE=0.132 P=0.004	>50
Smoking Status **(n=861)**	HR=-0.679, -0.301, 0.144 SE=0.157, 0.122, 0.129 P<0.0001, 0.001, 0.014, 0.263	Current, Missing > Never > Ex
Grade **(n=859)**	HR=-0.520, -0.399, -0.320 SE=0.158, 0.162, 0.106 P=0.002, 0.001, 0.014, 0.002	Poorly-differentiated > Moderately differentiated, Well-differentiated, Missing
T Stage **(n=861)**	HR=-0.507, -1.412, -0.775, -0.501 SE=0.136, 0.154, 0.118, 0.133 P<0.0001, 0.0001, 0.0001, 0.0001, 0.0001	T4>Missing, T3>T2>T1
N Stage **(n=861)**	HR=-0.287, -0.651, -0.189 SE=0.132, 0.107, 0.134	N3/N2, N1, Missing> N0

	P<0.0001, 0.030, 0.0001, 0.158	
M Stage (n=861)	HR=-0.899, -0.975 SE=0.177, 0.176 P<0.0001, 0.0001, 0.0001	M1>M0, Missing
TNM (n=861)	HR=-0.412, -1.195, -0.673, -0.616 SE=0.139, 0.158, 0.138, 0.133 P<0.0001, 0.003, 0.0001, 0.0001, 0.0001	Stage IV>Missing>Stage III, Stage II>Stage I
Marital Status (n=830)	HR=-0.185, -0.255, 0.225 SE=0.190, 0.133, 0.145 P<0.0001, 0.329, 0.056, 0.121	Single > Married, Divorced/Separated

Due to evident and expected confounding between T, N, M, and TNM stages, TNM stage was used as a representative variable for each cancer stage in the multivariate model. Results of this model are presented in the Table below.

Table 6.12 Variables significantly predictive of cancer-specific survival by multivariate cox proportional hazard model. The initial model included all variables univariately significant except for T, N, and M stage (n=830).

Variable/Factor	Statistic	Increased Risk of Death
Age (Continuous)	HR=0.025 SE=0.005 P<0.0001	Older age
Smoking Status	HR=-0.672, -0.384, 0.016 SE=0.158, 0.126, 0.139 P<0.0001, 0.0001, 0.002, 0.907	Current, Missing > Never > Ex
TNM Stage	HR=-0.541, -1.207, -0.683, -0.606	Stage IV>Missing, Stage III, Stage II>Stage I

	SE=0.146, 0.161, 0.144, 0.136 P<0.0001, 0.0001, 0.0001, 0.0001, 0.0001	
Marital Status	HR=-0.167, -0.077, 0.232 SE=0.199, 0.138, 0.150 P=0.021, 0.401, 0.574, 0.124	Single > Married, Divorced/Separated

A sensitivity analysis was conducted on all of the above results (from Section 6.5 onwards) whereby all tests were repeated excluding all cases for which data was missing in relevant tests. The results were broadly similar, suggesting that though bias did exist in the kinds of patients for whom data was unknown/missing, the data still reflected the same significant relationships.

6.9 Discussion

The aims of this chapter were;

- to determine the distribution of the present population by patient and tumour characteristics
- to assess the comparability of the study population to the Irish oropharyngeal, oral cavity, and laryngeal SCC population and any selection bias using key patient and tumour characteristics
- to evaluate the correlation between patient and tumour characteristics
- and to determine predictors of overall and cancer-specific survival in this population.

To the first aim, Table 6.4 showcases the distribution of this population by all of the available patient and tumour characteristics provided by the NCRI. The mean age of the population was 63.30 (CI: 62.52, 64.08), and most of the population was aged above 50. The majority of the study cohort were current smokers, while oral cavity, laryngeal, and oropharyngeal sub-sites represented the largest, moderate, and smallest populations in the sample, respectively. Mean deprivation fell just above 3, at 3.39 (CI: 3.29, 3.49) and just under the majority of patients (44.1%) resided in Dublin. Most patients were married

though a quarter of them were single, and less than a fifth were widowed. The majority of cases with available data were diagnosed as moderately differentiated, without distant metastasis, and at TNM Stage IV. As detailed in Table 6.5 and Figure 6.4, 31.4% of patients treated within 12 months of diagnosis were treated with radiotherapy while almost a quarter of them were treated with surgery/radiotherapy. Chemotherapy treatments not included in surgery/radiotherapy/chemotherapy together were very rare, representing cumulatively only 0.8% of cases.

To the second aim, 6 of the 14 variables available for analysis indicated that the population was representative of all oropharyngeal, oral cavity, and laryngeal SCC in Ireland and was not disproportionately affected by selection bias. In terms of age, sex, grade, M stage, social deprivation, and marital status, the current population, which was sampled entirely blind of any of these characteristics as described in full in Chapters 3 and 4, mirrored the oropharyngeal, oral cavity, and laryngeal SCC population in Ireland according to NCRI data and did not suffer from selection bias. The distribution of cases by: year of diagnosis indicated a slight under-sampling of patients diagnosed before 2004; sub-site saw slight over-sampling of the oropharynx; smoking indicated slight under-sampling of ex-smokers; T stage highlighted an over-sampling of T1; N stage revealed slight over-sampling of N2 patients; TNM showed over-sampling of Stage IV; geographic location signaled over-sampling from patients with urban and/or Dublin addresses; and treatment suggested slight over-sampling of patients receiving chemotherapy in any capacity.

The study population was established without prior knowledge of any patient characteristics but was determined by the ability to organize sample retrieval for different hospitals. Most hospitals for which this was possible were in urban centers, and despite hospitals like St. James' University Hospital being huge referral centers for head and neck patients around the country, this access-driven sampling likely influenced the aforementioned results. This said, it is extremely encouraging that the sample population did not appear to suffer disproportionately from missing data when compared to the national statistics available at the national level from the NCRI.

Given the slight under-sampling of patients from before 2004, it might preliminarily be suggested that the results from Figure 6.5, showing increasing raw incidence for oropharyngeal, oral cavity, and laryngeal SCC overall may only be attributed to the larger number of cases sampled from 2004 onwards. However, data from the NCRI shows that incidence of HNSCC overall has been increasing at a rate of 1.1% between 2001 and 2013[1], which validates the increasing number of cases seen amongst all oropharyngeal, oral cavity, and laryngeal SCC and indicates relatively successful random sampling of the Irish population from 7 different hospital sites across the country.

That there was slight selection bias for the oropharyngeal sub-site might initially suggest that both the individual raw incidences and the proportional contribution of each sub-site to overall incidence are not entirely accurate by these simple figures. However, comparisons to the proportional contributions to incidence at the national level suggest that this selection bias had little effect on the overall representative nature of the sample. A closer look at proportional incidence contributions from NCRI data supports the accuracy of the present raw incidence, even after the end of the present study[1–3]. In fact, oropharyngeal incidence in the ECHO study population and the national population contributed an average of 19.6% and approximately 20% to overall oropharyngeal, oral cavity, and laryngeal SCC incidence respectively between 1994 and 2013[2]. The same statistics in the oral cavity were 43.8% and 50% and in the larynx were 36.7% and 30%[22].

The raw incidence rates produced revealed increased incidence in all three sub-sites over time (Figure 6.5), with the oral cavity and larynx showing the largest contribution overall and the oropharynx contributing the least to all oropharyngeal, oral cavity, and laryngeal SCC. The present data therefore indicates that unlike in other North American and European countries where emerging OPSCC incidence is beginning to overshadow other HNSCC sub-types[4–10], OPSCC remains a slightly less significant burden on the Irish population. This said, the oropharynx saw the largest annual percentage increase over the time period of any sub-site at 9.4% ($p<0.0001$). This suggests that the incidence of these cancers is increasing at a faster rate than their counterparts, likely due to increasingly HPV-related tumours. This rate is not yet fast enough however to render oropharyngeal cancer the dominant contributor to all oropharyngeal, oral cavity, and laryngeal SCC incidence for the time being.

Results geared towards the third aim of this chapter also indicate that despite over-/under-sampling for some variables, the current sample population mirrors relationships between patient and tumour characteristics seen at the national level and across Europe and North America. Table 6.7 showcases all of the significant relationships emanating from comparisons between all tumour and patient characteristics.

In particular, sex was significantly associated with sub-site from which tumours originated, with 55.0% of female cases originating in the oral cavity compared to 33.40% of male cases, something highlighted in Figure 6.6. This is promising as the current population appears to be reflective of the noted resurgence of mouth cancers amongst women that has been noted since 2004 in the Irish population[2]. Interestingly, there was a gradation in age at diagnosis between sub-sites, with oropharyngeal, oral cavity, and laryngeal cases presenting at the youngest, middling, and oldest ages respectively (Figure 6.9). This is again reflective of the literature which definitively indicates that due to many oropharyngeal cases being driven by HPV-related carcinogenesis, patients are significantly younger at diagnosis than those that present in both the oral cavity and the larynx[4,11–13]

That T, N, and TNM stage were all individually related to sub-site is unsurprising, as each area of the head and neck has been consistently shown to not only have a differing relationship to TNM stage as a whole, but also to T[14,15] and N[15,16] stages separately. Almost half of all oropharyngeal cancers presented at the highest nodal stage which is consistent with their anatomical proximity and inclusivity of regions of the lymphatic system. This is especially true given that the oropharyngeal site includes the tonsil, an integral part of the lymphatic system itself which lends to rapid metastasis to surrounding lymph nodes[17,18]. The extremely disproportionate presentation of oropharyngeal cancers at Stage IV (Figure 6.8) is thus likely explained by their greater presentation at later N stage. Cumulatively, this might suggest that oropharyngeal cancers are more aggressive than those of the other sub-sites and escape diagnosis until Stage IV has already been reached. Over-sampling in Stage IV in this population could also be said to contribute to this relationship. However, the relevance of nodal extent for cancers originating in this region due to their HPV-related nature has

been discredited to the extent that these cancers have now been down-graded on the basis of HR HPV positive status in AJCC guidelines introduced in 2016[19].

That oropharyngeal tumours were likely to be treated more aggressively than their counterparts is unsurprising on the basis of their overwhelming presentation at Stage IV TNM stage. This is only compounded by the significant relationship of TNM stage to treatment which showed that Stage IV tumours were more likely to be treated with surgery/radiotherapy/chemotherapy in the first year after diagnosis. However, given the noted skew in TNM stage due to nodal category in the oropharynx reflected in the new AJCC guidelines, the mere distribution of this population decidedly agrees with the literature that these cancers might have historically been over-treated in the clinic, and could benefit from de-escalation of treatment[20–22].

Social deprivation's relationships to age, smoking status, and geographic location also indicate the population's representative nature of findings in the literature. That age at diagnosis decreased by one year consistently between scores 1 and 4 is indicative of the earlier ages at which lower socio-economic groups in Ireland are exposed to HNSCC carcinogens, including smoking and HR HPV as a result of persistent smoking habits[23] and more risky sexual behaviors[24]. Indeed, the current sample is reflective of the former, where current smokers were more likely to come from higher social deprivation categories. Similarly, that both social extremes (highest and lowest social deprivation scores) were more likely to arise in urban centers, and middle-level scores were likely to come from outside Dublin, is reflective of the Irish population's concentration of inequality in larger cities[25].

To the fourth aim of the study, 10 of the 14 available patient and tumours characteristics showed to be significant predictors of overall survival by Kaplan-Meier analysis (Figures 6.11 through 6.20). 8 of the 14 were predictive of cancer-specific survival by Kaplan-Meier analysis (Figures 6.21 through 6.28). For both overall and cancer-specific survival: age younger than or equal to 50; ex-, never-, and current smokers (in that order); earlier T stage; earlier N stage; no distant metastasis; earlier TNM stage; well- and moderately-differentiated grade; and non-single status predicted better survival. Where male sex saw

worse overall survival, no such difference was seen at the cancer-specific level. Treatment was predictive of overall survival, with surgery seeing the best survival of all treatments.

To note are the converging survival trends amongst T2 and T3 patients, and those amongst non-N0 N stage for both cancer-specific and overall survival (Figures 6.15, 6.16, 6.24 and 6.25). TNM Stages II and III also showed similar survival patterns (Figures 6.18 and 6.27). Indeed, findings have been similar amongst other studies of oropharyngeal, oral cavity, and laryngeal SCC, suggesting that various T, N, and TNM groupings might be converged at least for the purposes of analysis[11]. However, given the importance of analyzing this population at the sub-site level, and the large sample size at hand, it is prudent to keep staging categories as they are to ensure a full scope of understanding of further results.

In multivariate analysis, older age, current and never smoking status, later TNM stage, and single marital status were predictive of greater risk of death. At the cancer-specific level, older age, current smoking status, later TNM stage, and single marital status were all predictive of increased risk of death. It is particularly interesting to note that treatment was not a significant predictor of survival in the context of other variables for all oropharyngeal, oral cavity, and laryngeal SCC. This is a testament to the importance of various risk factors in driving survival trends, including diagnosis at early stage[11,26], immune-suppressing behaviours like smoking status[27], and age[28] as an indicator of ability to cope with treatment, no matter the type.

This said, where in univariate cox models and Kaplan-Meier analysis, treatment was a significant in predicting overall survival, it was not at the cancer-specific level. This suggests that treatment in oropharyngeal, oral cavity, and laryngeal SCC may not be a determinant for surviving the cancer itself, but inevitably determines long-term prosperity. Surgery was predictive of best overall survival by both univariate cox model and Kaplan-Meier analysis, suggesting that harsher treatments have long-term consequences for all oral generalized SCC patients that ultimately lead to co-morbidities. Indeed, harsher treatments have been shown to increase risk of heart disease and failure[29,30], risk of another (non-recurrence) primary tumour at another site[31–34], and complications due to immunosuppression.

Another notable finding was the role of marital status in determining both overall and cancer-specific survival (Tables 6.10 and 6.12). Single patients had significantly worse survival than those who were divorced/separated or married. It is difficult to assess precisely why these patients are disproportionately impacted in this sub-site, but it is possible that care and support outside the clinical context have a role to play. Other studies have also supported the idea that single and widowed patients do not have the same levels of emotional and physical support outside the hospital due to lack of spousal (in the case of married patients) or family (in the case of divorced patients) help[35,36].

With respect to missing data, it should be emphasized that the ECHO study cohort did not suffer more than all oropharyngeal, oral cavity, and laryngeal SCC patients recorded in the registry from missing data. This said, it is evident that missing data did not occur randomly, and more often affected patients with particular characteristics. The last row of Table 6.8 alludes to those characteristics for which patients are most likely to have missing data. Older, later stage, laryngeal, well-differentiated, current smokers were likely to be missing data from corresponding smoking, T, N, M, TNM, and grade variables. Survival analysis confirms these insinuations. Survival patterns for unknown data followed closely to the trends seen amongst current smokers (Figures 6.13 and 6.22), well-differentiated tumours (Figures 6.14 and 6.23), T2 and T3 stages (Figures 6.15 and 6.24), non-N0 status (Figures 6.16 and 6.25), M0 stage (Figures 6.17 and 6.25), and Stage II and III TNM stages (Figures 6.18 and 6.27).

That unknown cases saw worse survival than current smokers despite following their survival trend most closely suggests that current smokers for which data was available were biased for better survival (Figures 6.13 and 6.22). The opposite is true of well-differentiated cases who likely should have had better survival than showcased (Figures 6.14 and 6.25). In addition to this, the finding that there was slight over-sampling of never smokers and under-sampling of ex-smokers implies that never-smoking patient survival is underestimated in the dataset where ex-smoking patient survival is overestimated given the persistently best survival amongst ex-smokers which would not be expected to be superior to that for never smokers.

The literature posits that these kinds of selection biases and non-random missing data are typical of registry-based data due to the nature of both available information for these kinds of patients and the manner in which information is recorded[37,38]. In the case of missing stage data, it is unlikely that this is an artefact of recording methods in the registry, but rather that staging was simply not recorded for older patients with evident symptoms of cancer, with advanced cancer, or those who were not likely to not live long or were very frail and unsuitable for treatment. In these instances, clinicians may have opted not to perform biopsies or other investigations to confirm stage. With respect to smoking, available information entirely depends on whether the data is recorded in patient-clinician interactions. Whether or not clinicians ask will depend on a large extent on the individual. Interestingly, more laryngeal patients were missing smoking data in this population, which is atypical, as more often than not cancers more commonly associated with smoking tend to have available data. This may reflect a tendency to simply assume (correctly) based on the sub-site rather than record that these cancers are smoking related.

There are many ways to deal with non-random missing data and selection bias, including excluding all missing cases for relevant analyses, and statistical imputation (e.g. predicting the likely status of missing data points based on other variables for which data is not missing)[37,39–41]. Every method has advantages and disadvantages, but it is convention to simply include missing/unknown as a category in and of itself for variables for which data is missing in more than 10% of cases. The present analysis has done this and thus accounts for non-random missing data, making evident whether or not significance emanates from genuine relationships between variables or from missing data.

The present sample (n=861) population's distribution has thus been established. Many of its tumour and patient characteristics are reflective of that of the overall oropharyngeal, oral cavity, and laryngeal SCC population in Ireland between 1994 and 2013, which yields added value and significance for the results of the coming analysis. Additionally, the correlations between the patient and tumour characteristics foreshadow both potential confounding that will likely be adjusted for in relevant multivariate analyses, but also suggest differential characteristics amongst cancers of different sub-sites, most especially those of the oropharynx in relation to HPV. On the basis of the current population therefore, the coming

Chapters assess the role of HPV status in oropharyngeal, oral cavity, and laryngeal SCC with respect to epidemiological statistics, risk factors, treatment, and survival.

References

1. National Cancer Registry of Ireland. *Cancer Trends: Cancers of the Head and Neck*. (2014).
2. National Cancer Registry Ireland. National Cancer Registry Ireland. *National Cancer Registry* (2019). Available at: https://www.ncri.ie/. (Accessed: 30th April 2019)
3. National Cancer Registry of Ireland. *Cancer Trends: Cancers of the Head and Neck*. (2011).
4. Chaturvedi, A. K. *et al.* Worldwide Trends in Incidence Rates for Oral Cavity and Oropharyngeal Cancers. *J. Clin. Oncol.* **31**, 4550–4559 (2013).
5. Junor, E. J., Kerr, G. R. & Brewster, D. H. Oropharyngeal cancer. Fastest increasing cancer in Scotland, especially in men. *BMJ* **340**, c2512 (2010).
6. Wittekindt, C. *et al.* Increasing Incidence rates of Oropharyngeal Squamous Cell Carcinoma in Germany and Significance of Disease Burden Attributed to Human Papillomavirus. *Cancer Prev. Res.* (2019). doi:10.1158/1940-6207.CAPR-19-0098
7. Carlander, A.-L. F. *et al.* Continuing rise in oropharyngeal cancer in a high HPV prevalence area: A Danish population-based study from 2011 to 2014. *Eur. J. Cancer* **70**, 75–82 (2017).
8. Blomberg, M., Nielsen, A., Munk, C. & Kjaer, S. K. Trends in head and neck cancer incidence in Denmark, 1978-2007: Focus on human papillomavirus associated sites. *Int. J. Cancer* **129**, 733–741 (2011).
9. American Cancer Society. Larynx Statistics. *American Cancer Society Statistics Center* (2019). Available at: https://cancerstatisticscenter.cancer.org/?_ga=2.85526873.1727136116.1562600149-1669060168.1562600149#!/cancer-site/Larynx. (Accessed: 8th July 2019)
10. American Cancer Society. Oral cavity and pharynx Statistics. *American Cancer Society Statistics Center* (2019). Available at: https://cancerstatisticscenter.cancer.org/?_ga=2.85526873.1727136116.1562600149-1669060168.1562600149#!/cancer-site/Oral cavity and pharynx. (Accessed: 8th July

2019)

11. Woods, R. Human Papillomavirus-related Oropharyngeal Squamous Cell Carcinoma – Prevalence and Incidence in a Defined Population and Analysis of the Expression of Specific Cellular Biomarkers. (2016).
12. Marur, S., D'Souza, G., Westra, W. H. & Forastiere, A. A. HPV-associated head and neck cancer: a virus-related cancer epidemic. *Lancet Oncol.* **11**, 781–789 (2010).
13. Smith, E. M. *et al.* Age, sexual behavior and human papillomavirus infection in oral cavity and oropharyngeal cancers. *Int. J. Cancer* **108**, 766–772 (2004).
14. Fakhry, C. *et al.* Improved Survival of Patients With Human Papillomavirus-Positive Head and Neck Squamous Cell Carcinoma in a Prospective Clinical Trial. *JNCI J. Natl. Cancer Inst.* **100**, 261–269 (2008).
15. Krane, J. F. Role of Cytology in the Diagnosis and Management of HPV-Associated Head and Neck Carcinoma. *Acta Cytol.* **57**, 117–126 (2013).
16. Hafkamp, H. C. *et al.* Marked differences in survival rate between smokers and nonsmokers with HPV 16-associated tonsillar carcinomas. *Int. J. Cancer* **122**, 2656–2664 (2008).
17. Paz, I. B., Cook, N., Odom-Maryon, T., Xie, Y. & Wilczynski, S. P. Human papillomavirus (HPV) in head and neck cancer. An association of HPV 16 with squamous cell carcinoma of Waldeyer's tonsillar ring. *Cancer* **79**, 595–604 (1997).
18. Joo, Y.-H. *et al.* High-risk human papillomavirus and cervical lymph node metastasis in patients with oropharyngeal cancer. *Head Neck* **34**, 10–14 (2012).
19. American Joint Committee on Cancer & American Cancer Society. *AJCC Cancer Staging Manual*. (Springer Publishing, 2016).
20. Mirghani, H. & Blanchard, P. Treatment de-escalation for HPV-driven oropharyngeal cancer: Where do we stand? *Clin. Transl. Radiat. Oncol.* **8**, 4–11 (2018).
21. Eriksen, J. G. & Lassen, P. Human Papilloma Virus as a Biomarker for Personalized Head and Neck Cancer Radiotherapy. *Recent Results Cancer Res.* **198**, 143–161
22. Wierzbicka, M., Szyfter, K., Milecki, P., Składowski, K. & Ramlau, R. The rationale for HPV-related oropharyngeal cancer de-escalation treatment strategies. *Contemp. Oncol. (Poznan, Poland)* **19**, 313–22 (2015).
23. Ramsey, T. *et al.* Laryngeal cancer: Global socioeconomic trends in disease burden and smoking habits. *Laryngoscope* **128**, 2039–2053 (2018).

24. Layte Hannah McGee, R. & Rundle Gráinne Cousins Claire Donnelly Fiona Mulcahy Ronán Conroy, K. *The Irish Study of Sexual Health and Relationships*. (2006).
25. Central Office of Statistics. SAPS for Aggregate Urban Rural Areas. *Central Office of Statistics* (2019). Available at: https://www.cso.ie/en/census/census2006smallareapopulationstatisticssaps/sapsforaggregateurbanruralareas/. (Accessed: 3rd July 2019)
26. Woods, R. *et al.* Role of human papillomavirus in oropharyngeal squamous cell carcinoma: A review. *World J. Clin. cases* **2**, 172–93 (2014).
27. Sharp, L., McDevitt, J., Carsin, A.-E., Brown, C. & Comber, H. Smoking at diagnosis is an independent prognostic factor for cancer-specific survival in head and neck cancer: findings from a large, population-based study. *Cancer Epidemiol. Biomarkers Prev.* **23**, 2579–90 (2014).
28. Sommers, L. W. *et al.* Survival Patterns in Elderly Head and Neck Squamous Cell Carcinoma Patients Treated With Definitive Radiation Therapy. *Int. J. Radiat. Oncol.* **98**, 793–801 (2017).
29. Lyon, A. R. Heart failure resulting from cancer treatment: still serious but an opportunity for prevention. *Heart* **105**, 6–8 (2019).
30. Aleman, B. M. P. *et al.* Cardiovascular disease after cancer therapy. *Eur. J. Cancer Suppl.* **12**, 18–28 (2014).
31. Atienza, J. A. S. & Dasanu, C. A. Incidence of second primary malignancies in patients with treated head and neck cancer: a comprehensive review of literature. *Curr. Med. Res. Opin.* **28**, 1899–1909 (2012).
32. Chuang, S.-C. *et al.* Risk of second primary cancer among patients with head and neck cancers: A pooled analysis of 13 cancer registries. *Int. J. Cancer* **123**, 2390–2396 (2008).
33. Elicin, O. *et al.* Incidence of second primary cancers after radiotherapy combined with platinum and/or cetuximab in head and neck cancer patients. *Strahlentherapie und Onkol.* **195**, 468–474 (2019).
34. Wong, S. J., Heron, D. E., Stenson, K., Ling, D. C. & Vargo, J. A. Locoregional Recurrent or Second Primary Head and Neck Cancer: Management Strategies and Challenges. *Am. Soc. Clin. Oncol. Educ. B.* **36**, e284–e292 (2016).
35. Niu, Q. *et al.* The effect of marital status on the survival of patients with bladder

urothelial carcinoma: A SEER database analysis. *Medicine (Baltimore).* **97**, e11378 (2018).

36. Li, Y., Zhu, M. & Qi, S. Marital status and survival in patients with renal cell carcinoma. *Medicine (Baltimore).* **97**, e0385 (2018).
37. Pedersen, A. B. *et al.* Missing data and multiple imputation in clinical epidemiological research. *Clin. Epidemiol.* **9**, 157–166 (2017).
38. Hardy, S. E., Allore, H. & Studenski, S. A. Missing data: a special challenge in aging research. *J. Am. Geriatr. Soc.* **57**, 722–9 (2009).
39. Sterne, J. A. C. *et al.* Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* **338**, b2393 (2009).
40. Papageorgiou, G., Grant, S. W., Takkenberg, J. J. M. & Mokhles, M. M. Statistical primer: how to deal with missing data in scientific research?†. *Interact. Cardiovasc. Thorac. Surg.* **27**, 153–158 (2018).
41. Swalin, A. How to Handle Missing Data. *Towards Data Science* (2018). Available at: https://towardsdatascience.com/how-to-handle-missing-data-8646b18db0d4. (Accessed: 19th July 2019)

CHAPTER 7

PREVALENCE, GENOTYPE DISTRIBUTION, AND INCIDENCE OF, AND RISK FACTORS FOR HPV DNA IN OROPHARYNGEAL, ORAL CAVITY, AND LARYNGEAL CANCER IN IRELAND BETWEEN 1994 AND 2013

7 CHAPTER 7: PREVALENCE, GENOTYPE DISTRIBUTION, AND INCIDENCE OF, AND RISK FACTORS FOR HPV DNA IN OROPHARYNGEAL, ORAL CAVITY, AND LARYNGEAL CANCER IN IRELAND BETWEEN 1994 AND 2013

7.1 Introduction

HNSCC is the sixth most common type of cancer worldwide, accounting for an estimated 633,000 new cases diagnosed annually accompanied by 355,000 deaths[1]. According to the NCRI, HNSCC, including the 5% of HNCs not diagnosed as SCC, is the ninth most common cancer in Ireland[2]. It accounts for approximately 1.7% of all invasive cancers in women, and 3.9% of all invasive cancers in men[2]. Ireland's estimated incidence of HNSCC is 8.3 per 100,000, falling in the lower half of incidence rates in Europe[2].

While the incidence of other HNSCCs has decreased over the past two decades, correlating with decreased tobacco use, the age-adjusted incidence of OPSCC has been increasing over the same period[3–5]. The incidence of these increased by 2% to 3% annually from 1974 to 2001, then by 5.22% from 2000 to 2004 in the United States[6]. As a consequence, OSSC is now the most common HNC in the world[7]. In Ireland, incidence of OPSCCs, including tonsillar SCC, has increased in the last 30 years whilst most of those of the rest of the head and neck have seen either no change, or a decrease in their incidence[8].

Several carcinogenic risk factors have been suggested to explain these epidemiological trends, the most significant of which is HPV. The epidemiological data strongly supports the involvement of HPV as a carcinogen in HNSCC. The population-level incidence of HPV positive OPSCC in the United States increased by 225% between 1988 and 2004, with a concomitant decline of 50% for HPV-negative OPSCC[9]. The same trends are found across Europe and Australia[5,10–12]. The fact that HPV negative OPSCC incidence has declined whilst overall OPSCC incidence has increased in the same time period suggests the significant impact of HPV positive OPSCCs in driving the overall trend.

Prevalence statistics reflect this impact. The prevalence of HPV in these cancers is well-documented in numerous European, North American, and South Asian countries[9,10,13–15].

Prevalence statistics however vary greatly depending on the population and sub-sites sampled and the technologies used to detect the virus. Prevalence has been recorded as low as 18.5%[14] to as high as 90%[16] in Europe.

No matter the sub-site of the head and neck however, genotype distribution of the virus shows that HPV16 is overwhelmingly the most prevalent in these cancers[17–21]. HPV18 and HPV33 appear to account for less than 10% of cases[22]. This represents a distinct divergence from the genotype distribution of HPV in cervical cancers which, despite identifying HPV16 as the highest-risk genotype, estimate up to 20% prevalence of HPV18 and HPV45 respectively[23–30].

With respect to risk factors, HNSCC has traditionally been causally linked to smoking and alcohol consumption[31], poor oral hygiene[32], a diet low in fruit and vegetable consumption[33,34], and chronic inflammatory disease in the oral cavity[35,36]. Previous studies show clearly that there is a dose-response relationship between the onset of these cancers and the frequency and direction of tobacco and alcohol exposure[31]. Chronic exposure to these carcinogens, smoking and alcohol in particular, are well-established as precursors to dysplasia, pre-malignant lesions, and the eventual onset of cancer.

The anatomical subsite of presentation differs between HPV-related and HPV-unrelated tumours. Areas of the head and neck with invaginated reticulated squamous epithelium are particularly vulnerable to persistent HPV infection as a result of a naturally porous and disrupted basal cell layer[37]. Sites lined with this particular squamous epithelium are limited to the tonsil and the base of tongue. That HPV-related tumours are significantly associated with these particular sub-sites is therefore sensical[16,18,19,38,39]. Smoking- and alcohol-related tumours also occur in the tonsil and base of tongue but are often found further towards the front of the oral cavity, the larynx, and other oropharyngeal regions with unreticulated, non-porous epithelia including the posterior pharyngeal wall and the salivary glands.

Patients presenting with smoking and alcohol-related head and neck cancers are generally older, being diagnosed particularly in the seventh decade of life. By contrast, HPV-related OPSCC generally presents at a younger age, averaging a few years lower than HPV-negative

tumours[38]. With respect to sex, men are at three times the risk of women for developing head and neck cancers, HPV-related or otherwise[19,40–43].

Patients with immunodeficiency or HIV infection are at greater risk of both HPV-related and HPV-unrelated HNSCC[44–46]. Patients with previous radiation exposure and with Betel nut chewing habits are also at greater risk for HNSCCs. Occupational exposures such as leather dust or asbestos are also well-defined as increasing risk, along with several underlying genetic factors like Fanconi Aneamia[47].

Presentation and symptoms in HPV-unrelated HNSCCs and HPV-related HNSCCs differ significantly. HPV-related OPSCC generally presents with a more advanced clinical stage, with a higher nodal category. Furthermore, as opposed to most HNSCCs, HPV-positive HNSCCs have been strongly associated in comparison to HPV-negative HNSCCs with number of lifetime sexual partners, number of vaginal, oral, and anal sex partners, young age at first intercourse/earlier sexual contact, and history of sexually transmitted diseases, including genital warts[48–52].

Socio-economically, HNSCCs are associated with economic deprivation[53]. In the case of HPV-related OPSCCs, it is patients from higher socio-economic groups and who have a better performance status, that are at higher risk[54,55]. White males seem to be particularly at risk, with a rise in incidence reported in this group alone[9,10,56]. HPV-positive OPSCC has a lower incidence and prevalence in African-Americans than in other racial groups, with poorer survival in this racial group from OPSCC given that a higher proportion of OPSCC in this group is related to tobacco and alcohol exposure[57,58].

The roles of tobacco and alcohol exposure in HPV-related OPSCC and in oral HPV infection are uncertain based on the literature[59]. Some studies suggest a positive association by way of smoking-induced immunosuppression, leaving patients more vulnerable to HPV infection or an inability to clear the virus before persistent infection transforms cells. Others report a role for both in the potentiation of carcinogenesis, suggesting an additive and/or multiplicative role for smoking and alcohol in the genetic transformation of cells into

malignancies[49,59–62]. It is thus possible that tobacco exposure potentiates the effects of HPV carcinogenesis[63].

The data generated to establish the generally opposing risk factors for HPV-related and HPV-unrelated HNSCCs has been drawn mostly from North America and continental Europe. Currently, there is no nationwide Irish data on the subject. Aside from one study on 200 cases of oropharyngeal cancer alone[18], the corresponding prevalence, genotype, and incidence data for the Irish population does not yet exist in the literature. The variation of prevalence and incidence by subsite within the oral region is also unavailable. Furthermore, most epidemiological studies regarding the virus' role in HNCs utilize different HPV indicators, ranging from DNA to mRNA to viral load to integration, and employ varying technologies to detect each indicator. As a consequence, very little standardized and scientifically comparable data exists.

Establishing the population-level penetration and risk factors of the virus in HNCs in general, but also by individual sub-site, is particularly important given the potential HPV-related HNC's better survival[55,64–66] represents for the development of HPV-specific treatment types and decreased patient morbidity. De-escalation of treatment for HPV-related head and neck cancers for instance has been discussed for the last decade. The ability to increase quality and length of life through differential or de-escalated treatment for these patients is promising, but no definitive action can be taken without the appropriate epidemiological data.

Furthermore, the HPV vaccine is a highly cost-effective way to prevent the onset of cervical lesions[67]. However, whether or not precisely the same vaccine will function optimally to prevent head and neck cancers cannot be fully evaluated in the Irish context without the baseline prevalence, raw incidence, and genotype data. This is particularly relevant given the intended implementation of the nona-valent Gardasil 9 rather than the quadra-valent Gardasil vaccine in September 2019[68]. The recent recommendation by HIQA to include boys into the health system's nationwide HPV vaccination program[69,70] and the intended inculcation of boys into the national vaccination program in 2019[71,72] makes the

dissemination and use of this data even more urgent. It is on this basis that the current study determines its aims.

7.2 Aims

- To estimate the prevalence of HPV DNA positivity in archival tumour specimens from patients diagnosed with oropharyngeal, oral cavity, and laryngeal SCC in Ireland in the period between 1994 and 2013.
- To describe the genotype distribution in HPV positive tumours in this population.
- To estimate the raw incidence of HPV-positive and HPV-negative oropharyngeal, oral cavity, and laryngeal cancers.
- To identify patient (e.g. sex, age at diagnosis, smoking status, socio-economic status, geographic location, marital status) and clinical (e.g. stage, grade) factors associated with HPV-positivity for oropharyngeal, oral cavity, and laryngeal SCC.

7.3 Materials and Methods

7.3.1 Study Population

The population for this study was that of the ECHO study as a whole. It was through the NCRI's database that this study identified relevant specimens to create its own databank, as described in Sections 3.1, 3.2, and 3.3. Ethical approval for use of archival tissue specimens was obtained from all relevant local hospital ethics committees detailed in Section 3.3.3.

The population comprised 861 cases of newly-diagnosed (between 1994 and 2013), primary, invasive, oral (oropharyngeal, oral cavity, and largyngeal) SCC retrieved from 7 different hospitals sites. The ICD10 codes included in the study can be found in Section 3.3.1.

Analysis of the different sub-sites in the population was based on the same classifications of the entire study population in Chapter 6. The definitions of oropharyngeal, oral cavity, and laryngeal sub-sites by generalized ICD10 codes are summarized in Table 7.1.

Table 7.1 Summary of the ICD10 codes represented in the study population and the classification under which they were placed for the analysis.

ICD10 Codes	Classification
1.0, 2.4, 5.1, 5.2, 9.0-9.9, 10.0-10.9, 14.2	*Oropharynx*
2.0, 2.1, 2.2, 2.3, 2.8, 2.9, 3.0-3.9, 4.0-4.9, 5.0, 5.8, 5.9, 6.0, 6.1, 6.2, 6.9, 14.0, 14.8	*Oral Cavity*
32.0-32.9	*Larynx*

7.3.2 The Definition of an HPV-related Case

As dissected in Chapter 2, most studies define HPV positivity differently, deeming a HPV positive case based on different indicators of the virus' presence (e.g. DNA, mRNA, viral load, and viral integration) and using different technologies to detect these indicators. Each indicator is appropriate for particular contexts of defining 'HPV positivity' given their respective advantages and disadvantages. In the epidemiological context, studies have taken to using HPV DNA as it provides a good estimate of the prevalence of the virus in a population, regardless of its involvement or lack of involvement in the carcinogenesis of the associated tumour.

The lack of Irish data in the literature regarding HPV's role in oropharyngeal, oral cavity, and laryngeal cancer drives the epidemiological demands of this study. It is necessary to gauge the general presence of the virus in the Irish population first if any further conclusions based on other more functional indicators of the virus are to be as indicative as they can be. The study therefore requires a definition for HPV positivity indiscriminate of its provable relationship to associated tumours.

The study thus defines a HPV positive sample as any case the Multiplex PCR Luminex technology identifies as positive for HPV DNA. The technology is extremely sensitive, with a lowest limit of detection at 10 copies of the virus and it detects one of the most extensive ranges of HR and LR HPV genotypes of any available platform. As a consequence, it is an ideal method to achieve the epidemiological goals of the study.

7.3.3 HPV DNA Detection

Three 10um sections were cut using a microtome for each FFPE block associated with a particular case as described in Section 3.6.2. The sections were cut using the "IARC sterility protocol" denoted in Section 3.6.2.

DNA was then extracted from these sections in accordance with the steps outlined in Section 3.7.3.2. Extracted DNA was amplified with a Multiplex PCR detailed in 3.7.3.3. The PCR detected 19 HR or probably HR HPV genotypes (16, 18, 26, 31, 33, 35, 39, 45, 51, 53, 56, 58, 59, 66, 68a, 68b, 70, 73, 82) and two LR HPV genotypes (6, 11). Detection limits ranged from 10 to 1,000 copies of the viral genome per reaction. PCR products were then hybridized (Section 3.7.3.5) to oligonucleotide probes previously coupled to fluorescent beads (Section 3.7.3.4) and analysed by a Luminex 200 Analyser reporting on internal bead colour and Strep-PE reporter fluorescence. Results were expressed as the MFI of at least 100 beads per bead set and cut-offs were set as described in Section 3.7.3.5.

7.3.4 Patient Characteristics

The NCRI provided the ECHO study with all available and relevant patient characteristics using anonymized study numbers linking all HPV analysis performed to the national database. Some variables were adjusted for the purposes of the analysis. Table 7.2 exhibits an exhaustive list of the variables available for analysis and notes on any manner of adjustment.

Table 7.2 Variables made available by the NCRI for the population of the ECHO study and notes on any adjustments made for the purposes of the analysis.

Variable Code	Meaning	Variable Definition	Notes on Adjustments
SEX	Sex of patient	Sex of patient	N/A
AGE	Integer age at date of diagnosis	Integer age at date of diagnosis	Age was assessed both continuously and based on age younger than or equal to, and older than, age 50. Only continuous age was brought forward for multivariate analysis where relevant.
SMOKER_ID	Smoking status	Indication of current, ex-, or never-smoked behavior	N/A
GRADE	Grade of primary tumour	Poorly-, moderately, well-, or un-differentiated grade of tumour	Only 2 undifferentiated cases were detected in the population. These were excluded for all grade statistics generated to avoid skew in results.
T5	T stage	T category of stage (5th edition for cases diagnosed up to 2013) derived from best available clinical or pathological T data	Due to low frequencies for sub-stages, these were combined to yield the following T stage categories: T1, T2, T3, T4.
N5	N Stage	N category of stage (5th edition for cases diagnosed up to 2013) derived from best available clinical or pathological N data	Due to low frequencies for sub-stages, these were combined to yield the following N stage categories: N0, N1, N2, N3. N2 and N3 categories were also combined due to very low N3 frequencies.
M5	M Stage	M category of stage (5th edition for cases diagnosed up to 2013) derived from best available clinical or pathological M data	N/A
TNM5	TNM Stage	TNM stage (5th edition for cases diagnosed up to 2013) derived from best available clinical or pathological data	Due to low frequencies for sub-stages of Stage IV, TNM stages were combined to yield the following categories: Stage I, II, III, IV.
COUNTY_RES	County of residence	County of residence of patient at time of diagnosis	Due to low frequencies for many counties, county was assessed based on both residence in counties with large urban centers (Dublin/Limerick/Cork) and residence in or outside Dublin.

DEPRIV_POBAL_2011	Socio-economic status/Social Deprivation Score	Pobal index of deprivation from 1 to 5 for 2011 patient's Electoral Division (ED) of residence at diagnosis re-expressed as quintiles of 2011 population	Social deprivation score was categorical on a scale of 1 to 5, with 5 being the most deprived. It was assessed both categorically and as a continuous variable.
MARITAL	Marital status	Indication of single, separated, widowed, or divorced status of patient	Separated and divorced status were combined due to similarity in classification and low numbers of divorced patients.

7.3.5 Statistical Analysis

The analysis was carried out using IBM® SPSS® Statistics Version 25, XLSTAT 2019.1.3, Joinpoint Regression 4.7.0, and Microsoft Excel Version 16.25. All statistical tests were performed for all oropharyngeal, oral cavity, and laryngeal cancer in the population and also for cases within each sub-site (oropharynx, oral cavity, and larynx) individually. For variables from which more than 10% of data was missing, "missing/unknown" was included as a category of its own as is convention in the literature to account or detect any bias responsible for significance. For those variables with between 0% to 10% missing data, cases with missing data were excluded for relevant analyses. The two cases for which grade was undifferentiated were also excluded due to extremely low frequencies. Association tests for discrete variables were performed using Chi-square and Fisher exact tests (where expected counts were less than 5) for independence. For continuous variables, T-tests and one-way ANOVAs or, if normality was violated, Mann-Whitney U and Kruskal-Wallis tests were performed. Raw incidence was represented by moving average, whereby incidence for a single year was the average of incidence for the year prior, the year itself, and the year forthcoming (e.g. incidence of 1995=(1994+1995+1996)/3) as is standard in the literature. Average annual percentage change was calculated using Joinpoint Regression. Univariate logistic regression, including OR calculation, was carried out to determine patient and tumour factors that were significantly associated with HPV-related and HPV-unrelated tumours. Multivariate logistic regression using varying combinations and sequences of variables was also employed to assess confounding and to further identify significant predictors of HPV positivity. All significant variables by univariate models were included in the initial multivariate model. The least significant predictor was then taken out, and the

model was run again. The least significant predictor was again taken out, and the model was run again. This continued until all variables remaining in the model proved significantly predictive of HPV status, or until taking another variable out rendered the model as a whole insignificant.

It should also be noted that where age was assessed by univariate analysis in both continuous form and by categorical variable (Age≤50), it was only used in its continuous form in multivariate analysis. A significance level of ≤0.05 was used for all tests which were also all two-sided.

All HPV positive cases, no matter their genotype, were analyzed for prevalence and genotype distribution statistics. Comparisons between sub-site and HPV status and genotype distribution, incidence, and all risk factor assessment solely included HR HPV cases, with the 2 LR cases present considered negative due to their lack of carcinogenic potential.

7.4 The Prevalence and Genotype Distribution of HPV DNA in Oropharyngeal, Oral Cavity, and Laryngeal Cancer in Ireland between 1994 and 2013

To begin, prevalence and genotype distribution for all oropharyngeal, oral cavity, and laryngeal cancer, and within its sub-sites, was determined using simple proportions. The relationship between prevalence and sub-site, and genotype distribution and sub-site was then conducted using Chi-square and Fisher's exact tests where appropriate.

7.4.1 Oropharyngeal, Oral Cavity, and Laryngeal Cancer

The overall HPV DNA prevalence detected in the oropharyngeal, oral cavity, and laryngeal cancer population was 17.1% (147/861) (CI: 14.6, 19.6) as represented in Figure 7.1 below.

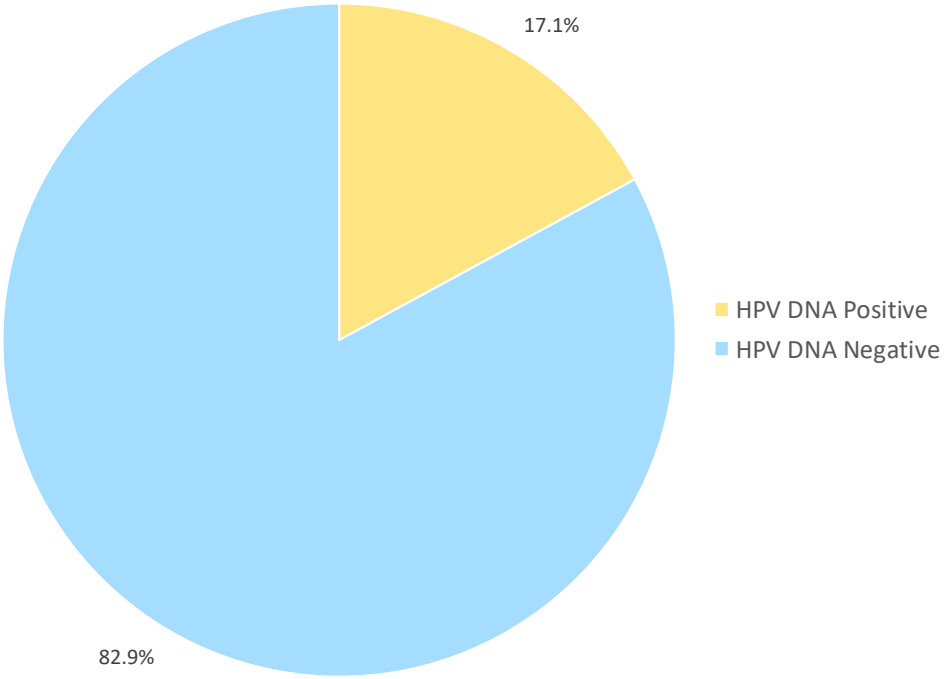

Figure 7.1 HPV DNA prevalence in oropharyngeal, oral cavity, and laryngeal cancer cases diagnosed in Ireland between 1994 and 2013 (n=861).

Amongst these HPV DNA positive cases (n=147), genotype distribution was as showcased in Figure 7.2 below.

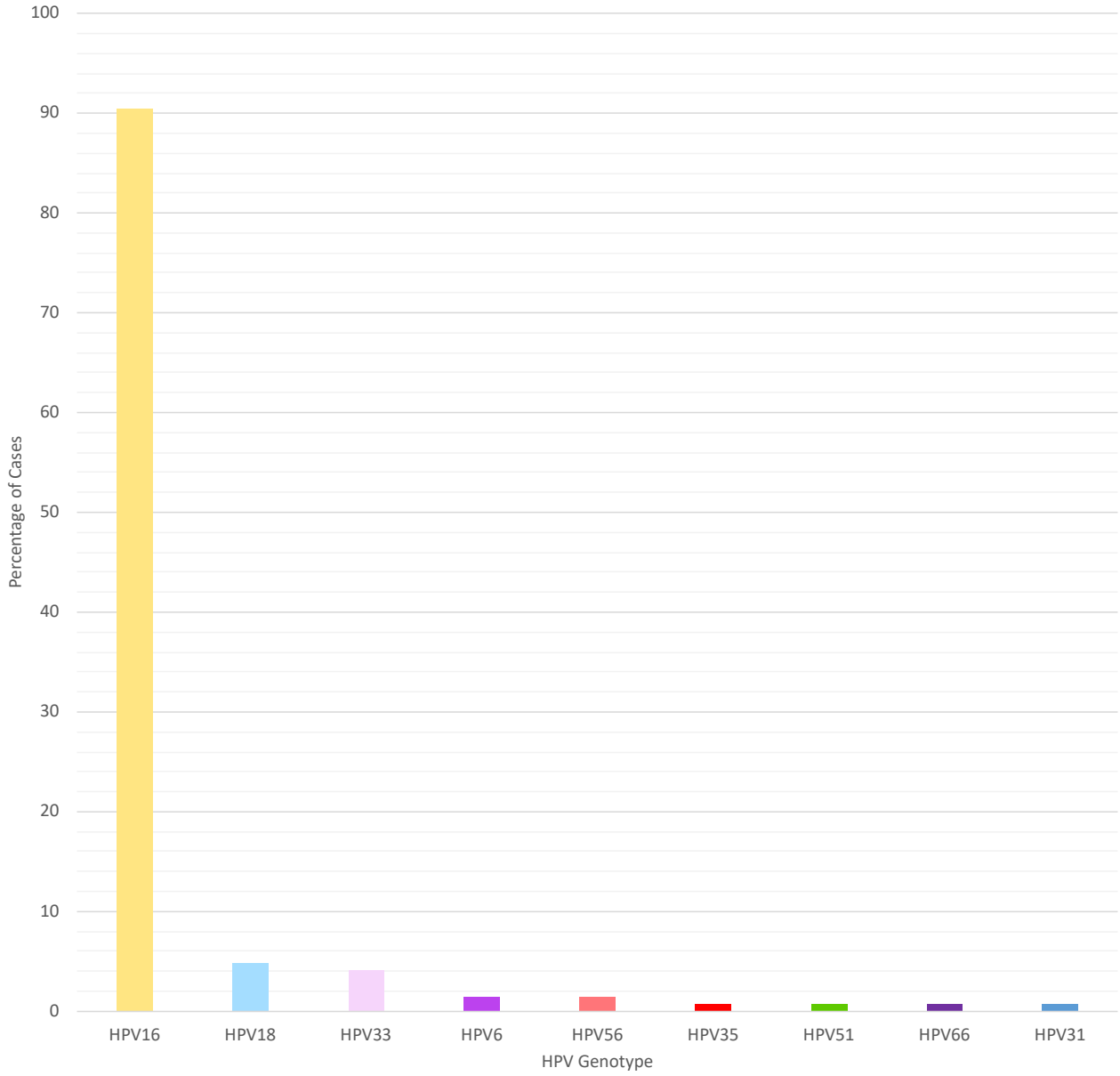

Figure 7.2 Distribution of HPV DNA genotypes in the population of HPV DNA positive cases of oropharyngeal, oral cavity, and laryngeal cancer diagnosed in Ireland between 1994 and 2013 (n=147).

Some HPV DNA positive cases in the population were dually infected with more than one genotype. Details of these cases can be found in Table 7.3.

Table 7.3 HPV DNA positive cases of oropharyngeal, oral cavity, and laryngeal cancer identified as having dual infections.

Cases are designated by anonymous case letters.

Anonymous Case Letter	Dual Infection Genotypes
A	16, 18
B	16, 33
C	16, 18
D	16, 18
E	16, 56
F	16, 33
G	16, 33

7.4.2 Oropharyngeal Cancer

The corresponding HPV DNA prevalence statistic in the oropharyngeal site was 41.1% (86/209) (CI: 34.5, 47.8). Figure 7.3 graphically represents this finding.

