Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Peter May; Prospective cohort study of hospital inpatients with advanced cancer: cost analysis of a palliative care consultation team intervention

Background: In a context of demographic change and shifting patterns of disease, models of health care provision for people with serious illness that improve patient and family outcomes in a financially sustainable manner are urgently sought. Research to date on palliative care programmes suggests improved outcomes at lower cost compared to conventional care, but the evidence base is small and further research on palliative care is needed.

Data: Demographic, socio-economic, clinical and hospital cost data collected as part of a prospective cohort study of patients admitted to four United States hospitals with an advanced cancer diagnosis between 2007 and 2011 (the 'Palliative Care for Cancer' ('PC4C') study). There are 1,023 patients in the final sample, 288 in the treatment (palliative care) group and 735 in the comparison group.

Methods: I performed a systematic literature review to provide a basis for my original research. Then using the 'PC4C' dataset I analysed the impact of the palliative care consultation team (PCCT) intervention on hospital costs and related outcomes. Propensity score matching is employed to control for observed confounding between treatment and comparison groups. Generalised linear models with a gamma distribution and a log link are used to estimate treatment effect on outcomes of interest. New methods that incorporate time-to-consult following hospital admission and define samples by patient complexity are applied.

Results: The primary analysis is presented in four sections with interconnected results.

I. An established practice in evaluations of PCCT impact on costs - using length of stay to control for unobserved heterogeneity - weakens the internal and external validity of treatment effect estimates. Methods incorporating time-to-consult following hospital admission deliver results that are more robust and more useful.
II. A palliative care consultation within six days of hospital admission is associated with lower cost of hospital care for patients with advanced cancer, and this effect is larger for earlier interventions. Cost-savings are achieved by reducing both length and intensity of hospital stay.
III. A palliative care consultation within 10 days of hospital admission is associated with lower cost of hospital care for patients with multimorbidity, and this effect is larger for more complex patients. These results are robust to the time-to-consult association: for any given sample defined by complexity, earlier treatment has a greater effect; for any given definition of the intervention according to timing, the treatment effect is largest for the most complex patients.
IV. For both patients who receive palliative care and those who do not, illness burden is a consistent driver of utilisation during hospitalisation.

Conclusion: Early palliative care should be made more widely available to patients admitted to United States hospitals with advanced cancer (given evidence-based assumptions about intervention efficacy). In a context where a minority of complex patients account for disproportionate levels of healthcare costs, early palliative care interventions may be an effective tool in curbing future cost growth. Controlling for length of stay in cost analysis of observational data is optimally avoided. Additional implications for practice, policy and future research are detailed.

Prospective cohort study of hospital inpatients with advanced cancer: cost analysis of a palliative care consultation team intervention

A dissertation submitted to the University of Dublin for the Degree of Doctor of Philosophy

Peter May
School of Medicine, Trinity College Dublin, 2015

Supervised by

Professor Charles Normand
Edward Kennedy Professor
Centre for Health Policy & Management
Trinity College Dublin
Ireland

Professor R. Sean Morrison, MD
Hermann Merkin Professor of Palliative Medicine
Icahn School of Medicine at Mount Sinai
New York, NY
United States

Declaration

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

I agree that the Library may lend or copy the thesis upon request. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

Signed: Peter May Date: 13/9/15

Peter May

Summary

The provision of care to people living and dying with serious chronic illness is a public health priority in most parts of the world. In a context of demographic change and shifting patterns of disease, the cost of care for this patient group is projected to grow substantially in the coming decades. Models of care that improve patient and family outcomes in a financially sustainable manner are urgently sought.

Palliative care programmes have expanded rapidly over the last 25 years to meet the needs of patients living and dying with serious illness and functional impairment. Research on these programmes suggests improved outcomes at lower cost compared to conventional care, but the evidence base is small and further research is needed.

This PhD thesis is concerned with cost analysis of the palliative care consultation team (PCCT) intervention for patients admitted to hospital with advanced cancer.

Prior to performing primary research I conduct a systematic literature review on the economic impact of PCCTs, identifying 12 prior studies that have found a consistent pattern of cost-saving effect from this intervention using a narrow (and in some cases flawed) methodological approach.