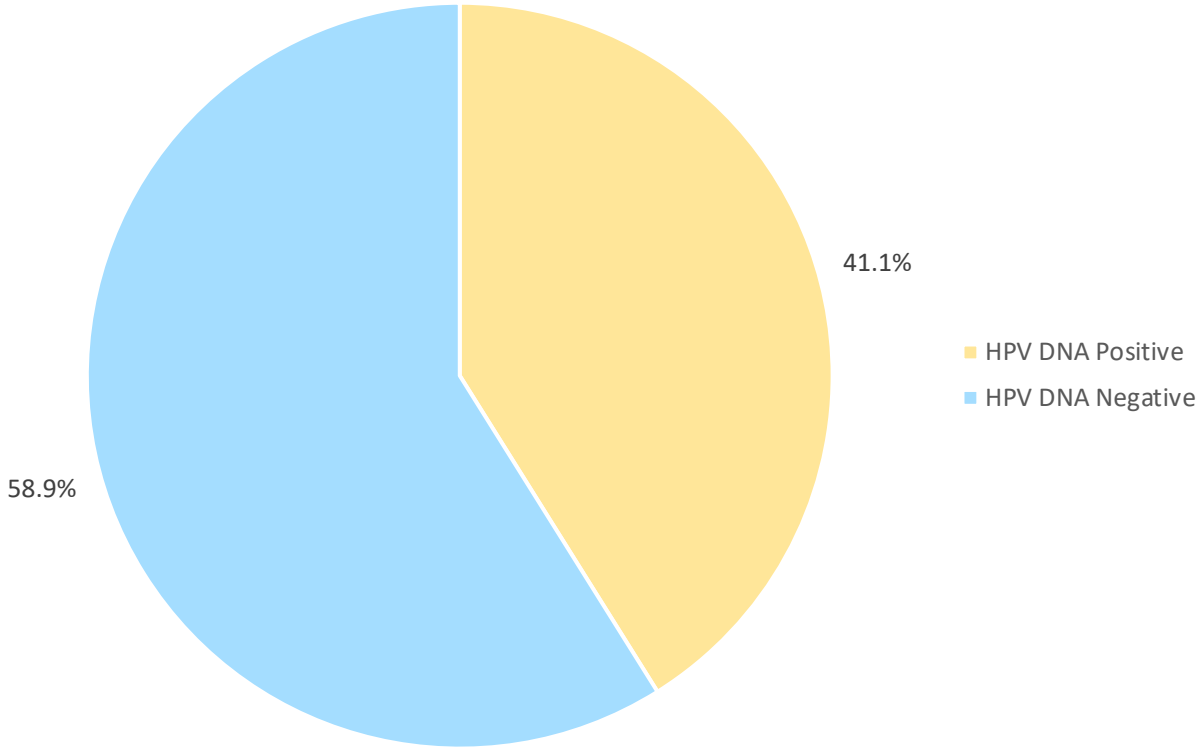

Figure 7.3 HPV DNA prevalence in oropharyngeal cancer cases diagnosed in Ireland between 1994 and 2013 (n=209).

The genotype distribution within the oropharyngeal sub-site for HPV DNA positive cases (n=86) is represented in Figure 7.4.

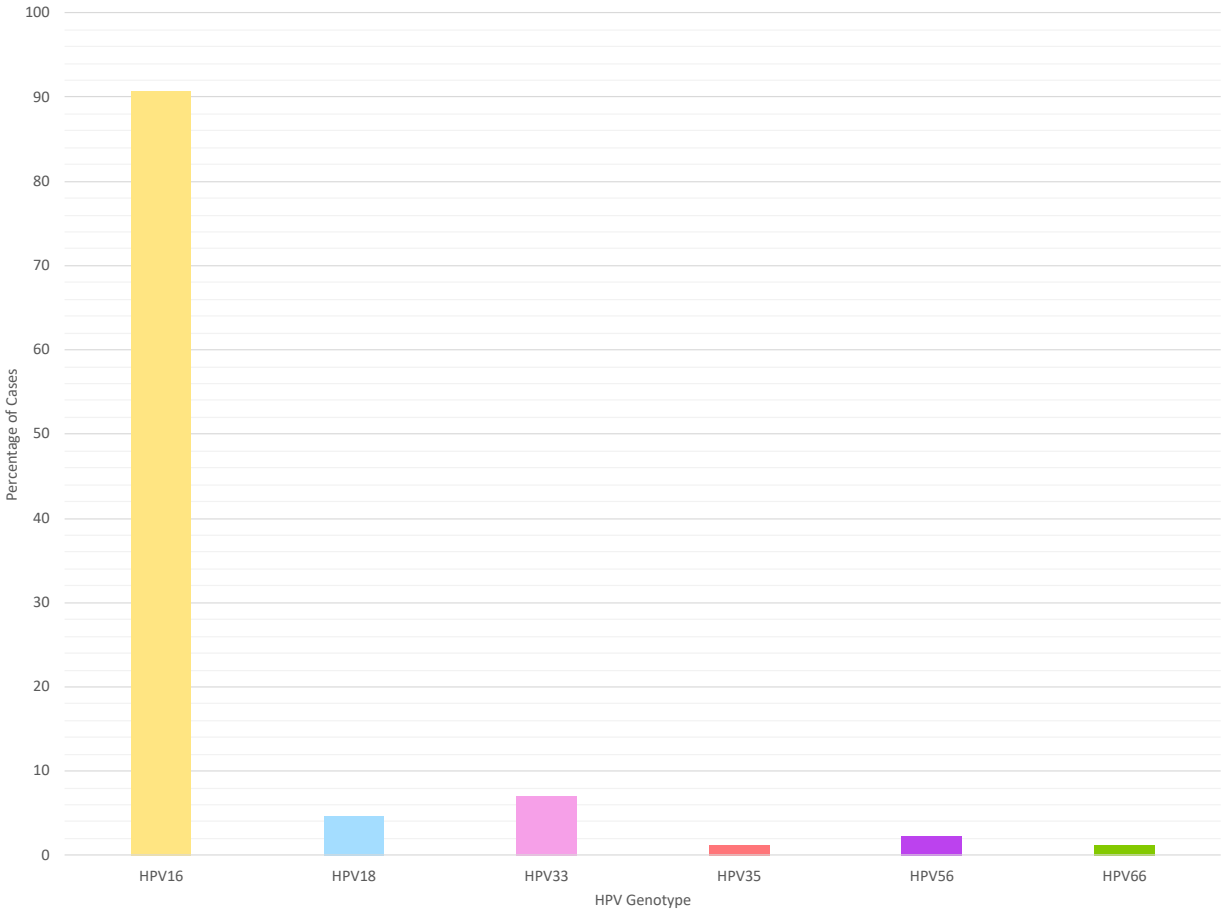

Figure 7.4 Distribution of HPV DNA genotypes in the population of HPV DNA positive cases of oropharyngeal cancer diagnosed in Ireland between 1994 and 2013 (n=86).

Some HPV DNA positive cases in the oropharynx were dually infected. Based on the same anonymous case letters described in Table 7.3, Table 7.4 summarizes these dual infections.

Table 7.4 HPV DNA positive cases of oropharyngeal cancer identified as having dual infections.

Cases are designated by anonymous case letters.

Anonymous Case Letter	Dual Infection Genotypes
A	16, 18
B	16, 33
C	16, 18
E	16, 56
F	16, 33
G	16, 33

7.4.3 Oral Cavity Cancer

HPV DNA prevalence in the oral cavity was 10.9% in this population (36/331) (CI: 7.5, 14.2). Figure 7.5 graphically represents this finding.

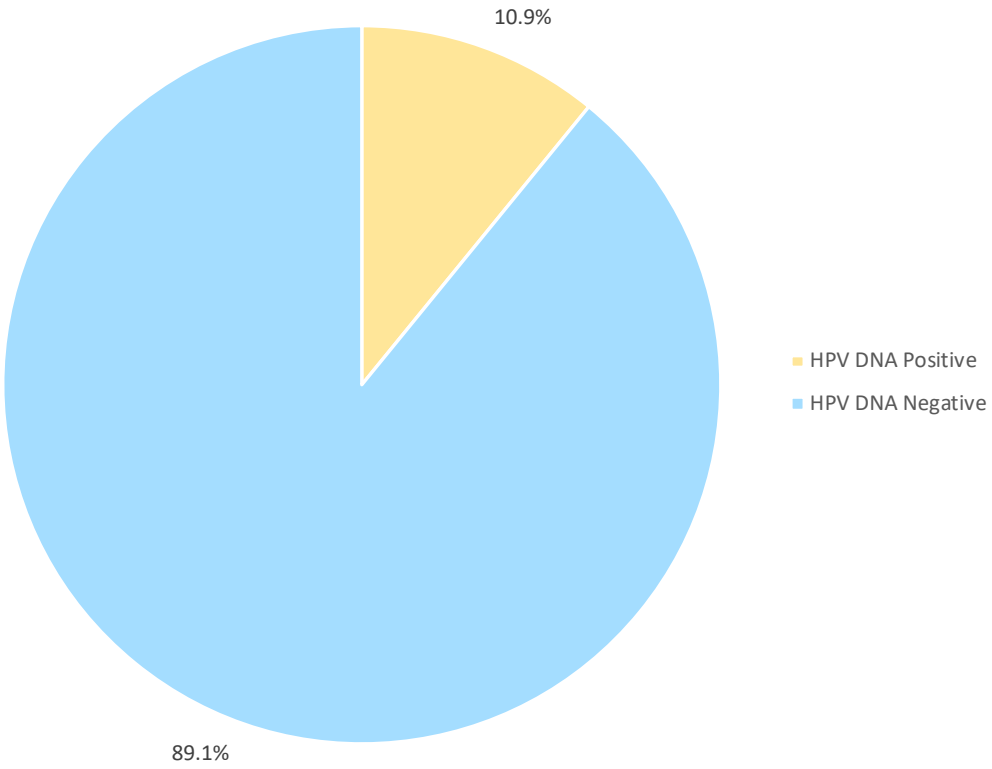

Figure 7.5 HPV DNA prevalence in oral cavity cancer cases diagnosed in Ireland between 1994 and 2013 (n=331).

Amongst HPV DNA positive cases (n=36) in this population, the genotype distribution is shown in Figure 7.6.

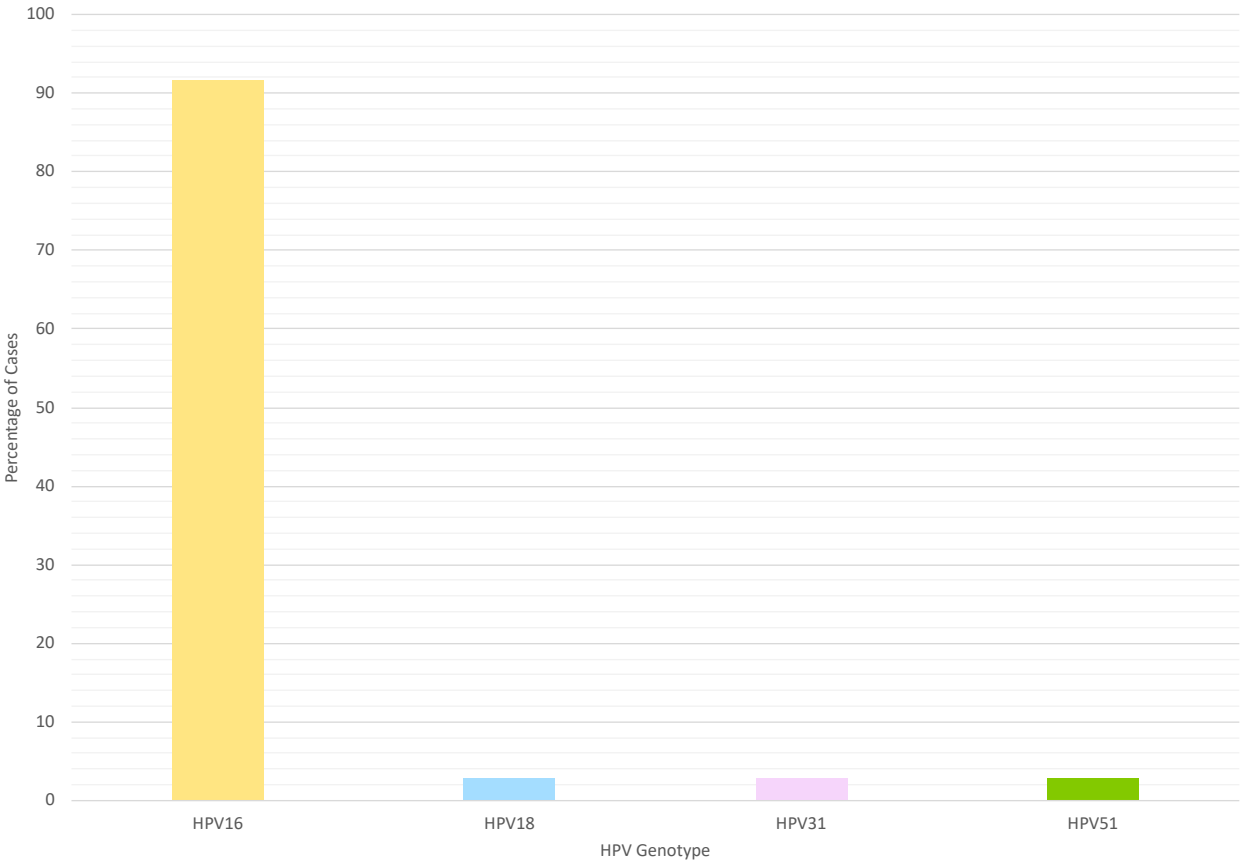

Figure 7.6 Distribution of HPV DNA genotypes in the population of HPV DNA positive cases of oral cavity cancer diagnosed in Ireland between 1994 and 2013 (n=36).

There were no dually-infected cases in the oral cavity.

7.4.4 Laryngeal Cancer

HPV DNA prevalence in the laryngeal sub-site was 7.8% (25/321) (CI: 4.9, 10.7). Figure 7.7 graphically represents this finding.

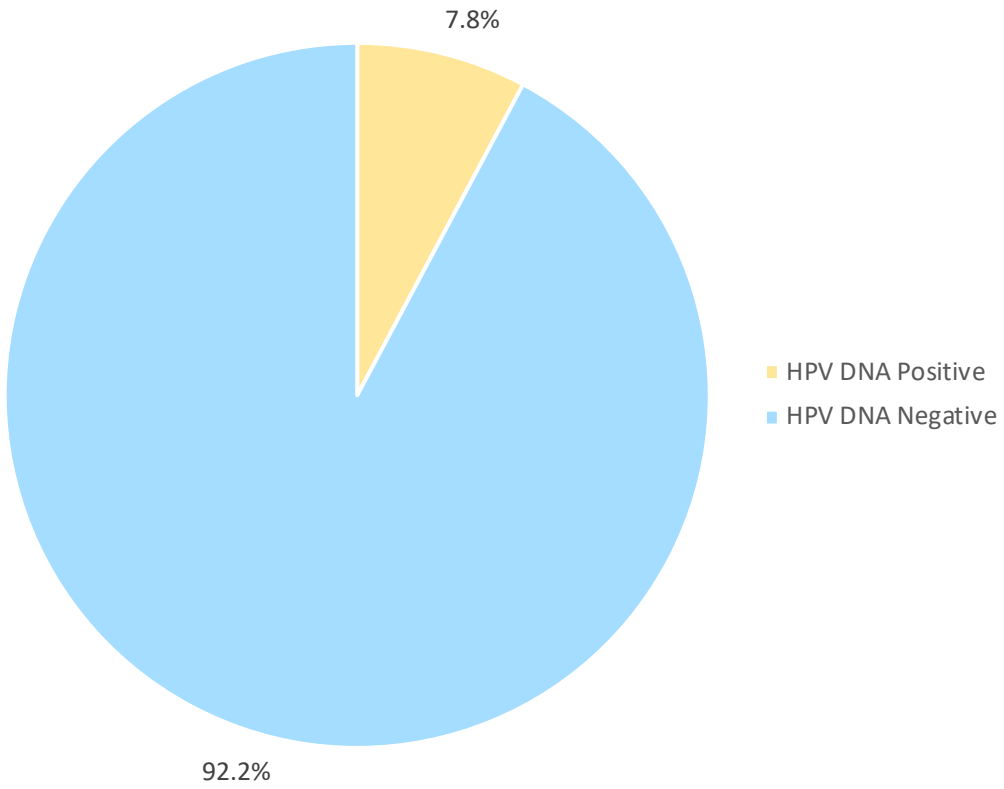

Figure 7.7 HPV DNA prevalence in laryngeal cancer cases diagnosed in Ireland between 1994 and 2013 (n=321).

The genotype distribution amongst HPV DNA positive cases (n=25) is detailed in Figure 7.8.

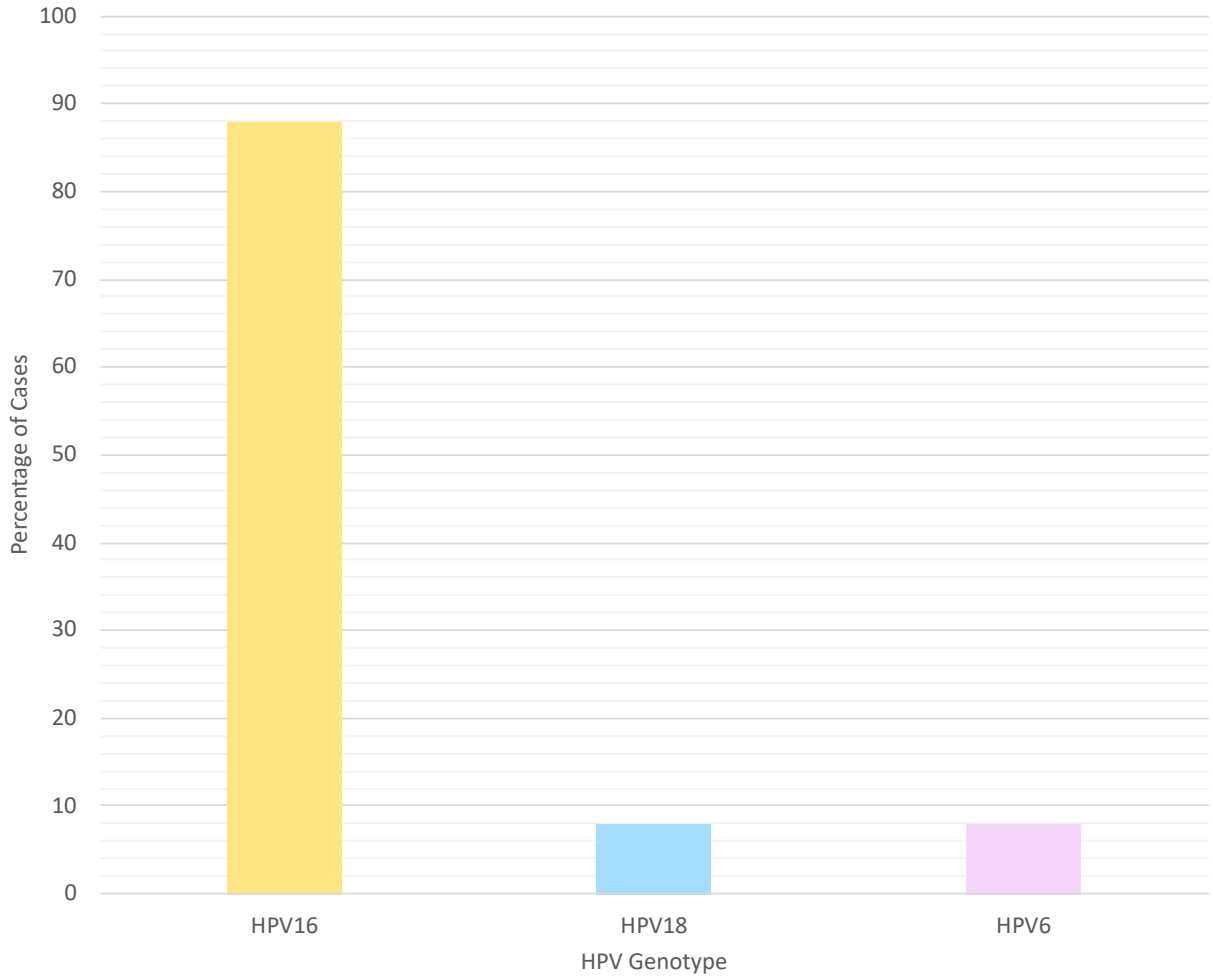

Figure 7.8 Distribution of HPV DNA genotypes in the population of HPV DNA positive cases of laryngeal cancer diagnosed in Ireland between 1994 and 2013 (n=25).

One single dually infected case arose in the larynx, as summarized in Table 7.5.

Table 7.5 HPV DNA positive cases of laryngeal cancer identified as having dual infections.

Cases are designated by anonymous case letters.

Anonymous Case Letter	Dual Infection Genotypes
D	16, 18

7.4.5 Prevalence and Genotype Comparison Between Sub-sites

The different HPV DNA prevalences detected for all cancers and within each subsite are shown in Table 7.6 and Figure 7.9 below.

Table 7.6 HPV DNA prevalence for oropharyngeal, oral cavity, and laryngeal cancer diagnosed in Ireland between 1994 and 2013.

Sub-site	Fraction	Prevalence
Oropharynx	86/209	41.1%
Oral Cavity	36/331	10.9%
Larynx	25/321	7.8%
All	147/861	17.1%

To assess the relationship between HR carcinogenic HPV status (thus considering the 2 cases of LR HPV in the population as HR HPV negative) and sub-site, Table 7.7 was generated and analyzed.

Table 7.7 Relationship between oropharyngeal, oral cavity, and laryngeal cancer sub-site and HR HPV status (n=861).

Sub-site	Count	HR HPV Status		Total
		HR HPV-	*HR HPV+*	
Oropharynx	*Count*	123	86	209
	% within Sub-site	58.9%	41.1%	100.0%
	% within HR HPV status	17.2%	59.3%	24.3%
	% of Total	14.3%	10.0%	24.3%
Oral Cavity	*Count*	295	36	331
	% within Sub-site	89.1%	10.9%	100.0%
	% within HR HPV status	41.2%	24.8%	38.4%
	% of Total	34.3%	4.2%	38.4%
Larynx	*Count*	298	23	321
	% within Sub-site	92.8%	7.2%	100.0%
	% within HR HPV status	41.6%	15.9%	37.3%
	% of Total	34.6%	2.7%	37.3%
Total	*Count*	716	145	861
	% within Sub-site	83.2%	16.8%	100.0%
	% within HR HPV status	100.0%	100.0%	100.0%
	% of Total	83.2%	16.8%	100.0%

Chi-square analysis revealed a significant association on the basis of Table 7.7 between cancer sub-site and HR HPV status (Chi-square=118.043, 2 d.f., $p<0.0001$). This was due to disproportionate representation of HPV positive cases in the oropharynx. 59.3% of HR HPV positive cases originated in the oropharyngeal sub-site compared to 24.8% in the oral cavity and 15.9% in the larynx. Figure 7.9 illustrates these findings.

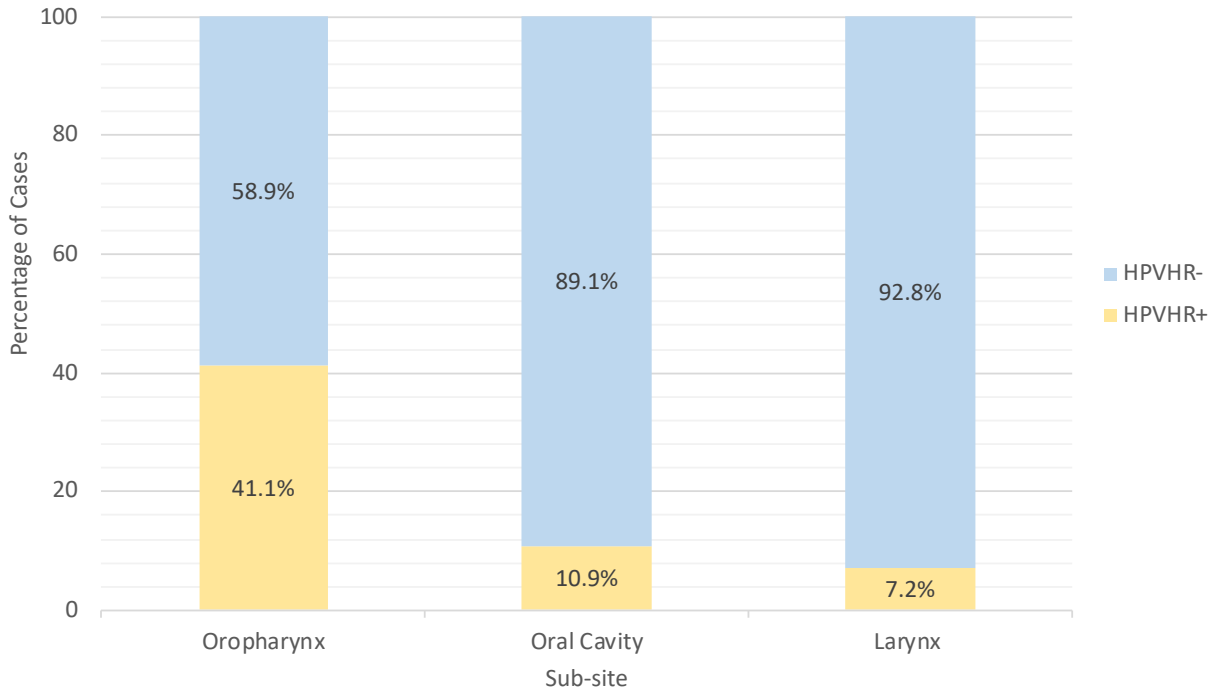

Figure 7.9 Percentage of cases HR HPV positive and negative in oropharyngeal, oral cavity, and laryngeal cancer by sub-site.

For oropharynx n=209, oral cavity n=331, and larynx n=321. For the total population, n=861.

A more detailed break-down of the sub-sites from which HR HPV prevalence emanated in the oropharynx is shown in Figure 7.10. The tonsillar sub-site (which included both the palatine and lingual tonsils) within the oropharynx had the highest HR HPV DNA prevalence, followed by the base of tongue and the rest of the oropharynx. The tonsil was thus the principle driver of overall HPV DNA prevalence reported in the oropharynx (41.1%).

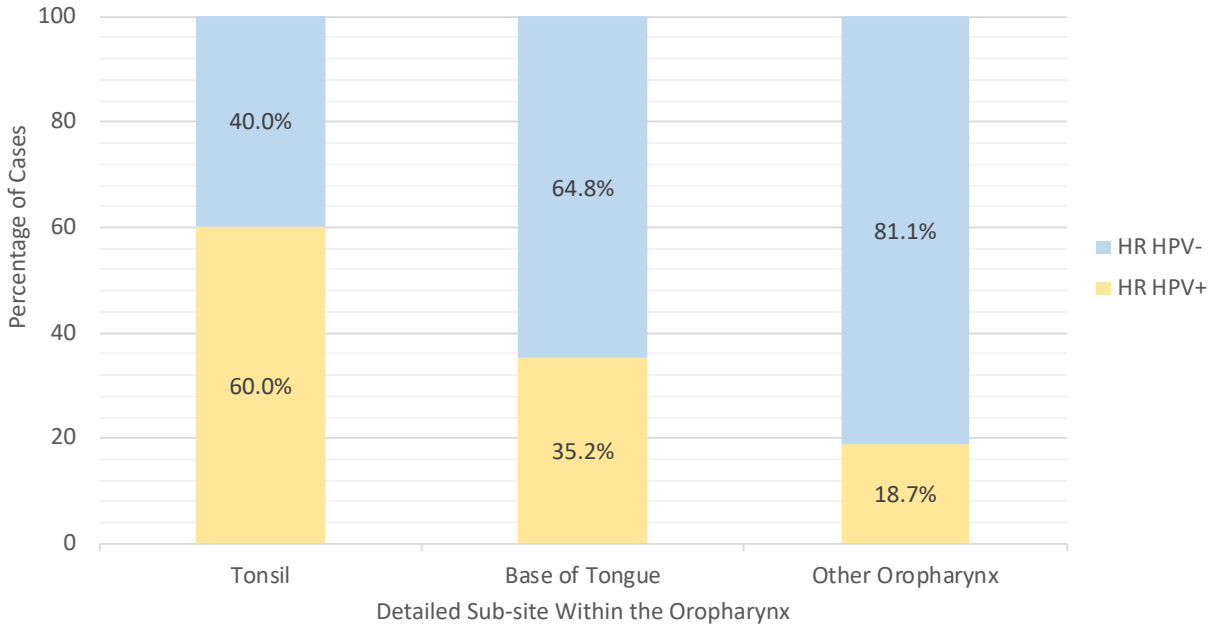

Figure 7.10 Percentage of cases HR HPV positive and negative in detailed sub-sites of the oropharynx alone.

Tonsillar sub-site here included both lingual and palatine tonsils. Other oropharynx included ICD10 codes 5.1, 5.2, and 10.0-10.9. For tonsil n=85, for base of tongue n=71, and for other oropharynx n=53.

To assess any relationship between cancer sub-site and HPV genotype distribution, Table 7.8 was generated. Though 147 cases were HPV positive (both HR and LR), dual infections were not combined, and the total population was thus 154 accounting for the 7 dual infections detected.

Table 7.8 Relationship between oropharyngeal, oral cavity, and laryngeal cancer sub-site and HPV genotype (n=154).

Sub-site	Count	HPV Genotype									Total
		6	*16*	*18*	*31*	*33*	*35*	*51*	*56*	*66*	
Oropharynx	*Count*	0	78	4	0	6	1	0	2	1	92
	% within Sub-site	0.0%	84.8%	4.3%	0.0%	6.5%	1.1%	0.0%	2.2%	1.1%	100.0%
	% within HR HPV status	0.0%	58.6%	57.1%	0.0%	100%	100%	0.0%	100%	100%	57.1%
	% of Total	0.0%	50.6%	2.6%	0.00%	3.9%	0.6%	0.00%	1.3%	0.6%	57.1%
Oral Cavity	*Count*	0	33	1	1	0	0	1	0	0	36
	% within Sub-site	0.0%	91.7%	2.8%	2.8%	0.0%	0.0%	2.8%	0.0%	0.0%	100.0%
	% within HR HPV status	0.0%	24.8%	14.3%	100%	0.0%	0.0%	100%	0.0%	0.0%	26.0%
	% of Total	0.0%	21.4%	0.6%	0.6%	0.0%	0.0%	0.6%	0.0%	0.0%	26.0%
Larynx	*Count*	2	22	2	0	0	0	0	0	0	26
	% within Sub-site	7.7%	84.6%	7.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%
	% within HR HPV status	100%	16.5%	28.6%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	16.9%
	% of Total	1.3%	14.3%	1.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	16.9%
Total	*Count*	2	133	7	1	6	1	1	2	1	154
	% within Sub-site	1.3%	86.4%	4.5%	0.65%	3.9%	0.65%	0.65%	1.3%	0.65%	100.0%
	% within HR HPV status	100%	100%	100%	100%	100%	100%	100%	100%	100%	100.0%
	% of Total	1.3%	86.4%	4.5%	0.6%	3.9%	0.6%	0.6%	1.3%	0.6%	100.0%

Indeed, on the basis of Table 7.8, there was no significant association between cancer sub-site and HPV genotype (Fisher's exact: p=0.146). Given the very low number of observed and expected cases for most genotypes however, the analysis was performed excluding the two LR HPV genotypes, and with the following genotype categories: HPV16, HPV18, and all other HR HPV genotypes. There remained no significant association between sub-site and HPV genotype (Fisher's exact: p=0.367).

7.5 Raw Incidence for HPV-related and HPV-unrelated Oropharyngeal, Oral Cavity, and Laryngeal Cancer Diagnosed Between 1994-2013 in Ireland

Raw incidence figures for HPV positive and HPV negative cases were generated for all cancer in the population and for each sub-site. The two LR HPV cases detected were considered

negative given their lack of designation as carcinogenic. Any coming reference to HPV positivity thus refers to HR HPV positive cases alone.

7.5.1 Oropharyngeal, Oral Cavity, and Laryngeal Cancer

Figure 7.11 illustrates the raw incidence represented by moving average of oropharyngeal, oral cavity, and laryngeal cancer in Ireland between 1994 and 2013 overall, and also broken down by HPV status.

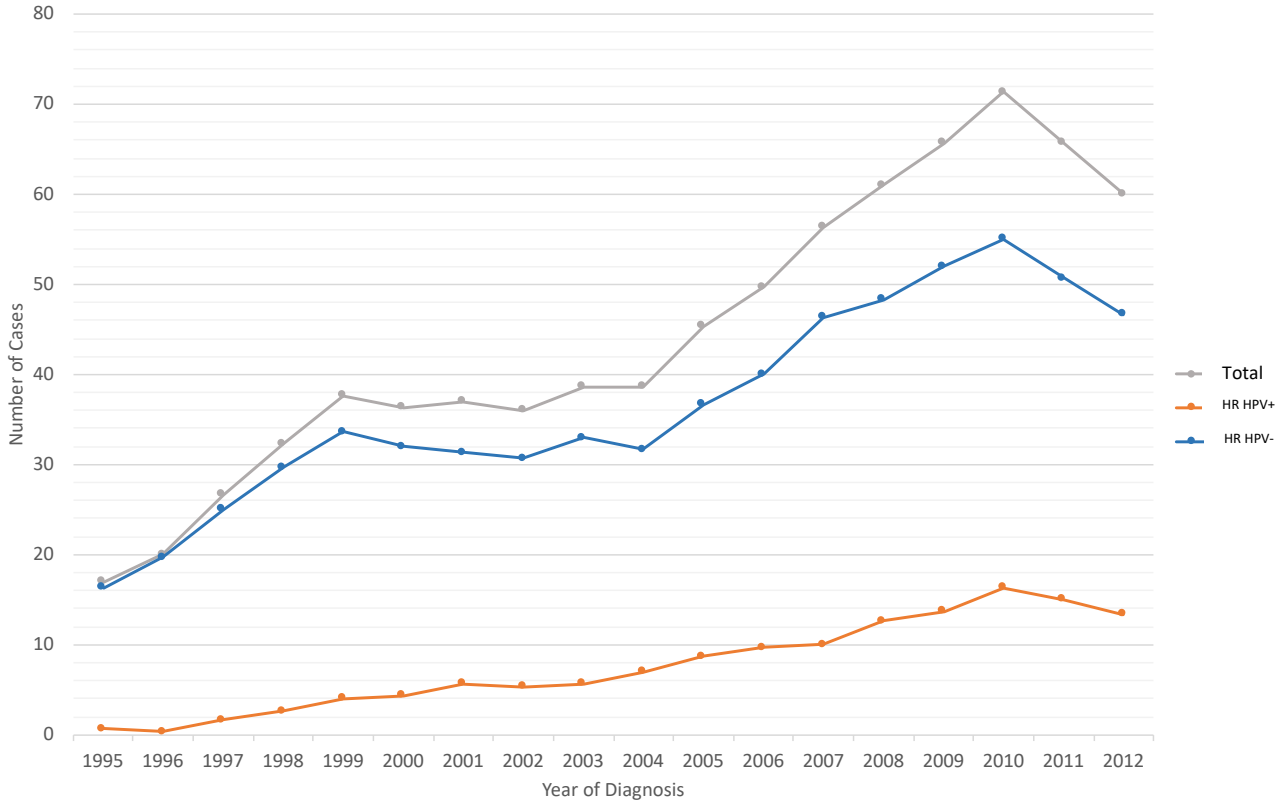

Figure 7.11 Raw incidence represented by moving average of oropharyngeal, oral cavity, and laryngeal cancer diagnosed between 1994 and 2013 in Ireland sub-divided by HPV status.

For total n=861, for HR HPV+ n=145, and for HR HPV- n=716.

The Figure indicates that the majority of the trend visible is attributable to HR HPV negative cancer, though HR HPV positive cancer appears to play a more significant role after 2004. The average annual percentage change for all oropharyngeal, oral cavity, and laryngeal SCC was 4.9 (CI: -1.3, 11.4)(p=0.300). For HPV-related cases it was 8.7 (CI: -6.3, 26.2)(p=0.300) and for HPV-unrelated cases it was 7.4 (CI: 3.2, 11.8)(p<0.0001).

7.5.2 Oropharyngeal Cancer

Figure 7.12 illustrates the raw incidence represented by moving average of oropharyngeal cancer in Ireland between 1994 and 2013 overall and broken down by HPV status.

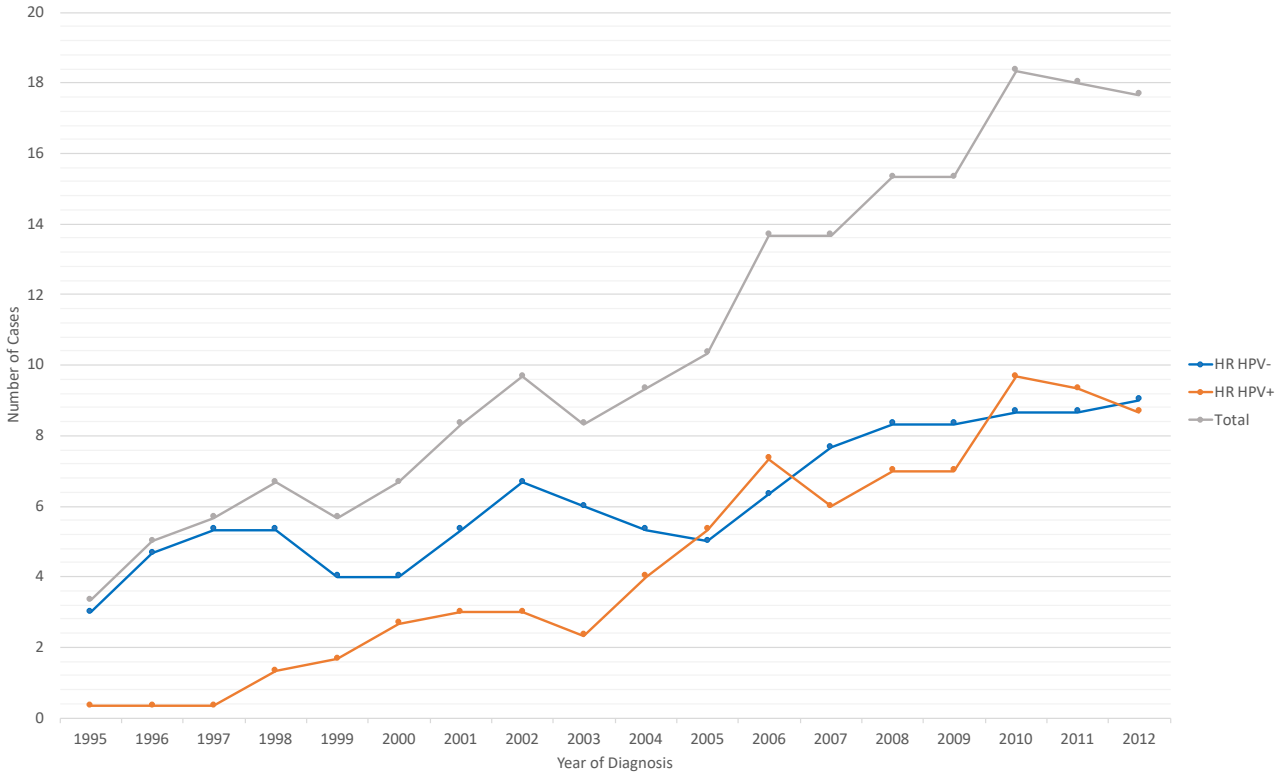

Figure 7.12 Raw incidence of oropharyngeal cancer diagnosed between 1994 and 2013 in Ireland sub-divided by HPV status represented by moving average.

For total n=209, for HR HPV+ n=86, and for HR HPV- n=123.

The Figure indicates that HR HPV negative cases comprised the majority of all oropharyngeal cases before 2004, with only one HPV positive case detected before 1998. After 2004 however, HR HPV positive cases progressively became an equal if not more significant contributor to the overall trend. The average annual percentage change for all oropharyngeal cases was 9.4 (CI: 5.6, 13.4)($p<0.0001$). For HPV-related cases it was 16.4 (CI: 11.0, 22.1)($p<0.0001$) and for HPV-unrelated cases it was 5.6 (CI: 2.1, 9.3)($p<0.0001$). The average proportion of all OPSCC attributable to HPV-related cases between 1994 and 2003 was 20.6%, a proportion that jumped to 47.4% between 2004 and 2013.

7.5.3 Oral Cavity Cancer

Figure 7.13 represents the raw incidence represented by moving average of oral cavity cancer in Ireland between 1994 and 2013 overall and broken down by HPV status.

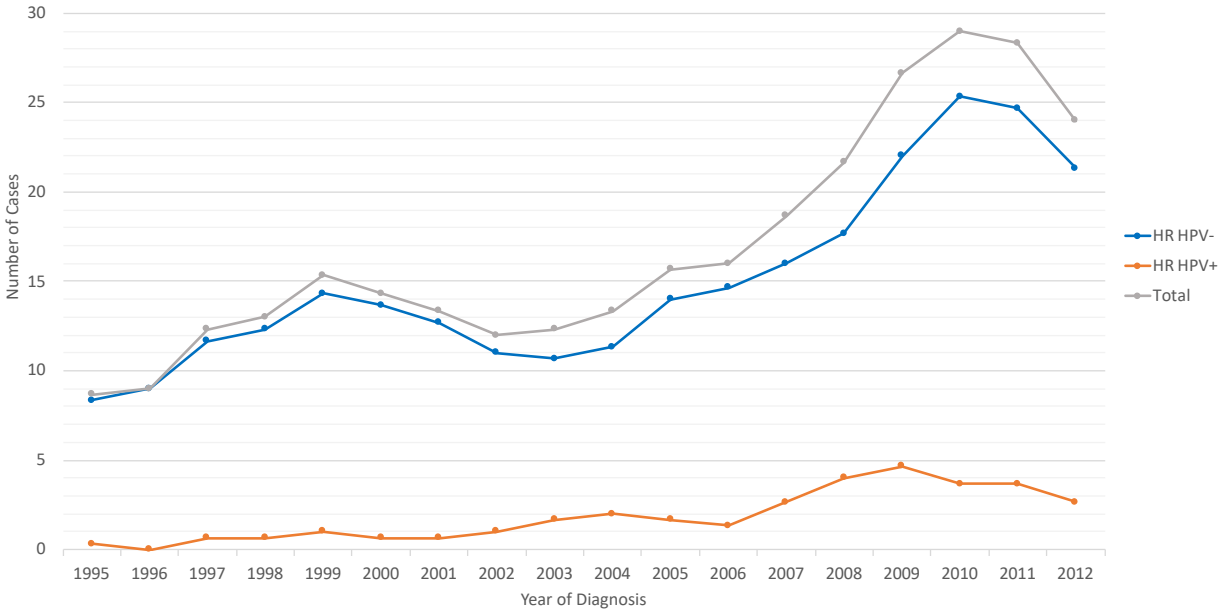

Figure 7.13 Raw incidence represented by moving average of oral cavity cancer diagnosed between 1994 and 2013 in Ireland sub-divided by HPV status.

For total n=331, for HR HPV+ n=36, and for HR HPV- n=295.

The Figure resoundingly suggests that despite a minor increase in HR HPV positive cases after 2004, HR HPV negative cases have historically and presently been the overwhelming majority of all oral cavity cases since 1994. The annual average percentage change for all oral cavity SCC was 6.6% (CI: 4.3, 6.3)($p<0.0001$). For HPV-related cases it was 9.9% (CI: 4.4, 15.7)($p<0.0001$) and for HPV-unrelated cases it was 5.9% (CI: 3.6, 8.3)($P<0.0001$). The proportion of all OSCC attributable to HPV-related cases never surpassed 22.8%.

7.5.4 Laryngeal Cancer

Figure 7.14 showcases the raw incidence represented by moving average of laryngeal cancer diagnosed between 1994 and 2013 in Ireland overall and broken down by HPV status.

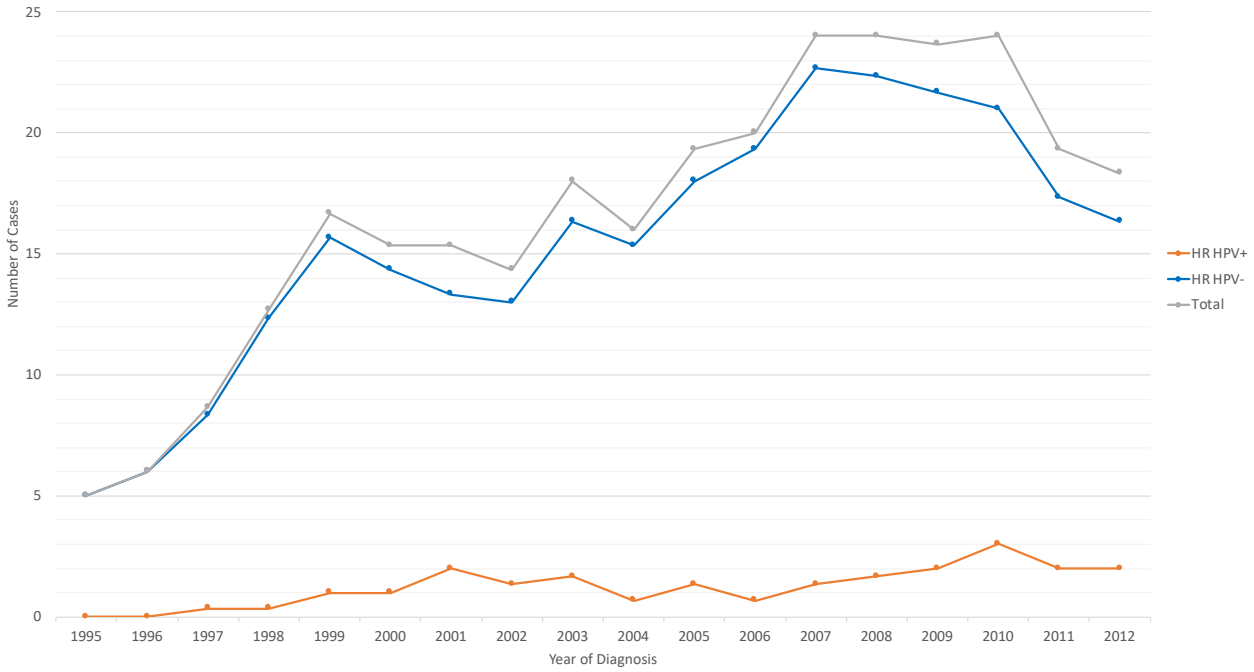

Figure 7.14 Raw incidence represented by moving average of laryngeal cancer diagnosed between 1994 and 2013 in Ireland sub-divided by HPV status (n=321).

The Figure, much like 7.13 for oral cavity cases, indicates that HR HPV negative cases have formed almost all of laryngeal cases diagnosed between 1994 and 2013 in Ireland. HPV positive cases, though slightly more prevalent since 2007, continue to be a relatively insignificant contributor to overall trends. The average annual percentage change in LSCC overall was 7.3% (CIL: 3.6, 11.2)($p<0.0001$). For HPV-related cases the average percentage change was 5.5% (CI: -0.5, 11.9)($p=0.100$) and for HPV-unrelated cases it was 7.0% (CI: 3.3, 10.7)($p<0.0001$). The proportion of LSCC attributable to HPV positive cases never surpassed 17% between 1994 and 2013, remaining at 6.25% in 2013.

7.6 Risk Factors for HPV-related and HPV-unrelated Oropharyngeal, Oral Cavity, and Laryngeal Cancer Diagnosed between 1994-2013 in Ireland

To assess risk factors for HPV-related and HPV-unrelated oropharyngeal, oral cavity, and laryngeal cancer and cancer within the sub-sites, Chi-square, Fisher's exact, T-test, Mann-Whitney, ANOVA, and Kruskal-Wallis analyses were conducted where appropriate using the available variables in Table 7.2. Tables and Figures representing the distribution of cases for each variable by HPV status were only presented for those variables which had a significant

relationship to HPV status. For the total cancer population and cancer within each sub-site, univariate and multivariate logistic regression was then performed using all of these variables to determine significant predictors of HPV positivity. The 2 LR HPV positive cases were considered HPV negative due to their lack of carcinogenic potential.

7.6.1 Oropharyngeal, Oral Cavity, and Laryngeal Cancer

7.6.1.1 Sex

Sex data was available for all 861 patients. There was no significant association between sex and HPV status (Chi-square=1.501, 1 d.f., p=0.220).

7.6.1.2 Age at Diagnosis

Age at diagnosis data was available for 861 patients. Figure 7.15 exhibits age at diagnosis by HR HPV status. The mean age at diagnosis for HPV negative cases was 64.29 (CI: 63.45, 65.14), where it was 58.39 (CI: 56.52, 60.52) for HPV positive patients. The median age at diagnosis for HPV negative cases was 64.00 and for HPV positive cases was 57.00. There was a significant difference between the age at diagnosis for HPV positive and HPV negative cases with positive cases presenting at younger age (Mann-Whitney: p<0.0001).

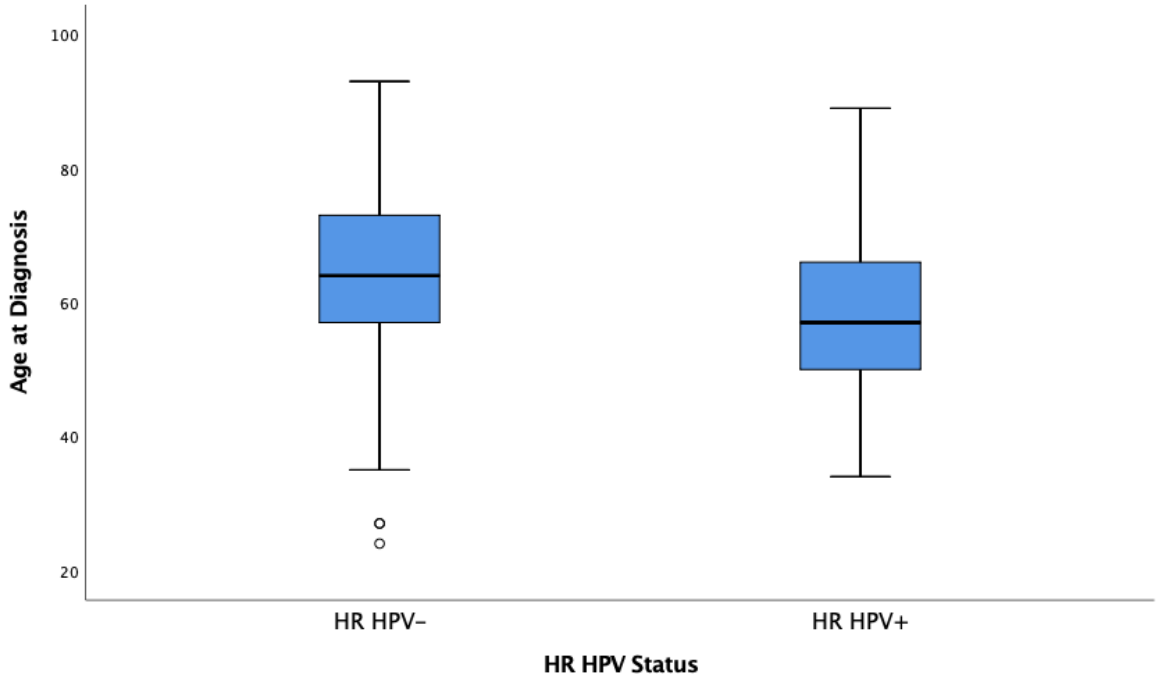

Figure 7.15 Age at diagnosis for oropharyngeal, oral cavity, and laryngeal cancer by HR HPV status.

For HR HPV+ n=145, and for HR HPV- n=716. For the total population, n=861.

Patients were then grouped into categories based on whether or not they were younger than or exactly age 50 at diagnosis, and older than age 50 at diagnosis.

Table 7.9 Relationship between age younger than or equal to 50, or older than 50, and HR HPV status for oropharyngeal, oral cavity, and laryngeal cancer (n=861).

Age	Count	HR HPV Status		Total
		HR HPV-	*HR HPV+*	
≤ 50	*Count*	81	40	121
	% within Age	66.9%	33.1%	100.0%
	% within HR HPV status	11.3%	27.6%	14.1%
	% of Total	9.4%	4.6%	14.1%
>50	*Count*	635	105	740
	% within Age	85.8%	14.2%	100.0%
	% within HR HPV status	88.7%	72.4%	85.9%
	% of Total	73.8%	12.2%	85.9%
Total	*Count*	716	145	861
	% within Age	83.2%	16.8%	100.0%
	% within HR HPV status	100.0%	100.0%	100.0%
	% of Total	83.2%	16.8%	100.0%

There was a significant association between age younger than or equal to 50, and older than 50, and HPV status (Chi-square=26.438, 1 d.f., p<0.0001). HPV positive cases were disproportionately represented in the younger than or equal to 50 age group. In fact, the proportion of HPV positive cases for the age group younger than or equal to 50 (27.6%) was more than twice as great as that for HPV negative cases (11.3%).

7.6.1.3 Sub-site

There was sub-site data available for 861 cases. Section 7.4.5 showcases the significant relationship between sub-site and HR HPV status (Table 7.7) status (Chi-square=118.043, 2

d.f., p<0.0001). Figures 7.9 and 7.10 indicate the disproportionate HR HPV positive status of cases within the oropharyngeal sub-site, specifically emanating from the palatine and lingual tonsils, and the corresponding HR HPV negativity of oral cavity and laryngeal cases.

7.6.1.4 Smoking Status

Data on smoking status was available for 745 patients, with the rest of the 861 cases having unknown smoking information. Table 7.10 showcases the relationship between smoking status and HR HPV status.

Table 7.10 Relationship between smoking status and HR HPV status for oropharyngeal, oral cavity, and laryngeal cancer (n=861).

Smoking Status	Count	HR HPV Status		Total
		HR HPV-	*HR HPV+*	
Current	*Count*	418	61	479
	% within Smoking status	87.3%	12.7%	100.0%
	% within HR HPV status	58.4%	42.1%	55.6%
	% of Total	48.5%	7.1%	55.6%
Ex	*Count*	79	31	110
	% within Smoking status	71.8%	28.2%	100.0%
	% within HR HPV status	11.0%	21.4%	12.8%
	% of Total	9.2%	3.6%	12.8%
Never	*Count*	125	31	156
	% within Smoking status	80.1%	19.9%	100.0%
	% within HR HPV status	17.5%	21.4%	18.1%
	% of Total	14.5%	3.6%	18.1%
Unknown	*Count*	94	22	116
	% within Smoking status	81.0%	19.0%	100.0%
	% within HR HPV status	13.1%	15.2%	13.5%
	% of Total	10.9%	2.6%	13.5%
Total	*Count*	716	145	861
	% within Smoking status	83.2%	16.8%	100.0%
	% within HR HPV status	100.0%	100.0%	100.0%
	% of Total	83.2%	16.8%	100.0%

There was a significant association between smoking status and HPV status (Chi-square=17.266, 3 d.f., p=0.001), with disproportionate numbers of HPV positive cases being ex-smokers, and a larger proportion of HPV positive cases being never smokers. HR HPV negative cases were also more often current smokers. Figure 7.16 illustrates these differences graphically.

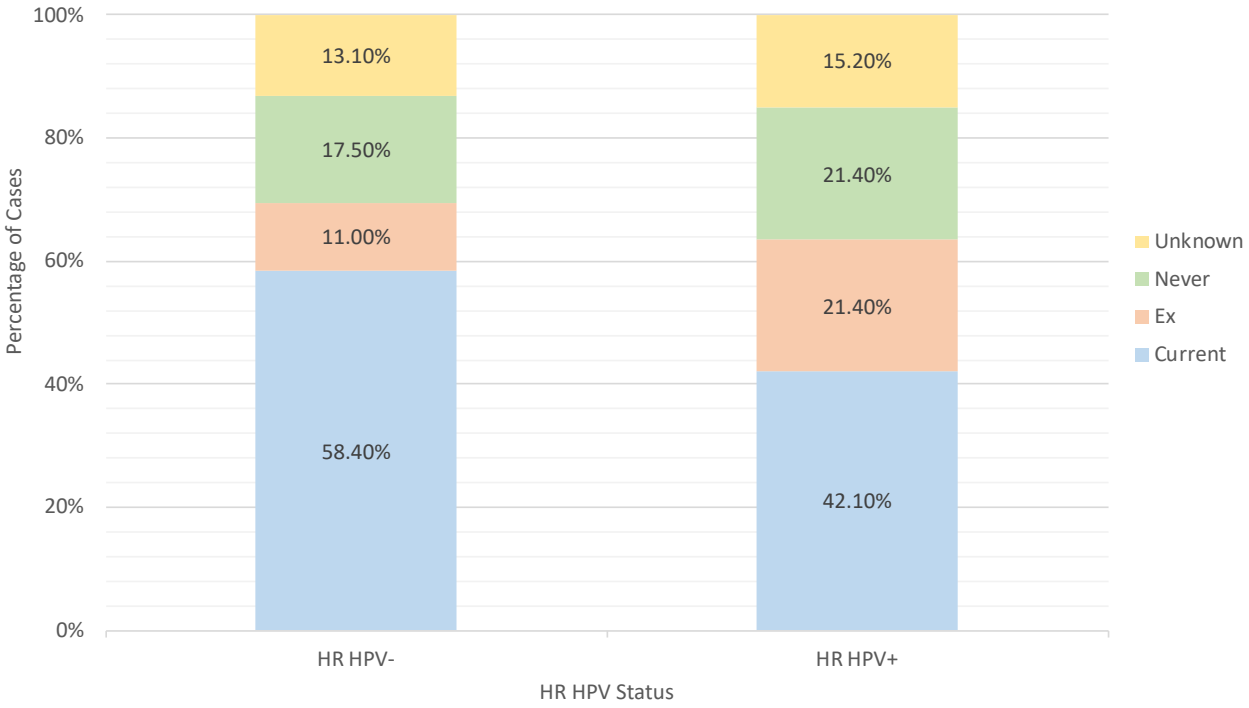

Figure 7.16 Smoking status by HR HPV status for oropharyngeal, oral cavity, and laryngeal cancer.

For HR HPV- n=716 and for HR HPV+ n=145. For the total population, n=861.

7.6.1.5 Grade

Data was available for 752 patients regarding grade of tumour at presentation, with the rest of the 861 cases having unknown grade information. The 2 cases that were un-differentiated were excluded from the analysis as previously stated, yielding a total population analyzed of (n=859). There was no significant association between grade and HPV status (Chi-square=6.623, 3 d.f., p=0.085).

7.6.1.6 T, N, M, and TNM Stage

Data was available for 732 cases for T stage, with the rest of the 861 cases having unknown T stage information. There was no significant association between T stage and HPV status (Chi-square=6.012, 4 d.f., p=0.198).

N stage data was available for 718 cases, with the rest of the 861 cases having unknown N stage information. The relationship between N stage and HR HPV status is showcased in Table 7.11.

Table 7.11 Relationship between N stage and HR HPV status for oropharyngeal, oral cavity, and laryngeal cancer (n=861).

N Stage	Count	HR HPV Status		Total
		HR HPV-	*HR HPV+*	
Unknown	*Count*	124	19	143
	% within N stage	86.7%	13.3%	100.0%
	% within HR HPV status	17.3%	13.1%	16.6%
	% of Total	14.4%	2.2%	16.6%
N0	*Count*	302	39	341
	% within N stage	88.6%	11.4%	100.0%
	% within HR HPV status	42.2%	26.9%	39.6%
	% of Total	35.1%	4.5%	39.6%
N1	*Count*	109	22	131
	% within N stage	83.2%	16.8%	100.0%
	% within HR HPV status	15.2%	15.2%	15.2%
	% of Total	12.7%	2.6%	15.2%
N2/N3	*Count*	181	65	246
	% within N stage	73.6%	26.4%	100.0%
	% within HR HPV status	25.3%	44.8%	28.6%
	% of Total	21.0%	7.5%	28.6%
Total	*Count*	716	145	861
	% within N stage	83.2%	16.8%	100.0%
	% within HR HPV status	100.0%	100.0%	100.0%
	% of Total	83.2%	16.8%	100.0%

There was a significant association between N stage and HPV status (Chi-square=24.528, 3 d.f., p<0.0001), with greater extent of nodal metastasis being related to HPV-positivity. In fact, 44.8% of HPV positive cases presented at N2/N3 compared to 25.3% for HPV negative cases. Only 26.9% of HPV positive cases presented at stage N0 compared to 42.2% of HPV negative cases. The same proportion of positive and negative cases presented at N1. Figure 7.17 represents these disparities.

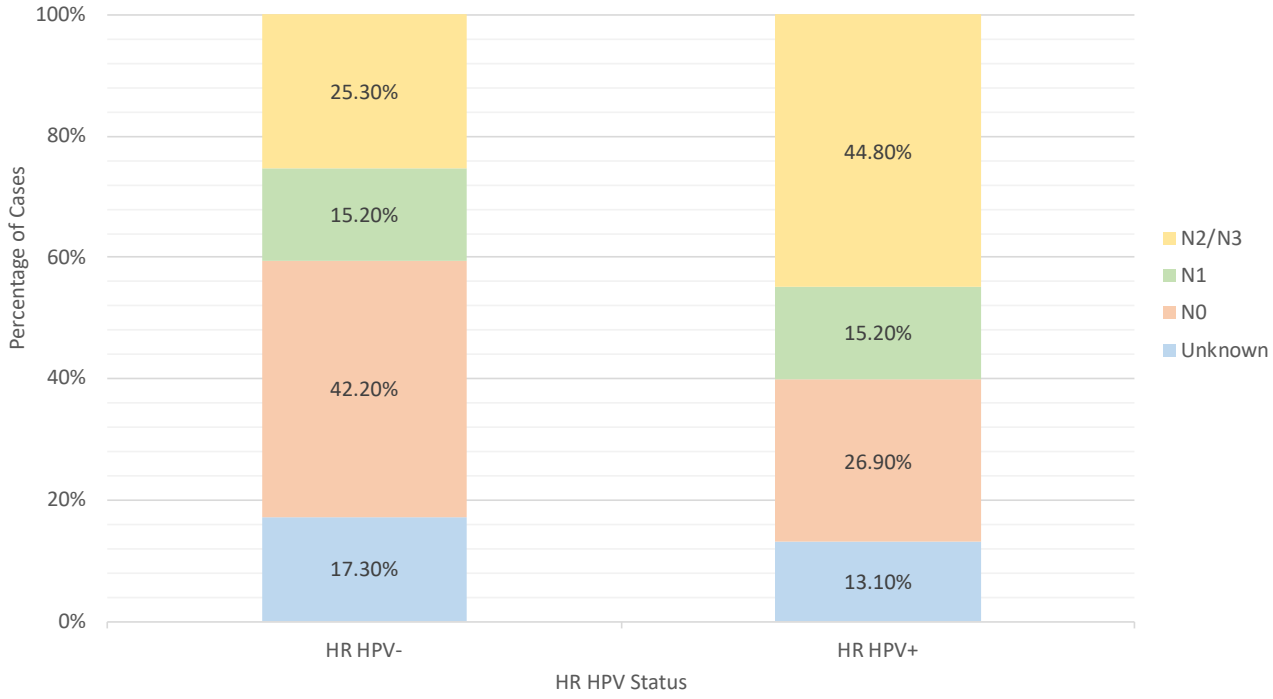

Figure 7.17 N stage by HR HPV status for oropharyngeal, oral cavity, and laryngeal cancer.

For HR HPV- n=716 and for HR HPV+ n=145. For the total population, n=861.

Data was available on M stage for 471 cases, with the rest of the 861 cases having unknown M stage data. The relationship between M stage and HPV status is presented in Table 7.12.

Table 7.12 Relationship between M stage and HR HPV status for oropharyngeal, oral cavity, and laryngeal cancer (n=861).

M Stage	Count	HR HPV Status		Total
		HR HPV-	*HR HPV+*	
Unknown	*Count*	340	50	390
	% within M stage	87.2%	12.8%	100.0%
	% within HR HPV status	47.5%	34.5%	45.3%
	% of Total	39.5%	5.8%	45.3%
M0	*Count*	340	85	425
	% within M stage	80.0%	20.0%	100.0%
	% within HR HPV status	47.5%	58.6%	49.4%
	% of Total	39.5%	9.9%	49.4%
M1	*Count*	36	10	46
	% within M stage	78.3%	21.7%	100.0%
	% within HR HPV status	5.0%	6.9%	5.3%
	% of Total	4.2%	1.2%	5.3%
Total	*Count*	716	145	861
	% within M stage	83.2%	16.8%	100.0%
	% within HR HPV status	100.0%	100.0%	100.0%
	% of Total	83.2%	16.8%	100.0%

There was a significant association between M stage and HPV status (Chi-square=8.318, 2 d.f., p=0.016). HPV negative cases were more likely to have missing M stage data than HPV positive cases, and HPV positive cases were more likely to present at M0.

TNM stage grouping information was available for 754 cases, with the rest of the 861 cases having unknown TNM stage information. Table 7.13 represents the relationship between TNM stage and HR HPV status.

Table 7.13 Relationship between TNM stage and HR HPV status in oropharyngeal, oral cavity, and laryngeal cancer (n=861).

TNM Stage	Count	HR HPV Status		Total
		HR HPV-	*HR HPV+*	
Unknown	*Count*	93	14	107
	% within TNM stage	86.9%	13.1%	100.0%
	% within HR HPV status	13.0%	9.7%	12.4%
	% of Total	10.8%	1.6%	12.4%
Stage I	*Count*	110	9	119
	% within TNM stage	92.4%	7.6%	100.0%
	% within HR HPV status	15.4%	6.2%	13.8%
	% of Total	12.8%	1.0%	13.8%
Stage II	*Count*	111	15	126
	% within TNM stage	88.1%	11.9%	100.0%
	% within HR HPV status	15.5%	10.3%	14.6%
	% of Total	12.9%	1.7%	14.6%
Stage III	*Count*	113	20	133
	% within TNM stage	85.0%	15.0%	100.0%
	% within HR HPV status	15.8%	13.8%	15.4%
	% of Total	13.1%	2.3%	15.4%
Stage IV	*Count*	289	87	376
	% within TNM stage	76.9%	23.1%	100.0%
	% within HR HPV status	40.4%	60.0%	43.7%
	% of Total	33.6%	10.1%	43.7%
Total	*Count*	716	145	861
	% within TNM stage	83.2%	16.8%	100.0%
	% within HR HPV status	100.0%	100.0%	100.0%
	% of Total	83.2%	16.8%	100.0%

There was a significant association between TNM stage grouping and HPV status (Chi-square=21.541, 4 d.f., $p<0.0001$). This significance was mostly due to the disproportionate diagnosis of HPV positive Stage IV tumours. Indeed, 40.4% of HPV negative tumours presented at Stage IV, but this was 20% less than the 60.0% of HPV positive cases presenting

at this late stage. HR HPV negative cases presented proportionally twice as often as HPV positive cases at Stage I. Figure 7.18 graphically represents this relationship.

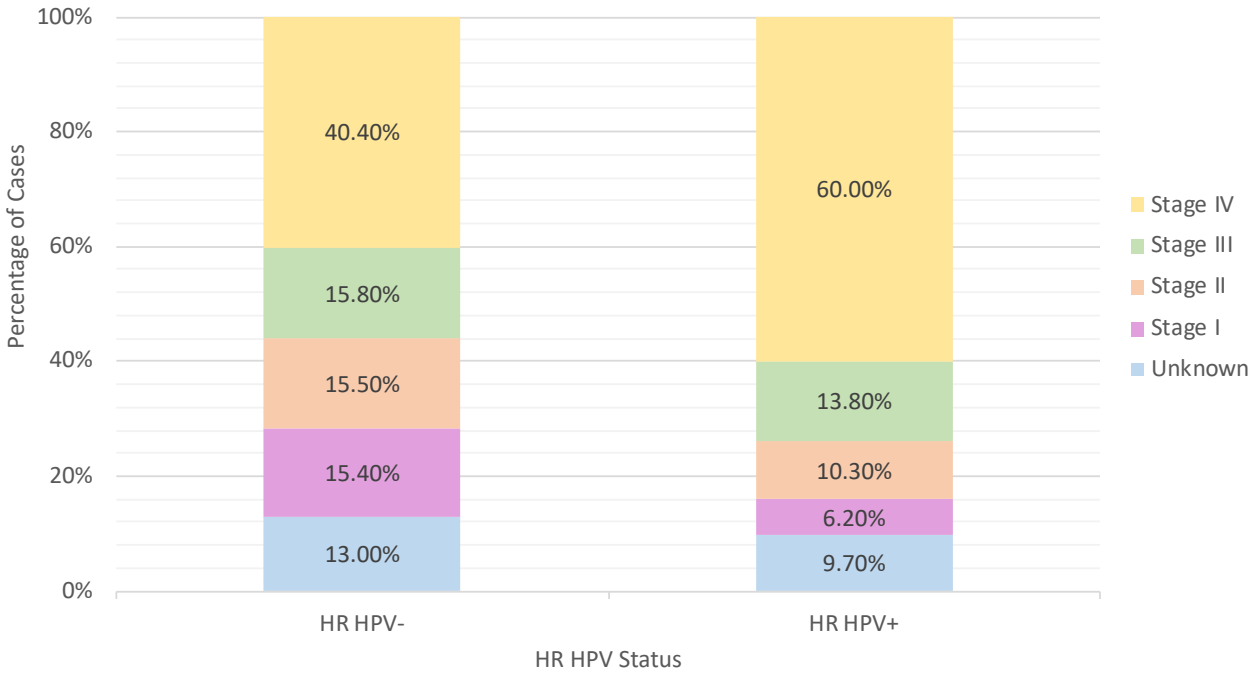

Figure 7.18 TNM stage by HPV status for oropharyngeal, oral cavity, and laryngeal cancer.

For HR HPV+ n=145 and for HR HPV- n=716. For the total population, n=861.

7.6.1.7 Socio-economic Status

Information regarding socio-economic status based on deprivation score was available for 794 cases. There was no significant association between social deprivation status and HPV status (Chi-square=1.868, 4 d.f., p=0.760). The mean deprivation score for HPV positive cases was 3.39 (CI: 3.14, 3.64) and for HPV negative cases it was 3.39 (CI: 3.28, 3.50). The median score for HPV negative cases was 4.00 and for HPV positive cases it was 4.00. There was also no significant difference between deprivation score for HPV positive and negative cases (Mann-Whitney: p=0.978).

7.6.1.8 Marital Status

Data on marital status was available for 830 patients. There was no significant association between marital status and HR HPV status (Fisher's exact=5.727, 3 d.f., p=0.126).

7.6.1.9 Geographic Location

Data was available on county of residence for 861 patients. There was no significant association between residence in counties with large urban centers (Dublin/Limerick/Cork) and residence outside these centers in rural areas (Chi-square=0.397, 1 d.f., p=0.0.529). There was no significant association between residence in or outside Dublin county and HPV status (Chi-square=0.135, 1 d.f., p=0.713).

7.6.1.10 Predictors of HR HPV Positivity: Univariate and Multivariate Analysis

Univariate logistic regression was performed using all foregoing variables to assess significant predictors of HPV positivity. The analysis identified the following variables as significantly predictive of HPV status presented in Table 7.14.

Table 7.14 Variables significantly predictive of HPV positivity by univariate logistic regression for oropharyngeal, oral cavity, and laryngeal cancer.

Variable/Factor	Statistic	Predictor of HPV Positivity
Age (Continuous) (n=861)	OR= -0.045 SE=0.008 P=<0.0001	Younger age
Age ≤ 50 (n=861)	OR=1.094 SE=0.220 P<0.0001	Age≤ 50
Sub-site (n=861)	OR=2.204, 0.458 SE=0.258, 0.279 P<0.0001, 0.0001, 0.101	Oropharynx>Larynx, Oral Cavity
Smoking Status (n=861)	OR=0.989, 0.530, 0.472 SE=0.252, 0.243, 0.274, 0.137 P=0.001, 0.0001, 0.029, 0.084	Never-smoked>Current smoker, Missing Ex-smoker>Current Smoker, Missing
N Stage (n=861)	OR=-0.852, -1.023, -0.576, -1.024 SE=0.286, 0.223, 0.275, 0.145 P<0.0001, 0.003, 0.0001, 0.036	N3/N2>Missing, N1>N0
M Stage (n=861)	OR=0.531, 0.636 SE=0.194, 0.388	M0>Missing

	P=0.017, 0.006, 0.101	
TNM Stage (n=861)	OR=-0.693, -1.303, -0.801, -0.531 SE=0.312, 0.368, 0.301, 0.272, 0.122 P<0.0001, 0.026, 0.0001, 0.008, 0.051	Stage IV, Stage III>Stage II, Stage I, Missing

Due to evident and expected confounding between T, N, M, and TNM stages, TNM stage was used as a representative for stage in the multivariate model. Results of this model are presented in Table 7.15 below.

Table 7.15 Variables significantly predictive of HPV positivity by multivariate analysis for oropharyngeal, oral cavity, and laryngeal cancer. The initial model included all variables deemed significant by univariate analysis except for T, N, and M stages (n=861).

Variable/Factor	Statistic	Predictor of HPV Positivity
Age (Continuous)	OR=-0.033 SE=0.009 P<0.0001	Younger age
Smoking Status	OR=1.262, 0.743, 0.747 SE=0.287, 0.271, 0.302 P<0.0001, 0.0001, 0.006, 0.013	Ex-smoker>Missing>Current smoker Never-smoker>Missing>Current smoker
Sub-site	OR=2.052, 0.274 SE=0.268, 0.289 P<0.0001, 0.0001, 0.341	Oropharynx>Larynx, Oral Cavity

7.6.2 Oropharyngeal Cancer

7.6.2.1 Sex

Sex data was available for all 209 oropharyngeal cases. There was no significant association between sex and HR HPV status (Chi-square=0.247, 1 d.f., p=0.619).

7.6.2.2 Age at Diagnosis

Age data was available for all 209 oropharyngeal cases. Figure 7.19 showcases the relationship between age at diagnosis and HR HPV status amongst these patients. The mean age at diagnosis for HPV negative cases was 60.72 (CI: 58.93, 62.52) and for HPV positive cases was (55.85 (CI: 53.84, 57.88). There was a significant difference between age at diagnosis for HPV positive and negative cases with HPV positive cases presenting at younger age (T-test=2.531, 207 d.f., p=0.001).

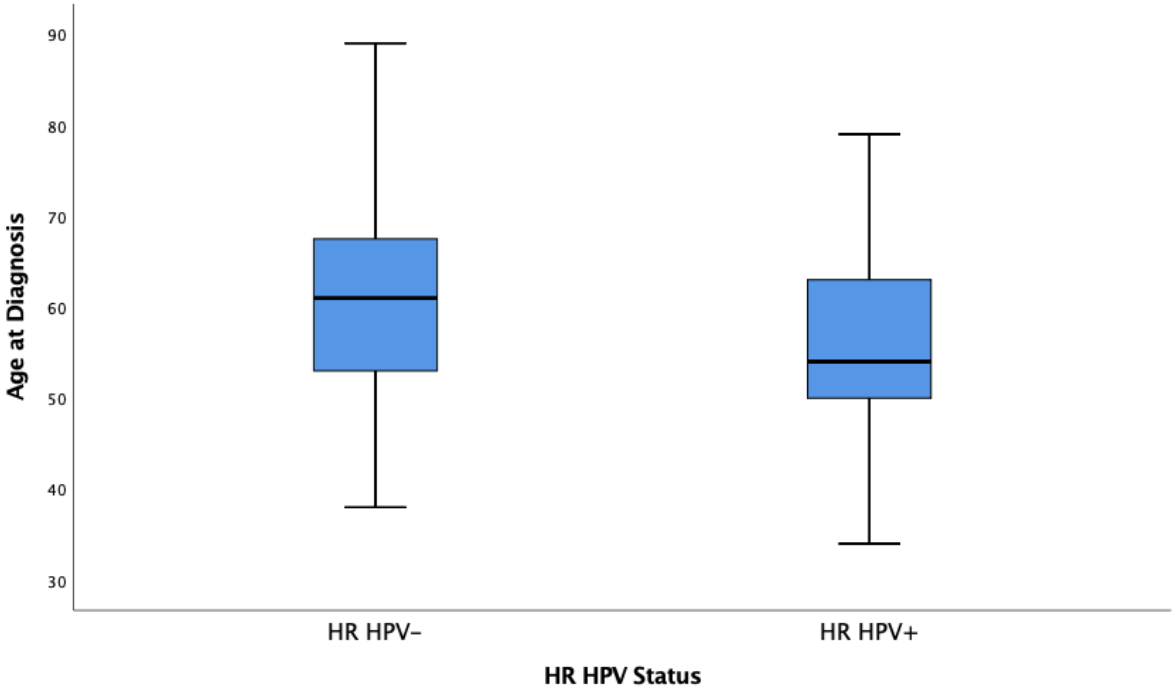

Figure 7.19 Age at diagnosis for oropharyngeal cancer by HR HPV status.

For HR HPV+ n=86, and for HR HPV- n=123. For the total population, n=209.

Patients were then grouped based on whether they were younger than or equal to age 50, or older than age 50. Table 7.16 showcases the distribution of cases based on this classification.

Table 7.16 Relationship between age younger than or equal to 50, or older than 50, and HR HPV status for oropharyngeal cancer (n=209).

Age ≤50	Count	HR HPV Status		Total
		HR HPV-	*HR HPV+*	
≤50	*Count*	21	26	47
	% within Age50	44.7%	55.3%	100.0%
	% within HR HPV Status	17.1%	30.2%	22.5%
	% of Total	10.0%	12.4%	22.5%
>50	*Count*	102	60	162
	% within Age50	63.0%	37.0%	100.0%
	% within HR HPV Status	82.9%	69.8%	77.5%
	% of Total	48.8%	28.7%	77.5%
Total	*Count*	123	86	209
	% within Age50	58.9%	41.1%	100.0%
	% within HR HPV status	100.0%	100.0%	100.0%
	% of Total	58.9%	41.1%	100.0%

There was a significant association between age younger than or equal to, or older than, 50 and HR HPV status with HPV positive cases disproportionately presenting at age younger than or equal to 50 (Chi square=5.028, 1 d.f., p=0.025).

7.6.2.3 Smoking Status

Smoking data was available for 184 patients with the rest of the 209 oropharyngeal cases having unknown smoking status. Table 7.17 showcases the relationship between smoking status and HR HPV status.

Table 7.17 Relationship between smoking status and HR HPV status for oropharyngeal cancer (n=209).

Smoking Status	Count	HR HPV Status		Total
		HR HPV-	*HR HPV+*	
Current	*Count*	88	33	121
	% within Smoking status	72.7%	27.3%	100.0%
	% within HR HPV status	71.5%	38.4%	57.9%
	% of Total	42.1%	15.8%	57.9%
Ex	*Count*	3	21	24
	% within Smoking status	12.5%	87.5%	100.0%
	% within HR HPV status	2.4%	24.4%	11.5%
	% of Total	1.4%	10.0%	11.5%
Never	*Count*	18	21	39
	% within Smoking status	46.2%	53.8%	100.0%
	% within HR HPV status	14.6%	24.4%	18.7%
	% of Total	8.6%	10.0%	18.7%
Unknown	*Count*	14	11	25
	% within Smoking status	56.0%	44.0%	100.0%
	% within HR HPV status	11.4%	12.8%	12.0%
	% of Total	6.7%	5.3%	12.0%
Total	*Count*	123	86	209
	% within Smoking status	58.9%	41.1%	100.0%
	% within HR HPV status	100.0%	100.0%	100.0%
	% of Total	58.9%	41.1%	100.0%

There was a significant association between smoking status and HR HPV status (33.593, 3 d.f., $p<0.0001$). 71.5% of HPV negative cases were current smokers compared to 38.4% of HPV positive cases. 24.4% of HPV positive cases were ex-smokers compared to 2.4% of HPV negative cases. There was also a 10% increase in the proportion of never-smokers between HPV negative and positive cases. Figure 7.20 highlights this relationship graphically.

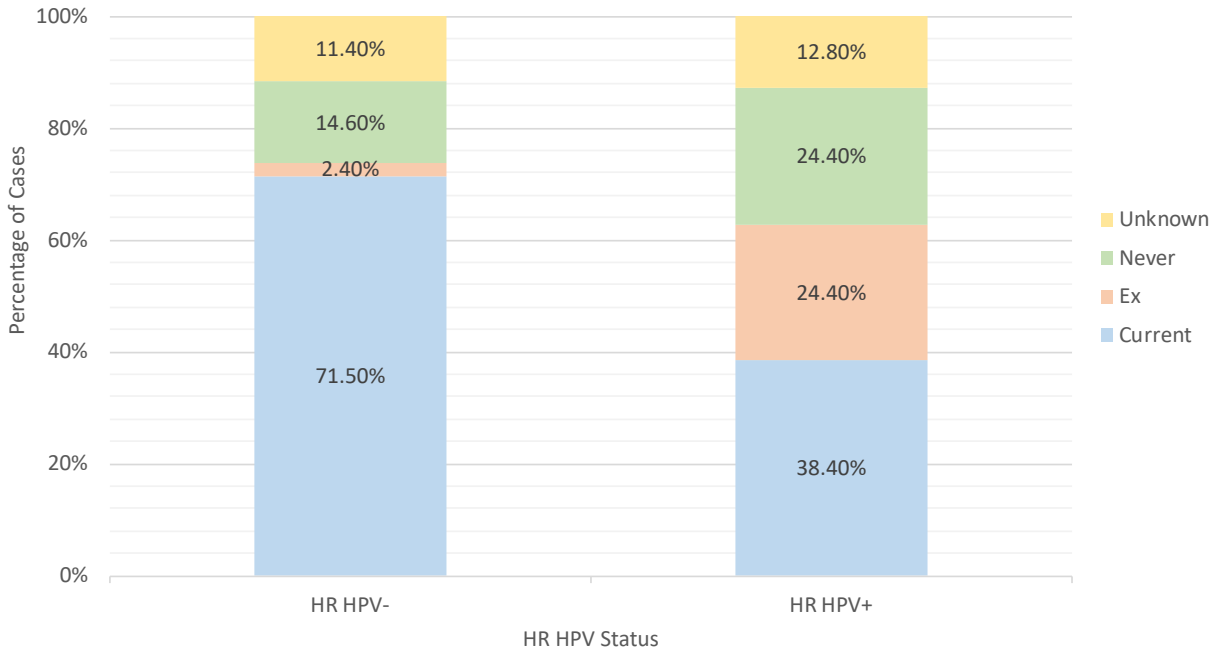

Figure 7.20 Smoking status by HR HPV status in oropharyngeal cancer.

For HR HPV- n =123 and for HR HPV+ n=86. For the total population, n=209.

7.6.2.4 Grade

Grade data was available for 180 oropharyngeal cases excluding the one case which was undifferentiated. The rest of the 209 cases had unknown grade status. There was no significant association between grade and HR HPV status (Fisher's exact=4.052, p=0.248).

7.6.2.5 T, N, M, and TNM Stage

T stage data was available for 175 patients, with the rest of the 209 patients having unknown T stage data. There was no significant association between T stage and HR HPV status (Chi-square=0.683, 4 d.f., p=0.953). N stage data was available for 185 cases with the rest of the 209 cases having unknown N stage data. There was a significant association between N stage and HR HPV status (Chi-square=10.706, 3 d.f., p=0.013). This relationship is presented in Table 7.18 below.

Table 7.18 Relationship between N stage and HR HPV status for oropharyngeal cancer (n=209).

N Stage	Count	HR HPV Status		Total
		HR HPV-	*HR HPV+*	
Unknown	*Count*	18	6	24
	% within N stage	75.0%	25.0%	100.0%
	% within HR HPV status	14.6%	7.0%	11.5%
	% of Total	8.6%	2.9%	11.5%
N0	*Count*	36	13	49
	% within N stage	73.5%	26.5%	100.0%
	% within HR HPV status	29.3%	15.1%	23.4%
	% of Total	17.2%	6.2%	23.4%
N1	*Count*	17	15	32
	% within N stage	53.1%	46.9%	100.0%
	% within HR HPV status	13.8%	17.4%	15.3%
	% of Total	8.1%	7.2%	15.3%
N2/N3	*Count*	52	52	104
	% within N stage	50.0%	50.0%	100.0%
	% within HR HPV status	42.3%	60.5%	49.8%
	% of Total	24.9%	24.9%	49.8%
Total	*Count*	123	86	209
	% within N stage	58.9%	41.1%	100.0%
	% within HR HPV status	100.0%	100.0%	100.0%
	% of Total	58.9%	41.1%	100.0%

The above Table indicates that HPV positive cases were more likely to present at later (N2/N3) nodal stages that HPV negative cases and more missing N stage data occurred in HPV negative cases. This is reflected in Figure 7.21 below.

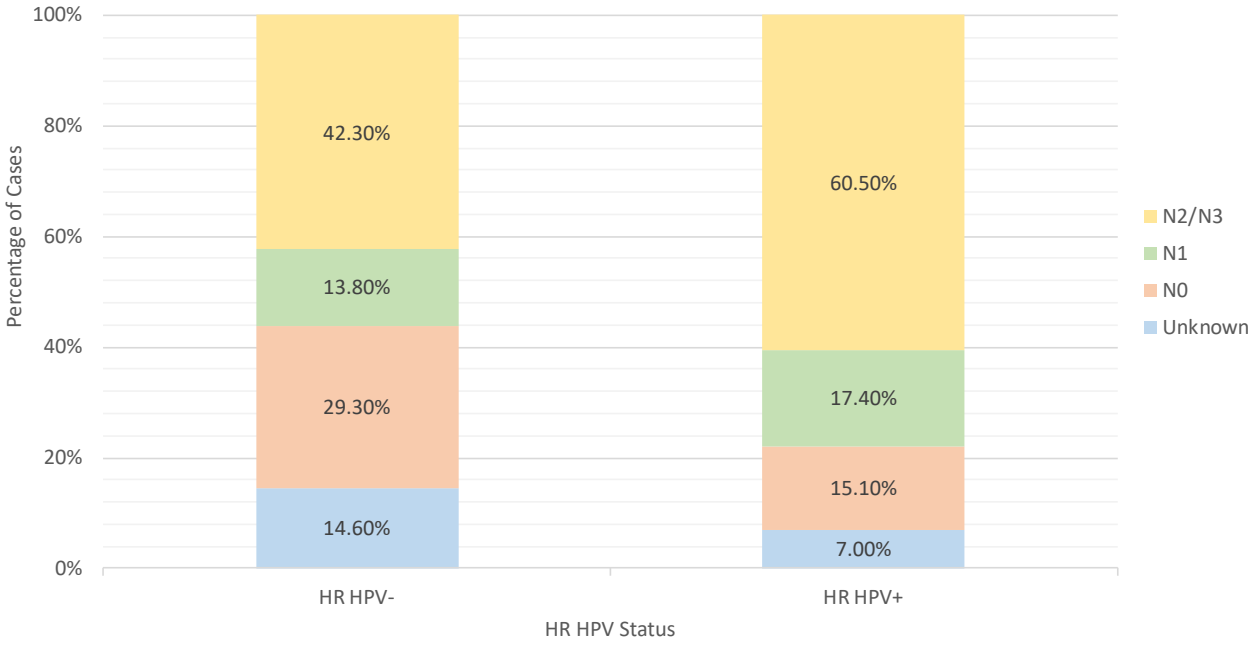

Figure 7.21 N stage by HR HPV status in oropharyngeal cancer.

For HR HPV- n =123 and for HR HPV+ n=86. For the total population, n=209.

M stage data was available for 122 cases, with the rest of the 209 cases having unknown M stage status. There was no significant association between M stage and HR HPV status (Chi-square=3.826, 2 d.f., p=0.148).

TNM stage data was available for 187 patients with the rest of the 209 cases having unknown TNM stage data. There was no significant association between TNM stage and HR HPV status (Fisher's exact=8.952, p=0.059). There was however a significant difference between the proportion of HPV positive tumours presenting at Stage IV (72.1%) and HPV negative tumours presenting at Stage IV (53.7%) (p<0.0001).

7.6.2.6 Socio-economic Status

Data on social deprivation was available for 193 cases of oropharyngeal cancer. The mean social deprivation score was 3.39 (CI: 3.11, 3.67) for HPV negative cases and 3.30 (CI:2.98, 3.62) for HPV positive cases. Median score for HPV negative cases was 4.00 and for HPV positive cases was 4.00. There was no significant difference between the median scores (Mann-Whitney: p=0.591) and when analyzed in tabular form, there was no significant association between score and HR HPV status (Chi-square=1.807, 4 d.f., p=0.771).

7.6.2.7 Marital Status

Marital status data was available for 200 patients. There was no significant association between marital status and HR HPV status (Chi-square=2.728, 3 d.f., p=0.435).

7.6.2.8 Geographic Location

County of residence data was available for all 209 oropharyngeal cases. There was no significant association between residence in counties with large urban centers (Dublin/Limerick/Cork) and residence in other counties (Chi-square=1.696, 1 d.f., p=0.193). The same was true when comparing residence in Dublin and all other counties (Chi-square=0.066, 1 d.f., p=0.797).

7.6.2.9 Predictors of HR HPV Positivity: Univariate and Multivariate Analysis

Univariate logistic regression analysis was performed using all of the above variables to assess their status as predictors of HPV positivity. Table 7.19 showcases those variables significant by univariate analysis.

Table 7.19 Variables significantly predictive of HPV positivity by univariate logistic regression for oropharyngeal cancer.

Variable/Factor	Statistic	Predictor of HPV Positivity
Age (Continuous) (n=209)	OR=-0.051 SE=0.015 P=0.001	Younger age
Age $\leq$ 50 (n=209)	OR=0.744 SE=0.335 P=0.027	Age≤50
Smoking Status (n=209)	OR=2.927, 1.135, 0.740 SE=0.650, 0.381, 0.452 P<0.0001, 0.0001, 0.003, 0.101	Ex-smoker> Current smoker, Missing Never-smoked>Current smoker, Missing

N Stage (n=209)	OR=-1.099, -1.019, -0.125 SE=0.511, 0.378, 0.405 P=0.016, 0.031, 0.007, 0.757	N2/N3, N1>N0, Missing

Due to evident and expected confounding between T, N, M, and TNM stages, TNM stage was used as a representative for stage in the multivariate model. Results of this model are presented in Table 7.20 below.

Table 7.20 Variables significantly predictive of HPV positivity by multivariate logistic regression for oropharyngeal cancer (n=209).

Variable/Factor	Statistic	Predictor of HPV Positivity
Age (Continuous)	OR= -0.061 SE=0.018 P=0.001	Younger age
Smoking Status	OR=3.112, 1.213, 0.942 SE=0.683, 0.395, 0.478 P<0.0001, 0.0001, 0.002, 0.049	Ex-smoker>Missing>Current Never-smoker>Missing>Current smoker

For interest, analysis was also conducted to determine significant predictors of the oropharyngeal sub-site itself. Multivariate analysis of those univariately significant predictors revealed that positive HR HPV status (OR=-1.605, SE=0.212, p<0.0001), younger age (OR=-0.035, SE=0.008, p<0.0001), poorly-differentiated grade (OR=-0.072, -1.280, -0.279, SE=0.310, 0.450, 0.214, p=0.033, 0.805, 0.004, 0.193), and later TNM stage (OR=-0.278, -1.524, -0.532, -0.647, SE=0.299, 0.285, 0.279, 0.268, p<0.0001, 0.353, 0.001, 0.057, 0.016) were significant predictors of oropharyngeal sub-site.

7.6.3 Oral Cavity Cancer

7.6.3.1 Sex

Sex data was available for all 331 oral cavity cancers. There was no significant association between sex and HR HPV status (Chi square=1.234, 1 d.f., p=0.267).

7.6.3.2 Age at Diagnosis

Age data was available for all 331 oral cavity cancers. The mean age at diagnosis for HPV negative cases was 63.87 (CI: 62.47, 65.26) and for HPV positive cases was 60.33 (CI: 55.86, 64.81). There was no significant difference between age at diagnosis amongst oral cavity cancers by HPV status (T-test=1.627, 329 d.f., p=0.105).

Cases were then classified according to age younger than or equal to 50 and age older than 50. Table 7.21 represents the relationship between age stratified in this way and HR HPV status.

Table 7.21 Relationship between age younger than or equal to 50, or older than 50, and HR HPV status for oral cavity cancer (n=331).

Age≤50	Count	HR HPV Status		Total
		HR HPV-	HR HPV+	
≤50	Count	37	10	47
	% within Age	78.7%	21.3%	100.0%
	% within HR HPV status	12.5%	27.8%	14.2%
	% of Total	11.2%	3.0%	14.2%
>50	Count	258	26	284
	% within Age	90.8%	9.2%	100.0%
	% within HR HPV status	87.5%	72.2%	85.8%
	% of Total	77.9%	7.9%	85.8%
Total	Count	295	36	331
	Expected Count	295.0	36.0	331.0
	% within Age	89.1%	10.9%	100.0%
	% within HR HPV status	100.0%	100.0%	100.0%
	% of Total	89.1%	10.9%	100.0%

There was a significant association between age younger than or equal to, and older than, 50 and HR HPV status (Chi-square=6.113, 1 d.f., p=0.013). Over twice the proportion of HR HPV positive cases were younger than or equal to age 50 compared to HR HPV negative cases.

7.6.3.3 Smoking Status

Smoking data was available for 281 oral cavity patients, with the rest of the 331 patients having unknown smoking status. There was no significant association between smoking status and HR HPV status (Fisher's exact=2.734, p=0.443).

7.6.3.4 Grade

Grade data was available for 302 oral cavity cases, with the remaining of the 331 cases having unknown grade status. There was no significant association between grade and HR HPV status (Fisher' exact=1.526, p=0.680).

7.6.3.5 T, N, M, and TNM Stage

T stage data was available for 293 oral cavity cases, with the rest of the 331 cases having unknown T stage. There was no significant association between T stage and HR HPV status (Fisher's exact=2.496, p=0.646). N stage data was available for 276 oral cavity cases, with the remaining of the 331 cases having unknown N stage status. There was no relationship between N stage and HR HPV status (Chi-square=2.441, 3 d.f., p=0.486). M stage data was available for 161 patients, with the remaining 331 patients having unknown M stage. There was no significant association between HR HPV status and M stage (Fisher's exact=5.527, p=0.051). TNM stage data was available for 297 oral cavity patients with the rest of the 331 cases having unknown TNM stage status. There was no significant association between TNM stage and HR HPV status (Fisher's exact=0.678, p=0.967).

7.6.3.6 Socio-economic Status

Social deprivation score was available for 303 oral cavity patients. Mean deprivation score for HPV negative cases was 3.30 (CI: 3.13, 3.48) and for HPV positive cases was 3.50 (CI: 2.97, 4.03). Median score for HPV negative cases was 4.00 and for positive cases was 3.50. There was no significant difference in median score by HPV status (Mann-Whitney: p=0.443). When assessed in tabular form, there was no association between social deprivation score and HR HPV status (Fisher's exact=6.910, p=0.129).

7.6.3.7 Marital Status

324 oral cavity patients had available marital status data. There was no significant association between marital status and HR HPV status (Fisher's exact=0.989, p=0.835).

7.6.3.8 Geographic Location

All 331 oral cavity cases had available county of residence information. There was no association between county of residence with large urban centers (Dublin/Limerick/Cork) and all other counties (Chi-square=0.131, 1 d.f., p=0.718). The same was true when comparing county of residence inside and outside of Dublin (Chi-square=0.693, 1 d.f., p=0.405).

7.6.3.9 Predictors of HR HPV Positivity: Univariate and Multivariate Analysis

Univariate logistic regression was performed using all of the foregoing variables. No variable was significantly predictive of HPV positivity in the oral cavity.

For interest, analysis was also conducted to determine significant predictors of the oral cavity sub-site itself. Multivariate analysis of those univariately significant predictors revealed that female sex (OR=0.748, SE=0.171, p<0.0001), negative HR HPV status (0.807, SE=0.217, p<0.0001), ex-smoking status (OR=0.696, -0.311, 0.183, SE=0.228, 0.207, 0.219, p=0.002, 0.002, 0.133, 0.404) and moderately differentiated grade (OR=0.386, 0.693, 0.379, SE=0.319, 0.244, 0.275, p=0.021, 0.227, 0.004, 0.168) were significantly predictive of oral cavity site.

7.6.4 Laryngeal Cancer

7.6.4.1 Sex

Sex data was available for all 321 laryngeal cases. There was no significant relationship between sex and HR HPV status (Fisher's exact: p=0.394).

7.6.4.2 Age at Diagnosis

Age data was available for all 321 laryngeal cases. The mean age at diagnosis for HPV negative cases was 66.19(CI: 64.94, 67.44) and for HPV positive cases was 64.78 (CI: 59.40, 70.17). There was no significant different between age at diagnosis in HPV positive and HPV negative patients (T-test=-0.589, 319 d.f., p=0.557). When assessed in tabular form for age younger than or equal to, and older than, 50, there was still no significant association between age and HR HPV status (Fisher exact: p=0.115).

7.6.4.3 Smoking Status

Smoking data was available for 280 laryngeal cases, with the rest of the 321 cases having unknown smoking status. There was no significant association between smoking status and HR HPV status (Fisher's exact=0.472, p=0.984).

7.6.4.4 Grade

Grade data was available for 269 laryngeal patients, while the remaining of the 321 cases had unknown grade status. Excluding the one case that was undifferentiated, there was no significant association between grade and HR HPV status (Fisher's exact=1.968, p=0.573).

7.6.4.5 T, N, M, and TNM Stage

T stage data was available for 264 patients with the rest of the 321 cases having unknown T stage. There was no significant association between T stage and HR HPV status (Fisher's exact=3.443, p=0.485). N stage data was available for 257 patients while the remaining of the 321 cases had unknown N stage. There was no significant association between N stage and HR HPV status (Fisher's exact=0.617, p=0.913). M stage data was available for 188 cases, with the rest of the 321 cases having unknown M stage status. There was no significant association between M stage and HR HPV status (Fisher's exact=0.181, p=0.933). TNM stage data was available for 270 cases while the remaining of the 321 cases had unknown TNM stage. There was no significant association between TNM stage and HR HPV status (Fisher's exact=2.828, p=0.590).

7.6.4.6 Socio-economic Status

Social deprivation score data was available for 298 laryngeal cases. When analyzed in tabular form, there was no association between deprivation score and HR HPV status (Fisher's exact=1.114, p=0.929). The mean deprivation score for HPV negative cases was 3.47 (CI: 3.31, 3.63) and for HPV positive cases was 3.61 (CI: 2.97, 4.25). Median score for HPV negative cases was 4.00 and for HPV positive cases was 4.00. There was no significant difference between mean deprivation score for HPV negative and positive groups (Mann-Whitney: p=0.725).

7.6.4.7 Geographic Location

County of residence data was available for all 321 laryngeal cases. There was no significant association between county of residence with large urban centers (Dublin/Limerick/Cork) and other county of residence (Chi-square=0.505, 1 d.f., p=0.477). There was also no association between county of residence in and outside Dublin (Chi-square=0.176, 1 d.f., p=0.675).

7.6.4.8 Marital Status

Data on marital status was available for 306 laryngeal cases. There was no significant association between marital status and HR HPV status (Fisher's exact=0.170, p=1.000).

7.6.4.9 Predictors of HR HPV Positivity: Univariate and Multivariate Analysis

Univariate logistic regression analysis was conducted using all of the above variables to assess predictors of HPV positivity in the larynx. No variables were significantly predictive of HPV status in the larynx.

For interest, analysis was also conducted to determine significant predictors of the laryngeal sub-site itself. Multivariate analysis of those univariately significant predictors revealed that male sex (OR=-0.839, SE=0.192, p<0.0001), older age (OR=0.029, SE=0.007, p<0.0001), HR HPV negative status (OR=1.116, SE=0.250, p<0.0001), earlier TNM stage (OR=0.593, 0.866, 0.783, 0.649, SE=0.245, 0.233, 0.225, 0.223, p<0.0001, 0.015, 0.0001, 0.001, 0.004), and well-differentiated grade (OR=0.485, 0.742, -0.070, SE=0.268, 0.285, 0.198, p=0.003, 0.l071, 0.009, 0.722) were predictive of laryngeal site.

7.7 Discussion

The aims of this chapter were to estimate the prevalence of HPV DNA positivity in archival tumour specimens from patients diagnosed with oropharyngeal, oral cavity, and laryngeal SCC in Ireland in the period between 1994 and 2013, to describe the genotype distribution in HPV positive tumours in this population, to estimate the raw incidence of HPV-positive and HPV-negative oropharyngeal, oral cavity, and laryngeal SCC, and to identify patient (e.g.

sex, age at diagnosis, smoking status, socio-economic status, geographic location, marital status) and clinical (e.g. stage, grade) factors associated with HPV-positivity for oropharyngeal, oral cavity, and laryngeal SCC.

Overall prevalence of HPV in oropharyngeal, oral cavity, and laryngeal SCC diagnosed in Ireland between 1994 and 2013 was estimated at 17.1% (CI: 14.6, 19.6). In the literature, HPV prevalence in similar populations has been estimated between 18.5% and 90%[13,14,16,73,74], yielding a comparatively low prevalence rate amongst these cancers in Ireland. Overall, Ireland remains relatively comparable to other European and North American countries with respect to HPV prevalence, but this result suggests that HPV may not play as large a role in the Irish oropharyngeal, oral cavity, and laryngeal SCC landscape as it does in other developed nations.

The distinct HPV prevalence of each generalized sub-site reveals that notably low HPV prevalence in what Figure 7.14 indicates is a very common cancer, LSCC, may be the root of this. The larynx HPV prevalence of 7.8% (CI: 4.9, 10.7) is relatively low in comparison to other recorded statistics that reach as high as 21% to 24% in the United States and Europe[74,75]. Ireland's laryngeal HPV prevalence appears to be more akin to that recorded in the Mediterranean[76]. Smoking rates in Ireland are indeed similar to those in Southern European countries[77], and given that smoking rates have been more resistant to change, this behavior may be more relevant than HPV in causing persistently high Irish laryngeal cancer rates.

By contrast, in the oropharynx, HPV prevalence was 41.1% (CI: 34.5, 47.8). This is similar to those statistics recorded in the United Kingdom[78], the United States[9], and Germany[79], but is significantly lower than those statistics emanating from Scandinavia[16]. This said, more detailed analysis showed that cases recorded as originating in the base of tongue and oropharyngeal (unspecified) sub-sites had lower prevalence than the tonsil, which was the largest source of positive cases with a prevalence of 60.0% (Figure 7.10). In the oral cavity, prevalence was expectedly lower at 10.9% (CI: 7.5, 14.2). This falls precisely within ranges found in the literature between 10% and 25%[74,80]. It also suggests that despite

heterogeneity in the definitions for the "oral cavity" in many studies, the role of HPV in the oral cavity in Ireland is no different from other developed nations.

Indeed, the oropharyngeal site appears to be the location within the head and neck most related to HPV given the significant association found between sub-site and HR HPV status (Table 7.7). Almost 60% of HR HPV positive cases originated in the oropharyngeal sub-site with almost half this rate in the oral cavity and less than a third of this rate in the larynx. Anatomically, this indicates that generalized MALT tissue is not the source of vulnerability to HR HPV infection that becomes persistent and carcinogenic. Instead, it indicates that it is the disrupted basal cell layer in the specialized MALT tissue in the crypts of the oropharynx, especially in the tonsil, that results in vulnerability to carcinogenic HPV. That Figure 7.10 showcases tonsillar HR HPV prevalence almost double that of the oropharynx (unspecified) and the base of tongue supports this conclusion.

This is not to say however that other areas of the oropharynx are not relevant to HPV's carcinogenic activity. Collectively, Table 7.7 and Figure 7.10 illustrate that HR HPV prevalence in non-tonsillar oropharyngeal regions is up to four times higher than in non-oropharyngeal sites. Disregarding the presence of an HR HPV infection in these sub-sites would be unwise without further elucidation and study of HR HPV's natural history outside the tonsillar crypts.

Despite prevalence disparities amongst sub-sites, there were no significant differences in genotype distributions between sub-sites (Table 7.8). No matter the extent of HPV's prevalence, HPV16 was overwhelmingly the most prevalent genotype in all sites together and individually. Over 90% of all HPV positive cases for HNSCC, OPSCC, and OSCC were infected with HPV16. The same was true for over 80% of LSCC HPV positive cases. HPV18 proved the second most prevalent genotype in all cases, and in laryngeal cases in particular. It was no more equivalent to HPV31 and HPV51 in the oral cavity. HPV18 was the third most prevalent genotype in oropharynx following HPV33. These distributions are in direct agreement with studies emerging from other European countries[16,40,74,81,82]. Where HPV16 appears to be a prime carcinogenic culprit, HPV18 lags significantly behind and is represented just as weakly, under 10%, as other HR HPV types[74,83].

This said, the oropharynx appears to harbour a greater diversity of HR HPV genotypes as showcased in Figure 7.4. Along with well-established HR HPV types 16 and 18, types 33, 35, 56, and 66 were also detected. This is in comparison to HPV16, 18, 31, and 51 in oral cavity cases and only HPV16 and 18 in laryngeal cases. This disparity mirrors the appearance of dual infections (Table 7.3), 6 of the total of 7 of which originated in the oropharynx (Table 7.4). The only other dual infection was an oral cavity SCC (Table 7.5). Every dual infection consisted of HPV16 and another genotype. Together, this is suggestive of the unique role HPV, especially HPV16, plays in carcinogenesis in the oropharyngeal region, highlighting an anatomical vulnerability so potent that not only a diversity of, but more than one type of HR HPV may persist and become dually-causal in the onset of the SCC.

Regardless of dual infection or sub-site, the genotype distribution in this Irish population is very promising in terms of the role of the HPV vaccine in preventing HPV-related HNSCC. Of the nine genotypes detected in the present population, the Gardasil-9 (nona-valent) vaccine (6, 11, 16, 18, 31, 33, 45, 52, and 58) protects against five (6, 16, 18, 31 and 33). These five genotypes accounted for over 95% of HPV positive cases, suggesting that the intended introduction of the nona-valent vaccine in 2019 would protect against the overwhelming majority of culpable carcinogenic infections in both boys and girls[68]. The vaccine would also seemingly protect against the onset of any issues related to the LR HPV6 detected in the population.

HR HPV genotypes 35, 51, 56, 66 were present in the current population and are not included in the nona-valent vaccine. This may indicate a divergence in Ireland from the HR genotype distribution in the United States, where the vaccine was formulated. Admittedly, these four missing genotypes represented less than 5% of HPV positive cases. On this basis, investing in expanding the protective range of the vaccine may not be worth the expense. This would be especially justifiable if these missing genotypes were only present in dual infections with HPV16, suggesting that protecting against the more virulent strain would be sufficient. However, HPV56 was the only missing genotype identified in a dual infection. HPV35, 51, and 66 were all single infections. Assuming their carcinogenic relationship to the

associated SCC, further development of a vaccine that protects against these types is worth investigation in the Irish context.

With respect to raw incidence, it appears that overall oropharyngeal, oral cavity, and laryngeal SCC incidence has been increasing since the mid-1990s in Ireland. Though the present population is not entirely representative of the oropharyngeal, oral cavity, and laryngeal SCC population (Chapter 6), data from the NCRI suggests the same overall trend though to a lesser extent[2,84,85]. This may suggest that the higher incidence in the second half of the 1994 to 2013 period is simply an artefact of the larger number of cases diagnosed after 2002 in the current population. Nonetheless, national data reflects the same conclusion.

Figure 7.11 is indicative of the fact that the overwhelming driver of this increase in incidence can be attributed to HPV negative cases, most of which emanated from the larynx and the oral cavity sub-sites. A closer look at raw incidence by sub-site further emphasizes this finding. HPV positive case incidence in the larynx and oral cavity sub-sites was negligible and remained consistently so from 1994 to 2013 (Figures 7.13 and 7.14). The proportion of both OSCC and LSCC attributable to HPV positive cases never surpassed 22% and 17%, respectively, in this time period.

However, the average annual percentage change statistics reported note that HPV positive cases have been increasing overall and within each sub-site. Evidently, this increase did not result in the domination of HPV positive cases in the larynx or oral cavity, failing to outcompete parallel increases in HPV negative cases, with a HPV positive average annual percentage change of 9.9% ($p<0.0001$) in the oral cavity and an insignificant 5.5% ($p=0.100$) in the larynx.

Instead, this overall increased contribution of HPV positive cases to all oropharyngeal, oral cavity, and laryngeal SCC was due to the steadily increasing proportion of OPSCC attributable to HPV positive cases. Figure 7.12 reveals that from 1994 to 2004, HPV negative OPSCC formed the overwhelming majority of all OPSCC diagnosed in that period. Since 2004, HPV positive cases have progressively been overtaking HPV negative cases in driving

the overall increasing trend. In fact, the average annual percentage change in HPV positive cases in the oropharynx was a notable 16.4% ($p<0.0001$). The average proportion of all OPSCC attributable to HPV-related cases between 1994 and 2003 was 20.6%, a proportion that jumped to 47.4% between 2004 and 2013.

This ever-increasing proportion of OPSCC alone driven by HPV in Ireland appears to be comparable to that experienced in Scandinavian and North American countries than it is to that seen by its neighbor, the United Kingdom. In the United Kingdom since the early 2000s, the proportion of OPSCC cases attributable to HPV has remained static[78]. By contrast, studies from Sweden, Denmark, and the United States have shown gradual increases in the proportions of OPSCC that are HPV positive in the same period, though this increase was noted in these populations beginning in the 1980s[9,16,83,86].

That the incidence of HPV-related OPSCC only started to significantly increase in the early- to mid-2000s in Ireland also indicates that HPV-related OPSCC has only recently begun to become the epidemiological phenomenon in Ireland that it has been in Scandinavia, North America, and continental Europe since the late 1980s. Ireland may therefore represent a unique opportunity to catch HPV-related OPSCC through prevention and targeted screening mechanisms before it begins to emerge as an even more significant driver of overall OPSCC incidence.

Changes in patterns of sexual behaviour including earlier age of sexual debut, number of sexual partners, and frequency of activity are likely responsible for the gradual emergence of HPV-related OPSCC[87], especially given that HPV is the most common sexually transmitted disease[88]. Reports in Ireland in the last 15 years show that where the median age of first intercourse amongst men and women aged under 25 in 2006 was 17, the corresponding figures for the 60-64 age group were 22 and 23[89]. The proportion of men and women experiencing first intercourse before age 17 was 11% and 2% for the 60-64 population where these figures rose to three times this amongst men (31%) and ten times this in women (22%) in the 18-24 age group[89]. There has also been a gradation in number of partners reported by age, with younger Irish people reporting higher rates of partner acquisition than older generations[89]. Given that many other studies have reported strong

association between lifetime sexual partners, number of sex partners, young age at first intercourse, and history of sexually transmitted disease as potential surrogate biomarkers for HPV16 exposure[48–51], and that gradually increasing rates of HPV infection across Europe mirror increased frequencies in these sexual behaviours, it is likely that they are the cause of present Irish OPSCC trends.

This said, OPSCC as a whole remains a less common cancer in Ireland than other, mostly HPV negative, LSCC and OSCC. This is not only true from the population-level statistics available from the NCRI[84,85], but also in the current population where the proportional representation of cancers from each sub-site of the oral region reflects this disparity. Oropharyngeal incidence in the present and national populations contributed an average of 19.6% and approximately 20% to overall oropharyngeal, oral cavity, and laryngeal SCC incidence respectively between 1994 and 2013[85]. The same statistics in the oral cavity were 43.8% and 50% and in the larynx were 36.7% and 30%[85]. OPSCC may therefore be increasingly predominantly driven by HPV carcinogenesis in Ireland, but its relative burden on the population is still slightly less significant than that of other HNSCC compared to other nations[9,10,90–92]. Should the current trends observed in raw incidence be allowed to continue however, the extremely fast-growing majority of OPSCC attributable to HPV will render OPSCC as a whole the majority contributor to all HNSCC.

Overall, significant predictors of HR HPV positivity by multivariate analysis for all oropharyngeal, oral cavity, and laryngeal SCC cases were younger age, oropharyngeal sub-site, and never-/ex-smoking status (Tables 7.15). For OPSCC, significant predictors of HPV positivity mirrored those predictive for all cases together with younger age and never-/ex-smoking status being the only remaining variables in multivariate analysis (Table 7.20). Age younger than or equal to 50 was significantly associated with HR HPV status in the oral cavity (Section 7.6.3.2), but univariate analysis revealed and that much like in the larynx, no variable was significantly predictive of HPV positivity.

HPV positive HNSCC and OPSCC presented at a significantly younger age than HPV negative counterparts. HPV positive cases were also significantly associated with age equal to or younger than 50 despite mean age being older than 50 (Tables 7.9 and 7.16). These findings

are consistent with the resounding results in the literature from other European[18,19] and North American nations[10] and are suggestive of HPV's natural history in oropharynx mirroring that of cervical SCC with respect to time lapsed between exposure to the virus and onset of carcinogenesis.

Smoking status proved a significant predictor of HPV positivity for all HNSCC and OPSCC cases, with Table 7.10 reflecting a disproportionate number of current smokers having HPV negative tumours, and never- and ex-smokers having HPV positive tumours. No information was available on pack years, number of years smoked, or number of years since quitting smoking, so it is difficult to definitively make conclusions regarding the relationship between the extent of smoking over time and HPV status. However, many studies show that smokers are likely to begin smoking in their third decade of life[77], and given that the average age of diagnosis was above 50 for both HPV positive and negative cases, it is logical to assume that current smokers are also long-term smokers. This suggests that smoking is likely the causal carcinogen in those cases that are HR HPV negative, and those cases that are HR HPV positive were likely caused by long-term persistent infection by the virus.

Nonetheless, some "hybrid" cases, potentially involving HPV-, smoking-, and other behaviourally-related carcinogenic pathways, are evident in the population. 7.1% of cases were both current smokers and HR HPV positive, and 14.5% of cases were both HR HPV negative and never-smokers (Table 7.10). That 14.5% of cases were HR HPV negative tumours and non-smokers is consistent with the continuing persistence of non-HPV carcinogens driving the incidence of HNSCC overall. However, it begs the question: If smoking and HPV are not to blame, then what carcinogen is involved? It is very likely that alcohol use has a distinct role to play in the carcinogenesis of non-smoker HPV-negative cases given the persistently high rates of consumption in Ireland[93]. Genetic predisposition[94] and occupational exposures[95] may also have a role to play in this sub-set.

The 7.1% of cases who were current smokers with HR HPV positive tumours are also indicative of an emerging class of HNSCC that is beginning to present a challenge in the clinic with respect to treatment. It is primarily difficult to determine whether or not smoking or HR HPV is the causal agent in these SCC, and more often than not these "hybrid" cases have

characteristics of both HPV positive- and smoking-related HNSCC. Some of these cases may be HR HPV positive but p16 negative due to parallel involvement of smoking-related carcinogenesis which has been found to methylate p16[96] and thus subvert some of HPV-related HNSCC's more self-destructive characteristics.

Where ex-smokers fit into this distinction is also unclear. Interestingly, univariate analysis revealed that though never-smokers were significantly more likely to have HPV positive tumours than current smokers, ex-smokers were also more likely to have HPV positive tumours than current smokers (Table 7.15 and 7.20). Given that ex-smokers were significantly older than current smokers (Chapter 6), this may be indicative of smoking's synergistic role in HPV-related HNSCC[97,98], causing long-term and lasting immunosuppression that allows the virus to become persistent and eventually causal in cancer[49,59–62]. This may also imply that ex-smokers represent an even more heterogeneous group of patients whose assessment as "HPV-related" or "smoking-related" or both will be more difficult even as the role of HPV in treatment determination becomes more defined. However, it should be said that inaccuracy surrounding self-reporting for smoking status may be the root of this finding.

It is also important to note the lack of HPV positive predictive power of any variable in the laryngeal and oral cavity sub-sites (Sections 7.6.3.9 and 7.6.4.9). The most severe interpretation of this finding is that HPV plays essentially no role in these regions, and points to the relatively unequivocal conclusion that OSCC and LSCC in Ireland are caused by smoking, alcohol, and other behavioral and occupational exposures. This is supported by the findings that HPV negativity, current/ex- smoking status, and classically older age at diagnosis for HPV-unrelated cases, were significant predictors of laryngeal and oral cavity sub-site (Sections 7.6.3.9 and 7.6.4.9). A more tempered interpretation might suggest that, synthesized with the fact that the larynx harboured the only 2 LR HPV types detected in the population, though HPV may be involved in carcinogenesis of some LSCC, the natural history of the virus and progression towards malignancy is not the same in the larynx and oral cavity as it is in the oropharynx. This may also hint that some "hybrid" cases emerging from the analysis of all cases are found in the laryngeal and oral cavity sub-sites.

Cumulatively, these smoking-related findings are likely the explanation for the persistent disparity between laryngeal and oral cavity, and oropharyngeal raw incidence previously remarked upon. The observation that laryngeal and oral cavity cancers have proportionally represented the overwhelming majority of oropharyngeal, oral cavity, and laryngeal cancer incidence in Ireland since 1994 (Chapter 6), is suggestive of the fact that smoking behaviour has been more resistant to change in Ireland than in other European and North American populations[77]. Indeed, smoking rates amongst young people in Ireland have not dropped as drastically as they have in other countries[77]. That the raw incidence numbers for HPV-unrelated oropharyngeal, oral cavity, and laryngeal SCC have not plateaued or decreased in the 1994 to 2013 time-period as distinctly as in, for instance, the United States[9] or Sweden[16] is thus unsurprising. For this reason, in the context of clearly rising HPV prevalence, Ireland is likely to already be more greatly impacted than less heterogenous populations in Sweden[16], Denmark[83], and Germany[79,99], by emerging "hybrid" cases driven by both HPV- and smoking-related carcinogenesis.

With respect to insignificant predictors of HPV-positivity for all cases in this analysis, these included sex, socio-economic status, county of residence, and marital status. That sex was not significantly predictive of HR HPV status may seem counter-intuitive at first given the 3:1 ratio at which males present with HPV-related HNSCC[19]. However, this ratio remains for HNSCC in general and is not specific to HPV-related cases[8]. Thus, there is actually no indication that the proportion of HNSCC that are HPV positive in males is three times larger than that in females. Instead, it is the incidence of HPV-related HNSCC in men that is three times that of the incidence in women. The current Irish population appears to be no different from the estimated 3:1 ratio for HNSCC overall and HPV-related HNSCC given the representative nature of the population with respect to sex determined in Chapter 6.

That oropharyngeal, oral cavity, and laryngeal SCC incidence continues to be disproportionately male whilst experiencing an increase, regardless of HPV status, is a testament to not only changing sexual behaviours throughout the entire population, but both the biological susceptibility of men to these diseases, and the role of homosexual activity in pre-disposing men to increased risk. That, for instance, oral HPV infection is more common in men than in women[40–43,97], implies that characteristics including hormonal

differences[100–102] and the potential protective immunity from seroconversion in response to cervical HPV infections among women[100,103,104] are responsible for male predominance. It is possible that the transmissibility of oral HPV may be higher for men performing oral sex on women, possibly due to a higher HPV copy number in the vagina and cervix[52]. Attitudes towards homosexual relations have also softened in Ireland in the last three decades[89], and frequency of encounters has increased as a result. Men who have sex with men (MSM) groups have concurrently been shown to have increased risk of oral HPV infection due to HIV positive status, more oral sex partners, and the vulnerability of the anal region to infection[87,105]. Indeed, HIV infected men have a higher incidence of both smoking- and HPV-related HNSCC than HIV un-infected men[106,107], insinuating that both innate and acquired immunosuppressive factors explain male predominance, regardless of HPV status. The anticipated introduction of boys into the Irish national vaccination scheme beginning in September 2019 is thus very welcome to tackle those HNSCC amongst men that are in fact HPV-related[71,72].

That socio-economic status was not predictive of HPV status overall and within each sub-site is also surprising, given the generally reported association between lower socio-economic status and greater extent of smoking's involvement in carcinogenesis[53,108–111]. The present insignificant socio-economic results suggest that HPV prevalence is similar across all sectors of the Irish population. This could be evidence of sexual behaviours being consistent across all socio-economic groups in Ireland, but also may be reflective of a balance being struck between declining smoking rates amongst higher socio-economic groups[112] and more HPV-exposing sexual behaviours amongst lower socio-economic groups[89]. Lower socio-economic groups could therefore be the origin of most "hybrid" oropharyngeal, oral cavity, and laryngeal SCC.

Indeed, overall HNSCC incidence is lower in higher socio-economic groups as a result of lower smoking rates[54,55]. Consequently, HPV's involvement in this group is more statistically evident. However, behaviours including a greater number of sexual partners and earlier age of sexual debut are significantly associated with lower socio-economic status. In fact, in Ireland, lower socio-economic groups are more likely to experience vaginal sex before the age of 17, with 29% and 14% of men and women in lower groups and 16% and 9% of men

and women in higher groups[89]. Use of contraception also falls drastically for lower socio-economic individuals[89]. These statistics are particularly important since it is highly likely that HPV-related HNSCC develop from persistent infections acquired in adolescence and in the third decade of life. Thus, lower socio-economic individuals may acquire HR HPV infections before smoking becomes a component of risk for HNSCC development. By the time both smoking- and HPV-related HNSCC develop in the sixth and seventh decades of life therefore, the virus' presence in lower and higher socio-economic groups may be almost equivalent. Overall, this indicates that lower socio-economic groups are simply more at risk than higher groups for both HPV-related and HPV-unrelated HNSCC. This is reflected in the relatively representative distribution of social deprivation score in the population with 28% of all cases in high groups (1-2), 15% in middle groups (3), and 49% in low groups (4-5).

Geographically speaking, there was no relationship between urban/rural or Dublin/non-Dublin residence and HPV status for all cases, or cases within each sub-site. Not many studies have been conducted with the aim of assessing comparisons between urban and rural place of residence, but oftentimes, these are reflective of socio-economic status and associated cultural attributes including sexual-, smoking-, and alcohol-related behaviour. Given that no significant association was detected between HPV status and socio-economic status in any sub-site and overall, it is not surprising that urban/rural or Dublin/non-Dublin residence does not play a significant role either.

Despite the fact that stage did not remain a significant predictor of HPV positivity for all cases in multivariate analysis, it was significantly related to HPV status in association tests and univariate analyses (Table 7.14). With respect to TNM stage, there was a significant association between HPV-positivity and diagnosis at Stage IV in particular for all cases (Table 7.13 and Figure 7.18). Indeed, almost 40% of all HPV negative tumours also presented at Stage IV, but this was 20% less than the 60.0% of HPV positive cases presenting at this late stage. In fact, as is evident from Table 7.13 there was an enormous jump in the proportion of HPV positive cases presenting at Stage III (13.8%) and Stage IV (60.0%).

The extremely large proportion of HPV positive cases diagnosed at Stage IV can be attributed almost entirely to their significant association with greater extent of nodal

metastasis (N stage) as showcased in Table 7.11 and Figure 7.17, and by univariate analysis (Table 7.14). Conversely, lesser nodal metastasis (N stage) was associated with HPV-negativity. Table 1.4 in Chapter 1 showcases the impact that greater N stage has on the overall later TNM stage determination using the AJCC's 5th edition criteria[113]. Cases assessed using this older edition were classified as Stage IV no matter their T stage, without distant metastasis, as long as nodal stage was N2 or greater. Thus, greater N stage in HPV positive cases disproportionately pushed overall TNM stage to Stage IV.

Whether or not this greater nodal metastasis is truly equivalent to a more aggressive, later-stage cancer however is not definitive. That HPV positive HNSCC sees better overall and cancer-specific survival[18,114] than HPV negative HNSCC suggests that either, HPV-related HNSCC is extremely aggressive, but very responsive to treatment, or, that nodal metastasis is not actually a sign of aggression in these cancers. Physiologically, the latter is most probable.

To begin, the only significant relationship of any stage variable to HPV status amongst individual sub-sites occurred between N stage and HPV in the oropharynx, with HPV positive oropharyngeal tumours showing greater nodal involvement than HPV negative tumours (Table 7.18 and Figure 7.21). Although TNM stage was not associated with HPV status in the oropharynx, missing data contributed to this insignificance and the proportion of HPV positive oropharyngeal tumours presenting at Stage IV (over 72.1%) was significantly different from the proportion of HPV negative oropharyngeal tumours presenting at Stage IV (53.7%). This clearly reflected the impact of higher nodal category for HPV positive OPSCC. Preliminarily, this indicates that HPV positive tumours behave differently than HPV negative tumours in the oropharynx as a whole. However, it is prudent to note that the majority of HPV positive cases in the oropharynx (51 of 86) were tonsillar tumours. More detailed analysis of the current population (p=0.571) and other studies find that when considering tonsillar sub-site alone, N stage no longer predicts HPV positivity[18,19]. This insinuates that it is the physiological features of the tonsillar sub-site itself, rather than the HPV-nature of tumours originating at this site, that determines greater nodal involvement, and confounds N stage when analyzing the oropharynx more generally.

Anatomically, the oropharynx is the site of the highest density of lymphatic tissue in the oral region of the head and neck due to the tonsillar sub-site. Early and speedy metastasis to other lymph nodes is thus sensical if the majority of HPV-related primary tumours already originate in the lymphatic system to begin with. This is especially true of the tonsil which, as showcased in Figure 1.11 in Chapter 1, is not only the site of cellular interface between pathogens entering by the mouth and the lymphatic system itself, but the site of unique anatomical and cellular vulnerability to carcinogenic HPV infection. Greater nodal involvement of HPV positive cases in the oropharynx as a whole is therefore a reflection of the easy access the most HPV-involved site, the tonsil, to lymph nodes, rather than an indication of a cancer that is quickly metastasizing due to aggressive HPV-related characteristics. Thus, as with all cancers, cells in HPV-related OPSCC accumulate in their closest surrounding tissue. However, by anatomical coincidence, the closest surrounding tissues in the case of most HPV-related OPSCC happen to be used as a measure of oncological aggressivity, inaccurately deeming these tumours more destructive than their HPV-unrelated counterparts.

These findings are sincere evidence that the move by the AJCC to change staging, especially for N stage, for the oropharynx based on HPV status is a clinically sensical one in the Irish context, though a narrower focus on the tonsil and perhaps base of tongue is implied (Figure 7.10). This is specifically true since, though the majority of HPV positive cases in the oropharynx come from the tonsil, HPV positive tonsillar SCC still maintains better overall and cancer-specific survival than HPV negative tonsillar SCC[115]. Based on p16 acting as a surrogate biomarker for HR HPV, details of the significant changes in T and N stage classifications individually can be found in the new AJCC 8th edition manual[116]. T stage has not been enormously impacted. In terms of N stage, one or multiple involved lymph nodes that are ipsilateral and less than 6cm in size are now included under the N1 umbrella. Contralateral and bilateral lymph nodes are now classified as N2, and lymph nodes greater than 6cm are now N3. The impact of this new N categorization on overall stage however is the most revealing aspect of the effort to downgrade aggressivity in HPV-related OPSCC regardless of nodal metastasis. Table 7.22 below showcases the new classification system.

Table 7.22 TNM stage classifications based on the new 8th edition AJCC staging manual published in 2016. *Any tumour presenting with distant metastasis is Stage IV.*

	N Stage			
T Stage	N0	N1	N2	N3
T0	N/A	I	II	III
T1	I	I	II	III
T2	I	I	II	III
T3	II	II	II	III
T4	III	III	III	III

In the new staging system, no p16-positive oropharyngeal cancer can be classified as Stage IV unless M stage is equivalent to M1, involving distant metastasis. N2 nodal metastasis is equivalent to Stage II no matter the T stage unless T stage is T4. In fact, p16-positive cancers can be as high as T2 and N1 and still be considered Stage I. In the older classification, any nodal involvement whatsoever would have immediately classified these cancers as Stage III.

The impact that this new staging might have on the relationship between stage and HPV-positivity is as yet unclear given that these guidelines were only introduced in 2016. This said, it is certain that these cancers will be less disproportionately diagnosed at later stages. However, the use of p16 as the sole determinant for this staging without HPV testing is questionable, given that HPV positive/p16 positive cases are truly the group that see the best overall survival[18]. Nonetheless, the present data is the first in Ireland to support the relevance of the new staging system to the Irish population, something that will hopefully be reflected in patient survival and morbidity in years to come.

In sum, the present analysis reveals three significant findings regarding HPV-related and HPV-unrelated oropharyngeal, oral cavity, and laryngeal SCC in the Irish population. First, it evidences that Ireland is experiencing a similar emergence of HPV positive tumours seen in other European and North American countries. These tumours are characterized by origin in the oropharyngeal sub-site, most especially the tonsil, overwhelming HPV16 genotype dominance, younger age at diagnosis, and never-smoker/ex-smoker status.

Second, it highlights that though these three patient characteristics appear to be representative and summative of others that are individually predictive of HPV positivity, it is important not to disregard those variables individually associated with HPV status that are themselves related to age and smoking status. For instance, age may be significantly predictive of HPV positivity accounting for any confounding with TNM stage, but the later stage at which HPV positive tumours present has crucial implications for treatment trajectories and eventual survival. Younger age as a predictor of HPV positivity alone, and better survival as a consequence, would be an unacceptable basis for differential treatment decisions. This remains true in the context of the new AJCC staging system[116], where, though uncommon, younger patients presenting with Stage IV tumours (tumours presenting with distant metastasis) would very likely be under-treated for unusually aggressive HPV-related cancers in the region.

Third, it reinforces the fact that despite the emerging significance of HPV-related OPSCC in particular, the Irish population appears to suffer from a more heterogeneous range of HNSCC overall. HPV-unrelated cases, especially those of the larynx and oral cavity, remain the most incidental, indicating that smoking- and alcohol-related behaviours are still pertinent in Ireland. "Hybrid" cases with carcinogenic features of both HR HPV and smoking in the oral cavity and larynx are also a more significant concern, especially amongst lower socio-economic groups due to increasingly diverse sexual behaviour amongst all social strata[89] and higher smoking- and alcohol-related behaviours in low socio-economic groups alone[112].

Cumulatively, these findings point to the extreme practicality of prophylaxis in preventing the evidently increasing relevance of HPV-related HNSCC in Ireland. The nona-valent Gardasil vaccine which will be introduced in the national scheme in September 2019 has the potential to protect against at least 95% of oropharyngeal, oral cavity, and laryngeal SCC cases. Vaccination is also hugely relevant to the Irish population given that HPV-related OPSCC incidence appears to have only begun increasing in the early- to mid-2000s. Ireland could be a unique case where the contribution of HPV-related OPSCC to overall HNSCC incidence can be caught before it becomes as aggressive as it is in other European nations and the United States. The vaccine may even reduce the emerging significance of "hybrid"

cases which currently represent an unknown with respect to tumour characteristics but known aggressivity with respect to behavior. Furthermore, despite the fact that sex was not a significant predictor of HPV positivity, men still proportionally represented the majority of OPSCC cases, suggesting that the intended expansion of the public vaccination scheme in 2019[69,71,72] in Ireland to boys is very merited.

Lastly, the foregoing analysis supports the relevance of the 8th edition AJCC staging criteria[116] for oropharyngeal cancer in Ireland, but implies that focus should perhaps shift to the tonsillar sub-site alone. However, it also posits that HR HPV-based staging cannot rely solely on p16 detection, nor can it convincingly rely on HR HPV status if only patients with clinically significant infections accompanied by all of the characteristics of purely HPV-related disease are to be identified. Only then will those patients who will certainly continue to develop HPV-related OPSCC at greater rates every year before the introduction and hopeful long-term impact of vaccination be diagnosed and treated most appropriately to maximize both their survival and morbidity, and that of their HPV-unrelated counterparts.

References

1. Ferlay, J. *et al.* Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer* **127**, 2893–2917 (2010).
2. National Cancer Registry of Ireland/Northern Ireland Cancer Registry. *All Ireland Cancer Atlas 1995-2007*. (2011).
3. Auluck, A. *et al.* Trends in oropharyngeal and oral cavity cancer incidence of human papillomavirus (HPV)-related and HPV-unrelated sites in a multicultural population. *Cancer* NA-NA (2010). doi:10.1002/cncr.25087
4. Ernster, J. A. *et al.* Rising Incidence of Oropharyngeal Cancer and the Role of Oncogenic Human Papilloma Virus. *Laryngoscope* **117**, 2115–2128 (2007).
5. Hocking, J. S. *et al.* Head and neck cancer in Australia between 1982 and 2005 show increasing incidence of potentially HPV-associated oropharyngeal cancers. *Br. J. Cancer* **104**, 886–891 (2011).
6. Shiboski, C. H., Schmidt, B. L. & Jordan, R. C. K. Tongue and Tonsil Carcinoma. *Cancer* **103**, 1843–1849 (2005).

7. de Martel, C. *et al.* Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. *Lancet Oncol.* **13**, 607–615 (2012).
8. National Cancer Registry of Ireland. *Cancer Trends: Cancers of the Head and Neck*. (2011).
9. Chaturvedi, A. K. *et al.* Human Papillomavirus and Rising Oropharyngeal Cancer Incidence in the United States. *J. Clin. Oncol.* **29**, 4294–4301 (2011).
10. Chaturvedi, A. K. *et al.* Worldwide Trends in Incidence Rates for Oral Cavity and Oropharyngeal Cancers. *J. Clin. Oncol.* **31**, 4550–4559 (2013).
11. Gillison, M. L. Human papillomavirus-related diseases: oropharynx cancers and potential implications for adolescent HPV vaccination. *J. Adolesc. Health* **43**, S52-60 (2008).
12. Ramqvist, T. & Dalianis, T. Oropharyngeal cancer epidemic and human papillomavirus. *Emerg. Infect. Dis.* **16**, 1671–7 (2010).
13. Hernandez, B. Y. *et al.* Human papillomavirus prevalence in invasive laryngeal cancer in the United States. *PLoS One* **9**, e115931 (2014).
14. Laskaris, S. *et al.* Prevalence of human papillomavirus infection in Greek patients with squamous cell carcinoma of the larynx. *Anticancer Res.* **34**, 5749–53 (2014).
15. Gheit, T. *et al.* Role of mucosal high-risk human papillomavirus types in head and neck cancers in central India. *Int. J. Cancer* **141**, 143–151 (2017).
16. Näsman, A. *et al.* Incidence of human papillomavirus (HPV) positive tonsillar carcinoma in Stockholm, Sweden: An epidemic of viral-induced carcinoma? *Int. J. Cancer* **125**, 362–366 (2009).
17. El-Mofty, S. K. & Lu, D. W. Prevalence of human papillomavirus type 16 DNA in squamous cell carcinoma of the palatine tonsil, and not the oral cavity, in young patients: a distinct clinicopathologic and molecular disease entity. *Am. J. Surg. Pathol.* **27**, 1463–70 (2003).
18. Woods, R. Human Papillomavirus-related Oropharyngeal Squamous Cell Carcinoma – Prevalence and Incidence in a Defined Population and Analysis of the Expression of Specific Cellular Biomarkers. (2016).
19. Woods, R. *et al.* Role of human papillomavirus in oropharyngeal squamous cell carcinoma: A review. *World J. Clin. cases* **2**, 172–93 (2014).
20. Combes, J.-D. & Franceschi, S. Human papillomavirus genome variants and head and

neck cancers: a perspective. *Infect. Agent. Cancer* **13**, 13 (2018).

21. de Martel, C., Plummer, M., Vignat, J. & Franceschi, S. Worldwide burden of cancer attributable to HPV by site, country and HPV type. *Int. J. Cancer* **141**, 664–670 (2017).
22. Evans, M. & Powell, N. G. The Changing Aetiology of Head and Neck Cancer: the Role of Human Papillomavirus. *Clin. Oncol.* **22**, 538–546 (2010).
23. Wentzensen, N., Schiffman, M., Palmer, T. & Arbyn, M. Triage of HPV positive women in cervical cancer screening. *J. Clin. Virol.* **76**, S49–S55 (2016).
24. Schiffman, M. *et al.* A Long-term Prospective Study of Type-Specific Human Papillomavirus Infection and Risk of Cervical Neoplasia Among 20,000 Women in the Portland Kaiser Cohort Study. *Cancer Epidemiol. Biomarkers Prev.* **20**, 1398–1409 (2011).
25. de Sanjose, S. *et al.* Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. *Lancet Oncol.* **11**, 1048–1056 (2010).
26. Srivastava, S., Shahi, U. P., Dibya, A., Gupta, S. & Roy, J. K. Distribution of HPV Genotypes and Involvement of Risk Factors in Cervical Lesions and Invasive Cervical Cancer: A Study in an Indian Population. *Int. J. Mol. Cell. Med.* **3**, 61–73 (2014).
27. Chinchai, T., Chansaenroj, J., Swangvaree, S., Junyangdikul, P. & Poovorawan, Y. Prevalence of Human Papillomavirus Genotypes in Cervical Cancer. *Int. J. Gynecol. Cancer* **22**, 1063–1068 (2012).
28. LeConte, B. A. *et al.* Differences in the viral genome between HPV-positive cervical and oropharyngeal cancer. *PLoS One* **13**, e0203403 (2018).
29. Cerasuolo, A. *et al.* Comparative analysis of HPV16 gene expression profiles in cervical and in oropharyngeal squamous cell carcinoma. *Oncotarget* **8**, 34070–34081 (2017).
30. Adams, A. K., Wise-Draper, T. M. & Wells, S. I. Human papillomavirus induced transformation in cervical and head and neck cancers. *Cancers (Basel).* **6**, 1793–820 (2014).
31. Lee, Y.-C. A. *et al.* Smoking addiction and the risk of upper-aerodigestive-tract cancer in a multicenter case-control study. *Int. J. Cancer* **133**, n/a-n/a (2013).
32. Guha, N. *et al.* Oral Health and Risk of Squamous Cell Carcinoma of the Head and Neck and Esophagus: Results of Two Multicentric Case-Control Studies. *Am. J. Epidemiol.* **166**, 1159–1173 (2007).

33. Franceschi, S. *et al.* Food groups, oils and butter, and cancer of the oral cavity and pharynx. *Br. J. Cancer* **80**, 614–620 (1999).
34. Fioretti, F., Bosetti, C., Tavani, A., Franceschi, S. & La Vecchia, C. Risk factors for oral and pharyngeal cancer in never smokers. *Oral Oncol.* **35**, 375–8 (1999).
35. Ahn, J., Segers, S. & Hayes, R. B. Periodontal disease, Porphyromonas gingivalis serum antibody levels and orodigestive cancer mortality. *Carcinogenesis* **33**, 1055–1058 (2012).
36. Tezal, M. *et al.* Chronic Periodontitis and the Incidence of Head and Neck Squamous Cell Carcinoma. *Cancer Epidemiol. Biomarkers Prev.* **18**, 2406–2412 (2009).
37. Pai, S. I. & Westra, W. H. Molecular pathology of head and neck cancer: implications for diagnosis, prognosis, and treatment. *Annu. Rev. Pathol.* **4**, 49–70 (2009).
38. Marur, S., D'Souza, G., Westra, W. H. & Forastiere, A. A. HPV-associated head and neck cancer: a virus-related cancer epidemic. *Lancet Oncol.* **11**, 781–789 (2010).
39. Fonmarty, D. *et al.* Study of the concordance between p16 immunohistochemistry and HPV-PCR genotyping for the viral diagnosis of oropharyngeal squamous cell carcinoma. *Eur. Ann. Otorhinolaryngol. Head Neck Dis.* **132**, 135–9 (2015).
40. Kreimer, A. R. *et al.* The Epidemiology of Oral HPV Infection among a Multinational Sample of Healthy Men. *Cancer Epidemiol. Biomarkers Prev.* **20**, 172–182 (2011).
41. Miller, C. S. & White, D. K. Human papillomavirus expression in oral mucosa, premalignant conditions, and squamous cell carcinoma: a retrospective review of the literature. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod.* **82**, 57–68 (1996).
42. Gillison, M. L. *et al.* Prevalence of Oral HPV Infection in the United States, 2009-2010. *JAMA* **307**, 693 (2012).
43. Kreimer, A. R. *et al.* Oral Human Papillomavirus in Healthy Individuals: A Systematic Review of the Literature. *Sex. Transm. Dis.* **37**, 1 (2010).
44. Engels, E. A. *et al.* Cancer risk in people infected with human immunodeficiency virus in the United States. *Int. J. Cancer* **123**, 187–194 (2008).
45. Gillison, M. L. Oropharyngeal cancer: a potential consequence of concomitant HPV and HIV infection. *Curr. Opin. Oncol.* **21**, 439–444 (2009).
46. Beachler, D. C., D'Souza, G., Sugar, E. A., Xiao, W. & Gillison, M. L. Natural history of anal vs oral HPV infection in HIV-infected men and women. *J. Infect. Dis.* **208**, 330–9 (2013).

47. Kutler, D. I. *et al.* High incidence of head and neck squamous cell carcinoma in patients with Fanconi anemia. *Arch. Otolaryngol. Head. Neck Surg.* **129**, 106–12 (2003).

48. D'Souza, G. *et al.* Case–Control Study of Human Papillomavirus and Oropharyngeal Cancer. *N. Engl. J. Med.* **356**, 1944–1956 (2007).

49. Smith, E. M. *et al.* Age, sexual behavior and human papillomavirus infection in oral cavity and oropharyngeal cancers. *Int. J. Cancer* **108**, 766–772 (2004).

50. Gillison, M. L. *et al.* Distinct Risk Factor Profiles for Human Papillomavirus Type 16-Positive and Human Papillomavirus Type 16-Negative Head and Neck Cancers. *JNCI J. Natl. Cancer Inst.* **100**, 407–420 (2008).

51. Syrjanen, S. *et al.* Oral HPV infection: current strategies for prevention and therapy. *Curr. Pharm. Des.* **18**, 5452–69 (2012).

52. D'Souza, G. & Dempsey, A. The role of HPV in head and neck cancer and review of the HPV vaccine. *Prev. Med. (Baltim).* **53**, S5–S11 (2011).

53. O'Hanlon, S., Forster, D. P. & Lowry, R. J. Oral cancer in the North-East of England: incidence, mortality trends and the link with material deprivation. *Community Dent. Oral Epidemiol.* **25**, 371–6 (1997).

54. Boscolo-Rizzo, P. *et al.* New insights into human papillomavirus-associated head and neck squamous cell carcinoma. *Acta Otorhinolaryngol. Ital.* **33**, 77–87 (2013).

55. Fakhry, C. *et al.* Improved Survival of Patients With Human Papillomavirus-Positive Head and Neck Squamous Cell Carcinoma in a Prospective Clinical Trial. *JNCI J. Natl. Cancer Inst.* **100**, 261–269 (2008).

56. Colevas, A. D. Population-based evaluation of incidence trends in oropharyngeal cancer focusing on socioeconomic status, sex, and race/ethnicity. *Head Neck* **36**, 34–42 (2014).

57. Jiron, J. *et al.* Racial disparities in Human Papillomavirus (HPV) associated head and neck cancer. *Am. J. Otolaryngol.* **35**, 147–53 (2014).

58. Settle, K. *et al.* Racial Survival Disparity in Head and Neck Cancer Results from Low Prevalence of Human Papillomavirus Infection in Black Oropharyngeal Cancer Patients. *Cancer Prev. Res.* **2**, 776–781 (2009).

59. Asvadi Kermani, I. *et al.* Human papilloma virus in head and neck squamous cell cancer. *Iran. J. cancer Prev.* **5**, 21–6 (2012).

60. Esquenazi, D., Bussoloti Filho, I., Carvalho, M. da G. da C. & Barros, F. S. de. The frequency of human papillomavirus findings in normal oral mucosa of healthy people by PCR. *Braz. J. Otorhinolaryngol.* **76**, 78–84 (2010).
61. Silva, K. C. *et al.* Risk factors associated with human papillomavirus infection in two populations from Rio de Janeiro, Brazil. *Mem. Inst. Oswaldo Cruz* **104**, 885–91 (2009).
62. Alam, S., Conway, M. J., Chen, H.-S. & Meyers, C. The Cigarette Smoke Carcinogen Benzo[a]pyrene Enhances Human Papillomavirus Synthesis. *J. Virol.* **82**, 1053–1058 (2008).
63. Sinha, P., Logan, H. L. & Mendenhall, W. M. Human papillomavirus, smoking, and head and neck cancer. *Am. J. Otolaryngol.* **33**, 130–6 (2012).
64. Ritchie, J. M. *et al.* Human papillomavirus infection as a prognostic factor in carcinomas of the oral cavity and oropharynx. *Int. J. Cancer* **104**, 336–344 (2003).
65. Sanders, E. A Comparison of Clinical Outcomes Between HPV Positive and HPV Negative Squamous Cell Carcinomas of the Oropharynx. *ORL. Head. Neck Nurs.* **34**, 11–4 (2016).
66. Liu, S. Z., Zandberg, D. P., Schumaker, L. M., Papadimitriou, J. C. & Cullen, K. J. Correlation of p16 expression and HPV type with survival in oropharyngeal squamous cell cancer. *Oral Oncol.* **51**, 862–9 (2015).
67. Dee, A. & Howell, F. A cost–utility analysis of adding a bivalent or quadrivalent HPV vaccine to the Irish cervical screening programme. *Eur. J. Public Health* **20**, 213–219 (2010).
68. HSE. About the HPV Vaccine - HSE.ie. *HSE* (2019). Available at: https://www.hse.ie/eng/health/immunisation/pubinfo/schoolprog/hpv/about/. (Accessed: 21st July 2019)
69. Health Information and Quality Authority (HIQA). PRESS RELEASE: HIQA advises changing to a more effective HPV vaccine and extending the vaccine to boys. *HIQA News Updates* (2018). Available at: https://www.hiqa.ie/hiqa-news-updates/hiqa-advises-changing-more-effective-hpv-vaccine-and-extending-vaccine-boys. (Accessed: 27th February 2019)
70. Health Information and Quality Authority. *HTA of extending the HPV vaccination to boys*. (2018).
71. Hilliard, M. HPV vaccination programme for boys to proceed in September. *The Irish*

Times (2019). Available at: https://www.irishtimes.com/news/health/hpv-vaccination-programme-for-boys-to-proceed-in-september-1.3933725. (Accessed: 21st July 2019)

72. Libreri, S. HPV vaccine to be extended to boys. *RTE News* (2018). Available at: https://www.rte.ie/news/health/2018/1207/1015750-hiqa-hpv/. (Accessed: 21st July 2019)
73. D'Souza, G. *et al.* Six-month natural history of oralversus cervical human papillomavirus infection. *Int. J. Cancer* **121**, 143–150 (2007).
74. Kreimer, A. R., Clifford, G. M., Boyle, P. & Franceschi, S. Human Papillomavirus Types in Head and Neck Squamous Cell Carcinomas Worldwide: A Systematic Review. *Cancer Epidemiol. Biomarkers Prev.* **14**, 467–475 (2005).
75. Hernandez, B. Y. *et al.* Human papillomavirus prevalence in invasive laryngeal cancer in the United States. *PLoS One* **9**, e115931 (2014).
76. Gungor, A. *et al.* Human papilloma virus prevalence in laryngeal squamous cell carcinoma. *J. Laryngol. Otol.* **121**, 772–774 (2007).
77. WHO | Prevalence of tobacco smoking. *WHO* (2016).
78. Schache, A. G. *et al.* HPV-Related Oropharynx Cancer in the United Kingdom: An Evolution in the Understanding of Disease Etiology. *Cancer Res.* **76**, 6598–6606 (2016).
79. Wittekindt, C. *et al.* Increasing Incidence rates of Oropharyngeal Squamous Cell Carcinoma in Germany and Significance of Disease Burden Attributed to Human Papillomavirus. *Cancer Prev. Res.* (2019). doi:10.1158/1940-6207.CAPR-19-0098
80. Hübbers, C. U. & Akgül, B. HPV and cancer of the oral cavity. *Virulence* **6**, 244–8 (2015).
81. Gillison, M. L. *et al.* Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. *J. Natl. Cancer Inst.* **92**, 709–20 (2000).
82. St Guily, J. L. *et al.* Human papillomavirus genotype distribution in oropharynx and oral cavity cancer in France—The EDiTH VI study. *J. Clin. Virol.* **51**, 100–104 (2011).
83. Carlander, A.-L. F. *et al.* Continuing rise in oropharyngeal cancer in a high HPV prevalence area: A Danish population-based study from 2011 to 2014. *Eur. J. Cancer* **70**, 75–82 (2017).
84. National Cancer Registry of Ireland. *Cancer Trends: Cancers of the Head and Neck*.

(2014).

85. National Cancer Registry Ireland. National Cancer Registry Ireland. *National Cancer Registry* (2019). Available at: https://www.ncri.ie/. (Accessed: 30th April 2019)
86. Blomberg, M., Nielsen, A., Munk, C. & Kjaer, S. K. Trends in head and neck cancer incidence in Denmark, 1978-2007: Focus on human papillomavirus associated sites. *Int. J. Cancer* **129**, 733–741 (2011).
87. Gillison, M. L. *et al.* Distinct Risk Factor Profiles for Human Papillomavirus Type 16–Positive and Human Papillomavirus Type 16–Negative Head and Neck Cancers. *JNCI J. Natl. Cancer Inst.* **100**, 407–420 (2008).
88. Center for Disease Control. STD Facts - Human papillomavirus (HPV). *Center for Disease Control* (2017). Available at: https://www.cdc.gov/std/hpv/stdfact-hpv.htm. (Accessed: 2nd February 2017)
89. Layte Hannah McGee, R. & Rundle Gráinne Cousins Claire Donnelly Fiona Mulcahy Ronán Conroy, K. *The Irish Study of Sexual Health and Relationships*. (2006).
90. American Cancer Society. Larynx Statistics. *American Cancer Society Statistics Center* (2019). Available at: https://cancerstatisticscenter.cancer.org/?_ga=2.85526873.1727136116.1562600149-1669060168.1562600149#!/cancer-site/Larynx. (Accessed: 8th July 2019)
91. American Cancer Society. Oral cavity and pharynx Statistics. *American Cancer Society Statistics Center* (2019). Available at: https://cancerstatisticscenter.cancer.org/?_ga=2.85526873.1727136116.1562600149-1669060168.1562600149#!/cancer-site/Oral cavity and pharynx. (Accessed: 8th July 2019)
92. Diz, P. *et al.* Oral and pharyngeal cancer in Europe. *Transl. Res. Oral Oncol.* **2**, 2057178X1770151 (2017).
93. Ritchie, H. & Roser, M. Alcohol Consumption. *OurWorldData* (2019). Available at: https://ourworldindata.org/alcohol-consumption.
94. Kutler, D. I. *et al.* High incidence of head and neck squamous cell carcinoma in patients with Fanconi anemia. *Arch. Otolaryngol. Head. Neck Surg.* **129**, 106–12 (2003).
95. Langevin, S. M. *et al.* Occupational dust exposure and head and neck squamous cell carcinoma risk in a population-based case-control study conducted in the greater

Boston area. *Cancer Med.* **2**, 978–986 (2013).

96. O'Regan, E. M. *et al.* p16INK4A genetic and epigenetic profiles differ in relation to age and site in head and neck squamous cell carcinomas. *Hum. Pathol.* **39**, 452–458 (2008).
97. D'Souza, G., McNeel, T. S. & Fakhry, C. Understanding personal risk of oropharyngeal cancer: risk-groups for oncogenic oral HPV infection and oropharyngeal cancer. *Ann. Oncol.* **28**, 3065–3069 (2017).
98. Gillison, M. L. *et al.* Prevalence of oral HPV infection in the United States, 2009-2010. *JAMA* **307**, 693–703 (2012).
99. Buttmann-Schweiger, N., Deleré, Y., Klug, S. J. & Kraywinkel, K. Cancer incidence in Germany attributable to human papillomavirus in 2013. *BMC Cancer* **17**, 682 (2017).
100. Warnakulasuriya, S. Global epidemiology of oral and oropharyngeal cancer. *Oral Oncol.* **45**, 309–316 (2009).
101. Gillison, M. L. *et al.* Eurogin Roadmap: Comparative epidemiology of HPV infection and associated cancers of the head and neck and cervix. *Int. J. Cancer* **134**, 497–507 (2014).
102. Shatalova, E. G., Klein-Szanto, A. J. P., Devarajan, K., Cukierman, E. & Clapper, M. L. Estrogen and Cytochrome P450 1B1 Contribute to Both Early- and Late-Stage Head and Neck Carcinogenesis. *Cancer Prev. Res.* **4**, 107–115 (2011).
103. Markowitz, L. E., Sternberg, M., Dunne, E. F., Mcquillan, G. & Unger, E. R. Seroprevalence of Human Papillomavirus Types 6, 11, 16, and 18 in the United States: National Health and Nutrition Examination Survey 2003–2004. *J. Infect. Dis.* **200**, (2009).
104. Safaeian, M. *et al.* Epidemiological study of anti-HPV16/18 seropositivity and subsequent risk of HPV16 and -18 infections. *J. Natl. Cancer Inst.* **102**, 1653–62 (2010).
105. Read, T. R. H. *et al.* Oral human papillomavirus in men having sex with men: risk-factors and sampling. *PLoS One* **7**, e49324 (2012).
106. D'souza, G. *et al.* Epidemiology of head and neck squamous cell cancer among HIV-infected patients. *J. Acquir. Immune Defic. Syndr.* **65**, 603–10 (2014).
107. Beachler, D. C. & D'Souza, G. Oral human papillomavirus infection and head and neck cancers in HIV-infected individuals. *Curr. Opin. Oncol.* **25**, 503–10 (2013).

108. Ramsey, T. *et al.* Laryngeal cancer: Global socioeconomic trends in disease burden and smoking habits. *Laryngoscope* **128**, 2039–2053 (2018).
109. Khalil, D. *et al.* Does Socioeconomic Status Affect Stage at Presentation for Larynx Cancer in Canada's Universal Health Care System? *Otolaryngol. Neck Surg.* **160**, 488–493 (2019).
110. Groome, P. A. *et al.* Explaining Socioeconomic Status Effects in Laryngeal Cancer. *Clin. Oncol.* **18**, 283–292 (2006).
111. Auluck, A. *et al.* Socio-economic deprivation: a significant determinant affecting stage of oral cancer diagnosis and survival. *BMC Cancer* **16**, 569 (2016).
112. Sheridan, A., Evans, D. & Kavanagh, P. *Adult Smoking in Ireland: A Special Analysis of the Healthy Ireland Survey and The Irish Longitudinal Study on Ageing (TILDA)*.
113. American Joint Committee on Cancer & American Cancer Society. Head and Neck Sites. in *AJCC Cancer Staging Manual* 24–59 (Lippincott-Raven Publishers, 1997).
114. Ang, K. K. *et al.* Human papillomavirus and survival of patients with oropharyngeal cancer. *N. Engl. J. Med.* **363**, 24–35 (2010).
115. Ahmadi, N., Chan, M., Huo, Y. R., Sritharan, N. & Chin, R. Y. Survival outcome of tonsillar squamous cell carcinoma (TSCC) in the context of human papillomavirus (HPV): A systematic review and meta-analysis. *Surg.* **17**, 6–14 (2019).
116. American Joint Committee on Cancer & American Cancer Society. *AJCC Cancer Staging Manual*. (Springer Publishing, 2016).

CHAPTER 8

SURVIVAL, PROGNOSIS, AND TREATMENT FOR HPV-RELATED AND HPV-UNRELATED OROPHARYNGEAL, ORAL CAVITY, AND LARYNGEAL CANCER

8 CHAPTER 8: SURVIVAL, PROGNOSIS, AND TREATMENT FOR HPV-RELATED AND HPV-UNRELATED OROPHARYNGEAL, ORAL CAVITY, AND LARYNGEAL CANCER

8.1 Introduction

Prognosis and survival for HNSCC is generally poor. Approximately half of all patients with HNSCC have advanced stage disease at the time of diagnosis, with an expected 5-year survival rate between 10% to 40%[1]. This is mostly attributed to the fact that diagnosis of HNSCC is frequently delayed because symptoms for which patients will seek medical attention such as pain, dysphagia, and shortness of breath occur late in the stage of disease[2].

Despite treatments that may consist of mutilating surgery, radiotherapy and/or chemotherapy, overall long-term survival remains low due to uncontrollable persistent or recurrent disease[2]. In fact, the median overall survival for recurrent or metastatic HNC remains less than 1 year despite modern chemotherapy and targeted agents[3]. Patients with locoregional and distant recurrences also have very low overall long-term survival.

Palliative chemotherapy and cetuximab, the EGFR inhibitor, constitute the backbone of treatment for patients with recurrent and metastatic HNSCC. Phase III trials for platinum doublets include cisplatin/5-FU, cisplatin/paclitaxel, and cisplatin/pemetrexed. Platinum chemotherapy in combination with 5-fluorouracil and cetuximab has resulted in the longest median overall survival for HNSCC[3].

HPV-positive HNSCC, and more specifically OPSCC, has a unique relationship to diagnosis, prognosis, and treatment. These tumours generally present with a more advanced clinical stage, with a higher nodal category[4,5], despite lower tumour extent[5,6], and have different tendencies for extracapsular spread and perineural invasion[7]. They often present with early lymph node metastases[8,9], which are sometimes confused with branchial cleft cysts[10]. However, tonsil SCCs in general are known to present with early lymph node metastases[11] and it may simply be that the anatomy of the site itself facilitates the early spread and depth of invasion[9].

Despite more advanced presentation, improved survival, which is consistently higher than 30% [12], is evident in HPV-related OPSCC[9,13–17], irrespective of treatment modality[6,18–24]. The improved prognosis and response to treatment holds true for all indicators of HPV-positivity including seropositivity, mRNA, oncoprotein expression, and viral load and copy number[25]. It also remains salient in the case of HPV-positive OPSCC biomarkers, including p16, p53, EGFR, and Bcl-xL[25,26].

For most patients with high-risk, resected HNSCC, the standard treatment constitutes adjuvant radiation therapy with high doses of cisplatin. This course of treatment appears to work well for HPV-positive tumours. Adjuvant chemoradiation therapy with one dose of weekly cisplatin had 3-year overall survival rates of 86% and 91% and 3-year recurrence free survival of 82% and 84% in one study, suggesting that cisplatin is a good treatment for HPV-positive OPSCC to preserve survival and minimize toxicity[27,28].

Given this positive response to therapies[6,29–32], it is possible that de-escalation of therapy would be appropriate for these HPV-positive HNSCC. This is particularly important given the long-term consequences and associated morbidities amongst those patients who do survive. Though patients express gratitude for the success of their treatments, many suffer from difficulty swallowing, breathing, and speaking, chronic pain, osteoradionecrosis, hypertension, pneumonia, dysphagia, weight loss, malnutrition, dental issues, and third-degree burns. These are acute hindrances to the quality of the rest of their lives.

Despite extensive reports in the literature, most studies regarding the differential prognosis and treatment modalities of HPV-related and HPV-unrelated HNSCC analyze fewer than 300 cases[16,17,33]. Furthermore, when pooled for meta-analysis[24], the definition of what constitutes an HPV positive case, heavily dependent upon HPV indicator chosen and technology used, varies by study. Though one study reached 720 patients[18], the present ECHO study presents the unique opportunity to analyze a large population (n=861) of HPV-related and HPV-unrelated cases drawn from a population relatively representative of an entire European country. It also lends to standardized definitions of an HPV positive case, using DNA alone detected by an extremely sensitive Multiplex PCR Luminex technology. Furthermore, there is an evident gap in the literature regarding the Irish population's HPV-

related and HPV-unrelated HNSCC survival. The coming analysis intends to fill this chasm with one of the largest HPV-related HNC survival studies conducted to date.

8.2 Aims

- To compare treatment received and survival in patients with HPV-positive and HPV-negative tumours in oropharyngeal, oral cavity, and laryngeal SCC.
- To identify significant predictors of survival in oropharyngeal, oral cavity, and laryngeal SCC.

8.3 Material and Methods

8.3.1 Study Population

The population for this study was that of the ECHO study as a whole. It was through the NCRI's database that this study identified relevant specimens to create its own databank, as described in Sections 3.1, 3.2 and 3.3. Ethical approval for use of archival tissue specimens was obtained from all relevant local hospital ethics committees detailed in Section 3.3.3.

The population comprised 861 cases of primary, invasive, oropharyngeal, oral cavity, and laryngeal SCC retrieved from 7 different hospitals sites around Ireland newly diagnosed between 1994 and 2013. The ICD10 codes included in the study can be found in Sections 3.1 and 3.2.

Analysis of the different sub-sites in the population was based on the same classifications of the entire study population as in Chapters 6 and 7. The definitions of oropharyngeal, oral cavity, and laryngeal sub-sites by generalized ICD10 codes are summarized in Table 8.1.

Table 8.1 Summary of ICD10 codes represented in the study population and the classification under which they were placed for the analysis.

ICD10 Codes	Classification
1.0, 2.4, 5.1, 5.2, 9.0-9.9, 10.0-10.9, 14.2	*Oropharynx*
2.0, 2.1, 2.2, 2.3, 2.8, 2.9, 3.0-3.9, 4.0-4.9, 5.0, 5.8, 5.9, 6.0, 6.1, 6.2, 6.9, 14.0, 14.8	*Oral Cavity*
32.0-32.9	*Larynx*

The patient and tumour characteristics for the study population used in the analysis were also the same as those described in detail in Section 7.3.4.

8.3.2 The Definition of an HPV-related Case

As dissected in Chapter 2, most studies define HPV positivity differently, deeming an HPV positive case based on different indicators of the virus' presence (e.g. DNA, mRNA, viral load, and viral integration) and using different technologies to detect these indicators. Each indicator is appropriate for particular contexts of defining 'HPV positivity' given their respective advantages and disadvantages. In the epidemiological context, studies have taken to using HPV DNA as it provides a good estimate of the prevalence of the virus in a population, regardless of its involvement or lack of involvement in the carcinogenesis of the associated tumour.

The lack of Irish data in the literature regarding HPV's role in oropharyngeal, oral cavity, and laryngeal cancer drives the epidemiological demands of this study. It is necessary to gauge the general presence of the virus in the Irish population first if any further conclusions based on other more functional indicators of the virus are to be as indicative as they can be. The study therefore requires a definition for HPV positivity indiscriminate of its provable relationship to associated tumours.

The study thus defines an HPV positive sample as any case the Multiplex PCR Luminex technology identifies as positive for HPV DNA. The technology is extremely sensitive, with a lowest limit of detection at 10 copies of the virus and it detects one of the most extensive

ranges of HR and LR HPV genotypes of any available platform. As a consequence, it is an ideal method to achieve the epidemiological goals of the study. For the purposes of this analysis however, only HR HPV cases were considered HPV positive given their implicated role in carcinogenesis.

8.3.3 HPV DNA Detection

Three 10um sections were cut using a microtome for each FFPE block associated with a particular case as described in Section 3.6.2. The sections were cut using the "IARC sterility protocol" denoted in Section 3.6.2.

DNA was then extracted from these sections in accordance with the steps outlined in Section 3.7.3.2. Extracted DNA was amplified with a Multiplex PCR detailed in 3.7.3.3. The PCR detected 19 HR or probably HR HPV genotypes (16, 18, 26, 31, 33, 35, 39, 45, 51, 53, 56, 58, 59, 66, 68a, 68b, 70, 73, 82) and two LR HPV genotypes (6, 11). Detection limits ranged from 10 to 1,000 copies of the viral genome per reaction. PCR products were then hybridized (Section 3.7.3.5) to oligonucleotide probes previously coupled to fluorescent beads (Section 3.7.3.4) and analysed by a Luminex 200 Analyser reporting on internal bead colour and Strep-PE reporter fluorescence. Results were expressed as the MFI of at least 100 beads per bead set and cut-offs were set as described in Section 3.7.3.5.

8.3.4 Survival and Treatment Data

Data regarding patient survival and treatment was obtained from the NCRI database. The NCRI provided anonymized study numbers to the Researcher that linked all HPV analysis the Researcher performed to the associated characteristics in the national database. The survival and treatment information were broken down into a variety of variables. Tables 8.2 and 8.3 detail the available characteristics. These variables were then used to compare and contrast survival, prognosis, and treatment administered between HPV-related and HPV-unrelated groups.

Table 8.2 Variables regarding patient treatment provided by the NCRI. These variables were used individually and in combination with one another for the analysis.

Variable Code	Meaning	Definition
Chemo_1y	Chemotherapy	Binary indication of whether or not patient was treated with chemotherapy targeting the cancer within 1 year of diagnosis.
Radio_1y	Radiotherapy	Binary indication of whether or not patient was treated with radiotherapy targeting the cancer within 1 year of diagnosis.
Surg_1y	Surgery	Binary indication of whether or not patient was treated with surgery targeting the cancer within 1 year of diagnosis

Table 8.3 Variables regarding patient survival provided by the NCRI.

Variable Code	Meaning	Variable Definition
VITAL_STAT	Overall survival	All-cause vital status of patient (0 alive or 1 dead) at common censoring date based mainly on death-certificate matching.
VITAL_CAN	Disease-specific (cancer-specific) survival	Cause-specific vital status (0 alive or died of other cause or different cancer or 1 died from the cancer of interest) at common censoring date.
SURVIVAL_MONTHS	Survival in months	Number of complete months from diagnosis of a specific tumour to common censoring date.

8.3.5 Statistical Analysis

Statistics were generated using IBM SPSS Statistics Version 25, XLSTAT 2019.1.3, and Microsoft Excel Version 16.25. Overall and cancer-specific survival for all oropharyngeal, oral cavity, and laryngeal cancer, and within each sub-site, based on HR HPV status was assessed by Kaplan-Meier curves and log-rank test. Additional cox proportional hazard statistics were generated to confirm Kaplan-Meier results. The relationship between treatment and HPV status for all cases and for each sub-site was evaluated using Chi-square statistics and Fisher's exact tests in cases where expected counts fell below 5. The cohort of patients assessed for treatment was limited to those who received treatment of any kind within 12 months of diagnosis as these were patients of interest. Predictors of overall and cancer-

specific survival were evaluated individually by univariate cox proportional hazard models. For variables from which more than 10% of data was missing, “missing” was included as a category of its own as is convention in the literature to account or detect any bias responsible for significance. For those variables with between 0% to 10% missing data, cases with missing data were excluded for univariate and multivariate analyses. Those variables significant in univariate analysis were brought forward for multivariate analysis. All significant variables by univariate models were included in the initial multivariate model. The least significant predictor was then taken out, and the model was run again. The least significant predictor was again taken out, and the model was run again. This continued until all variables remaining in the model proved significantly predictive of survival and risk of death, or until taking another variable out rendered the model as a whole insignificant.

8.4 HPV Status and Survival for Oropharyngeal, Oral Cavity, and Laryngeal Cancers

To begin, Kaplan-Meier survival analysis using the log-rank statistic was performed on all cases within the study stratified by HR HPV status for both overall and cancer-specific survival. This was then repeated for cancers originating in each of the three key sub-sites.

8.4.1 Oropharyngeal, Oral Cavity, and Laryngeal Cancer

Figure 8.1 below shows the result of the Kaplan-Meier analysis of overall survival for the population stratified by HPV status. There was significantly worse survival for the HPV negative group than the HPV positive group (Log-rank: Chi-square=12.593, 1 d.f., $p<0.0001$). Cox proportional hazard model for HPV status and overall survival confirmed the increased risk of death for HPV negative patients (HR=0.372, 1 d.f., $p<0.0001$).

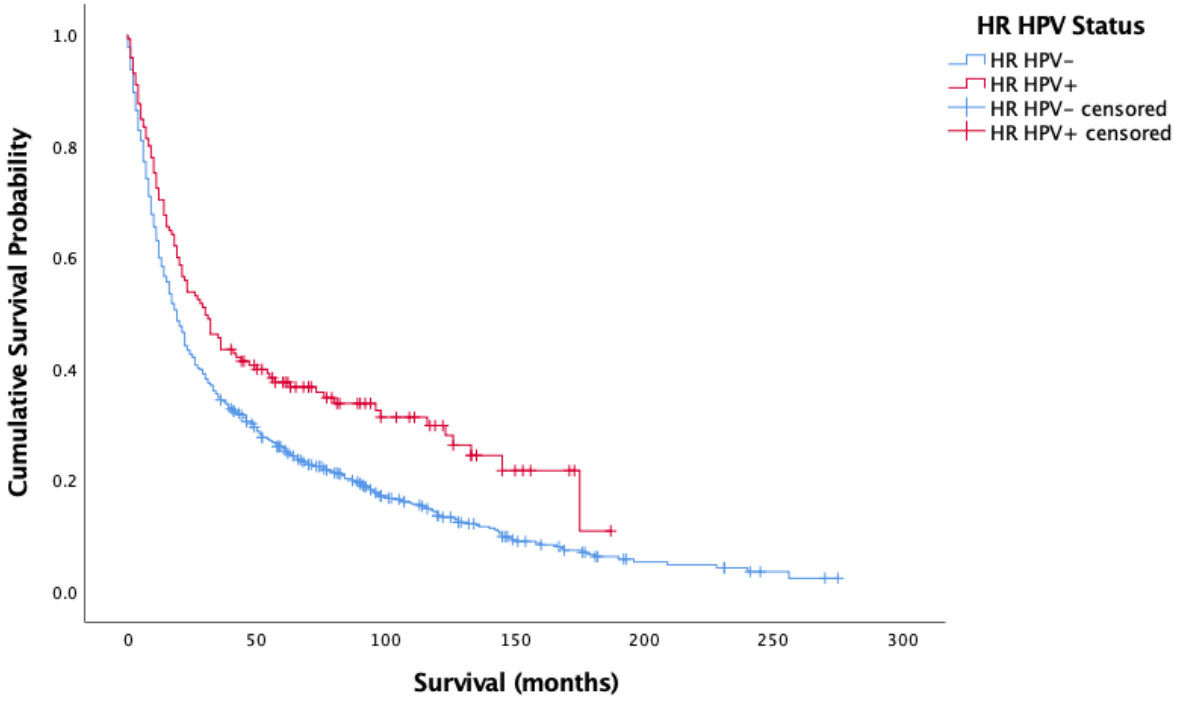

Figure 8.1 Kaplan-Meier analysis of overall survival in months based on HR HPV status for oropharyngeal, oral cavity, and laryngeal cancer (n=861).

Figure 8.2 is the Kaplan-Meier result for cancer-specific survival in the population stratified by HPV status. The analysis was a mirror of the findings in Figure 8.1, showing better survival for HPV positive cases than HPV negative cases (Log-rank: Chi-square=4.582, d.f.=1, p=0.032). Cox proportional hazard model seconded the significantly increased risk of cancer-specific death in the HPV negative group (HR=0.257, SE=0.122, p=0.035).

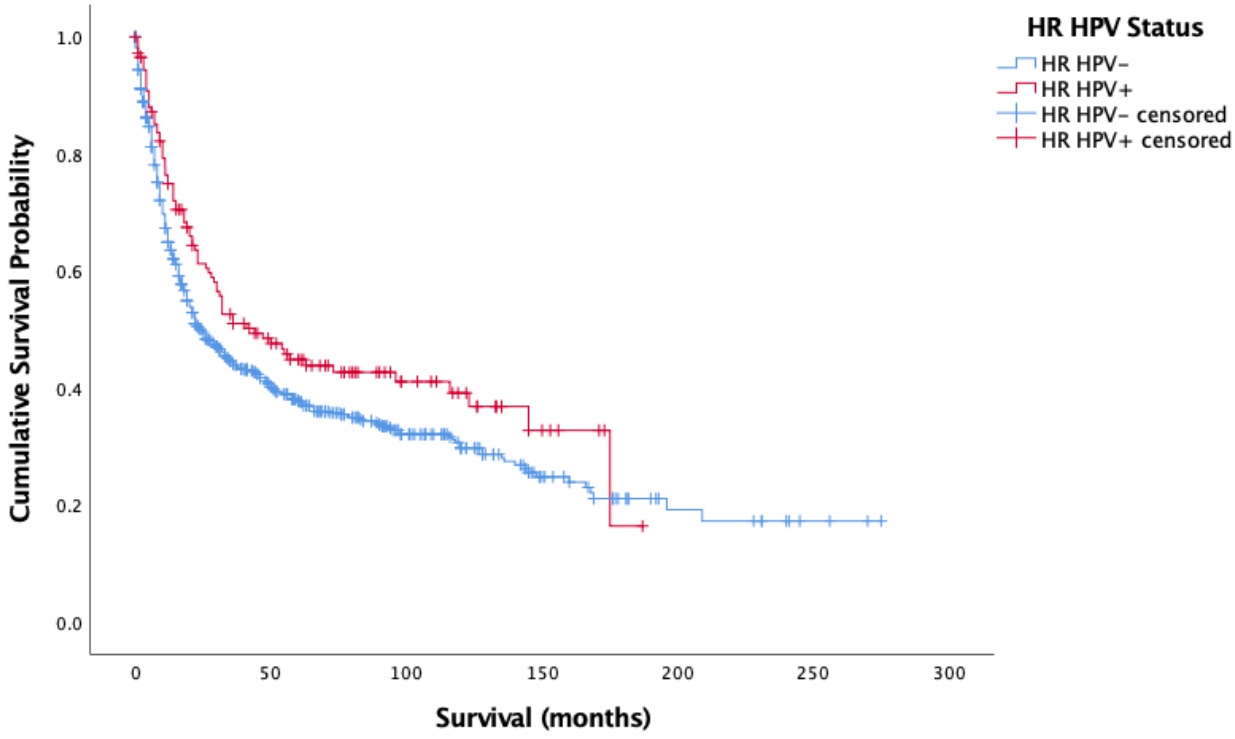

Figure 8.2 Kaplan-Meier analysis of disease-specific (cancer-specific) survival in months based on HR HPV status for oropharyngeal, oral cavity, and laryngeal cancer (n=861).

8.4.2 Oropharyngeal Cancer

Figure 8.3 showcases the Kaplan-Meier analysis of overall survival stratified by HPV status for the oropharyngeal sub-site alone. Much like for all cases, there was significantly worse prognosis for HPV negative cases (Log-rank: Chi-square=17.017, 1 d.f., p<0.0001). Cox proportional hazard model confirmed this finding (HR=0.659, SE=0.165, p<0.0001).

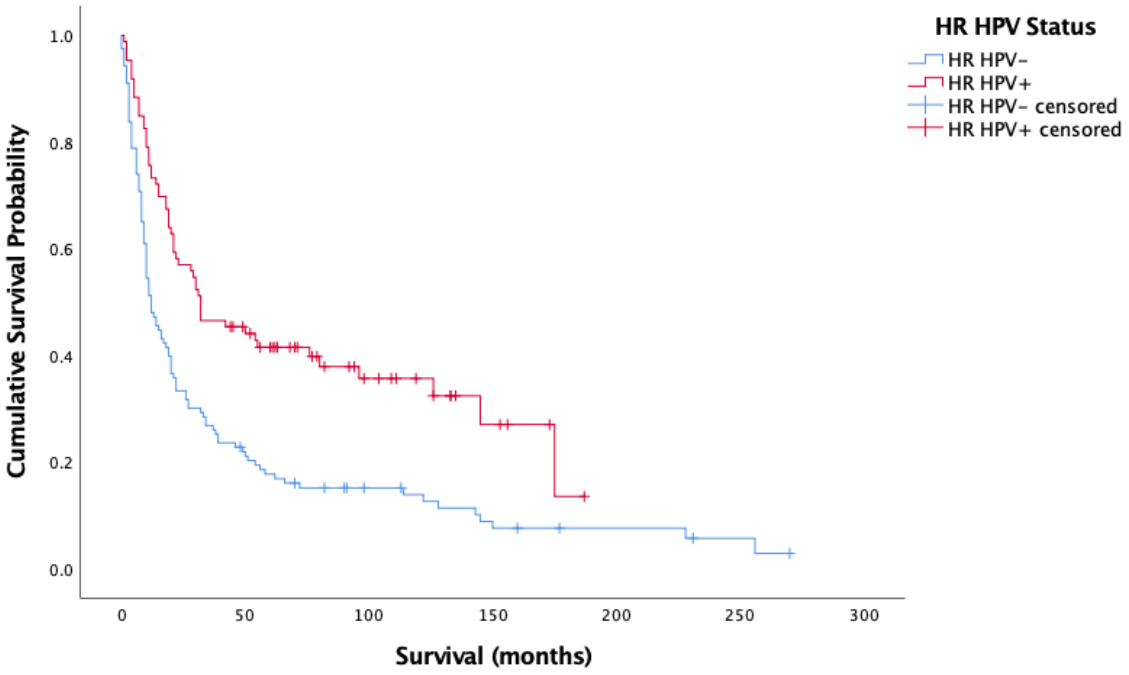

Figure 8.3 Kaplan-Meier analysis of overall survival in months based on HR HPV status for oropharyngeal cancer (n=209).

Figure 8.4 is the Kaplan-Meier result for cancer-specific survival stratified by HPV status for the oropharyngeal sub-site alone. Mirroring results for all cases, there was significantly improved cancer-specific survival amongst HPV positive patients (Log-rank: Chi-square 11.902, 1 d.f., p=0.001). Cox proportional hazard model agreed with this finding, seeing increased risk of cancer-specific death amongst HPV negative patients (HR=00.620, SE=0.185, p=0.001).

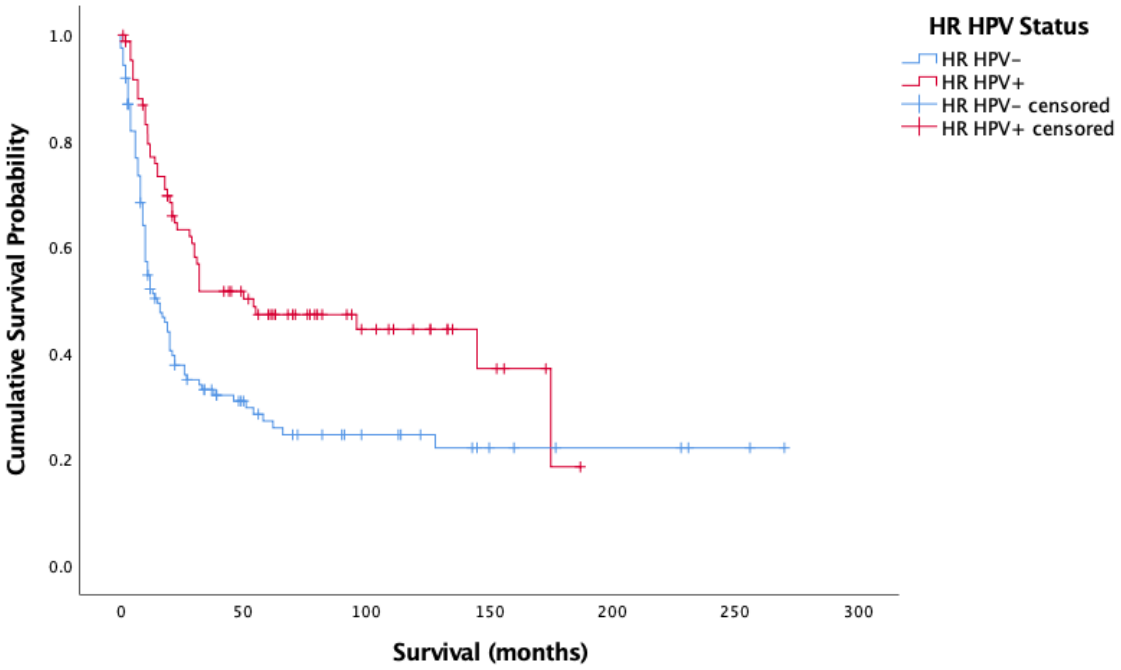

Figure 8.4 Kaplan-Meier analysis of disease-specific (cancer-specific) survival in months based on HR HPV status for oropharyngeal cancer(n=209).

8.4.3 Oral Cavity Cancer

Figure 8.5 shows the insignificant results of overall survival stratified by HPV status for oral cavity cancers by Kaplan-Meier analysis (Log-rank: Chi-square=0.872, 1 d.f., p=0.351). Cox proportional hazard model showed no significantly increased risk of death for HPV negative or positive cases (HR=0.183, SE=0.199, p=0.358).

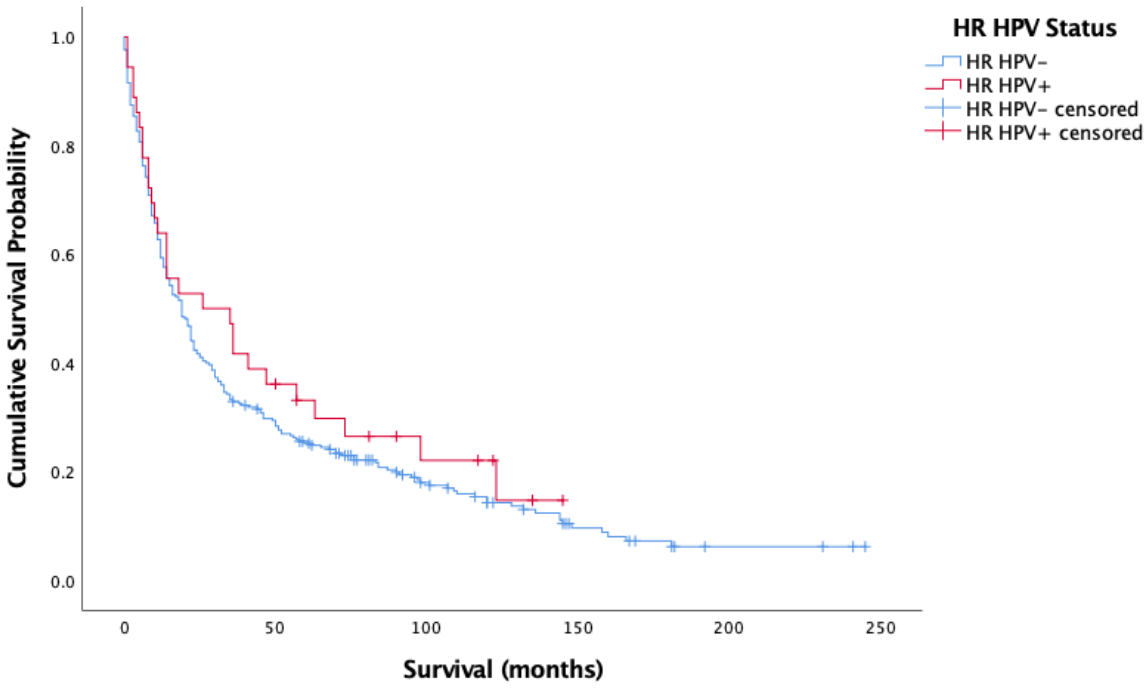

Figure 8.5 Kaplan-Meier analysis of overall survival in months based on HR HPV status for oral cavity cancer (n=331).

Figure 8.6 reveals no significant difference between HPV positive and negative cases for cancer-specific survival amongst oral cavity cases alone (Log-rank: Chi-square=0.051, 1 d.f., p=0.821). Cox proportional hazard model confirmed this finding (HR=0.048, SE=0.217, p=0.824).

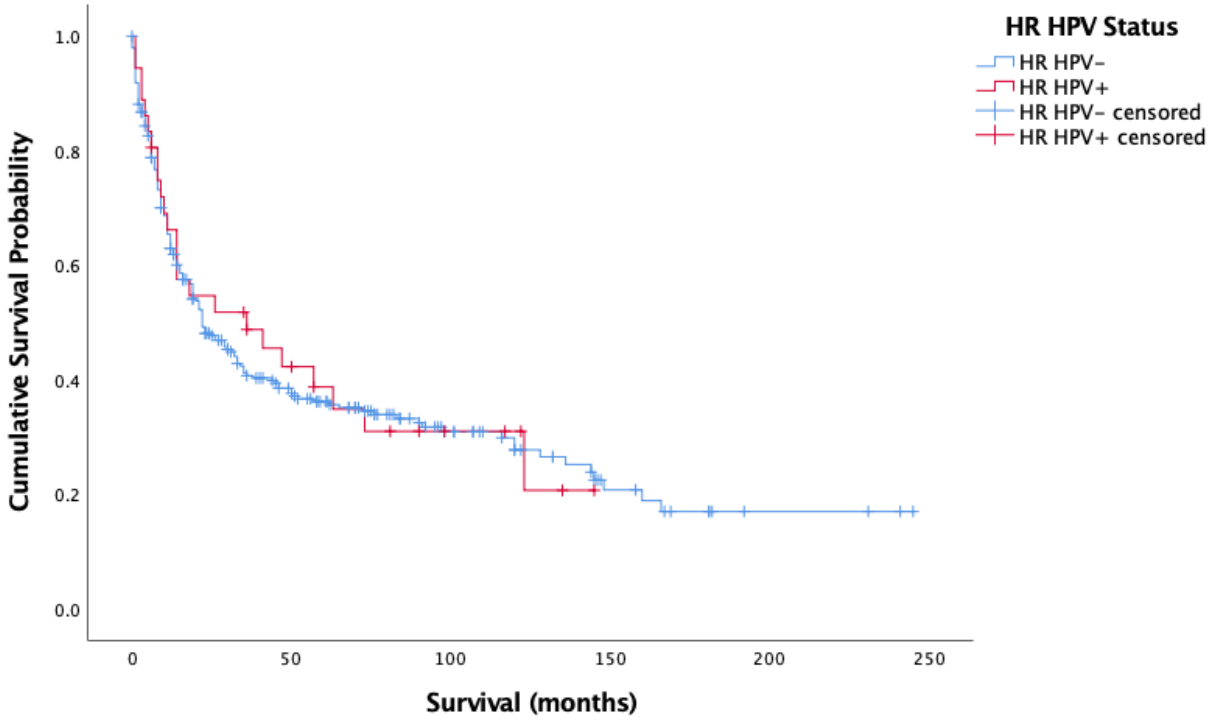

Figure 8.6 Kaplan-Meier analysis of disease-specific (cancer-specific) survival in months based on HR HPV status for oral cavity cancer (n=331).

8.4.4 Laryngeal Cancer

Figure 8.7 reveals no significant difference between HPV positive and negative laryngeal cases for overall survival by Kaplan-Meier analysis (Log-rank: Chi square=0.216, 1 d.f., p=0.642). Cox proportional hazard model also saw no significantly increased risk of death overall for HPV positive or negative cases (HR=0.112, 1 d.f., p=0.645).

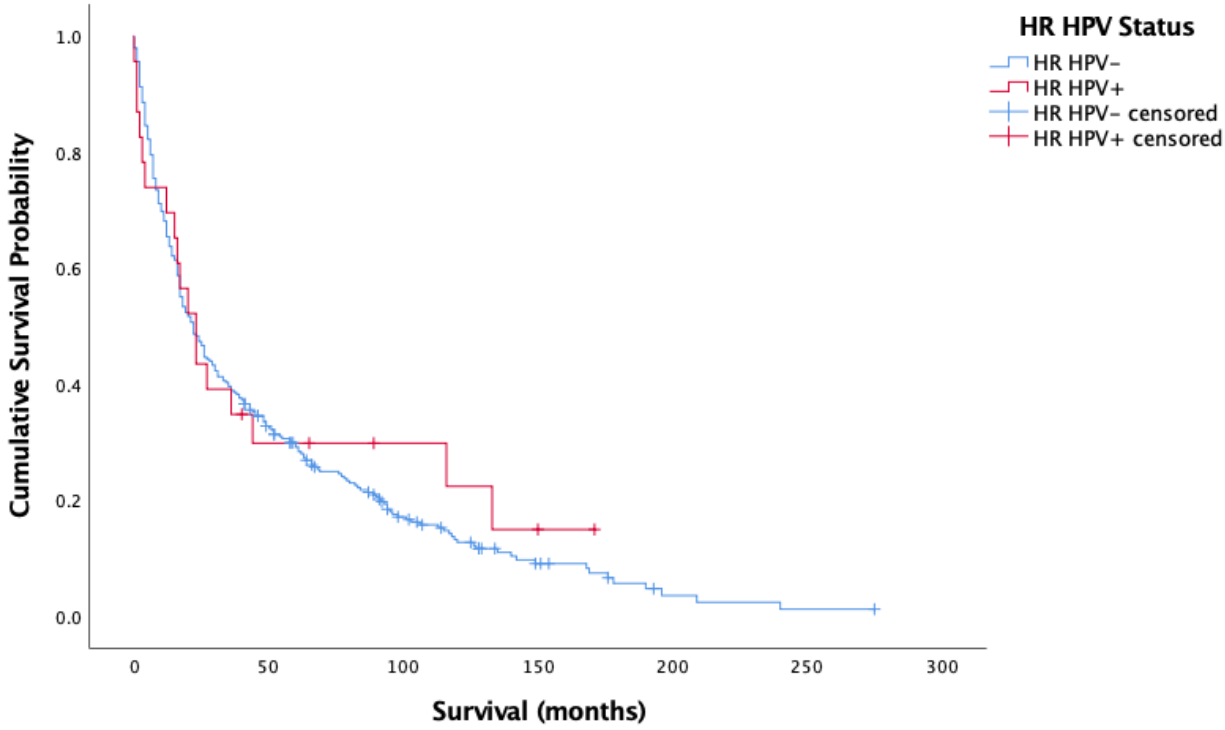

Figure 8.7 Kaplan-Meier analysis of overall survival in months based on HR HPV status for laryngeal cancer (n=321).

Figure 8.8 similarly reveals no significant difference in cancer-specific survival probability between HPV positive and negative cases in the larynx by Kaplan-Meier analysis (Log-rank: Chi-square=0.172, 1 d.f., p=0.678). Cox proportional hazard model confirmed this insignificance in the laryngeal sub-site (HR=0.128, SE=0.312, p=0.681).

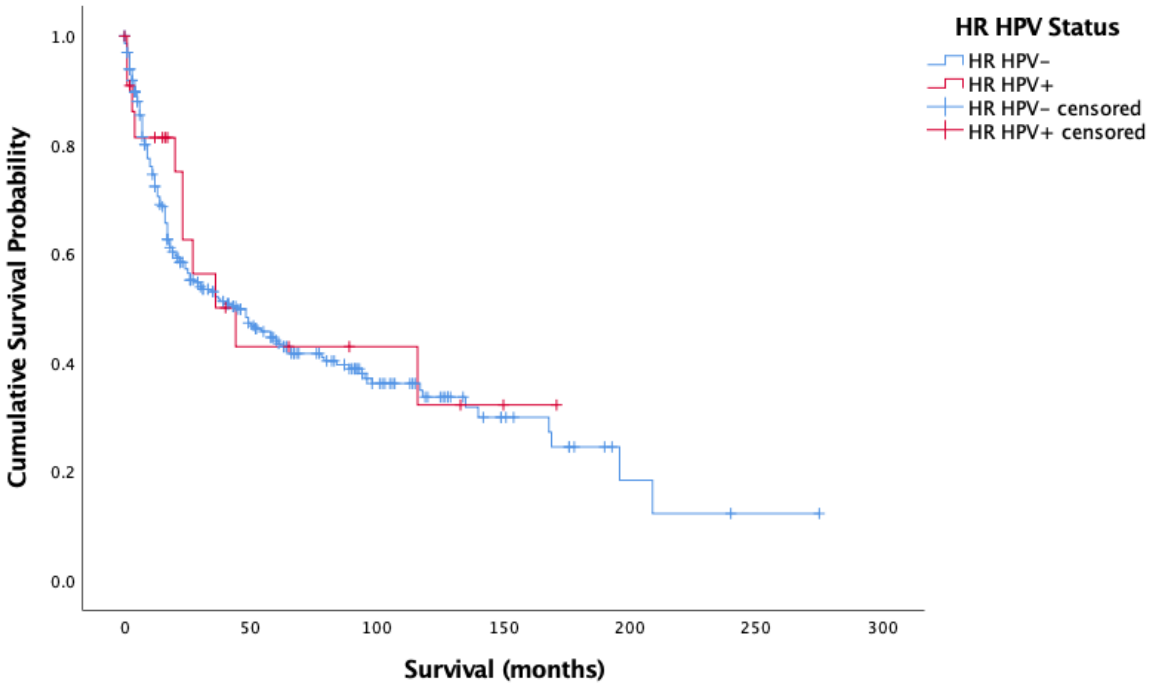

Figure 8.8 Kaplan-Meier analysis of disease-specific (cancer-specific) survival in months based on HR HPV status for laryngeal cancer (n=321).

8.5 Treatment and HPV Status in Oropharyngeal, Oral Cavity, and Laryngeal Cancer

To assess the relationship between treatment administered and HR HPV status, Chi-square and Fisher's exact analyses were conducted where appropriate for all cases in the population and then individually for each key sub-site.

8.5.1 Oropharyngeal, Oral Cavity, and Laryngeal Cancer

Table 8.4 showcases the distribution of treatment administered to patients according to HR HPV status for all cancer patients in the population who were treated within 12 months of diagnosis.

Table 8.4 Treatment administered by HPV status for oropharyngeal, oral cavity, and laryngeal cancer (n=758).

Treatment	Count	HR HPV Status		Total
		HR HPV-	*HR HPV+*	
Chemotherapy	*Count*	4	0	4
	% within Treatment	100.0%	0.0%	100.0%
	% within HR HPV status	0.6%	0.0%	0.5%
	% of Total	0.5%	0.0%	0.5%
Radiotherapy	*Count*	214	24	238
	% within Treatment	89.9%	10.1%	100.0%
	% within HR HPV status	34.1%	18.3%	31.4%
	% of Total	28.2%	3.2%	31.4%
Radiotherapy/ Chemotherapy	*Count*	76	43	119
	% within Treatment	63.9%	36.1%	100.0%
	% within HR HPV status	12.1%	32.8%	15.7%
	% of Total	10.0%	5.7%	15.7%
Surgery	*Count*	116	17	133
	% within Treatment	87.2%	12.8%	100.0%
	% within HR HPV status	18.5%	13.0%	17.5%
	% of Total	15.3%	2.2%	17.5%
Surgery/ Chemotherapy	*Count*	2	0	2
	% within Treatment	100.0%	0.0%	100.0%
	% within HR HPV status	0.3%	0.0%	0.3%
	% of Total	0.3%	0.0%	0.3%
Surgery/ Radiotherapy	*Count*	159	24	183
	% within Treatment	86.9%	13.1%	100.0%
	% within HR HPV status	25.4%	18.3%	24.1%
	% of Total	21.0%	3.2%	24.1%
Surgery/ Radiotherapy/ Chemotherapy	*Count*	56	23	79
	% within Treatment	70.9%	29.1%	100.0%
	% within HR HPV status	8.9%	17.6%	10.4%
	% of Total	7.4%	3.0%	10.4%
Total	*Count*	627	131	758
	% within Treatment	82.7%	17.3%	100.0%
	% within HR HPV status	100.0%	100.0%	100.0%
	% of Total	82.7%	17.3%	100.0%

Due to very low frequencies amongst patients treated with chemotherapy and chemotherapy/surgery alone, these patients were excluded for the analysis. There was significant relationship between HPV status and treatment administered (Chi-square=49.732, 4 d.f., $p<0.0001$). HPV positive patients were almost twice as likely to be treated more harshly than HPV negative cases with all three treatment modalities. More HPV negative patients were treated with surgery or radiotherapy alone and almost three times as many HPV positive patients were treated chemically with radiotherapy/chemotherapy than HPV negative patients.

8.5.2 Oropharyngeal Cancer

Table 8.5 shows the distribution of treatment administered to HPV positive and negative cases in the oropharynx alone for patients treated within 12 months of diagnosis.

Table 8.5 Treatment administered by HPV status for oropharyngeal cancer (n=191).

Treatment	Count	HR HPV Status		Total
		HR HPV-	*HR HPV+*	
Chemotherapy	*Count*	2	0	2
	% within Treatment	100.0%	0.0%	100.0%
	% within HR HPV status	1.8%	0.0%	1.0%
	% of Total	1.0%	0.0%	1.0%
Radiotherapy	*Count*	42	16	58
	% within Treatment	72.4%	27.6%	100.0%
	% within HR HPV status	37.8%	20.0%	30.4%
	% of Total	22.0%	8.4%	30.4%
Radiotherapy/ Chemotherapy	*Count*	26	30	56
	% within Treatment	46.4%	53.6%	100.0%
	% within HR HPV status	23.4%	37.5%	29.3%
	% of Total	13.6%	15.7%	29.3%
Surgery	*Count*	13	5	18
	% within Treatment	72.2%	27.8%	100.0%
	% within HR HPV status	11.7%	6.3%	9.4%
	% of Total	6.8%	2.6%	9.4%
Surgery/ Radiotherapy	*Count*	18	12	30
	% within Treatment	60.0%	40.0%	100.0%
	% within HR HPV status	16.2%	15.0%	15.7%
	% of Total	9.4%	6.3%	15.7%
Surgery/ Radiotherapy/ Chemotherapy	*Count*	10	17	27
	% within Treatment	37.0%	63.0%	100.0%
	% within HR HPV status	9.0%	21.3%	14.1%
	% of Total	5.2%	8.9%	14.1%
Total	*Count*	111	80	191
	% within Treatment	58.1%	41.9%	100.0%
	% within HR HPV status	100.0%	100.0%	100.0%
	% of Total	58.1%	41.9%	100.0%

Due to low expected counts within the chemotherapy group, the analysis was conducted without these 2 cases. There was a significant association between treatment and HPV status (Chi-square=14.401, 4 d.f., p=0.006). Proportionally, twice as many HPV positive cases were treated with surgery/radiotherapy/chemotherapy than HPV negative cases. The reverse was true for both radiotherapy alone and surgery alone, with twice as many HPV

negative patients being treated with these less aggressive approaches within 12 months of diagnosis.

8.5.3 Oral Cavity Cancer

Table 8.6 reveals the distribution of treatment administered by HR HPV status for the oral cavity sub-site alone for cases who were treated within 12 months of diagnosis.

Table 8.6 Treatment administered by HPV status for oral cavity cancer (n=290).

Treatment	Count	HR HPV Status		Total
		HR HPV-	HR HPV+	
Chemotherapy	Count	2	0	2
	% within Treatment	100.0%	0.0%	100.0%
	% within HR HPV status	0.8%	0.0%	0.7%
	% of Total	0.7%	0.0%	0.7%
Radiotherapy	Count	45	3	48
	% within Treatment	93.8%	6.3%	100.0%
	% within HR HPV status	17.4%	9.7%	16.6%
	% of Total	15.5%	1.0%	16.6%
Radiotherapy/ Chemotherapy	Count	23	8	31
	% within Treatment	74.2%	25.8%	100.0%
	% within HR HPV status	8.9%	25.8%	10.7%
	% of Total	7.9%	2.8%	10.7%
Surgery	Count	81	8	89
	% within Treatment	91.0%	9.0%	100.0%
	% within HR HPV status	31.3%	25.8%	30.7%
	% of Total	27.9%	2.8%	30.7%
Surgery/ Radiotherapy	Count	82	10	92
	% within Treatment	89.1%	10.9%	100.0%
	% within HR HPV status	31.7%	32.3%	31.7%
	% of Total	28.3%	3.4%	31.7%
Surgery/ Radiotherapy/ Chemotherapy	Count	26	2	28
	% within Treatment	92.9%	7.1%	100.0%
	% within HR HPV status	10.0%	6.5%	9.7%
	% of Total	9.0%	0.7%	9.7%
Total	Count	259	31	290
	% within Treatment	89.3%	10.7%	100.0%
	% within HR HPV status	100.0%	100.0%	100.0%
	% of Total	89.3%	10.7%	100.0%

Due to low frequencies of chemotherapy-treated patients, the analysis conducted without these cases and showed no significant association between treatment and HPV status (Chi-square=7.837, 4 d.f., p=0.098).

8.5.4 Laryngeal Cancer

Table 8.7 shows the distribution of treatment administered by HR HPV status in the laryngeal sub-site for cases who were treated within 12 months of diagnosis.

Table 8.7 Treatment administered by HPV status for laryngeal cancer (n=277).

Treatment	Count	HR HPV Status		Total
		HR HPV-	*HR HPV+*	
Radiotherapy	*Count*	127	5	132
	% within Treatment	96.2%	3.8%	100.0%
	% within HR HPV status	49.4%	25.0%	47.7%
	% of Total	45.8%	1.8%	47.7%
Radiotherapy/ Chemotherapy	*Count*	27	5	32
	% within Treatment	84.4%	15.6%	100.0%
	% within HR HPV status	10.5%	25.0%	11.6%
	% of Total	9.7%	1.8%	11.6%
Surgery	*Count*	22	4	26
	% within Treatment	84.6%	15.4%	100.0%
	% within HR HPV status	8.6%	20.0%	9.4%
	% of Total	7.9%	1.4%	9.4%
Surgery/ Chemotherapy	*Count*	2	0	2
	% within Treatment	100.0%	0.0%	100.0%
	% within HR HPV status	0.8%	0.0%	0.7%
	% of Total	0.7%	0.0%	0.7%
Surgery/Radiotherapy	*Count*	59	2	61
	% within Treatment	96.7%	3.3%	100.0%
	% within HR HPV status	23.0%	10.0%	22.0%
	% of Total	21.3%	0.7%	22.0%
Surgery/ Radiotherapy/ Chemotherapy	*Count*	20	4	24
	% within Treatment	83.3%	16.7%	100.0%
	% within HR HPV status	7.8%	20.0%	8.7%
	% of Total	7.2%	1.4%	8.7%
Total	*Count*	257	20	277
	% within Treatment	92.8%	7.2%	100.0%
	% within HR HPV status	100.0%	100.0%	100.0%
	% of Total	92.8%	7.2%	100.0%

The analysis was conducted excluding the very low number of patients treated with surgery/chemotherapy and showed a significant association between treatment and HPV status (Fisher's exact=12.423, p=0.007). It is very clear from the HPV negative column of Table 8.7 that these differences emanated from the disproportionate treatment of HPV negative cases with radiotherapy alone, and the more frequent treatment of HPV positive patients with all three modalities, surgery, and radiotherapy/chemotherapy.

8.6 HPV Status, Treatment, and Survival

To evaluate the relationship between HPV status, treatment, and survival, Kaplan-Meier assessment accompanied by cox proportional hazard models were performed on both HPV positive and HPV negative cases for all oropharyngeal, oral cavity, and laryngeal SCC and each individual sub-site for both overall and cancer-specific survival. Due to extremely low frequencies of patients treated using only Chemotherapy or Chemotherapy/Surgery (n=6), these cases were excluded.

8.6.1 Oropharyngeal, Oral Cavity, and Laryngeal Cancer

Amongst HPV positive cases, there was a significant difference between overall survival rates by treatment types (Log-rank=11.194, 4 d.f., p=0.024). Figure 8.9 represents this relationship. Cox proportional hazard model showed a significant difference in risk of death between treatments, with surgery/radiotherapy/chemotherapy showing least risk of death, followed by significantly increased risk of death amongst radiotherapy/chemotherapy, surgery, and surgery/radiotherapy, followed by significantly increased risk of death amongst radiotherapy patients (HR=1.032, 0.340, 0.264, 0.321, SE=0.371, 0.358, 0.427, 0.392, p=0.033, 0.005, 0.343, 0.538, 0.413).

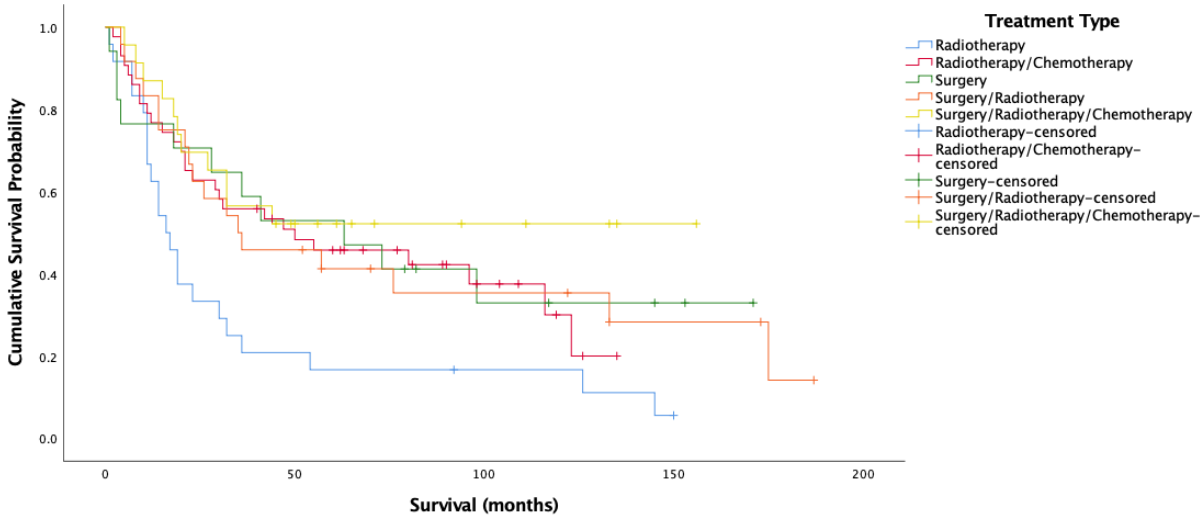

Figure 8.9 Kaplan-Meier analysis for overall survival amongst HPV positive oropharyngeal, oral cavity, and laryngeal cancer stratified by treatment type (n=131).

For cancer-specific survival amongst HPV positive patients, there was no significant difference between treatment types (Log-rank=4.751, 4 d.f., p=0.314), something also reflected in cox proportional hazard model (HR=0.748, 0.179, 0.171, 0.141, SE=0.410, 0.385, 0.461, 0.429, p=0.338, 0.068, 0.642, 0.711, 0.642).

For HPV negative cases, there was a significant difference in overall survival by treatment type exhibited in Figure 8.10 (Log-rank=17.868, 4 d.f., p=0.001). Cox proportional hazard model reflected this difference indicating lowest risk of death amongst patients treated with surgery, with significantly increased risk of death for patients treated with surgery/radiotherapy, radiotherapy, and all three treatment modalities, and highest risk of death amongst radiotherapy/chemotherapy patients (HR=-0.069, 0.212, =0.411, -0.202, SE=0.158, 0.185, 0.176, 0.164, p=0.002, 0.661, 0.252, 0.019, 0.219).

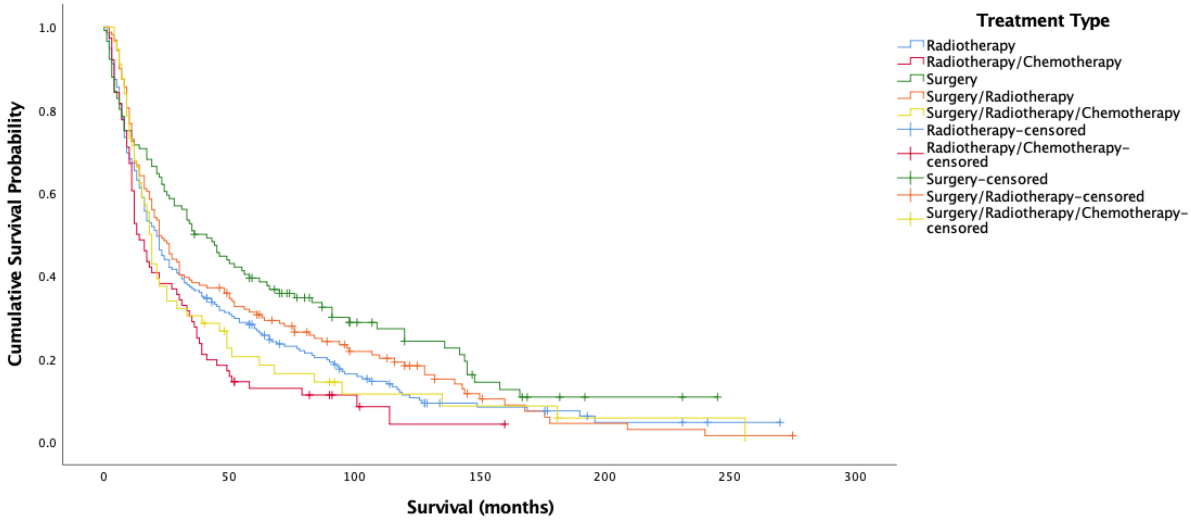

Figure 8.10 Kaplan-Meier analysis for overall survival amongst HPV negative oropharyngeal, oral cavity, and laryngeal cancer stratified by treatment type (n=621).

For HPV negative cases, there was also a significant difference in cancer-specific survival by treatment type showcased in Figure 8.11 (Log-rank=10.691, 4 d.f., p=0.030). Cox proportional hazard model reflected this difference with least risk of death amongst patients treated with surgery, followed by significantly increased risk of death amongst radiotherapy and surgery/radiotherapy patients, and even further increased death amongst patients treated with all three modalities and radiotherapy/chemotherapy (HR=-0.236, -0.043, -0.564, -0.287, SE=0.180, 0.216, 0.204, 0.186, p=0.035, 0.190, 0.844, 0.006, 0.123).

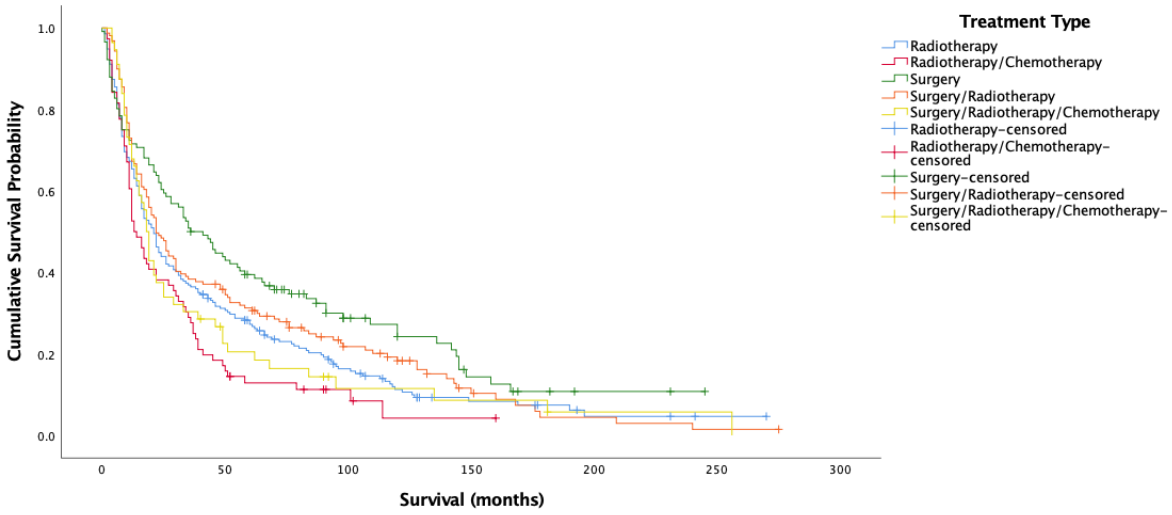

Figure 8.11 Kaplan-Meier analysis for cancer-specific survival amongst HPV negative oropharyngeal, oral cavity, and laryngeal cancer stratified by treatment type (n=621).

8.6.2 Oropharyngeal Cancer

For HPV positive cases in the oropharynx, there was a significant difference in overall survival by treatment type (Log-rank=10.481, 4 d.f., p=0.033) (n=80) (Figure 8.12). Cox proportional hazard model reflected this significance for all treatments (HR=1.166, 0.579, -0.055, 0.096, SE=0.460, 0.443, 0.803, 0.559, p=0.049, 0.011, 0.192, 0.946, 0.864). Surgery alone and all three treatment modalities maximized overall survival for HPV positive OPSCC patients. Radiotherapy minimized survival.

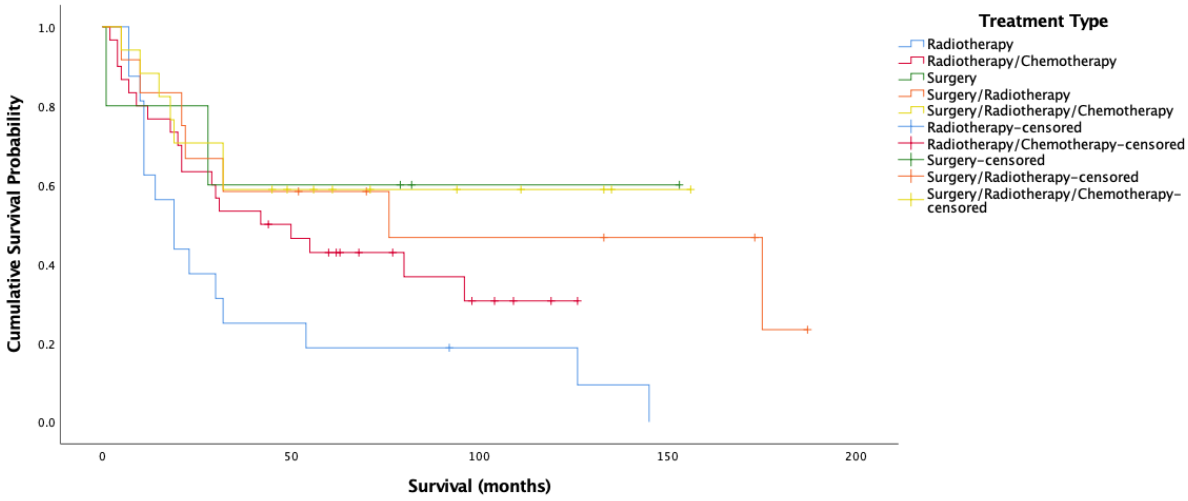

Figure 8.12 Kaplan-Meier analysis for overall survival amongst HPV positive oropharyngeal cancer stratified by treatment type (n=80).

For cancer-specific survival amongst HPV positive oropharyngeal cases, there was a significant difference by treatment type (Log-rank=11.398, 4 d.f., p=0.022). Figure 8.13 showcases this difference. Indeed, this significance was reflected by cox proportional hazard model, with surgery having the lowest risk of death, followed by surgery/radiotherapy and surgery/radiotherapy/chemotherapy, followed by radiotherapy/chemotherapy, with radiotherapy having the worst risk of death (HR=1.186, 0.490, -0.567, -0.137, SE=0.496, 0.484, 1.082, 0.649, p=0.039, 0.017, 0.312, 0.600, 0.833).

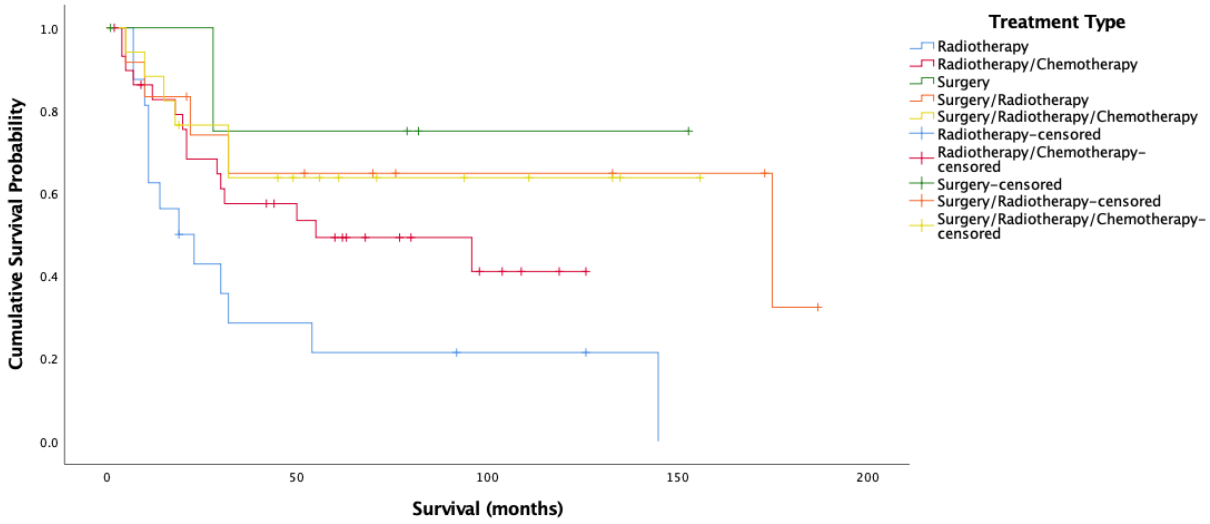

Figure 8.13 Kaplan-Meier analysis for cancer-specific survival amongst HPV positive oropharyngeal cancer stratified by treatment type (n=80).

For HPV negative oropharyngeal cases, there was no difference in overall survival by treatment type (HR=1.983, 4 d.f., p=0.739) (n=109), something also reflected by cox proportional hazard model (HR=0.227, 0.198, -0.208, 0.035, SE=0.373, 0.395, 0.465, 0.416, p=0.754, 0.543, 0.615, 0.654, 0.932). There was also no significant difference in cancer-specific survival by treatment type amongst HPV negative oropharyngeal cases (Log-rank=1.189, 4 d.f., p=0.880) (n=109). This was again seconded by cox proportional hazard model (HR=0.057, -0.139, -0.315, -0.151, SE=0.400, 0.434, 0.503, 0.450, p=0.887, 0.886, 0.748, 0.531, 0.737).

8.6.3 Oral Cavity Cancer

For HPV positive oral cavity cases, there was no significant difference between overall survivals by treatment type (Log-rank=5.013, 4 d.f., p=0.286) (n=31). This was reflected by cox proportional hazard model (HR=1.043, -0.405, -0.293, 0.306, SE=1.163, 1.117, 1.098, 1.066, p=0.350, 0.370, 0.717, 0.789, 0.774). Amongst HPV positive oral cavity cancer, there was also no difference in cancer-specific survival by treatment type (Log-rank=3.030, 4 d.f., p=0.553). Cox proportional hazard showed the same result (HR=0.617, -0.646, -0.462, 0.161, SE=1.236, 1.146, 1.116, 1.075, p=0.593, 0.618, 0.573, 0.679, 0.881).

For HPV negative oral cavity cancer, there was a significant difference between overall survivals by treatment type (Log-rank=41.454, 4 d.f., p<0.0001) (n=257). Cox proportional hazard model reflected this with surgery seeing the lowest risk of death, following by surgery/radiotherapy and surgery/radiotherapy/chemotherapy, followed by radiotherapy/chemotherapy, and radiotherapy having the highest risk of death (HR=0.801, 0.544, -0.385, 0.010, SE=0.263, 0.304, 0.250, 0.245, p<0.0001, 0.002, 0.074, 0.123, 0.966). Figure 8.14 illustrates this relationship.

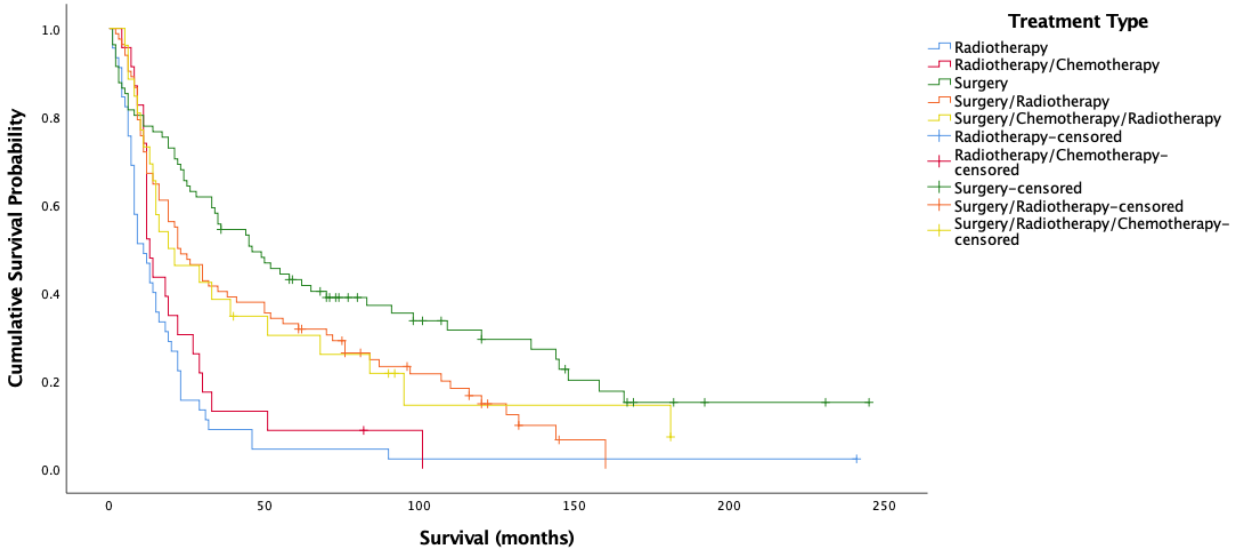

Figure 8.14 Kaplan-Meier analysis for overall survival amongst HPV negative oral cavity cancer stratified by treatment type (n=257).

Amongst HPV negative oral cavity cancers, there was also a significant difference between cancer-specific survivals by treatment type (Log-rank=41.897, 4 d.f., p<0.001). Cox proportional hazard reflected this with surgery seeing the least risk of death, followed by surgery/radiotherapy and all three treatment modalities, followed by radiotherapy/chemotherapy, and radiotherapy having the greatest risk of death by a significant margin (HR=0.976, 0.521, -0.347, -0.008, SE=0.306, 0.362, 0.303, 0.298, p<0.0001, 0.001, 0.150, 0.251, 0.978). Figure 8.15 illustrates this relationship.

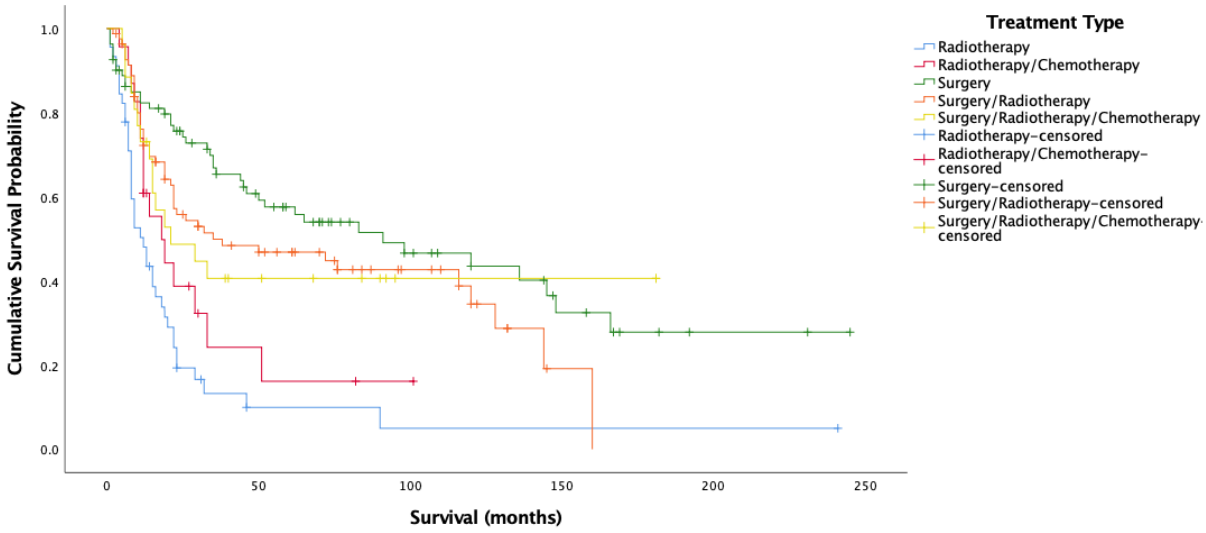

Figure 8.15 Kaplan-Meier analysis for cancer-specific survival amongst HPV negative oral cavity cancer stratified by treatment type (n=257).

8.6.4 Laryngeal Cancer

Amongst HPV positive laryngeal cases, there was no significant difference in overall survival by treatment type (Log-rank=1.161, 4 d.f., p=0.885) which was also reflected by cox proportional hazard model (HR=0.559, -0.214, 0.218, 0.004, SE=0.789, 0.820, 0.833, 0.932, SE=0.891, 0.479, 0.794, 0.794, 0.996) (n=20). The same was true for cancer-specific survival (Log-rank=3.124, 4 d.f., p=0.537) (HR=-12.660, -0.702, 0.157, -0.700, SE=395.453, 0.923, 0.844, 1.184, p=0.871, 0.974, 0.447, 0.853, 0.554).

For HPV negative laryngeal cases, there was a significant difference in overall survival by treatment type (Log-rank=13.224, 4 d.f., p=0.010), something mirrored by cox proportional hazard model (HR=-0.665, -0.177, -0.278, -0.695, SE=0.246, 0.307, 0.314, 0.270, p=0.014, 0.007, 0.564, 0.375, 0.010). As represented in Figure 8.16, radiotherapy and surgery/radiotherapy had the best survival overall, followed by surgery alone, radiotherapy/chemotherapy, and all three treatment modalities.

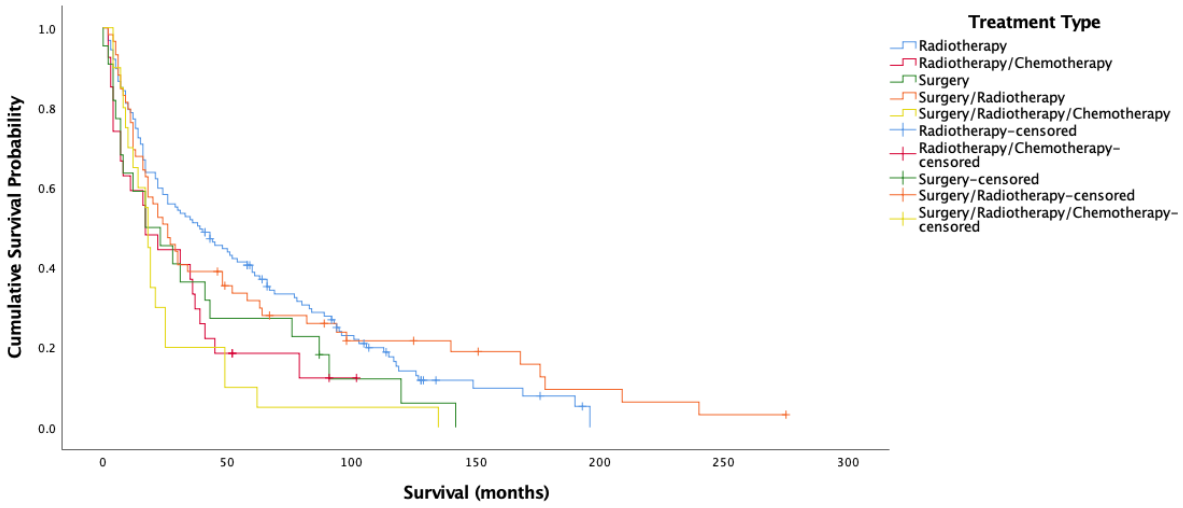

Figure 8.16 Kaplan-Meier analysis for overall survival amongst HPV negative laryngeal cancer stratified by treatment type (n=255).

For HPV negative laryngeal cases, there was a significant different between survival at the cancer-specific level by treatment type (Log-rank=16.474, 4 d.f., p=0.002) This was replicated by cox proportional hazard model (HR=-1.043, -0.605, -0.805, -0.692, SE=0.273, 0.365, 0.396, 0.290, p=0.004, 0.0001, 0.097, 0.042, 0.017). As illustrated by Figure 8.17, all three treatment modalities had significantly worse survival than other treatments. Radiotherapy maximized survival most consistently, but data makes it unclear as to whether this remained true after 200 months.

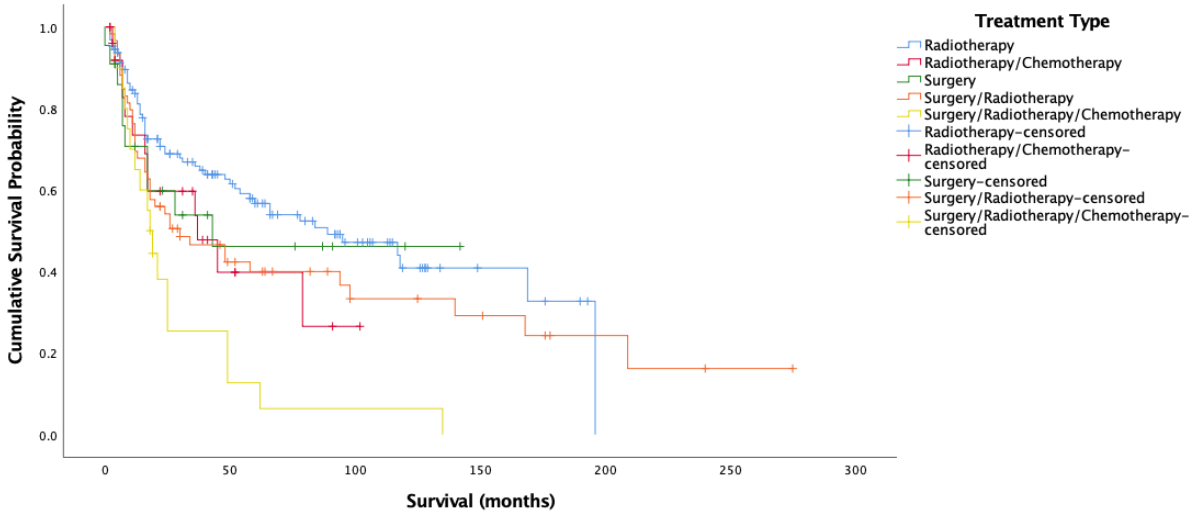

Figure 8.17 Kaplan-Meier analysis for cancer-specific survival amongst HPV negative laryngeal cancer stratified by treatment type (n=255).

8.7 Predictors of Survival

Univariate and multivariate cox proportional hazard models for risk of death based on patient and tumour characteristics were conducted for all cancers in the population, oropharyngeal cancer, oral cavity cancer, and laryngeal cancer for both overall and cancer-specific survival. It should also be noted that where age was assessed by univariate analysis in both continuous form and by categorical variable ($Age \leq 50$), it was only used in its continuous form in multivariate analyses. Furthermore, the 6 patients treated with either chemotherapy or surgery/chemotherapy were excluded from analysis given that it would be extremely risky to make determinations regarding survival on the basis of so few patients.

8.7.1 Oropharyngeal, Oral Cavity, and Laryngeal Cancer

Table 8.8 summarizes the patient and tumour characteristics significantly predictive of overall survival by univariate analysis for all cases. These statistics were presented in Chapter 6 except for that including HPV status but are shown here again to assure fluency to the multivariate model.

Table 8.8 Patient and tumour characteristics significantly predicting overall survival amongst oropharyngeal, oral cavity, and laryngeal cancer patients by univariate cox proportional hazard models.

Variable/Factor	Statistic	Increased Risk of Death
HR HPV Status (n=861)	HR=0.372 SE=0.107 P<0.0001	HPV negativity
Age (n=861)	HR=0.028 SE=0.003 P<0.0001	Older age
Age$\leq$ 50 (n=861)	HR=-0.659 SE=0.122 P<0.0001	>50
Sex (n=861)	HR=-0.232 SE=0.091 P=0.011	Male>Female
Smoking Status (n=861)	HR=-0.559, -0.189, 0.130 SE=0.125, 0.101, 0.112 P<0.0001, 0.0001, 0.060, 0.247	Current smoker, Missing > Ex-smoker
T Stage (n=861)	HR=-0.436, -1.192, -0.638, -0.495 SE=0.118, 0.124, 0.101, 0.117 P<0.0001, 0.001, 0.0001, 0.0001, 0.0001	T4>T3, T2, Missing>T1
N Stage (n=861)	HR=-0.161, -0.484, -0.097 SE=0.112, 0.091, 0.116 P<0.0001, 0.151, 0.0001, 0.404	N2/3, N1, Missing>N0
M Stage (n=861)	HR=-0.792, -0.872 SE=0.162, 0.161 P<0.0001, 0.0001, 0.0001	M1>M0, Missing
TNM (n=861)	HR=-0.294, -0.970, -0.468, -0.503 SE=0.117, 0.124, 0.112, 0.111 P<0.0001, 0.012, 0.0001, 0.0001, 0.0001	Stage IV>Missing, Stage III, Stage II> Stage I
Marital Status (n=830)	HR=-0.168, -0.291, 0.193 SE=0.159, 0.111, 0.123 P<0.0001, 0.291, 0.009, 0.116	Single > Married

Treatment Type (n=752)	HR=0.193, 0.152, -0.197, -0.004 SE=0.144, 0.163, 0.162, 0.150 P=0.017, 0.193, 0.351, 0.223, 0.977	Radiotherapy, Radiotherapy/Chemotherapy > Surgery

Due to evident and expected confounding between T, N, M, and TNM stages, TNM stage was used as a representative variable for each cancer stage in the multivariate model. Results of this model are presented in the Table below.

Table 8.9 Patient and tumour characteristics significantly predicting overall survival amongst oropharyngeal, oral cavity, and laryngeal cancer patients by multivariate cox proportional hazard model. The initial model contained all variables univariately significant except for T, N, and M stage (n=727).

Variable/Factor	Statistic	Increased Risk of Death
Age (Continuous)	HR=0.027 SE=0.004 P<0.0001	Older age
HR HPV Status	HR=0.356 SE=0.126 P=0.005	HPV negativity
Smoking Status	HR=-0.529, -0.164, -0.017 SE=0.141, 0.113, 0.136 P=0.002, 0.0001, 0.146, 0.900	Current smoker, Missing > Never smoker > Ex-smoker
TNM Stage	HR=0.-0.530, -0.990, -0.493, -0.506 SE=0.142, 0.135, 0.126, 0.121 P<0.0001, 0.0001, 0.0001, 0.0001, 0.0001	Stage IV > Stage III, Stage II, Missing > Stage I
Marital Status	HR=0.034, -0.035, 0.310 SE=0.179, 0.127, 0.140 P=0.008, 0.848, 0.784, 0.026	Single > Separated/Divorced, Married

Table 8.10 reveals the variables significantly predictive of cancer-specific survival by univariate cox proportional hazard model for all cases. These statistics were presented in Chapter 6 except for that including HPV status but are shown here again to assure fluency to the multivariate model.

Table 8.10 Patient and tumour characteristics significantly predicting disease-specific (cancer-specific) survival amongst oropharyngeal, oral cavity, and laryngeal cancer patients by univariate cox proportional hazard models.

Variable/Factor	Statistic	Increased Risk of Death
HR HPV Status (n=861)	HR=0.257 SE=0.122 p=0.035	HPV negativity
Age$\leq$ 50 (n=861)	HR=-0.385 SE=0.132 P=0.004	>50
Age (Continuous) (n=861)	HR=0.018 SE=0.004 P<0.0001	Older age
Smoking Status (n=861)	HR=-0.679, -0.301, 0.144 SE=0.157, 0.122, 0.129 P<0.0001, 0.0001, 0.014, 0.263	Current smoker, Missing > Never smoker > Ex-smoker
Grade (n=859)	HR=-0.520, -0.399, -0.320 SE=0.158, 0.162, 0.106 P=0.002, 0.001, 0.014, 0.002	Poorly-differentiated > Well-differentiated, Moderately differentiated, Missing
T Stage (n=861)	HR=-0.507, -1.412, -0.775, -0.501 SE=0.136, 0.154, 0.118, 0.133 P<0.0001, 0.0001, 0.0001, 0.0001, 0.0001	T4>T3, Missing>T2>T1
N Stage (n=861)	HR=-0.287, -0.651, -0.189 SE=0.132, 0.107, 0.134 P<0.0001, 0.030, 0.0001, 0.158	N2/N3, N1,Missing>N0
M Stage (n=861)	OR=-0.899, -0.975 SE=0.177, 0.176 P<0.0001, 0.0001, 0.0001	M1>M0, Missing
TNM Stage (n=861)	HR=-0.412, -1.195, -0.673, -0.616 SE=0.139, 0.158, 0.138, 0.133 P<0.0001, 0.003, 0.0001, 0.0001, 0.0001	Stage IV > Missing > Stage III, Stage II > Stage I
Marital Status	HR=-0.185, -0.255, 0.225	Single

(n=830)	SE=0.190, 0.133, 0.145 P<0.0001, 0.329, 0.056, 0.121	> Divorced/Separated, Married

Due to evident and expected confounding between T, N, M, and TNM stages, TNM stage was used as a representative variable for each cancer stage in the multivariate model. Results of this model are presented in the Table below.

Table 8.11 Patient and tumour characteristics significantly predicting disease-specific (cancer-specific) survival amongst oropharyngeal, oral cavity, and laryngeal cancer patients by multivariate cox proportional hazard model. The original model included all significant variables from univariate analysis but excluded T, N, and M stage (n=828).

Variable/Factor	Statistic	Increased Risk of Death
Age (Continuous)	HR=0.025 SE=0.005 P<0.0001	Older age
Marital Status	HR=-0.167, -0.077, 0.232 SE=0.199, 0.138, 0.150 P=0.021, 0.401, 0.574, 0.124	Single > Divorced/Separated, Married
Smoking Status	HR=-0.672, -0.384, -0.016 SE=0.158, 0.126, 0.139 P<0.0001, 0.0001, 0.002, 0.907	Current smoker, Missing > Never smoker > Ex-smoker
TNM Stage	HR=-0.541, -1.207, -0.683, -0.606 SE=0.146, 0.161, 0.144, 0.136 P<0.0001, 0.0001, 0.0001, 0.0001, 0.0001	Stage IV > Missing, Stage III, Stage II > Stage I

8.7.2 Oropharyngeal Cancer

Table 8.12 summarizes the variables significantly predictive of overall survival amongst all oropharyngeal cases in the population by univariate cox proportional hazard models.

Table 8.12 Patient and tumour characteristics significantly predicting overall survival amongst oropharyngeal cancer patients by univariate cox proportional hazard models.

Variable/Factor	Statistic	Increased Risk of Death
HR HPV Status (n=209)	HR=0.659 SE=0.165 P<0.0001	HPV negativity
Age (Continuous) (n=209)	HR=0.034 SE=0.008 P<0.0001	Older age
Age$\leq$ 50 (n=209)	HR=-0.709 SE=0.211 P=0.001	>50
Smoking Status (n=209)	HR=-0.730, -0.589, -0.080 SE=0.285, 0.223, 0.240 P=0.008, 0.010, 0.008, 0.738	Current smoker, Missing > Never smoker, Ex-smoker
T Stage (n=209)	HR=-0.291, -1.529, -0.557, -0.546 SE=0.229, 0.362, 0.203, 0.232 P<0.0001, 0.203, 0.0001, 0.006, 0.019	T4>Missing>T3, T2>T1
M Stage (n=209)	HR=-0.769, -0.951 SE=0.288, 0.286 P=0.004, 0.008, 0.001	M1>M0, Missing
Urban v Rural (n=209)	HR=-0.373 SE=0.164 P=0.023	Urban center (Dublin/Limerick/Cork) > Rural (all other counties)
Treatment (n=189)	HR=0.796, 0.453, 0.188, 0.315 SE=0.286, 0.294, 0.383 P=0.031, 0.005, 0.124, 0.624, 0.331	Radiotherapy > Surgery, Surgery/Radiotherapy/Chemotherapy

For the multivariate analysis using all of the variables in Table 8.12, Table 8.13 summarizes results.

Table 8.13 Patient and tumour characteristics significantly predicting overall survival amongst oropharyngeal cancer patients by multivariate cox proportional hazard model (n=189).

Variable/Factor	Statistic	Increased Risk of Death
Age (Continuous)	HR=0.020 SE=0.009 P=0.029	Older age
HR HPV Status	HR=0.737 SE=0.190 P<0.0001	HPV negativity
T Stage	HR=-0.709, -1.548, -0.604, -0.593 SE=0.273, 0.273, 0.221, 0.247 P<0.0001, 0.009, 0.0001, 0.006, 0.016	T4>T3, T2, Missing>T1
M Stage	HR=-1.049, -1.198 SE-0.316, 0.313 P=0.001, 0.001, 0.0001	M1>M0, Missing

Table 8.14 elucidates the variables significantly predictive of cancer-specific survival in oropharyngeal cases by univariate cox proportional hazard models.

Table 8.14 Patient and tumour characteristics significantly predicting disease-specific (cancer-specific) survival amongst oropharyngeal cancer patients by univariate cox proportional hazard models.

Variable/Factor	Statistic	Increased Risk of Death
HR HPV Status (n=209)	HR=0.620 SE=0.185 P=0.001	HPV negativity
Age$\leq$ 50 (n=209)	HR=-0.515 SE=0.227 P=0.023	>50
Age (Continuous) (n=209)	HR=0.025 SE=0.009 P=0.006	Older age
T Stage (n=209)	HR=-0.395, -1.671, -0.629, -0.491 SE=0.261, 0.437, 0.228, 0.254 P=0.001, 0.130, 0.0001, 0.006, 0.053	T4, Missing>T3, T2>T1
M Stage (n=209)	HR=-0.847, -1.121 SE=0.302, 0.302 P=0.001, 0.005, 0.0001	M1>M0, Missing
TNM Stage (n=209)	HR=-0.553, -1.178, -0.379, -0.541 SE=0.320, 0.588, 0.297, 0.288 P=0.047, 0.084, 0.045, 0.202, 0.060	Stage IV> III, II, Missing>Stage I
Social Deprivation (n=209)	HR=0.135 SE=0.064 P=0.035	More deprived

Due to evident and expected confounding between T, N, M, and TNM stages, TNM stage was used as a representative variable for each cancer stage in the multivariate model. Results of this model are presented in the Table below.

Table 8.15 Patient and tumour characteristics significantly predicting disease-specific (cancer-specific) survival amongst oropharyngeal cancer patients by multivariate cox proportional hazard model (n=209).

Variable/Factor	Statistic	Increased Risk of Death
Age (Continuous)	HR=0.039 SE=0.012 P=0.002	Older age
HR HPV Status	HR=0.937 SE=0.247 P<0.0001	HPV negativity
Deprivation Score	HR=0.165 SE=0.064 P=0.010	More deprived
TNM Stage	HR=-0.694, -1.615, -0.564, -0.540 SE=0.327, 0.598, 0.305, 0.292 P=0.006, 0.034, 0.007, 0.065, 0.065	T4>T3, T2, Missing>T1

It should be noted that as showcased in Table 8.14, TNM stage was only very minorly significant. The analysis was conducted again in the multivariate model using T and M stage separately instead of TNM stage, and results remained the same, with T and M stage simply replacing TNM stage in Table 8.15 above. Deprivation score however became only minorly significant (HR=0.129, SE=0.065, p=0.046).

8.7.3 Oral Cavity Cancer

Univariate analysis for overall survival of patients with oral cavity cancer revealed those variables in Table 8.16 as significantly predictive.

Table 8.16 Patient and tumour characteristics significantly predicting overall survival amongst oral cavity cancer patients by univariate cox proportional hazard models.

Variable/Factor	Statistic	Increased Risk of Death
Sex (n=331)	HR=-0.391 SE=0.131 P=0.003	Male>Female
Age (Continuous) (n=331)	HR=0.031 SE=0.005 P<0.0001	Older age
Age$\leq$ 50 (n=331)	HR=-0.893 SE=0.200 P<0.0001	<50
Smoking Status (n=331)	HR=-0.760, -0.311, -0.186 SE=0.180, 0.180, 0.177 P<0.0001, 0.0001, 0.084, 0.292	Current smoker, Missing > Ex-smoker
Grade (n=331)	HR=-0.767, -0.262, -0.368 SE=0.249, 0.227, 0.151 P=0.010, 0.002, 0.248, 0.010	Poorly-differentiated > Well-differentiated, Moderately differentiated > Missing
T Stage (n=331)	HR=-0.313, -1.181, -0.625, -0.508 SE=0.210, 0.192, 0.156, 0.201 P<0.0001, 0.136, 0.0001, 0.0001, 0.011	T4, Missing>T3, T2>T1
N Stage (n=331)	HR=-0.178, -0.653, -0.025 SE=0.182, 0.150, 0.173 P<0.0001, 0.326, 0.0001, 0.884	N3/2, N1, Missing>N0
TNM (n=331)	HR=-0.023, -0.985, -0.567, -0.522 SE=0.203, 0.190, 0.186, 0.177 P<0.0001, 0.909, 0.0001, 0.002, 0.003	Stage IV, Missing > Stage III, Stage II > Stage I
Marital Status (n=324)	HR=-0.220, -0.283, 0.240 SE=0.260, 0.178, 0.190 P=0.003, 0.398, 0.112, 0.208	Single > Married, Divorced/Separated
Treatment Type	HR=0.808, 0.254, -0.378, 0.025	Radiotherapy

(n=288)	SE=0.257, 0.285, 0.243, 0.238 P<0.0001, 0.002, 0.373, 0.120, 0.915	> Surgery/Radiotherapy/Chemotherapy, Surgery

Due to evident and expected confounding between T, N, M, and TNM stages, TNM stage was used as a representative variable for each cancer stage in the multivariate model. Results of this model are presented in the Table below.

Table 8.17 Patient and tumour characteristics significantly predicting overall survival amongst oral cavity cancer patients by multivariate cox proportional hazard model (n=282).

Variable/Factor	Statistic	Increased Risk of Death
Age (Continuous)	HR=0.039 SE=0.007 P<0.0001	Older age
Sex	HR=-0.514 SE=0.153 P=0.001	Male>Female
Smoking Status	HR=-0.634, -0.371, -0.246 SE=0.196, 0.196, 0.219 P=0.007, 0.001, 0.058, 0.260	Current smoker, Missing > Ex-smoker
Treatment	HR=0.577, 0.222, -0.638, -0.126 SE=0.272, 0.292, 0.256, 0.246 P<0.0001, 0.034, 0.447, 0.013, 0.608	Radiotherapy > Surgery/Radiotherapy/Chemotherapy, Surgery

Table 8.18 showcases the variables significantly predictive of cancer-specific survival amongst oral cavity cases by univariate cox proportional hazard models.

Table 8.18 Patient and tumour characteristics significantly predicting disease-specific (cancer-specific) survival amongst oral cavity cancer patients by univariate cox proportional hazard models.

Variable/Factor	Statistic	Increased Risk of Death
Sex **(n=331)**	HR=-0.341 SE=0.149 P=0.023	Male>Female
Age (Continuous) **(n=331)**	HR=0.027 SE=0.006 P<0.001	Older age
Age ≤50 **(n=331)**	HR=-0.756 SE=0.222 P=0.001	>50
Smoking Status **(n=331)**	HR=-0.995, -0.334, -0.128 SE=0.227, 0.205, 0.195 P<0.0001, 0.0001, 0.104, 0.513	Current smoker, Missing, Never smoker > Ex-smoker
Marital Status **(n=324)**	HR=-0.677, -0.317, -0.253, 0.210 SE=0.729, 0.329, 0.203, 0.217 P=0.042, 0.353, 0.336, 0.214, 0.333	Single, Widowed > Married, Separated/Divorced
Grade **(n=331)**	HR=-0.741, -0.278, -0.418 SE=0.285, 0.257, 0.169 P=0.028, 0.009, 0.280, 0.013	Poorly-differentiated > Moderately differentiated, Missing
T Stage **(n=331)**	HR=-0.315, -1.310, -0.811, -0.390 SE=0.233, 0.224, 0.179, 0.214 P<0.0001, 0.176, 0.0001, 0.0001, 0.068	T4, Missing, T3>T2>T1
N Stage **(n=331)**	HR=-0.123, -0.663, -0.091 SE=0.204, 0.172, 0.199 P=0.001, 0.546, 0.0001, 0.649	N3/N2, N1, Missing>N0
TNM Stage **(n=331)**	HR=0.010, -1.048, -0.665, -0.597 SE=0.227, 0.224, 0.219, 0.207 P<0.0001, 0.965, 0.0001, 0.002, 0.004	Stage IV, Missing> Stage III, II>Stage I
Treatment Type **(n=288)**	HR=0.956, 0.288, -0.336, 0.017 SE=0.297, 0.337, 0.291, 0.285 P<0.0001, 0.001, 0.392, 0.248, 0.952	Radiotherapy > Radiotherapy/Chemotherapy, Surgery/Radiotherapy, Surgery/Radiotherapy/Chemotherapy > Surgery

Due to evident and expected confounding between T, N, M, and TNM stages, TNM stage was used as a representative variable for each cancer stage in the multivariate model. Results of this model are presented in the Table below.

Table 8.19 Patient and tumour characteristics significantly predicting disease-specific (cancer-specific) survival amongst oral cavity cancer patients by multivariate cox proportional hazard model. The initial model included all those variables significant by univariate analysis but excluded T, N, and M stage (n=282).

Variable/Factor	Statistic	Increased Risk of Death
Sex	HR=-0.459 SE=0.179 P=0.010	Male>Female
Age (Continuous)	HR=0.028 SE=0.008 P<0.0001	Older age
Smoking Status	HR=-0.880, -0.247, -0.086 SE=0.255, 0.226, 0.243 P=0.005, 0.001, 0.125, 0.725	Current smoker, Never smoker, Missing > Ex-smoker
Treatment Type	HR=0.840, 0.274, -0.458, -0.051 SE=0.312, 0.341, 0.303, 0.292 P<0.0001, 0.007, 0.422, 0.131, 0.860	Radiotherapy > Radiotherapy/Chemotherapy, Surgery/Radiotherapy, Surgery/Radiotherapy/Chemotherapy > Surgery

8.7.4 Laryngeal Cancer

Table 8.20 lists the variables significantly predictive of overall survival amongst laryngeal cancers by univariate cox proportional hazard models.

Table 8.20 Patient and tumour characteristics significantly predicting overall survival amongst laryngeal cancer patients by univariate cox proportional hazard models.

Variable/Factor	Statistic	Increased Risk of Death
Age (Continuous) (n=321)	HR=0.024 SE=0.006 P<0.0001	Older age
Smoking Status (n=321)	HR=-0.145, 0.143, 0.628 SE=0.229, 0.148, 0.184	Missing >

	P=0.005, 0.525, 0.332, 0.001	All other smoking types
T Stage (n=321)	HR=-0.600, -1.192, -0.765, -0.434 SE=0.196, 0.194, 0.178, 0.188 P<0.0001, 0.001, 0.0001, 0.0001, 0.021	T4>Missing,T3, T2>T1
N Stage (n=321)	HR=-0.474, -0.769, -0.310 SE=0.194, 0.168, 0.229 P<0.0001, 0.015, 0.0001, 0.175	N3/N2> N1,Missing> N0
M Stage (n=321)	HR=-1.401, -1.462 SE=0.282, 0.278 P<0.0001, 0.0001, 0.0001	M1>M0, Missing
TNM Stage (n=321)	HR=-0.643, -1.268, -0.758, -0.681 SE=0.180, 0.190, 0.178, 0.181 P<0.0001, 0.001, 0.0001, 0.0001, 0.0001	Stage IV>Stage III, II, Missing>Stage I
Marital Status (n=306)	HR=-0.081, -0.387, 0.243 SE=0.283, 0.166, 0.193 P<0.0001, 0.775, 0.020, 0.209	Single, Widowed > Married

Due to evident and expected confounding between T, N, M, and TNM stages, TNM stage was used as a representative variable for each cancer stage in the multivariate model. Results of this model are presented in the Table below.

Table 8.21 Patient and tumour characteristics significantly predicting overall survival amongst laryngeal cancer patients by multivariate cox proportional hazard model (n=306).

Variable/Factor	Statistic	Increased Risk of Death
Age (Continuous)	HR=0.030 SE=0.007 P<0.0001	Older age
TNM Stage	HR=-0.704, -1.260, -0.790, -0.653 SE=0.190, 0.197, 0.189, 0.188 P<0.0001, 0.001, 0.0001, 0.0001, 0.001	T4>T3, T2, Missing>T1
Marital Status	HR=0.341, -0.171, 0.340 SE=0.294, 0.172, 0.200 P=0.008, 0.246, 0.318, 0.090	Single > Married

Table 8.22 summarizes the variables significantly predictive of cancer-specific survival for laryngeal cancer by univariate cox proportional hazard model.

Table 8.22 Patient and tumour characteristics significantly predicting disease-specific (cancer-specific) survival amongst laryngeal cancer patients by univariate cox proportional hazard models.

Variable/Factor	Statistic	Increased Risk of Death
Smoking Status (n=321)	HR=-0.415, -0.062, 0.643 SE=0.318, 0.194, 0.221 P=0.007, 0.191, 0.750, 0.004	Missing > All other smoking types
Grade (n=320)	HR=-0.541, -0.610, -0.411 SE=0.255, 0.247, 0.189 P=0.042, 0.034, 0.013, 0.030	Poorly-differentiated > Moderately differentiated, Missing, Well-differentiated
Marital Status (n=306)	HR=0.214, -0.298, 0.359 SE=0.329, 0.218, 0.247 P=0.005, 0.514, 0.173, 0.145	Single > Married
T Stage (n=321)	HR=-0.638, -1.512, -0.889, -0.582 SE=0.225, 0.256, 0.217, 0.231 P<0.0001, 0.002, 0.0001, 0.0001, 0.012	T4>T3, T2, Missing>T1
N Stage (n=321)	HR=-0.648, -0.906, -0.285 SE=0.239, 0.202, 0.266 P<0.0001, 0.007, 0.0001, 0.284	N2/3, N1>N0, Missing
M Stage (n=321)	HR=-1.605, -1.656 SE=0.309, 0.304 P<0.0001, 0.0001, 0.0001	M1>M0, Missing
TNM Stage (n=321)	HR=-0.742, -1.565, -0.961, -0.786 SE=0.222, 0.257, 0.228, 0.225 P<0.0001, 0.001, 0.0001, 0.0001, 0.001	Stage IV>Stage III, II, Missing>Stage I
Treatment Type (n=275)	HR=-0.981, -0.634, -0.637, -0.615 SE=0.257, 0.339, 0.355, 0.275 P=0.005, 0.0001, 0.061, 0.072, 0.025	Surgery/Radiotherapy/Chemotherapy > Surgery, Surgery/Radiotherapy, Radiotherapy/Chemotherapy

>

Radiotherapy

Due to evident and expected confounding between T, N, M, and TNM stages, TNM stage was used as a representative variable for each cancer stage in the multivariate model. Results of this model are presented in the Table below.

Table 8.23 Patient and tumour characteristics significantly predicting disease-specific (cancer-specific) survival amongst laryngeal cancer patients by multivariate cox proportional hazard model. The initial model included all significant variables by univariate analysis but excluded T, N, and M stage (n=262).

Variable/Factor	Statistic	Increased Risk of Death
TNM Stage	HR=-0.717, -1.423, -0.818, -0.757 SE=0.264, 0.276, 0.248, 0.248 P<0.0001, 0.007, 0.0001, 0.001, 0.002	Stage IV> Stage III, II, Missing>Stage I

8.8 Discussion

The aims of this chapter were to compare treatment received and survival in patients with HPV positive and HPV negative oropharyngeal, oral cavity, and laryngeal SCC in Ireland, and to identify significant predictors of survival for oropharyngeal, oral cavity, and laryngeal SCC in Ireland.

In the present 1994-2013 Irish population, HPV-related oropharyngeal, oral cavity, and laryngeal SCC had significantly increased cancer-specific survival in comparison to HPV-unrelated oropharyngeal, oral cavity, and laryngeal SCC (Figure 8.2). This relationship was primarily driven by HPV positive cases originating in the oropharynx (Figure 8.4), and HPV negative cases originating in the larynx and oral cavity. Overall survival for HPV-related oropharyngeal, oral cavity, and laryngeal SCC was also significantly better than for HPV-unrelated oropharyngeal, oral cavity, and laryngeal SCC (Figure 8.1). These relationships were mirrored in the oropharynx (Figure 8.5), but ceased to remain in the larynx and the oral cavity (Figures 8.5 and 8.7).

HPV positive oropharyngeal, oral cavity, and laryngeal SCC was significantly more likely to be treated harshly (with surgery, radiotherapy, and chemotherapy in the first 12 months after diagnosis) than HPV negative oropharyngeal, oral cavity, and laryngeal SCC (Table 8.4). This was again driven by the oropharyngeal site (Table 8.5), where both the larynx and oral cavity showed a relatively even distribution of treatment type by HPV status, though HPV negative tumours in the larynx were much more likely to be treated with radiotherapy and HPV positive tumours were more likely to be treated with harsher treatment schemes including all three modalities (Tables 8.7).

Broken down by HPV status, there was a significant difference in survival by treatment type at the overall and cancer-specific levels (Figures 8.12 and 8.13). Surgery, followed closely by all three treatment modalities and surgery/radiotherapy saw the best survival. Radiotherapy minimized survival in HPV positive oropharyngeal cases. HPV negative cases in the oropharynx saw no significant differences by treatment at either level of survival. For HPV positive oral cavity cancers, there was no difference in survival at all by treatment type, where for the HPV negative cases surgery alone saw best overall and cancer-specific outcomes within the first 10 years (Figures 8.14 and 8.15). In the larynx, only HPV negative cases saw differences by treatment, with radiotherapy and surgery/radiotherapy alternating and seeing the best overall and cancer-specific survival at difference points in time (Figures 8.16 and 8.17). These break-downs by sub-site revealed the key drivers of HPV positive and negative overall and cancer-specific trends for all cases (Section 8.6.1).

With respect to overall survival and increased risk of death, significant predictors remaining in multivariate analysis for all cases were older age, HPV negativity, late TNM stage, current vs never smoker, current vs ex-smoker, and single marital status (Table 8.9); for OPSCC, the equivalent predictors were older age, HPV negativity, late T stage, and distant metastasis (Table 8.13); for OSCC predictors were older age, male sex, current vs ex-smoker, and radiotherapy or chemotherapy/radiotherapy treatment (Table 8.17); and for LSCC, predictors were older age, later TNM stage, and single marital status (Table 8.21).

For cancer-specific survival, significant predictors of increased risk of death remaining in multivariate analysis for all cases were older age, current vs never smoking, current vs ex-smoker, late TNM stage, and single marital status (Table 8.11); for OPSCC the equivalent predictors were older age, HPV negativity, more deprived social status, and later TNM stage (or T and M alone) (Table 8.15); for OSCC these were older age, current vs ex-smoker, male sex, and radiotherapy treatment (Table 8.19); and for LSCC the only significant predictor was late TNM stage (Table 8.23).

The primary key finding of the present analysis was that both overall and cancer-specific survival were significantly improved for HPV positive cases in all oropharyngeal, oral cavity, and laryngeal SCC and in OPSCC. This is reflective of most studies in the literature[17,33,34]. It suggests that the Irish HNSCC and OPSCC populations are no different than their European and North American counterparts with respect to survival behavior by HPV status. This is likely given that HPV-related patients, being younger (Chapter 7), are less likely to have had significant exposure to tobacco, marijuana, alcohol, diabetes, chronic obstructive pulmonary disease, anxiety disorders, and major depression. The most at-risk populations are thus those with the best immune ability to combat HPV-related disease. Furthermore, the current results support the notion that the viral origins of HPV positive tumours, accompanied by their expression of viral oncoproteins and related HPV positive tumour antigens at sites of huge immune and lymphatic activity likely attracts a more aggressive and specific immune response that improves both overall and cancer-specific survival[35,36]. Younger patients are also more likely to better survive harsh treatments and their potent side-effects, which is particularly important in the case of HPV-positive tumours in this population given that they were more likely to be treated harshly than HPV-negative patients (Tables 8.4 and 8.5).

Another significant finding of the current analysis was that when broken down by sub-site and HPV status, oropharyngeal HPV positive cases a significant difference in survival by treatment type at both the overall and cancer-specific levels. Figures 8.12 and 8.13 show very clearly that survival amongst HPV positive OPSCC patients was maximized by surgery alone by an enormous margin in comparison to radiotherapy and radiotherapy/chemotherapy, but also in comparison to other cancer-specific survival rates

amongst the majority of laryngeal and oral cavity cases (those that were HPV negative). In fact, Figure 8.13 shows over 70% cancer-specific survival rates after 10 years for patients treated with surgery alone, with those treated by all three modalities and surgery/radiotherapy following closely behind. Cancer-specific survival amongst the majority of laryngeal and oral cavity patients (Figures 8.15 and 8.17) never exceeded 45% after 10 years even with the best treatment types.

These results first highlight the importance of surgical intervention for HPV positive OPSCC, with treatment approaches not involving surgery seeing very poor survival. Secondly, the findings are extremely promising in terms of the potential of de-escalation of treatment for these patients in the Irish clinic. Indeed, all three treatment modalities and surgery/radiotherapy saw similar survival rates overall compared to surgery alone, but ultimately, it was surgery that saw the best outcomes. This is indicative of the chance that HPV-related OPSCC presents for drastically improving quality of life for patients by avoiding the administration of extremely harsh treatments and the long-term side-effects that accompany them. These associated and debilitating side-effects range from difficulties swallowing, breathing, and speaking, to chronic pain, osteoradionecrosis, hypertension, pneumonia, dysphagia, weight loss, malnutrition, dental issues, and third-degree burns[37,38]. Harsher treatments also increase risk of heart disease and failure[39,40], risk of another (non-recurrence) primary tumour at another site[41–44], and complications due to immunosuppression.

The present Irish data thus supports the notion that robotic trans-oral resection (TOR) alone yields extremely good results for HPV-related patients[45–47] regardless of stage and posits that this kind of non-chemical curative approach giving patients better functional outcomes[48–50] may be the way forward in the Irish context. Other studies are in agreement where TOR without adjuvant therapy is often adequate treatment for HPV-related OPSCC, with anywhere between 48% to 74% of patients not requiring chemotherapy after TORs[45,46,51]. This said, it is understandable that patients may feel more comfortable being treated with more than just surgery, with studies showing that nearly 70% of patients are not willing to risk a 5% or less drop in survival likelihood to switch from chemoradiation to radiation alone after surgery[52]. In the present population, this 5% drop in survival is not

evident amongst HPV positive OPSCC patients, with surgery alone seeing better survival than surgery/radiotherapy, and the margin between surgery/radiotherapy and all three treatments being minimal (Figure 8.12). This is something that may give patients more incentive to opt for less harsh schemes. Nonetheless, many trials currently underway are based on the suggestion that surgery with de-escalated radiotherapy yields maximal survival with decreased morbidity and associated side-effects[53–55], a scheme that might satisfy survival outcomes, minimize side-effects, and ensure patient peace of mind simultaneously.

Indeed, this population is supportive of de-escalation amongst HPV positive OPSCC. However, two caveats should be noted. First, there were smaller sample sizes available when sub-dividing all 861 cases into their sub-site, HPV status, and treatment groups. Targeted sampling of OPSCC alone is needed for further confirmation of these promising findings. Second, the analysis also emphasizes that in terms of potential de-escalation, it would be unethical to make treatment decisions for these patients, or their negative counterparts, based solely on HPV status. For oropharyngeal, oral cavity, and laryngeal SCC, multivariate predictors of overall risk of death did include HPV negativity, but HPV status was not confounded by other patient characteristics including older age and current smoker status (Tables 8.9). For OPSCC, HPV positivity was predictive of decreased risk of death at the overall and cancer-specific levels (Tables 8.13 and 8.15). However, HPV was not confounded by age or social deprivation. It also did not predict risk of death for any survival in LSCC and OPSCC (Tables 8.17, 8.19, 8.21, 8.23).

Cumulatively, these findings imply two key concepts. First, they support the notion that the oropharynx is the sub-site in which HPV-related tumours occur and that it is therefore the region for which any HPV-related treatment alterations should be made. Second, they indicate that though HPV-related tumours are already significantly associated with younger aged patients[34,55,56] and never-/ex-smokers[57,58] (Chapter 7), it would be extremely prudent to select patients who might benefit from de-escalation based on not only HPV positive status but also on other survival-maximizing characteristics at both the cancer-specific and overall levels. Thus, those patient characteristics that not only indicate stereotypically HPV-driven tumours, but that the present multivariate analyses indicate might optimize survival and morbidity with de-escalated treatment are summarized in Table 8.24 below.

Table 8.24 Patient characteristics indicative of stereotypically HPV-driven oropharyngeal, oral cavity, and laryngeal SCC that may be the basis for the precise selection of patients for whom treatment de-escalation is possible.

Characteristic
HR HPV Positive
Oropharyngeal sub-site
Younger age or ≤ 50
Never- or ex-smoker

This collection of patient characteristics has recently been recognized in the literature as the only group of oropharyngeal, oral cavity, and laryngeal SCC patients for which de-escalation of treatment is acceptable. In fact, several of the ongoing trials regarding de-escalation only include patients meeting these criteria to assure no jeopardizing of patient safety[53–55,59], but also to target the group that will likely benefit most from less severe treatment. Long-term data on overall survival has yet to be published from these trials, but the current population suggests that de-escalation will be successful in optimizing survival and quality of life for these patients.

It might be suggested that the inclusion of "higher socio-economic" status in Table 8.24 is also justifiable on the basis of the present findings (Table 8.15), as more deprived patients survived OPSCC less often when adjusting for other variables, something reflected in the HNSCC literature many times over due to later stage at diagnosis and other multiplicative factors like current smoking status (Chapter 6)[60–62]. However, social deprivation did not remain significant in predicting overall survival in OPSCC, and its significance for cancer-specific survival was severely hindered when replacing barely significant TNM stage with univariately significant T and M stages.

With respect to Table 8.24, it should also be noted that there is still a need to distinguish clinically significant HR HPV infections from transient ones. Where in this analysis HR HPV DNA was used to determine HPV-related status, many trials only use p16 as a representative biomarker of an active HPV infection[55]. Neither of these alone is entirely satisfactory in the

clinical context given the potential for transient HR HPV infections, and the expression of p16 regardless of HPV status. In fact, HPV DNA may be misleading even if other patient characteristics are suggestive of a classically HPV-related case. In the clinic, these kinds of risks resulting in the potential under-treatment of patients cannot be taken. Further specification of 'HPV positivity' as a necessary characteristic for de-escalation will likely make treatment decisions and thus survival determinations even more accurate. Pairing p16 with HR HPV DNA[63,64], or simply using HPV mRNA[65], represent mechanisms to refine this process in the clinic, though the present HR HPV DNA data is a resounding start.

Despite positive indications of de-escalation potential in HPV positive OPSCC, the present analysis indicates that HPV positive HNSCC and OPSCC were more likely to be treated harshly than their HPV negative counterparts (Tables 8.4 and 8.5). The population that might have benefited most from less severe treatment schemes was thus the population being treated most severely. The present data and the literature explain that this irony is due to the later stage at which HPV-related OPSCC are diagnosed[12,34,56,66]. Specifically, they disproportionately present at Stage IV due to late N stage (Chapter 7) according to the 5th edition AJCC guidelines relevant to this population between 1994 and 2013[67]. The current analysis posits therefore that the new 2017 8th edition AJCC guidelines[68] updated for the oropharyngeal sub-site alone, reflecting the role of HR HPV, are very highly relevant to the Irish context. This is especially true since neither N stage nor TNM stage were significant predictors of overall survival in OPSCC in either univariate or multivariate analysis (Tables 8.12 and 8.13). TNM stage was barely significant in predicting survival in OPSCC at the cancer-specific level (Table 8.14), and N stage remained insignificant. This implies that the nodal and cumulative staging of the older staging systems were not accurate assessors of the aggressivity of these tumours, likely due to the unique features of HPV-related tumours in this region. Those HPV-related cases diagnosed as Stage IV before 2017 will now be downgraded to at least Stage III if not even Stage I due to adjustments in N stage relating to nodal metastasis. It is very likely that the consequent down-grading of stage in OPSCC will act as a de-escalation mechanism of its own, implicating less severe treatment requirements from the moment the cancer is diagnosed.

With respect to laryngeal cases in the current population, the relevance and significance of HPV status to survival, even adjusting for co-variates, ranges from questionable to determinedly inconsequential. To begin, there was no significant difference in survival by HPV status in laryngeal cases for either overall or cancer-specific survival (Figures 8.7 and 8.8). The insignificance of HPV status in the larynx, in tandem with very low HR HPV prevalence and the concentration of LR HPV genotypes in the region (Chapter 7), implies that HPV is likely irrelevant to the carcinogenesis of SCC at this site. The fact that even HR HPV positive laryngeal cases do not behave survival-wise in the same way as their counterparts in the oropharynx supports the overwhelming smoking- and/or alcohol-driven carcinogenesis in this site (Figures 8.7 and 8.8). That HR HPV is still detected in the larynx may instead be a product of smoking-related immunosuppression leading to HR HPV infection that begins after carcinogenesis has started. It may also suggest that hybrid cases, a consequence of both smoking- and HPV-related carcinogenesis, exist at this site and behave far more aggressively than their more purely HPV-related counterparts.

Univariate and multivariate results for overall and cancer-specific survival analyses further solidify these conclusions, with no sign of HPV being a significant predictor of survival in the laryngeal sub-site (Tables 8.20, 8.21, 8.22, and 8.23). Instead, tumour-related characteristics appear to be the determinants of survival in LSCC, with later TNM stage alone being the significant predictor of cancer-specific survival adjusting for other variables (Table 8.23). These results highlight the importance of early detection and prevention in these likely smoking- and alcohol-related cancers. Once these tumours are diagnosed at later stages, overall and cancer-specific survival become very difficult to salvage, especially since laryngeal patients tend to be older at diagnosis and single (Chapter 7), and these were the only other factors determining overall survival (Table 8.21). The literature recognizes the unfortunate nature of LSCC and HPV-unrelated SCC with most late-stage LSCC and HPV-unrelated patients treated using a palliative approach that prolongs life slightly and attempts to minimize side-effects, but does not yield successful remission[2,3].

With respect to treatment in the laryngeal sub-site, univariate analysis suggested that radiotherapy was most successful in prolonging cancer-specific survival. This was further supported by Figures 8.16 and 8.17 which demonstrated that HPV negative laryngeal cases

(comprising the vast majority of all laryngeal cases) saw the best overall and cancer-specific survival rates with radiotherapy alone and surgery/radiotherapy. Indeed, this is supported in the literature where laryngeal cancers treated with radiotherapy alone, the majority of which are diagnosed at stages I and II (Chapter 6), show 90% cancer-specific survival rates after 10 years, the highest of any other treatment type[69].

To note regarding LSCC however is that the majority of these cancers (those that were HPV negative), even treated with radiotherapy, still showed the lowest overall survival rate after 10 years in comparison to the other sub-sites (Figure 8.16). This may suggest that these patients may be more at risk of other co-morbidities associated with their personal characteristics (summarized in Chapter 6), including older age, smoking-related cardiovascular disease or the emergence of another primary tumour at another site, risks that are increased by the treatments already administered to them[62,70,71]. This is indeed reflected in the literature where older patients[72] and smoking patients[62] are more likely to suffer co-morbidities even after surviving the present cancer.

OSCC represents a more heterogenous group of cancers with respect to survival and treatment. Indeed, HPV status made no difference to overall and cancer-specific survival in this group (Figures 8.5 and 8.6). Instead, male sex, older age, current smoking status, and treatments other than surgery predicted worse survival (Tables 8.17 and 8.19). The absence of HPV from any of these analyses suggests the irrelevance of the virus at this site. This said, characteristics that were not predictive of LSCC survival including male sex and current smoking status were significantly predictive of survival in OSCC, implying differential carcinogenesis, presentation, and risk factors in this region.

Treatment-wise, Table 8.19 reveals that oral cavity cancers saw the best overall and cancer-specific survival by surgery alone. This was further solidified in Figures 8.14 and 8.15 where the majority of oral cavity cancers (those that were HPV negative) saw the best overall and cancer-specific survival rates above 30% after 10 years when treated with surgery alone. These findings may on the one hand support differential carcinogenesis at this site, showing tumours that respond better to surgical treatment than radiotherapy treatment in the larynx. On the other, they may be an indication that the earlier stage at which they present

(Chapter 6) and the nature of their physiological position that makes them highly-resectable, allow for a more tempered treatment approach. Indeed, reviews have shown that oral cavity cancers generally only need additional radiotherapy and chemotherapy if diagnosed at T3 and T4 T stage[73,74].

In tandem, the survival and treatment results from the analysis suggest that those treatments that optimize survival for patients in different sub-sites are those that are most often administered to them. For instance, almost 50% of laryngeal patients in the present population were treated with radiotherapy alone (Table 8.7), while the largest proportion of oral cavity patients were treated with surgery (Table 8.6). That patients diagnosed with these cancers between 1994 and 2013 mostly received the treatments that ultimately maximized their survival (accounting for variation based on TNM stage) is a testament to the quality of care for oropharyngeal, oral cavity, and laryngeal SCC patients in the Irish clinic.

However, the oropharyngeal sub-site represents somewhat of a divergence as a result of increasingly HPV-related carcinogenesis. Indeed, it is sensical that most HPV positive OPSCC would be treated aggressively with all three treatment modalities given the overwhelmingly Stage IV TNM stage at which they present (Chapter 7) according to the 5th edition AJCC system. This said, it is clear from the present analysis that this harsh approach is not necessary for HPV-related OPSCC alone, and on the basis of pending results of current trials[55], the Irish clinic may need to adapt. HPV-unrelated OPSCC presents more of a challenge, with no treatment types optimizing either cancer-specific or overall survival, suggesting that these growths behave differently, even in comparison to their HPV-unrelated counterparts in the larynx and oral cavity.

To note regarding all foregoing analyses is the role of non-random missing data on some findings due to the registry-based nature of data sourcing. Though relationships between individual patient characteristics and survival remain regardless of missing data, it is evident that the majority of patients with missing smoking data came from current and never smokers. Missing data often showed the same if not worse survival patterns than current smokers (Tables 8.9, 8.11), and though never smoking was predictive of better survival than

current smoking in most cases (Tables 8.9, 8.11, 8.12), this difference was minimized in the oral cavity (Tables 8.17 and 8.19). In fact, missing cases had significantly worse survival than all other smoking statuses in the larynx by univariate analysis (Tables 8.20 and 8.22). Thus, though the absence of smoking in any multivariate analyses in the larynx is indicative of the poor prognosis of these patients regardless of smoking or other non-HPV-related carcinogenesis, it is prudent to note that the survival of current smokers in given data limitation is overestimated and that of never-smokers is underestimated.

Despite diverging treatment characteristics and predictors of survival in each sub-site by HPV status, several factors do unite all the SCC analysed in this population. The first is that the Irish population continues to be proof of the suggestion that HPV positive oropharyngeal, oral cavity, and laryngeal SCC, and more specifically, OPSCC alone, is a better-surviving cancer than its HPV negative counterparts.

The second is that though more than simply surgery is and will continue to be necessary to treat some late-stage patients (Figure 8.9), the benefit of treating patients with all three modalities is questionable when oropharyngeal, oral cavity, and laryngeal SCC is evaluated sub-divided by sub-site. Figures 8.12 through 8.15 and multivariate models for all sub-sites highlight this uncertainty. Radiotherapy, surgery, and radiotherapy/surgery seemingly have important roles to play in maximizing survival amongst the majority of laryngeal, oral cavity, and oropharyngeal cases. However, the addition of chemotherapy to either one or both other treatment schemes does not significantly improve 10-year survival in any sub-site by comparison. The literature indicates that this is likely a reflection of both the responsiveness of these tumours to chemotherapy, and the consequences of harsh treatments including increased risk of heart disease and failure, along with occurrence of second-primaries[40–42,70,71,75]. This conclusion is not definitive however given the very few patients receiving chemotherapy alone and surgery/chemotherapy (n=6) in this study and their consequent exclusion from this analysis.

The third uniting factor is the significant role that marital status appears to play in determining survival for all oropharyngeal, oral cavity, and laryngeal SCC patients, especially those of the larynx where single marital status remained a significant predictor of poorest

overall survival adjusting for other variables. It is difficult to assess precisely why single patients are disproportionately impacted, but it is possible that care and support outside the clinical context have a role to play. Other studies have also supported the idea that single patients do not have the same levels of emotional and physical support outside the hospital due to lack of spousal (in the case of married patients) or family (in the case of divorced/separated patients) help[76,77]. This is sensical in the present population given that LSCC has a higher incidence in older men (Chapter 7) and that marital status was significantly associated with sex, with the vast majority of single individuals being men (Chapter 7). Older single men developing oropharyngeal, oral cavity, and laryngeal SCC, specifically LSCC, may therefore be at greatest risk of death due to the less responsive nature of their cancers, their age, and their single status. The need to assess patient needs including at-home social care based on risk factors including marital status is evident from this finding.

The fourth uniting factor is the essential role that p16 and other indicators of clinically relevant HPV infections including patient characteristics will play in determining treatment options for oropharyngeal, oral cavity, and laryngeal SCC **in addition to**, and **not instead of** sensitive HR HPV detection. Indeed, the new AJCC 8th edition and 2017 CAP guidelines[78] already recognize p16's role[68] but not HR HPV's role, for oropharyngeal cancers alone. The need to merge approaches for identifying truly HPV-related SCC is clear if the new AJCC staging system is to avoid mistakenly down-grading OPSCC that only appears to be HPV-related and if similar mistakes are to be prevented in the administration of de-escalated treatment.

Fifth, the diverse treatment, survival, and HPV characteristics observed in the present data converge and point to the crucial nature of prevention and early detection in oropharyngeal, oral cavity, and laryngeal SCC if survival, overall and cancer-specific, is to be maximized. All oropharyngeal, oral cavity, and laryngeal SCC are overwhelmingly behaviourally-driven cancers, whether by exposure to HPV or to smoking (and likely alcohol), or by exposure to both in the case of "hybrid" tumours. Though cancer-specific survival tends to be better amongst HPV-related SCC, the need to even consider cancer-related survival statistics could be entirely eliminated with the use of an appropriate

prophylactic vaccine for both boys and girls[79–81]. Data is still emerging on the impact the quadra- and nona-valent Gardasil vaccines on HR HPV in the oral cavity[82], though preliminary data from cervical trials testing oral rinses shows that HPV16/HPV18 prevalence is lower in vaccinated groups compared to control groups, with an estimated efficacy of 93.3% for HPV16/18[83]. Predictive modelling studies also suggest that with a 50% vaccination uptake and 50% vaccine efficacy, the vaccination of young boys for the prevention of HPV-related OPSCC would be cost-effective[82,84]. The need for more data is evident, but the systemic nature of vaccines logically suggests that the administration of the vaccine in early adolescence should be as effective in preventing HNSCC as it is in the cervical context. That the Irish government has received encouragement from HIQA to expand the public vaccination scheme to boys and intends to do so beginning in 2019 is hugely promising[85,86].

The need for public health schemes to combat re-emerging smoking habits amongst Irish people and to encourage smokers to quit is also clear. This is not only due to the overwhelming indications that LSCC, as the most common HNSCC in Ireland[87], is almost solely caused by smoking- and non-HPV carcinogens, but also given the very positive survival outcomes in ex-smokers particularly evident in Tables 8.9, 8.11, and 8.13. Anti-smoking schemes have tapered in the last 10 years as public consensus on the risks of smoking has become well-established. However, data from the NCRI suggests that as a result of increasing smoking rates, especially amongst young women, incidence of mouth cancers is increasing[87,88].

Lastly, the present analysis underlines the urgent need for effective and systematic HNC screening tools. Early detection of those SCC that do go on to develop despite preventative measures is tantamount to prolonging overall survival, no matter how promising or poor cancer-specific survival is and regardless of HPV status. For HPV-unrelated HNSCC in this analysis, especially in the larynx, diagnosis at later TNM stage was the only predictor of cancer-specific survival after adjustment for other variables (Table 8.23), and no treatment, harsh or otherwise, was ultimately successful in maintaining patient overall survival above 20% after 10 years (Figures 8.7 and 8.16) . This makes death almost a certainty for late-stage HPV-unrelated patients of the larynx. Additionally, not only was late TNM stage a significant predictor of overall and cancer-specific survival adjusting for other variables in all

oropharyngeal, oral cavity, and laryngeal SCC, but for mostly HPV-related HNSCC in the oropharynx, later TNM, and T and M stages separately remained significant predictors at the overall and cancer-specific levels.

Efforts are currently being made to investigate the best ways to sample tissue from the oral site, but it is made difficult by the region's confined nature and the dense, complex network of MALT tissues that line it[89,90]. This is especially true in the cases of the tonsillar crypts. Recently, mobile microscopy with a simple brush biopsy has shown to be an effective screening mechanism for oral cavity cancer, even in low-resource areas[91], but such a sampling method is not ideal for the deep, hidden, crypts of the oropharynx. The role that HPV might play in this screening is also uncertain, though monitoring systems like those established in the cervical case[92,93] are a promising way of catching HR HPV patients who, perhaps even after vaccination, go on to develop lesions. In all, pairing early detection with preventative mechanisms and curative approaches suitable to tumour **and** patient characteristics will render oropharyngeal, oral cavity, and laryngeal SCC an imminently manageable and rare disease. These public health and clinical measures will ultimately mean huge cost-savings, and more importantly, the difference between life and death for potential and current oropharyngeal, oral cavity, and laryngeal SCC patients.

References

1. Goldberg, H. I., Lockwood, S. A., Wyatt, S. W. & Crossett, L. S. Trends and differentials in mortality from cancers of the oral cavity and pharynx in the United States, 1973-1987. *Cancer* **74**, 565–72 (1994).
2. Thomas, G. R., Nadiminti, H. & Regalado, J. Molecular predictors of clinical outcome in patients with head and neck squamous cell carcinoma. *Int. J. Exp. Pathol.* **86**, 347–63 (2005).
3. Price, K. A. R. & Cohen, E. E. Current Treatment Options for Metastatic Head and Neck Cancer. *Curr. Treat. Options Oncol.* **13**, 35–46 (2012).
4. Hafkamp, H. C. *et al.* Marked differences in survival rate between smokers and nonsmokers with HPV 16-associated tonsillar carcinomas. *Int. J. Cancer* **122**, 2656–2664 (2008).

5. Krane, J. F. Role of Cytology in the Diagnosis and Management of HPV-Associated Head and Neck Carcinoma. *Acta Cytol.* **57**, 117–126 (2013).
6. Fakhry, C. *et al.* Improved Survival of Patients With Human Papillomavirus-Positive Head and Neck Squamous Cell Carcinoma in a Prospective Clinical Trial. *JNCI J. Natl. Cancer Inst.* **100**, 261–269 (2008).
7. Mroz, E. A., Forastiere, A. A. & Rocco, J. W. Implications of the Oropharyngeal Cancer Epidemic. *J. Clin. Oncol.* **29**, 4222–4223 (2011).
8. Paz, I. B., Cook, N., Odom-Maryon, T., Xie, Y. & Wilczynski, S. P. Human papillomavirus (HPV) in head and neck cancer. An association of HPV 16 with squamous cell carcinoma of Waldeyer's tonsillar ring. *Cancer* **79**, 595–604 (1997).
9. Joo, Y.-H. *et al.* High-risk human papillomavirus and cervical lymph node metastasis in patients with oropharyngeal cancer. *Head Neck* **34**, 10–14 (2012).
10. McHugh, J. B. Association of cystic neck metastases and human papillomavirus-positive oropharyngeal squamous cell carcinoma. *Arch. Pathol. Lab. Med.* **133**, 1798–803 (2009).
11. Thompson, L. D. & Heffner, D. K. The clinical importance of cystic squamous cell carcinomas in the neck: a study of 136 cases. *Cancer* **82**, 944–56 (1998).
12. Psyrri, A., Sasaki, C., Vassilakopoulou, M., Dimitriadis, G. & Rampias, T. Future directions in research, treatment and prevention of HPV-related squamous cell carcinoma of the head and neck. *Head Neck Pathol.* **6 Suppl 1**, S121-8 (2012).
13. Chien, C.-Y. *et al.* Lower prevalence but favorable survival for human papillomavirus-related squamous cell carcinoma of tonsil in Taiwan. *Oral Oncol.* **44**, 174–179 (2008).
14. Fischer, C. A. *et al.* p16 expression in oropharyngeal cancer: its impact on staging and prognosis compared with the conventional clinical staging parameters. *Ann. Oncol.* **21**, 1961–1966 (2010).
15. Marur, S. & Forastiere, A. A. Head and Neck Cancer: Changing Epidemiology, Diagnosis, and Treatment. *Mayo Clin. Proc.* **83**, 489–501 (2008).
16. Kimple, R. J. & Harari, P. M. The prognostic value of HPV in head and neck cancer patients undergoing postoperative chemoradiotherapy. *Ann. Transl. Med.* **3**, S14 (2015).
17. Nichols, A. C. *et al.* Does HPV type affect outcome in oropharyngeal cancer? *J. Otolaryngol. - Head Neck Surg.* **42**, 9 (2013).

18. Ang, K. K. *et al.* Human papillomavirus and survival of patients with oropharyngeal cancer. *N. Engl. J. Med.* **363**, 24–35 (2010).
19. Gillison, M. L. HPV and prognosis for patients with oropharynx cancer. *Eur. J. Cancer* **45**, 383–385 (2009).
20. Lassen, P. *et al.* The influence of HPV-associated p16-expression on accelerated fractionated radiotherapy in head and neck cancer: Evaluation of the randomised DAHANCA 6&7 trial. *Radiother. Oncol.* **100**, 49–55 (2011).
21. Posner, M. R. *et al.* Survival and human papillomavirus in oropharynx cancer in TAX 324: a subset analysis from an international phase III trial. *Ann. Oncol.* **22**, 1071–1077 (2011).
22. Rischin, D. *et al.* Prognostic Significance of p16INK4A and Human Papillomavirus in Patients With Oropharyngeal Cancer Treated on TROG 02.02 Phase III Trial. *J. Clin. Oncol.* **28**, 4142–4148 (2010).
23. Shaw, R. & Robinson, M. The increasing clinical relevance of human papillomavirus type 16 (HPV-16) infection in oropharyngeal cancer. *Br. J. Oral Maxillofac. Surg.* **49**, 423–429 (2011).
24. Albers, A. E., Qian, X., Kaufmann, A. M. & Coordes, A. Meta analysis: HPV and p16 pattern determines survival in patients with HNSCC and identifies potential new biologic subtype. *Sci. Rep.* **7**, 16715 (2017).
25. Kumar, B. *et al.* EGFR, p16, HPV Titer, Bcl-xL and p53, sex, and smoking as indicators of response to therapy and survival in oropharyngeal cancer. *J. Clin. Oncol.* **26**, 3128–37 (2008).
26. Dahlstrom, K. R. *et al.* HPV Serum Antibodies as Predictors of Survival and Disease Progression in Patients with HPV-Positive Squamous Cell Carcinoma of the Oropharynx. *Clin. Cancer Res.* **21**, 2861–9 (2015).
27. Wierzbicka, M., Szyfter, K., Milecki, P., Składowski, K. & Ramlau, R. The rationale for HPV-related oropharyngeal cancer de-escalation treatment strategies. *Contemp. Oncol. (Poznan, Poland)* **19**, 313–22 (2015).
28. Geiger, J. L. *et al.* Adjuvant chemoradiation therapy with high-dose versus weekly cisplatin for resected, locally-advanced HPV/p16-positive and negative head and neck squamous cell carcinoma. *Oral Oncol.* **50**, 311–318 (2014).
29. Gillison, M. L. *et al.* Evidence for a causal association between human papillomavirus

and a subset of head and neck cancers. *J. Natl. Cancer Inst.* **92**, 709–20 (2000).

30. Kimple, R. J. *et al.* Enhanced Radiation Sensitivity in HPV-Positive Head and Neck Cancer. *Cancer Res.* **73**, 4791–4800 (2013).
31. Rieckmann, T. *et al.* HNSCC cell lines positive for HPV and p16 possess higher cellular radiosensitivity due to an impaired DSB repair capacity. *Radiother. Oncol.* **107**, 242–246 (2013).
32. Vu, H. L., Sikora, A. G., Fu, S. & Kao, J. HPV-induced oropharyngeal cancer, immune response and response to therapy. *Cancer Lett.* **288**, 149–155 (2010).
33. Woods, R. Human Papillomavirus-related Oropharyngeal Squamous Cell Carcinoma – Prevalence and Incidence in a Defined Population and Analysis of the Expression of Specific Cellular Biomarkers. (2016).
34. Woods, R. *et al.* Role of human papillomavirus in oropharyngeal squamous cell carcinoma: A review. *World J. Clin. cases* **2**, 172–93 (2014).
35. Balermpas, P. *et al.* Tumour-infiltrating lymphocytes predict response to definitive chemoradiotherapy in head and neck cancer. *Br. J. Cancer* **110**, 501–509 (2014).
36. Jung, A. C. *et al.* CD8-alpha T-cell infiltration in human papillomavirus-related oropharyngeal carcinoma correlates with improved patient prognosis. *Int. J. Cancer* **132**, E26–E36 (2013).
37. Cancer.Net. Side Effects of Radiation Therapy. *ASCO* (2019). Available at: https://www.cancer.net/navigating-cancer-care/how-cancer-treated/radiation-therapy/side-effects-radiation-therapy. (Accessed: 4th July 2019)
38. Tolentino, E. de S. *et al.* Oral adverse effects of head and neck radiotherapy: literature review and suggestion of a clinical oral care guideline for irradiated patients. *J. Appl. Oral Sci.* **19**, 448–54 (2011).
39. Lyon, A. R. Heart failure resulting from cancer treatment: still serious but an opportunity for prevention. *Heart* **105**, 6–8 (2019).
40. Aleman, B. M. P. *et al.* Cardiovascular disease after cancer therapy. *Eur. J. Cancer Suppl.* **12**, 18–28 (2014).
41. Atienza, J. A. S. & Dasanu, C. A. Incidence of second primary malignancies in patients with treated head and neck cancer: a comprehensive review of literature. *Curr. Med. Res. Opin.* **28**, 1899–1909 (2012).
42. Chuang, S.-C. *et al.* Risk of second primary cancer among patients with head and neck

cancers: A pooled analysis of 13 cancer registries. *Int. J. Cancer* **123**, 2390–2396 (2008).

43. Elicin, O. *et al.* Incidence of second primary cancers after radiotherapy combined with platinum and/or cetuximab in head and neck cancer patients. *Strahlentherapie und Onkol.* **195**, 468–474 (2019).
44. Wong, S. J., Heron, D. E., Stenson, K., Ling, D. C. & Vargo, J. A. Locoregional Recurrent or Second Primary Head and Neck Cancer: Management Strategies and Challenges. *Am. Soc. Clin. Oncol. Educ. B.* **36**, e284–e292 (2016).
45. White, H. N. *et al.* Transoral Robotic-Assisted Surgery for Head and Neck Squamous Cell Carcinoma. *Arch. Otolaryngol. Neck Surg.* **136**, 1248 (2010).
46. Iseli, T. A. *et al.* Functional Outcomes after Transoral Robotic Surgery for Head and Neck Cancer. *Otolaryngol. Neck Surg.* **141**, 166–171 (2009).
47. Genden, E. M. The Role for Surgical Management of HPV-Related Oropharyngeal Carcinoma. *Head Neck Pathol.* **6**, 98–103 (2012).
48. Golusiński, W. Functional Organ Preservation Surgery in Head and Neck Cancer: Transoral Robotic Surgery and Beyond. *Front. Oncol.* **9**, 293 (2019).
49. Mahmoud, O., Sung, K., Civantos, F. J., Thomas, G. R. & Samuels, M. A. Transoral robotic surgery for oropharyngeal squamous cell carcinoma in the era of human papillomavirus. *Head Neck* **40**, 710–721 (2018).
50. Weinstein, G. S. *et al.* Transoral Robotic Surgery Alone for Oropharyngeal Cancer. *Arch. Otolaryngol. Neck Surg.* **138**, 628 (2012).
51. Genden, E. M. *et al.* Transoral robotic resection and reconstruction for head and neck cancer. *Laryngoscope* **121**, 1668–1674 (2011).
52. Brotherston, D. C. *et al.* Patient preferences for oropharyngeal cancer treatment de-escalation. *Head Neck* **35**, 151–159 (2013).
53. Chen, A. M. *et al.* Reduced-dose radiotherapy for human papillomavirus-associated squamous-cell carcinoma of the oropharynx: a single-arm, phase 2 study. *Lancet Oncol.* **18**, 803–811 (2017).
54. Villaflor, V. M. *et al.* Response-adapted volume de-escalation (RAVD) in locally advanced head and neck cancer. *Ann. Oncol.* **27**, 908–913 (2016).
55. Mirghani, H. & Blanchard, P. Treatment de-escalation for HPV-driven oropharyngeal cancer: Where do we stand? *Clin. Transl. Radiat. Oncol.* **8**, 4–11 (2018).

56. Smith, E. M. *et al.* Age, sexual behavior and human papillomavirus infection in oral cavity and oropharyngeal cancers. *Int. J. Cancer* **108**, 766–772 (2004).
57. Boscolo-Rizzo, P. *et al.* New insights into human papillomavirus-associated head and neck squamous cell carcinoma. *Acta Otorhinolaryngol. Ital.* **33**, 77–87 (2013).
58. D'Souza, G., McNeel, T. S. & Fakhry, C. Understanding personal risk of oropharyngeal cancer: risk-groups for oncogenic oral HPV infection and oropharyngeal cancer. *Ann. Oncol.* **28**, 3065–3069 (2017).
59. Brisson, R. J. *et al.* De-escalation in HPV-negative locally advanced head and neck squamous cell cancer (LA-HNSCC) in patients after induction chemotherapy: A retrospective case series. *J. Clin. Oncol.* **36**, e18090–e18090 (2018).
60. Auluck, A. *et al.* Socio-economic deprivation: a significant determinant affecting stage of oral cancer diagnosis and survival. *BMC Cancer* **16**, 569 (2016).
61. Khalil, D. *et al.* Does Socioeconomic Status Affect Stage at Presentation for Larynx Cancer in Canada's Universal Health Care System? *Otolaryngol. Neck Surg.* **160**, 488–493 (2019).
62. Sharp, L., McDevitt, J., Carsin, A.-E., Brown, C. & Comber, H. Smoking at diagnosis is an independent prognostic factor for cancer-specific survival in head and neck cancer: findings from a large, population-based study. *Cancer Epidemiol. Biomarkers Prev.* **23**, 2579–90 (2014).
63. Vermorken, J. B. *et al.* Impact of tumor HPV status on outcome in patients with recurrent and/or metastatic squamous cell carcinoma of the head and neck receiving chemotherapy with or without cetuximab: retrospective analysis of the phase III EXTREME trial. *Ann. Oncol.* **25**, 801–7 (2014).
64. Gheit, T. *et al.* Role of mucosal high-risk human papillomavirus types in head and neck cancers in central India. *Int. J. Cancer* **141**, 143–151 (2017).
65. Jung, A. C. *et al.* Biological and clinical relevance of transcriptionally active human papillomavirus (HPV) infection in oropharynx squamous cell carcinoma. *Int. J. Cancer* **126**, NA-NA (2010).
66. Husain, N. & Neyaz, A. Human papillomavirus associated head and neck squamous cell carcinoma: Controversies and new concepts. *J. oral Biol. craniofacial Res.* **7**, 198–205 (2017).
67. American Joint Committee on Cancer & American Cancer Society. Head and Neck

Sites. in *AJCC Cancer Staging Manual* 24–59 (Lippincott-Raven Publishers, 1997).

68. American Joint Committee on Cancer & American Cancer Society. *AJCC Cancer Staging Manual*. (Springer Publishing, 2016).
69. Salvador-Coloma, C. & Cohen, E. Multidisciplinary Care of Laryngeal Cancer. *J. Oncol. Pract.* **12**, 717–24 (2016).
70. Okoye, C. C. *et al.* Cardiovascular Risk and Prevention in Head and Neck Cancer Patients Treated with Radiotherapy. *Head Neck* **39**, 527 (2017).
71. Wei, M. *et al.* Cardiovascular disease risks among head and neck cancer survivors in a large, population-based cohort study. *J. Clin. Oncol.* **36**, 6051–6051 (2018).
72. Sommers, L. W. *et al.* Survival Patterns in Elderly Head and Neck Squamous Cell Carcinoma Patients Treated With Definitive Radiation Therapy. *Int. J. Radiat. Oncol.* **98**, 793–801 (2017).
73. Wolff, K.-D., Follmann, M. & Nast, A. The diagnosis and treatment of oral cavity cancer. *Dtsch. Arztebl. Int.* **109**, 829–35 (2012).
74. Funk, G. F. *et al.* Presentation, treatment, and outcome of oral cavity cancer: A national cancer data base report. *Head Neck* **24**, 165–180 (2002).
75. Abola, M. V, Prasad, V. & Jena, A. B. Association between treatment toxicity and outcomes in oncology clinical trials. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **25**, 2284–9 (2014).
76. Niu, Q. *et al.* The effect of marital status on the survival of patients with bladder urothelial carcinoma: A SEER database analysis. *Medicine (Baltimore).* **97**, e11378 (2018).
77. Li, Y., Zhu, M. & Qi, S. Marital status and survival in patients with renal cell carcinoma. *Medicine (Baltimore).* **97**, e0385 (2018).
78. College of American Pathologists. Human Papillomavirus Testing in Head and Neck Carcinomas. *College of American Pathologists* (2017). Available at: https://www.cap.org/protocols-and-guidelines/cap-guidelines/current-cap-guidelines/human-papillomavirus-testing-in-head-and-neck-carcinomas. (Accessed: 4th June 2019)
79. Guo, T., Eisele, D. W. & Fakhry, C. The potential impact of prophylactic human papillomavirus vaccination on oropharyngeal cancer. *Cancer* **122**, 2313–23 (2016).
80. Wang, C. *et al.* Targeting Head and Neck Cancer by Vaccination. *Front. Immunol.* **9**,

830 (2018).

81. D'Souza, G. & Dempsey, A. The role of HPV in head and neck cancer and review of the HPV vaccine. *Prev. Med. (Baltim).* **53**, S5–S11 (2011).
82. Guo, T., Eisele, D. W. & Fakhry, C. The potential impact of prophylactic human papillomavirus vaccination on oropharyngeal cancer. *Cancer* **122**, 2313–23 (2016).
83. Herrero, R. *et al.* Reduced Prevalence of Oral Human Papillomavirus (HPV) 4 Years after Bivalent HPV Vaccination in a Randomized Clinical Trial in Costa Rica. *PLoS One* **8**, e68329 (2013).
84. Graham, D. M. *et al.* A cost-effectiveness analysis of human papillomavirus vaccination of boys for the prevention of oropharyngeal cancer. *Cancer* **121**, 1785–1792 (2015).
85. Health Information and Quality Authority (HIQA). PRESS RELEASE: HIQA advises changing to a more effective HPV vaccine and extending the vaccine to boys. *HIQA News Updates* (2018). Available at: https://www.hiqa.ie/hiqa-news-updates/hiqa-advises-changing-more-effective-hpv-vaccine-and-extending-vaccine-boys. (Accessed: 27th February 2019)
86. Health Information and Quality Authority. *HTA of extending the HPV vaccination to boys*. (2018).
87. National Cancer Registry Ireland. National Cancer Registry Ireland. *National Cancer Registry* (2019). Available at: https://www.ncri.ie/. (Accessed: 30th April 2019)
88. National Cancer Registry of Ireland. *Cancer Trends: Cancers of the Head and Neck*. (2014).
89. Fakhry, C., Rosenthal, B. T., Clark, D. P. & Gillison, M. L. Associations between Oral HPV16 Infection and Cytopathology: Evaluation of an Oropharyngeal "Pap-Test Equivalent" in High-Risk Populations. *Cancer Prev. Res.* **4**, 1378–1384 (2011).
90. Kreimer, A. R. & Chaturvedi, A. K. HPV-associated Oropharyngeal Cancers--Are They Preventable? *Cancer Prev. Res.* **4**, 1346–1349 (2011).
91. Skandarajah, A. *et al.* Mobile microscopy as a screening tool for oral cancer in India: A pilot study. *PLoS One* **12**, e0188440 (2017).
92. Health Service Executive. Cervical Check. *Health Service Executive* (2019). Available at: https://www.hse.ie/eng/cervicalcheck/. (Accessed: 19th May 2019)
93. Wentzensen, N., Schiffman, M., Palmer, T. & Arbyn, M. Triage of HPV positive women

in cervical cancer screening. *J. Clin. Virol.* **76**, S49–S55 (2016).

CHAPTER 9

DISCUSSION

9 CHAPTER 9: DISCUSSION

9.1 The Former State of the Literature

HNC has stereotypically been a sub-set of cancer driven by smoking-, alcohol[1]-, occupation[2,3]-, and genetic[4]-related carcinogenesis. Up until the 1980s in fact, these were the sole known risk factors associated with HNC, no matter the sub-site within the head and neck from which they emerged. The first papers published regarding HPV as a major etiologic factor in HNSCCs in 1983 and 1985 respectively[5,6] irrevocably changed the understanding of oncogenesis in the region. Later, when HPV's particular relevance to the subset of HNSCCs that arise mainly in the oropharynx was clarified[7,8], the need for differential clinical assessment and treatment of these cancers became evident.

Since, data from the United States[9,10], Canada[11], Scandinavia[12–14], the United Kingdom[15], and Germany[16,17] have demonstrated the ever-increasing proportion of oropharyngeal, oral cavity, and laryngeal SCC, specifically OPSCC, attributable to HPV. The virus' role in driving these trends is so great that there are suggestions that it may result in the annual numbers of HPV-related OPSCC surpassing those of cervical cancer in the near future[18]. HPV-related OPSCC has even been designated as an epidemic in scientific nomenclature[13,19].

Most importantly, the unveiling of persistent HPV infection as a carcinogen in oropharyngeal, oral cavity, and laryngeal SCC has resulted in the emergence of a completely different carcinogenic pathway in these cancers, driven by the virus' oncogenes E6[20–22] and E7[23,24]. These interfere with the p53 and pRb tumour suppressor genes respectively, inducing de-regulation of DNA damage repair, genomic instability, telomerase proliferation, and cell proliferation. Most significantly, they also result in the up-regulation of tumour suppressor protein p16 and the down-regulation of EGFR[25,26], strangely sabotaging characteristics that play a role in the better overall and cancer-specific survival of these diseases[19,27–30].

De-escalation of treatment on the basis of the HPV-related nature of many OPSCC has been posited in the literature since 2010[31], but findings are not resounding even with the available evidence. Trials are ongoing to investigate the potential of de-escalation in these

patients[32–34], with a great extent of hope in the potential for surgery followed by decreased dose adjuvant radiotherapy. However, the definition of what constitutes a 'HPV positive'case and thus HPV-related patient is still unclear and pending outcomes of these trials are likely to be influenced by the HPV-associated characteristics, including surrogate biomarker p16 and younger patient age, which are chosen to represent HPV status and patients least likely to be under-treated with de-escalation.

Though expansive and ever-growing, the current literature regarding HPV-related oropharyngeal, oral cavity, and laryngeal SCC is still limited to a confined number of countries with relevant epidemiological data, which for the most part are small in scale. Though exceptions do exist, including a study with over 1,000 cases from the United Kingdom[15], most investigations regarding HPV's role in oropharyngeal, oral cavity, and laryngeal SCC range between 100 and 300 cases. In addition to this, the vast majority of these studies employ differing HPV-detecting technologies and as a consequence, define 'HPV positivity' differently, making it difficult to meaningfully compare results between different studies and populations, and definitively delineate HPV-related cases of oropharyngeal, oral cavity, and laryngeal SCC in the clinic.

9.2 The Initial Value of the ECHO Study

It was on the basis of the aforementioned gaps in the literature that the ECHO study found the source of its value. The ECHO study, analysing 861 cases, is one of the largest studies ever conducted regarding oropharyngeal, oral cavity, and laryngeal SCC. It is also the first of its kind to describe the epidemiology of oropharyngeal, oral cavity, and laryngeal SCC in Ireland, setting Ireland apart from large European countries on the continent whose comparable data has not yet been published.

The ECHO study is also providing the WHO, which is conducting a worldwide meta-analysis on HPV-related HNSCC called the HPV-AHEAD study[35,36], with the first ever Irish data. To be included in the HPV-AHEAD study, the ECHO study was required to follow a standardized sectioning, processing, and testing protocol which is currently being used to test samples from around the world for the same meta-analysis. The homogenized nature of the

methodology used in the ECHO study as a part of the larger HPV-AHEAD study is extremely valuable given the extent of heterogeneity of technologies and principles used to test for HPV throughout the literature. The consequences of varying detection methods and differing definitions for what constitutes an 'HPV-positive' case for the validity, feasibility, and reliability of results are significant (Chapter 2). In fact, not only did the ECHO study provide the opportunity to review the currently available technologies for detecting DNA (Chapter 2), but the pilot study for the ECHO project provided an additional chance to assess the validity of two different HPV DNA-detecting methodologies. Until now, only a few comparisons between technologies optimized for cervical samples had been carried out[37,38], with no such studies existing for HNCs.

Finally, the size and retrieval methods for samples included in the ECHO study made it possible to evaluate and make constructive suggestions regarding the manner in which Irish clinical research is carried out. The significance of this demonstrated the effect and impact that the current hospital and registry systems of collating and sorting data and material for cancer cases in Ireland have on implementing large population-based research studies. No appraisal of this nature has yet been conducted. Thus, the essential nature of the ECHO study for the current gaps in the literature was evident at the out-set.

9.3 Summary of Findings

The aims of the ECHO study were three-tiered, as outlined in Section 1.11. With respect to the first tier of aims, Chapter 7 unveiled the estimated prevalence of HPV DNA positivity in archival tumour specimens from patients diagnosed with oropharyngeal, oral cavity, and laryngeal SCC between 1994 and 2013 in Ireland. Prevalence amongst all oropharyngeal, oral cavity, and laryngeal SCC was 17.1% (CI: 14.6, 19.6), while in the oropharyngeal sub-site it was 41.1% (CI: 34.5, 47.8), in the oral cavity it was 10.9% (CI: 7.5, 14.2), and in the larynx it was 7.8% (CI: 4.9, 10.7). Within the oropharynx, the high prevalence was mainly driven by the tonsillar sub-site (60.0%). The ECHO study also described the genotype distribution in this population as overwhelmingly dominated by HR HPV16, with over 80% if not 90% of HPV positive cases overall within each sub-site being HPV16 positive. Other HR HPV genotypes including HPV18, 31, 33, 35, 51, 56, and 66 were also detected, but all presented

prevalence below 10%. Raw incidence of HPV positive and negative oropharyngeal, oral cavity, and laryngeal SCC was seen to increase over the 1994 to 2013 time-period, but proportionally, HPV negative cases from the oral cavity and the larynx dominated the landscape. Expectedly, the oropharynx was the only sub-site where raw incidence of HPV positive cases began to proportionally outweigh the influence of HPV negative cases over time, and average annual percentage increase was a notable 16.4% ($p<0.0001$) in these cases.

In Chapter 7, the study also described the predictors of HPV positivity for all oropharyngeal, oral cavity, and laryngeal SCC in Ireland which were younger age, oropharyngeal sub-site, and never- and ex-smoker status. The only sub-site for which there were significant predictors of HPV positivity was the oropharynx, which included included younger age and ex- and never-smoking status. That there were no predictors of HPV positivity in the larynx or oral cavity highlighted that the oropharynx was the only sub-site for which HPV likely played a role in carcinogenesis. This was further emphasized by additional analyses conducted showing that HPV positivity was a significant predictor of oropharyngeal sub-site while HPV negativity was a significant predictor of both oral cavity and laryngeal cancers. Later TNM stage and younger age were also predictors of oropharyngeal sub-site while ex-smoking status and female sex predicted oral cavity sub-site, and current smoker status, well-differentiated grade, and earlier TNM stage predicted laryngeal sub-site. Chapter 7 also highlighted that the only sub-site for which any stage variable was associated with HR HPV status was the oropharynx in relation to N stage, likely due to the physiological features of the tonsillar sub-site, which contributed the majority of HPV positive cases to the pool of positive cases in the oropharynx.

Chapter 8 established that both overall and cancer-specific survival were significantly improved amongst HPV positive oropharyngeal, oral cavity, and laryngeal SCC patients, and detailed that this relationship emanated strictly from the oropharyngeal sub-site. Where TNM stage, age, and smoking status were significant predictors of survival across most sub-sites, HR HPV status was only a significant predictor of survival in the oropharyngeal sub-site. Most importantly, the study found that when analysed by sub-site and HPV status individually, the most optimal treatment schemes for each sub-site were divergent. For HPV

positive OPSCC, surgery alone saw the best overall and cancer-specific survival. There was no significant evidence to suggest significant additional benefit with the addition of radiotherapy and chemotherapy. The majority of laryngeal tumours (those that were HPV negative) responded best to radiotherapy, while the same was true of most oral cavity cancers (those that were HPV negative) and surgery alone. HPV positive cases, again driven by those in the oropharynx, were more likely to be treated with surgery/radiotherapy/chemotherapy than their HPV negative counterparts due to disproportionate diagnosis at later TNM stage. HPV negative patients, mostly driven by those in the larynx, were more likely to be treated with radiotherapy.

The initial detection technology pilot study within ECHO (Chapter 5) also determined the significant influence of the type of HPV-detection technology and sample processing protocol employed on the results of studies in the head and neck space. It highlighted, by comparing SPF10 PCR to Multiplex PCR using varying sterility protocols, that highly sensitive technologies with extremely strict and standardized protocols are those that should be prioritized for HPV detection, if not only in the current epidemiological context, but in the clinic should HPV testing become required in addition to p16.

Finally, the assessment of procedures, parties, and lengths of time required for sample collection for the ECHO study found that 45 unique parties and procedures were involved in the identification and retrieval of cases (Chapter 4). An average of 273.67 (CI: 121.04, 426.30) days were required to organize the review of reports and block retrieval, compared to an average of 7.17 (CI: 1.35, 12.99) days to carry out the review and retrieval of blocks. Where a total of 3426 samples were originally eligible in the Irish population, 861 were ultimately included due to attrition of ineligible cases at the review, collection, and histopathological evaluation steps of sample collection and assessment. Reasons for elimination ranged from simple inaccessibility, to time delays due to new protocols introduced after the introduction of GDPR, to samples originating in irrelevant parts of the body, indicating a diversity and disconnect in the management of patient samples and information for the purposes of Irish research.

9.4 The Irish Population in Context

The ECHO study proved the hypotheses addressed in Chapter 1 to be overwhelmingly accurate. First, the Irish population was comparable to other European and North American countries with respect to the extent of HPV's role in carcinogenesis in oropharyngeal, oral cavity, and laryngeal SCC. The overall prevalence of HPV in the population was 17.1%, falling just outside the bounds of some of the lowest already recorded in these regions[12,39–42].

The slightly low prevalence in Ireland was evidently attributable to the extremely low HPV prevalence recorded in the larynx. Though the role of HPV in laryngeal SCC has shown to be insignificant in previous studies with prevalence as low as 7%[43], others suggest that prevalence of the virus in the region can reach as high as 62%[44]. In the United States the figure rests just above 20%[40]. The 7.8% of LSCC that were HPV positive in this study, and the fact that both LR genotypes detected overall originated in the larynx, resoundingly indicated that in Ireland, the larynx is an almost determinedly irrelevant site of HPV carcinogenesis in the Irish population.

By contrast, the prevalence of HPV in the oropharynx of 41.1% was directly comparable to other high rates reported in Europe[12,16,45]. This is especially true of SCC in the tonsil, whose prevalence has been shown to be the highest of any other detailed sub-site within the oropharynx[46–48]. Furthermore, the contribution of HPV positive cases of OPSCC to the overall OPSCC raw incidence has been steadily increasing in Ireland since the 1990s, but particularly since the mid-2000s. In fact, though the Irish population might be more similar to the British population with respect to the influences of smoking- and alcohol-related carcinogenesis remaining extremely significant for the incidence of all oropharyngeal, oral cavity, and laryngeal SCC, the trend in proportional contribution of HPV-related OPSCC to all OPSCC in Ireland was far more similar to those shown to be increasing rather than static in Scandinavia and the United States.

Nonetheless, the study highlighted the more heterogeneous nature of SCC arising in the oral cavity, with prevalence resting at 10.9%, a percentage falling within the ranges detected in other countries[40,49]. Risk factors for HPV positivity and survival amongst these cancers indicated possible "hybrid" tumours in the region, involving HPV- and smoking-

carcinogenesis, likely disproportionately originating from socially deprived patients whose exposure to the virus and smoking/alcohol consumption occurs at younger ages. Whether this is due to synergistic effects of HPV and other carcinogens[10,50] or simply to the diverse nature of the more detailed sub-sites included in the oral cavity (e.g. mouth, tongue, other mouth/pharynx) is unclear.

9.5 The Irish Population and Prevention: The HPV Vaccine, Sexual Education, and Smoking Prevention

The ECHO study's findings emphasize that the HPV vaccine, especially the nona-valent Gardasil-9 vaccine, could potentially have a crucial role to play in the prevention of oropharyngeal, oral cavity, and laryngeal SCC from now on, something that will be reflected in Irish public health policy in September 2019[51]. The vaccine is currently only licensed for prevention of cervical, vulva, vaginal, and anal cancers. However, the present results, should HPV DNA detection be taken as proof of carcinogenic involvement, suggest that the vaccine has the potential to protect against over 90% of HR HPV infections involved in oropharyngeal, oral cavity, and laryngeal SCC, and all LR HPV infections identified in this population. The recent decline in uptake of the HPV vaccine amongst girls included in the public scheme is therefore worrying[52,53], but the recovery of these rates in the last two years[54] suggests that the future HNSCC-free future of young women in Ireland has been at least somewhat salvaged.

Importantly, the sampling of the population with respect to sex was representative of the entire oropharyngeal, oral cavity, and laryngeal SCC population in Ireland between 1994 and 2013. Men comprised over two thirds of the population, both at the overall and sub-site level. This determined that men are more at risk of developing both HPV-related and HPV-unrelated SCC in Ireland, much like in other developed countries. The present data therefore highlights the urgency of the expansion of the public vaccination scheme to young boys in Ireland. In fact, the data compliments recent studies indicating that gender neutral vaccination in particular is crucial to the extermination of HPV-related disease in general[55]. For herd immunity to be effective throughout the Irish population, gender neutral vaccination is a necessity, accompanied by a high rate of uptake by both girls and boys[55]. An

uptake rate of at least 50% for boys is necessary should herd immunity cease the propagation of the virus[56]. This number rises to over 80% should the majority of the decline in HPV-related disease be attributable to mostly male maladies, HNSCC being proportionally the largest of these[56]. Should uptake be any lower in boys, most gains will remain attributable to declines in cervical cancer. That HIQA has already recommended this move for policy at the national level[56,57] and that boys will be included in the scheme in September 2019[58,59] is an indication on the basis of this study that this expansion will unequivocally be a long-term investment in the health of Irish people.

Given that many sexual behaviours have been shown to be surrogate biomarkers for HPV16 positivity[60], the present study emphasizes the increasing importance of ever-diversifying, earlier, and more frequent sexual behaviours in driving the emerging dominance of HPV-related OPSCC in the overall incidence of cancers in the sub-site[61]. Ireland's sexual education in national schools has never been lauded for its direct, fact-based, evidence-driven approach. In fact, a 2019 report from the Joint Committee on Education and Skills concluded that the current state of sexual education in Ireland is actually doing young people a disservice[62]. The urgent need to pair potential expansions of vaccination with appropriate information delivery to students regarding the spread of STIs and how to prevent them is thus evident if both young women and men are to avoid the longer-term consequences of risky sexual behaviour. The need for policy changes on this matter are particularly relevant in the case of HPV given its status as the most common STI in the world[63]. That the Dáil passed a bill guaranteeing fact-based sexual education to students in 2018 is a promising step towards the better short- and long-term health of young Irish people in relation to HPV-related disease, including HNSCC[64].

The ECHO study has also reinforced the seemingly increasing carcinogenic involvement that smoking- and alcohol-related behaviours continue to have on the vast majority of oropharyngeal, oral cavity, and laryngeal SCC despite the widespread knowledge that smoking in particular is causal in a variety of different cancer types. Indeed, smoking rates have decreased in Ireland since the 1970s, but this decline has not been as severe as it has been in other countries[65], with resurgence seen amongst younger women in particular, and

the persistent dominance of mostly HPV-unrelated laryngeal and oral cavity cancers in driving oropharyngeal, oral cavity, and laryngeal SCC incidence[66–68].

Indeed, the HSE has invested in many public health endeavours to address these trends, including support for smokers with the intention to quit (QUIT)[69,70]. The ECHO study reveals the importance of these quitting schemes as ex-smoking status: often predicted greater risk of HPV positivity and the potential hybridity of tumours; is related to female gender for whom smoking rates have recently increased; and also had better survival than current smoking status. However, the need to catch young people before they begin smoking is urgent. The introduction of more anti-smoking education in schools, and a return to investing in public health campaigns emphasizing the consequences of the behaviour are required.

9.6 The Irish Population and HNC Screening Tools

The ECHO study results showcase the urgent need for HNC screening tools, no matter their HPV status. For HPV-unrelated HNSCC in this analysis, especially in the larynx, diagnosis at later TNM stage was the only predictor of cancer-specific survival after adjustment for other variables, and no treatment, harsh or otherwise, was ultimately successful in maintaining patient overall survival above 20% after 10 years. Later T and M, and TNM stage also remained significant predictors of overall and cancer-specific survival respectively in the oropharynx. For laryngeal patients in particular, diagnosis at later stage makes death almost a certainty. Given the good response early stages of LSCC have to radiotherapy alone[71], the ECHO study thus emphasizes the extent of imminently saveable life wasted in mostly HPV-unrelated SCC without appropriate early detection screening tools.

Efforts are currently being made to investigate the best ways to sample tissue from the oral site, but it is made difficult by the region's confined nature and the dense, complex network of MALT tissues that line it[72,73]. This is especially true in the cases of the tonsillar crypts. Recently, mobile microscopy with a simple brush biopsy has shown to be an effective screening mechanism for oral cavity cancer, even in low-resource areas[74], but such a sampling method is not ideal for the deep, hidden, crypts of the oropharynx. The role that

HPV might play in this screening is also uncertain, though monitoring systems like those established in the cervical case[75,76] are a promising way of catching HR HPV patients who, perhaps even after vaccination, go on to develop lesions.

9.7 The Irish Population and De-escalation of Treatment for HPV-related OPSCC

Trials regarding the de-escalation of treatment for HPV-related oropharyngeal, oral cavity, and laryngeal SCC are already under-way[32–34,77]. This said, the current data confirms that de-escalation, specifically of adjuvant radiotherapy, is likely to be just as if not more relevant to HPV-related OPSCC in the Irish population, given the optimal survival using surgery alone amongst these patients, and the confirmation that chemotherapy could be entirely eliminated for some patients. The significance of surgery, whether accompanied by other interventions or not, is clear for HPV-related OPSCC, and the promise of trans-oral resection (TOR) using robotics as a survival-maximizing technique with often better functional outcomes than radiotherapy for OPSCC is implied despite the fact that HPV positive tumours tend to be node positive under older AJCC staging systems.

The findings thus also underscore the extreme relevance of the new 8th edition AJCC staging system[78] to the Irish context, with HPV-related OPSCC being previously misinterpreted as more aggressive than their behaviour suggests as a result of both their anatomical proximity to and early involvement with the lymphatic system, and their HPV-specific behaviour/response to treatment. However, the ECHO study's suggestion that both down-staging and de-escalation are relevant to Irish HPV-related OPSCC, also warns that the combination of the two could result in the under-treatment of these cases. The earlier stages at which HPV-related OPSCC are now being diagnosed require less severe treatment schemes to begin with (those without chemotherapy), and further de-escalation could jeopardize patient safety, an unknown outcome that will be clarified with pending clinical trial results.

This said, the current data highlight the importance of nodal category to the determination of survival for patients, something that, given the statistical analysis generated, was evidently not reflected by the 5th AJCC criteria. That increased number of lymph nodes

involved at diagnosis shows congruently poorer survival has however been adopted within the 8th edition for HPV positive oropharyngeal disease, and the nuance associated with nodal category for these patients has become key to treatment decisions[79,80]. Specifically, N1 denotes any ipsilateral lymph node involvement; N2 indicates any bilateral or contralateral lymph node involvement; and N3 denotes lymph nodes larger than 6cm. Risk of distant failure is exponential when disease is diagnosed at N2b or higher, and despite chemotherapy being seemingly unnecessary for many early stage patients, any patient with N2b or higher nodal category will now receive concurrent chemo-radiation without surgery[79]. Thus, though patients in the present data set appeared not to benefit enormously from chemotherapy, the importance of individual assessment and personalized approaches to modality selection will continue to be crucial to the best patient outcomes.

9.8 Contributions to World-wide Epidemiological Data

The need for world-wide epidemiological data regarding HPV's relationship to oropharyngeal, oral cavity, and laryngeal SCC using standardized technologies that generate comparable and reliable results was established in Chapter 2. The HPV-AHEAD Study[35,36] from the IARC is the first of its kind to do exactly this, and that the ECHO study is Irish contributor to this meta-analysis enriches its value extending beyond informing Irish health and clinical policy. Specifically, the HPV-AHEAD study aims to provide important insights for the diagnosis, treatment, and prophylaxis of HPV-related HNSCC world-wide. Importantly, the study aims to, with the use of HPV DNA, mRNA, and p16, further elucidate the role that HPV may play in any eventual screening for HNSCC given the enormous populations to which it will have access. These large-scale analyses will indicate the proportion of HPV infections most likely to have been causal in each cancer, and whether or not screening for HPV-related HNSCC on the basis of these numbers is a feasible and cost-effective policy.

9.9 Limitations of the ECHO Study

The ECHO study was not limited in its findings by its sample size, whether for all oropharyngeal, oral cavity, and laryngeal SCC or for each individual sub-site for risk factor analysis. With respect to sub-divisions by survival and treatment, sample sizes for particular combinations of characteristics did limit sample size, which, given that the majority of

patients were deceased, did not discredit findings, but simply suggests that further analysis targeting particular patients treated in certain ways should be conducted to confirm results.

The study was also not disproportionately impacted by missing data in comparison to missing data in the NCRI database for all eligible cases. Nonetheless, the impact of non-random missing data was evident in relation to some variables including stage, grade, and smoking status. Survival analysis and association tests in Chapter 6 indicated that missing data was disproportionately relevant to patients likely to having the following characteristics: current smoker, well-differentiated tumour, and Stage II and III TNM stages. Current smoking survival was thus likely biased for better outcomes, and well-differentiated tumours were biased for worse outcomes. The over-sampling of never smokers also yielded bias for worse survival than expected in these patients and better survival than expected in ex-smokers. These are simply artefacts of registry-based data, and impact the vast majority of epidemiological studies discussed in Chapter 6. There is no perfect way to deal with non-random missing data[81–84], but the study employed a technique that is conventional in the literature, including a "missing/unknown" category for those variables with greater than 10% missing data to account for any significance emanating from missing data. Despite this, relationships between known data and HPV status and survival were not severely impacted by non-random missing data, with sensitivity analysis conducted excluding all missing cases for relevant tests achieving broadly similar results. That treatment administered was also limited to available information within the first year of diagnosis also meant that assessment of this variable was not as comprehensive as it could have been.

Additionally, though the epidemiological goals of the ECHO study were satisfied with testing for HPV DNA alone, the extrapolation of HPV DNA detection in a tumour specimen in a snap-shot of time (e.g. when the relevant tumours were excised) to the certain role of the virus in the tumour's on-set, was not unequivocal. This is particularly pertinent given that where the natural history of HPV in the cervix is well-described, the same is not true of HPV in the oral cavity. When HPV infections become persistent, and what the time scale is between persistence and the onset of carcinogenesis has not yet been elucidated. Thus, whether or not HPV DNA infections detected in the population were present before tumours originated and were determinedly causal in their growth cannot be confirmed. The

age at which administration of the HPV vaccine could be most optimal for OPSCC prevention is unclear as a consequence, though the current administration in early adolescence for cervical cancer prevention remains sensical due to first exposure to the virus occurring on average in teenage years.

9.10 Future Directions

Though the ECHO is the most comprehensive summary of oropharyngeal, oral cavity, and laryngeal SCC in Ireland to date, the findings of the study lend to the exciting potential of additional work not only on oropharyngeal, oral cavity, and laryngeal SCC in general, but for those cases included in the study especially. The need for compounding evidence for the carcinogenic nature of HPV infections in the clinical context requires the addition of p16 immunohistochemistry and/or mRNA detection to the current epidemiological data. Accompanying slides for this purpose were processed in anticipation of this throughout the ECHO study (Chapter 3). Furthermore, mRNA's potential for representing carcinogenic HPV activity alone is still unclear. The size of the ECHO study, and the preparation of mRNA samples for future analysis that formed a part of its protocol (Chapter 3), provides the perfect opportunity to perform mRNA testing on samples with already generated HPV DNA data. Together, the compilation of HPV DNA, mRNA, and p16 immunohistochemistry for 861 oropharyngeal, oral cavity, and laryngeal SCC will provide evidence on a large scale as to which indicators of HPV accurately assess carcinogenic activity, whether alone or in tandem with one another, not only in the Irish context but also as a part of the larger HPV-AHEAD study. The behaviour and characteristics of "hybrid" cases of oropharyngeal, oral cavity, and laryngeal SCC will also be further elucidated with the enormous population numbers the AHEAD study will have access to.

This said, even with the addition of p16 and mRNA data to the current data, the natural history of HPV in the oral cavity will remain a mystery as these 'snap-shot' indicators do not give the full picture of the activity of these infections in patients. This is important if the role that HPV may play, if any, in HNSCC screening is to be determined. Based on the findings of this study, it is likely that HNSCC screening should begin in the 40s, especially amongst men,

to account for HPV-related HNSCC's disproportionate occurrence in those under the age of 50.

However, without confirmation of the time-scale between infection and carcinogenesis, or the determinants of which HR HPV infections go on to become carcinogenic, it will be difficult to justify the inclusion of HPV in screening for HNSCC. For instance, it may not be necessary to screen for HPV preventatively if there is no established risk-assessment to act as a triage of patients who test positively for HR types, and screening could just rely on swabs for cells in the early stages of transformation for high-risk populations. This is especially true since the vast majority of HPV infections, even those that are HR, are cleared, and it would not be cost-effective to screen entire populations without an understanding of what an identified infection really means for risk of eventual SCC. mRNA could have an interesting role to play in screening for this reason, indicating an active infection, but the risk that positivity predicts for onset of HNSCC is also unknown for this indicator.

With respect to the clinic, the results of pending clinical trials will ultimately satisfy the necessary and currently missing data relating to de-escalation of treatment in HPV-related HNSCC. It is hoped that the evidence from the current epidemiological data for HPV-related OPSCC in particular will confirm that modern surgical techniques alone including TOR, or perhaps surgery with de-escalated radiotherapy yield the same if not better survival rates amongst these patients as harsher treatment schemes. Once the results of the ADEPT, SIRS, and ECOG E3311 in particular come to bear after 2020 however, it will be important to assess two phenomena.

The first is how the already down-graded nature of HPV-related OPSCC in the 8th edition AJCC guidelines will impact on the extent of de-escalation required for patients. It is possible that less harsh treatments applied to now earlier Stage patients may already account for moderate de-escalation, and avoidance of the under-treatment of patients is imperative. Second, there will be a need to evaluate how patients, whose treatment is becoming decreasingly paternalistic, react to the idea that de-escalated treatment may work just as effectively as current treatments. One of the only studies conducted on the matter found that patients are generally reluctant to accept de-escalation on the assumption that more

aggressive schemes have the best survival outcomes[85]. How to present these (potential) treatment options so that patients do not unwittingly sabotage their own quality of life, or experience undue anxiety about less harsh treatment will need to be investigated.

Furthermore, though the updated CAP and AJCC 8th edition guidelines for staging and diagnosing SCC represent a revolutionary and appropriate change to clinical practice for HPV-related OPSCC, the current study posits the need for further adjustment of the guidelines. It was already known in the literature that p16, though considered a surrogate biomarker for HPV positivity, does not always indicate a purely HPV-driven tumour[86,87]. Paired with the evidence in the literature suggesting that it is HPV positive and p16 positive tumours that see the best survival, significantly better than even p16 positive tumours alone[48,88,89], this highlights the urgency for defining the exact group of patients for whom any treatment decisions might be altered. In fact, testing for more than one HPV indicator or biomarker could yield a treatment triage-system for all HNSCC tumours, graded on the basis of the extent of their relationship to HPV. For instance, tumours positive for p16 but negative for HPV DNA and mRNA might benefit from slight de-escalation of treatment, but not to the extent that a tumour positive for all three markers might. Such a system could also account for the differential behaviours of "hybrid" cases of oropharyngeal, oral cavity, and laryngeal SCC.

On the subject of treatment changes however, immune- and tumour microenvironment-based approaches are also worth investigation. Recently, PD-1:PD-L1 immune checkpoint inhibitors have shown over 20% responsiveness in HPV positive patients, with some patients even exhibiting complete responses[90]. Attributes of the tumour microenvironment in HPV-related cases, especially given their high CD8+ and CD4+ infiltration, are also potentially exploitable[91,92]. Though anti-angiogenic agents show poor clinical response rates, the Society for Immunotherapy of Cancer recently published a consensus statement on the use of pembrolizumab for the treatment of recurrent or metastatic HNSCC on the basis of immune scoring of PD-L1 (>50%)[93]. The Food and Drug Association in the United States also approved the immunotherapy as a single agent for patients with HNSCC whose tumours express a PD-L1 combined positive score of 1 or higher[93]. Cancer associated fibroblasts prepared from HNSCC also differ transcriptionally from normal fibroblasts[94,95] in more than

500 genes encoding proteins such as IGF-2, IL-6, IL-8, and CXCL-1[96–98], factors that are crucial for the maintenance of stem cell properties of HNSCC cells. How these differential expressions can be targeted for cancer therapy however is as yet unknown.

In the lab, HNSCC cell lines (including SQD9, SCC61, Cal27, SC179, SC2763, JH011) have been identified to have an increased response to radiation therapy following exposure to demethylating agents in vitro[99]. Histone deacetylase inhibitors could also induce cell cycle arrest and promote apoptosis in HNSCC, increasing sensitivity to chemotherapy. For instance, combinations of 5-AZA-CdR or a histone deacetylase inhibitor with cisplatin enhanced cytotoxic effectiveness in HNSCC treatment in two studies[100,101]. DNA methyl transferases have also shown some synergistic effects with radiation by reducing HNSCC cell survival compared to singular treatments and by increasing radiation-induced apoptosis[99]. Trials focusing on histone deacetylase targets are currently ongoing, but in its current state for HNSCC, immune scoring and tumour microenvironment composition represent new and but singular steps towards personalized treatment for all HNSCC types, with HPV being only one biomarker upon which to base schemes. For this reason, the less elucidated area of combination treatments including immunotherapies, episomal targets, and standard courses of radiotherapy and chemotherapeutic treatments to avoid tumour resistance will continue to be extremely valuable areas of investigation for these cancers.

No matter, extensive work is still necessary to clarify HPV's role in the screening, diagnosis, and treatment of HPV-related HNSCC. However, what is clear is that the indicators on which these procedures are based need to be homogenous, standardized, and optimal for each context. This will require further investigation of other biomarkers, the development of new technologies suitable to maximal sensitivity and specificity in fresh, FFPE, and frozen tissue, and like in the cervical context, the clinical validation of these panels of biomarkers and platforms. For instance, in the clinic, in-situ hybridization has been used to detect HR HPV, but it is cumbersome, not maximally sensitive, and not universally available for HR HPV testing in the clinical context[102]. Highly sensitive and efficient technologies that will likely become more cost-effective with time including that used in the ECHO study (Multiplex PCR paired with Luminex® technology) may instead become standard in the clinic as a complement to p16 immunohistochemistry as discussed in Chapter 2.

Lastly, Chapter 4's evaluation of sample collection for the ECHO study revealed a quantitatively urgent need for a homogenized database and biobanking system for the Irish nation, given the noted relationship between the extent of centralization in the research process, and the level of efficiency, quality, and impact of the research on patients[103,104]. Since the ultimate goal of clinical research, including the ECHO study, is a realized improvement in patient outcomes, the faster meaningful results can be generated, and the greater the quality of these results, the more research serves its real purpose. Efforts to centralize management of patient samples are already underway with the work of BIOBANK: Ireland Trust, the country's most noted push towards establishing an Irish BioBank Network[105]. However, the BIOBANK is only present in two hospitals, and their expansion and funding is crucial if Ireland is to continue disproportionately contributing valuable research to the literature in comparison to other countries. This is particularly important in relation to HPV-related disease, with the emerging significance of other HPV-related ano-genital cancers yet to be analysed in Ireland.

Thus, there is still an enormous amount of work necessary to build on the value of the ECHO study. Nevertheless, the gap in the literature which the ECHO study has filled now provides the perfect platform on which these efforts can be based. Ultimately, the ECHO study will be a significant contribution to needed changes in policy and practice in Ireland and beyond.

References

1. Lee, Y.-C. A. *et al.* Smoking addiction and the risk of upper-aerodigestive-tract cancer in a multicenter case-control study. *Int. J. Cancer* **133**, n/a-n/a (2013).
2. Deeken, J. F. *et al.* The Rising Challenge of Non-AIDS-Defining Cancers in HIV-Infected Patients. *Clin. Infect. Dis.* **55**, 1228–1235 (2012).
3. Langevin, S. M. *et al.* Occupational dust exposure and head and neck squamous cell carcinoma risk in a population-based case-control study conducted in the greater Boston area. *Cancer Med.* **2**, 978–986 (2013).
4. Kutler, D. I. *et al.* High incidence of head and neck squamous cell carcinoma in patients with Fanconi anemia. *Arch. Otolaryngol. Head. Neck Surg.* **129**, 106–12

(2003).

5. Syrjänen, K., Syrjänen, S., Lamberg, M., Pyrhönen, S. & Nuutinen, J. Morphological and immunohistochemical evidence suggesting human papillomavirus (HPV) involvement in oral squamous cell carcinogenesis. *Int. J. Oral Surg.* **12**, 418–24 (1983).
6. Löning, T. *et al.* Analysis of oral papillomas, leukoplakias, and invasive carcinomas for human papillomavirus type related DNA. *J. Invest. Dermatol.* **84**, 417–20 (1985).
7. Paz, I. B., Cook, N., Odom-Maryon, T., Xie, Y. & Wilczynski, S. P. Human papillomavirus (HPV) in head and neck cancer. An association of HPV 16 with squamous cell carcinoma of Waldeyer's tonsillar ring. *Cancer* **79**, 595–604 (1997).
8. Gillison, M. L. *et al.* Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. *J. Natl. Cancer Inst.* **92**, 709–20 (2000).
9. Chaturvedi, A. K. *et al.* Human Papillomavirus and Rising Oropharyngeal Cancer Incidence in the United States. *J. Clin. Oncol.* **29**, 4294–4301 (2011).
10. Gillison, M. L. *et al.* Prevalence of oral HPV infection in the United States, 2009-2010. *JAMA* **307**, 693–703 (2012).
11. Forte, T., Niu, J., Lockwood, G. A. & Bryant, H. E. Incidence trends in head and neck cancers and human papillomavirus (HPV)-associated oropharyngeal cancer in Canada, 1992–2009. *Cancer Causes Control* **23**, 1343–1348 (2012).
12. Näsman, A. *et al.* Incidence of human papillomavirus (HPV) positive tonsillar carcinoma in Stockholm, Sweden: An epidemic of viral-induced carcinoma? *Int. J. Cancer* **125**, 362–366 (2009).
13. Ramqvist, T. & Dalianis, T. Oropharyngeal cancer epidemic and human papillomavirus. *Emerg. Infect. Dis.* **16**, 1671–7 (2010).
14. Carlander, A.-L. F. *et al.* Continuing rise in oropharyngeal cancer in a high HPV prevalence area: A Danish population-based study from 2011 to 2014. *Eur. J. Cancer* **70**, 75–82 (2017).
15. Schache, A. G. *et al.* HPV-Related Oropharynx Cancer in the United Kingdom: An Evolution in the Understanding of Disease Etiology. *Cancer Res.* **76**, 6598–6606 (2016).
16. Buttmann-Schweiger, N., Deleré, Y., Klug, S. J. & Kraywinkel, K. Cancer incidence in Germany attributable to human papillomavirus in 2013. *BMC Cancer* **17**, 682 (2017).
17. Wittekindt, C. *et al.* Increasing Incidence rates of Oropharyngeal Squamous Cell

Carcinoma in Germany and Significance of Disease Burden Attributed to Human Papillomavirus. *Cancer Prev. Res.* (2019). doi:10.1158/1940-6207.CAPR-19-0098

18. Chaturvedi, A. K. Epidemiology and Clinical Aspects of HPV in Head and Neck Cancers. *Head Neck Pathol.* **6**, 16–24 (2012).
19. Marur, S., D'Souza, G., Westra, W. H. & Forastiere, A. A. HPV-associated head and neck cancer: a virus-related cancer epidemic. *Lancet Oncol.* **11**, 781–789 (2010).
20. Rampias, T., Sasaki, C., Weinberger, P. & Psyrri, A. E6 and E7 Gene Silencing and Transformed Phenotype of Human Papillomavirus 16-Positive Oropharyngeal Cancer Cells. *JNCI J. Natl. Cancer Inst.* **101**, 412–423 (2009).
21. Thomas, M., Pim, D. & Banks, L. The role of the E6-p53 interaction in the molecular pathogenesis of HPV. *Oncogene* **18**, 7690–7700 (1999).
22. McMurray, H. R. & McCance, D. J. Human papillomavirus type 16 E6 activates TERT gene transcription through induction of c-Myc and release of USF-mediated repression. *J. Virol.* **77**, 9852–61 (2003).
23. Liu, X., Clements, A., Zhao, K. & Marmorstein, R. Structure of the Human Papillomavirus E7 Oncoprotein and Its Mechanism for Inactivation of the Retinoblastoma Tumor Suppressor. *J. Biol. Chem.* **281**, 578–586 (2006).
24. Moody, C. A. & Laimins, L. A. Human papillomavirus oncoproteins: pathways to transformation. *Nat. Rev. Cancer* **10**, 550–560 (2010).
25. Taberna, M. *et al.* The Use of HPV16-E5, EGFR, and pEGFR as Prognostic Biomarkers for Oropharyngeal Cancer Patients. *Front. Oncol.* **8**, 589 (2018).
26. Khaznadar, S. S. *et al.* EGFR overexpression is not common in patients with head and neck cancer. Cell lines are not representative for the clinical situation in this indication. *Oncotarget* **9**, 28965–28975 (2018).
27. Bose, S., Evans, H., Lantzy, L., Scharre, K. & Youssef, E. p16INK4A is a surrogate biomarker for a subset of human papilloma virus-associated dysplasias of the uterine cervix as determined on the Pap smear. *Diagn. Cytopathol.* **32**, 21–24 (2005).
28. El-Naggar, A. K. & Westra, W. H. p16 expression as a surrogate marker for HPV-related oropharyngeal carcinoma: A guide for interpretative relevance and consistency. *Head Neck* **34**, 459–461 (2012).
29. Schlecht, N. F. *et al.* A comparison of clinically utilized human papillomavirus detection methods in head and neck cancer. *Mod. Pathol.* **24**, 1295–1305 (2011).

30. Wang, H., Sun, R., Lin, H. & Hu, W. P16 ^{INK4A} as a surrogate biomarker for human papillomavirus-associated oropharyngeal carcinoma: Consideration of some aspects. *Cancer Sci.* **104**, 1553–1559 (2013).
31. Ang, K. K. *et al.* Human papillomavirus and survival of patients with oropharyngeal cancer. *N. Engl. J. Med.* **363**, 24–35 (2010).
32. Mirghani, H. & Blanchard, P. Treatment de-escalation for HPV-driven oropharyngeal cancer: Where do we stand? *Clin. Transl. Radiat. Oncol.* **8**, 4–11 (2018).
33. Villaflor, V. M. *et al.* Response-adapted volume de-escalation (RAVD) in locally advanced head and neck cancer. *Ann. Oncol.* **27**, 908–913 (2016).
34. Chen, A. M. *et al.* Reduced-dose radiotherapy for human papillomavirus-associated squamous-cell carcinoma of the oropharynx: a single-arm, phase 2 study. *Lancet Oncol.* **18**, 803–811 (2017).
35. HPV-AHEAD - Role of human papillomavirus infection and other co-factors in the aetiology of head and neck cancer in Europe and India. (2017). Available at: http://hpv-ahead.iarc.fr/. (Accessed: 22nd August 2017)
36. International Agency for Research on Cancer & World Health Organization. HPV-AHEAD Study. *International Agency for Research on Cancer* (2019). Available at: http://hpv-ahead.iarc.fr/about/index.php. (Accessed: 27th April 2019)
37. Safaeian, M. *et al.* Comparison of the SPF10-LiPA system to the Hybrid Capture 2 Assay for detection of carcinogenic human papillomavirus genotypes among 5,683 young women in Guanacaste, Costa Rica. *J. Clin. Microbiol.* **45**, 1447–54 (2007).
38. Hesselink, A. T. *et al.* Comparison of GP5+/6+-PCR and SPF10-Line Blot Assays for Detection of High-Risk Human Papillomavirus in Samples from Women with Normal Cytology Results Who Develop Grade 3 Cervical Intraepithelial Neoplasia. *J. Clin. Microbiol.* **46**, 3215–3221 (2008).
39. D'Souza, G. *et al.* Six-month natural history of oralversus cervical human papillomavirus infection. *Int. J. Cancer* **121**, 143–150 (2007).
40. Kreimer, A. R., Clifford, G. M., Boyle, P. & Franceschi, S. Human Papillomavirus Types in Head and Neck Squamous Cell Carcinomas Worldwide: A Systematic Review. *Cancer Epidemiol. Biomarkers Prev.* **14**, 467–475 (2005).
41. Hernandez, B. Y. *et al.* Human papillomavirus prevalence in invasive laryngeal cancer in the United States. *PLoS One* **9**, e115931 (2014).

42. Laskaris, S. *et al.* Prevalence of human papillomavirus infection in Greek patients with squamous cell carcinoma of the larynx. *Anticancer Res.* **34**, 5749–53 (2014).
43. Gungor, A. *et al.* Human papilloma virus prevalence in laryngeal squamous cell carcinoma. *J. Laryngol. Otol.* **121**, 772–774 (2007).
44. Tong, F. *et al.* Prevalence and Prognostic Significance of HPV in Laryngeal Squamous Cell Carcinoma in Northeast China. *Cell. Physiol. Biochem.* **49**, 206–216 (2018).
45. Khot, K. P., Deshmane, S. & Choudhari, S. Human Papilloma Virus in Oral Squamous Cell Carcinoma - The Enigma Unravelled. *Chin. J. Dent. Res.* **19**, 17–23 (2016).
46. Mourad, M. *et al.* Epidemiological Trends of Head and Neck Cancer in the United States: A SEER Population Study. *J. Oral Maxillofac. Surg.* **75**, 2562–2572 (2017).
47. Woods, R. Human Papillomavirus-related Oropharyngeal Squamous Cell Carcinoma – Prevalence and Incidence in a Defined Population and Analysis of the Expression of Specific Cellular Biomarkers. (2016).
48. Woods, R. *et al.* Role of human papillomavirus in oropharyngeal squamous cell carcinoma: A review. *World J. Clin. cases* **2**, 172–93 (2014).
49. Hübbers, C. U. & Akgül, B. HPV and cancer of the oral cavity. *Virulence* **6**, 244–8 (2015).
50. D'Souza, G., McNeel, T. S. & Fakhry, C. Understanding personal risk of oropharyngeal cancer: risk-groups for oncogenic oral HPV infection and oropharyngeal cancer. *Ann. Oncol.* **28**, 3065–3069 (2017).
51. HSE. About the HPV Vaccine - HSE.ie. *HSE* (2019). Available at: https://www.hse.ie/eng/health/immunisation/pubinfo/schoolprog/hpv/about/. (Accessed: 21st July 2019)
52. Health Service Executive. *HPV Vaccine Uptake in Ireland: 2015/16*. (2017).
53. Health Service Executive. *HPV Vaccine Uptake in Ireland: 2016/17*. (2018).
54. Cullen, P. HPV vaccine uptake among girls in Ireland rises to 70%. *The Irish Times* (2019). Available at: https://www.irishtimes.com/news/health/hpv-vaccine-uptake-among-girls-in-ireland-rises-to-70-1.3821784. (Accessed: 6th July 2019)
55. Hintze, J. M. & O'Neill, J. P. Strengthening the case for gender-neutral and the nonavalent HPV vaccine. *Eur. Arch. Oto-Rhino-Laryngology* **275**, 857–865 (2018).
56. Health Information and Quality Authority. *HTA of extending the HPV vaccination to boys*. (2018).

57. Health Information and Quality Authority (HIQA). PRESS RELEASE: HIQA advises changing to a more effective HPV vaccine and extending the vaccine to boys. *HIQA News Updates* (2018). Available at: https://www.hiqa.ie/hiqa-news-updates/hiqa-advises-changing-more-effective-hpv-vaccine-and-extending-vaccine-boys. (Accessed: 27th February 2019)
58. Libreri, S. HPV vaccine to be extended to boys. *RTE News* (2018). Available at: https://www.rte.ie/news/health/2018/1207/1015750-hiqa-hpv/. (Accessed: 21st July 2019)
59. Hilliard, M. HPV vaccination programme for boys to proceed in September. *The Irish Times* (2019). Available at: https://www.irishtimes.com/news/health/hpv-vaccination-programme-for-boys-to-proceed-in-september-1.3933725. (Accessed: 21st July 2019)
60. D'Souza, G. *et al.* Case–Control Study of Human Papillomavirus and Oropharyngeal Cancer. *N. Engl. J. Med.* **356**, 1944–1956 (2007).
61. Layte Hannah McGee, R. & Rundle Gráinne Cousins Claire Donnelly Fiona Mulcahy Ronán Conroy, K. *The Irish Study of Sexual Health and Relationships*. (2006).
62. Joint Committee on Education and Skills. *Report on Relationships and Sexuality Education*. (2019).
63. Center for Disease Control. STD Facts - Human papillomavirus (HPV). *Center for Disease Control* (2017). Available at: https://www.cdc.gov/std/hpv/stdfact-hpv.htm. (Accessed: 2nd February 2017)
64. Houses of the Oireachtas. Provision of Objective Sex Education Bill 2018 – No. 34 of 2018 – Houses of the Oireachtas. *Oireachtas* (2018). Available at: https://www.oireachtas.ie/en/bills/bill/2018/34/. (Accessed: 21st July 2019)
65. WHO | Prevalence of tobacco smoking. *WHO* (2016).
66. National Cancer Registry Ireland. National Cancer Registry Ireland. *National Cancer Registry* (2019). Available at: https://www.ncri.ie/. (Accessed: 30th April 2019)
67. National Cancer Registry of Ireland. *Cancer Trends: Cancers of the Head and Neck*. (2014).
68. National Cancer Registry of Ireland. *Cancer Trends: Cancers of the Head and Neck*. (2011).
69. Health Service Executive. QUIT. *Health Service Executive* (2019). Available at:

https://www.hse.ie/eng/health/hl/change/quit/about quit.html. (Accessed: 7th July 2019)

70. Health Service Executive. HSE National Standard for Tobacco Cessation Support Programme. *Health Service Executive* (2019). Available at: https://www.hse.ie/eng/about/who/tobaccocontrol/cessation/. (Accessed: 7th July 2019)
71. Salvador-Coloma, C. & Cohen, E. Multidisciplinary Care of Laryngeal Cancer. *J. Oncol. Pract.* **12**, 717–24 (2016).
72. Fakhry, C., Rosenthal, B. T., Clark, D. P. & Gillison, M. L. Associations between Oral HPV16 Infection and Cytopathology: Evaluation of an Oropharyngeal "Pap-Test Equivalent" in High-Risk Populations. *Cancer Prev. Res.* **4**, 1378–1384 (2011).
73. Kreimer, A. R. & Chaturvedi, A. K. HPV-associated Oropharyngeal Cancers--Are They Preventable? *Cancer Prev. Res.* **4**, 1346–1349 (2011).
74. Skandarajah, A. *et al.* Mobile microscopy as a screening tool for oral cancer in India: A pilot study. *PLoS One* **12**, e0188440 (2017).
75. Health Service Executive. Cervical Check. *Health Service Executive* (2019). Available at: https://www.hse.ie/eng/cervicalcheck/. (Accessed: 19th May 2019)
76. Wentzensen, N., Schiffman, M., Palmer, T. & Arbyn, M. Triage of HPV positive women in cervical cancer screening. *J. Clin. Virol.* **76**, S49–S55 (2016).
77. Wierzbicka, M., Szyfter, K., Milecki, P., Składowski, K. & Ramlau, R. The rationale for HPV-related oropharyngeal cancer de-escalation treatment strategies. *Contemp. Oncol. (Poznan, Poland)* **19**, 313–22 (2015).
78. American Joint Committee on Cancer & American Cancer Society. *AJCC Cancer Staging Manual*. (Springer Publishing, 2016).
79. O'Sullivan, B. *et al.* Development and validation of a staging system for HPV-related oropharyngeal cancer by the International Collaboration on Oropharyngeal cancer Network for Staging (ICON-S): a multicentre cohort study. *Lancet Oncol.* **17**, 440–451 (2016).
80. Haughey, B. H. *et al.* Pathology-based staging for HPV-positive squamous carcinoma of the oropharynx. *Oral Oncol.* **62**, 11–19 (2016).
81. Hardy, S. E., Allore, H. & Studenski, S. A. Missing data: a special challenge in aging research. *J. Am. Geriatr. Soc.* **57**, 722–9 (2009).

82. Pedersen, A. B. *et al.* Missing data and multiple imputation in clinical epidemiological research. *Clin. Epidemiol.* **9**, 157–166 (2017).
83. Sterne, J. A. C. *et al.* Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* **338**, b2393 (2009).
84. Swalin, A. How to Handle Missing Data. *Towards Data Science* (2018). Available at: https://towardsdatascience.com/how-to-handle-missing-data-8646b18db0d4. (Accessed: 19th July 2019)
85. Brotherston, D. C. *et al.* Patient preferences for oropharyngeal cancer treatment de-escalation. *Head Neck* **35**, 151–159 (2013).
86. Begum, S., Cao, D., Gillison, M., Zahurak, M. & Westra, W. H. Tissue distribution of human papillomavirus 16 DNA integration in patients with tonsillar carcinoma. *Clin. Cancer Res.* **11**, 5694–9 (2005).
87. Klingenberg, B. *et al.* p16INK4A overexpression is frequently detected in tumour-free tonsil tissue without association with HPV. *Histopathology* **56**, 957–967 (2010).
88. Hong, A. M. *et al.* Human papillomavirus, smoking status and outcomes in tonsillar squamous cell carcinoma. *Int. J. Cancer* **132**, 2748–2754 (2013).
89. Heath, S. *et al.* Clinically Significant Human Papilloma Virus in Squamous Cell Carcinoma of the Head and Neck in UK Practice. *Clin. Oncol.* **24**, e18–e23 (2012).
90. National Cancer Institute. Immunotherapy Drug Shows Promise against HPV-Related Cancers. *National Cancer Insitute* (2019). Available at: https://www.cancer.gov/news-events/cancer-currents-blog/2019/immunotherapy-y-trap-hpv-cancers. (Accessed: 16th September 2019)
91. Jung, A. C. *et al.* CD8-alpha T-cell infiltration in human papillomavirus-related oropharyngeal carcinoma correlates with improved patient prognosis. *Int. J. Cancer* **132**, E26–E36 (2013).
92. Wang, H.-F. *et al.* The Double-Edged Sword-How Human Papillomaviruses Interact With Immunity in Head and Neck Cancer. *Front. Immunol.* **10**, 653 (2019).
93. Cohen, E. E. W. *et al.* The Society for Immunotherapy of Cancer consensus statement on immunotherapy for the treatment of squamous cell carcinoma of the head and neck (HNSCC). *J. Immunother. Cancer* **7**, 184 (2019).
94. Dvořánková, B. *et al.* Cancer-associated fibroblasts are not formed from cancer cells by epithelial-to-mesenchymal transition in nu/nu mice. *Histochem. Cell Biol.* **143**,

463–469 (2015).

95. Álvarez-Teijeiro, S. *et al.* Factors Secreted by Cancer-Associated Fibroblasts that Sustain Cancer Stem Properties in Head and Neck Squamous Carcinoma Cells as Potential Therapeutic Targets. *Cancers (Basel).* **10**, 334 (2018).
96. Kolář, M. *et al.* Upregulation of IL-6, IL-8 and CXCL-1 production in dermal fibroblasts by normal/malignant epithelial cells *in vitro* : Immunohistochemical and transcriptomic analyses. *Biol. Cell* **104**, 738–751 (2012).
97. Gál, P. *et al.* How Signaling Molecules Regulate Tumor Microenvironment: Parallels to Wound Repair. *Molecules* **22**, 1818 (2017).
98. Plzák, J. *et al.* The Head and Neck Squamous Cell Carcinoma Microenvironment as a Potential Target for Cancer Therapy. *Cancers (Basel).* **11**, (2019).
99. Boscolo-Rizzo, P., Furlan, C., Lupato, V., Polesel, J. & Fratta, E. Novel insights into epigenetic drivers of oropharyngeal squamous cell carcinoma: role of HPV and lifestyle factors. *Clin. Epigenetics* **9**, 124 (2017).
100. Viet, C. T. *et al.* Decitabine Rescues Cisplatin Resistance in Head and Neck Squamous Cell Carcinoma. *PLoS One* **9**, e112880 (2014).
101. Diyabalanage, H. V. K., Granda, M. L. & Hooker, J. M. Combination therapy: Histone deacetylase inhibitors and platinum-based chemotherapeutics for cancer. *Cancer Lett.* **329**, 1–8 (2013).
102. Lydiatt, W. M. *et al.* Head and neck cancers-major changes in the American Joint Committee on cancer eighth edition cancer staging manual. *CA. Cancer J. Clin.* **67**, 122–137 (2017).
103. Ehrenstein, V., Nielsen, H., Pedersen, A. B., Johnsen, S. P. & Pedersen, L. Clinical epidemiology in the era of big data: new opportunities, familiar challenges. *Clin. Epidemiol.* **9**, 245–250 (2017).
104. Mooney, S. J., Westreich, D. J. & El-Sayed, A. M. Commentary: Epidemiology in the era of big data. *Epidemiology* **26**, 390–4 (2015).
105. Flanagan, C. & Gaffney, E. Cancer Research in Dublin Ireland | Biobank Ireland Trust. *Biobank Trust Ireland* (2019). Available at: http://www.biobankireland.com/. (Accessed: 30th April 2019)