The findings of the literature review provide a framework for my original research. Using a prospective cohort study of patients admitted to one of four United States hospitals with an advanced cancer diagnosis (the 'Palliative Care for Cancer' ('PC4C') study), I analyse the impact of the PCCT intervention on hospital costs. There are 1,023 patients in the final sample, 288 in the treatment (palliative care) group and 735 in the comparison group. Propensity score matching is employed to control for observed confounding between groups.

First I demonstrate that an established practice in previous economic evaluations of PCCTs - using length of stay (LOS) to control for unobserved heterogeneity - weakens the internal and external validity of cost-effect estimates. Palliative care research, including the 'PC4C' study, is dominated by observational designs, making unobserved heterogeneity a substantial concern. However, LOS is connected to both the intervention (long LOS is an indicator of patient need) and the dependent variable (LOS

is strongly correlated with cost), meaning that controlling for LOS exacerbates endogeneity concerns and undermines evaluation of the causal relationship.

Second I identify time-to-consultation following hospital admission as having a potentially important association with PCCT impact on cost, and I develop new methods for incorporating intervention timing into analyses that minimise endogeneity concerns.

Using these new methods I report the following results:

- A palliative care consultation within six days of hospital admission is associated with lower cost of hospital care for patients with advanced cancer, and this cost-saving effect is larger for earlier interventions. Cost-savings are achieved by reducing both length and intensity of hospital stay. Failure to incorporate consult timing into evaluations of PCCTs may disguise important relationships;

- A palliative care consultation within 10 days of hospital admission is associated with lower cost of hospital care for patients with multimorbidity, and this cost-saving effect is larger for more complex patients. These results are robust to the time-to-consult association: for any given sample defined by complexity, earlier treatment has a greater effect; for any given definition of the intervention according to timing, the effect is largest for the most complex patients. In a context where a small number of complex patients account for disproportionate levels of healthcare costs, palliative care may be a more effective tool than previously realised in curbing future cost growth;

- For both patients who receive palliative care and those who do not, illness burden is a consistent driver of utilisation during hospitalisation.

Taken together, these three findings highlight substantial scope to reduce the large and growing cost of care to seriously-ill patients admitted to United States hospitals.

These findings are summarised in a series of conclusions and recommendations that, contingent on evidence-based assumptions about the patient and family outcomes of the intervention, can inform practice, policy and future studies in the ongoing search for effective and sustainable approaches to providing care for people with serious illness.

Acknowledgements

I thank Charles Normand and Sean Morrison for three years of insightful and stimulating supervision. Sean provided an exceptional opportunity for developing a PhD enquiry, giving generously of his time and knowledge. Charles and I shared countless thought-provoking discussions that equipped me to convert interesting early work into distinctive and meaningful findings. Charles and Sean share a passion for scientific research in general and palliative care research in particular that elevated my PhD experience beyond the professional to something altogether more satisfying and inspiring.

I was very fortunate to work on the 'Palliative Care for Cancer' study courtesy of Sean, Diane Meier and colleagues at Mount Sinai School of Medicine in New York. Particular gratitude is owed to Melissa Garrido, a constant source of careful and incisive feedback.

At times it feels like undertaking a PhD in health sciences is itself a high-risk behaviour but I have benefitted from the friendship of many fellow travellers at Trinity College Dublin, including Geralyn Hynes, Rebecca Moore, Lorna Roe, Marianne Griffiths and Conor Keegan. Working between two institutions three thousand miles apart introduced an additional layer of complexity that I negotiated only thanks to the expertise and patience of Sheena Cleary in Trinity and Katie Madden at Mount Sinai.

The PhD was funded by an Economics of Cancer Fellowship from the Health Research Board of Ireland and the National Cancer Institute in the United States. I am grateful to them, and to the funders of the 'Palliative Care for Cancer' study, for their assistance.

Finally I thank my parents, to whom this work is dedicated, and Nóra, whose support and camaraderie were essential to completing another strange and interesting chapter in our lives.

Dedication

This thesis is dedicated to my parents, who have always met my efforts in education and work with support and encouragement, and to Dónal, who arrived just in time to approve the final manuscript.

Table of Contents

List of Tables

List of Figures

Glossary

ABFMI	After-Death Bereaved Family Member Interview survey
ADL	Activities of daily living
AIC	Akaike information criterion
ATE	Average treatment effect
CI	Confidence interval
CMA	Cost-minimisation analysis
CMSAS	Condensed Memorial Symptom Assessment Scale
CNS	Central nervous system
CTM	Care Transition Management survey
DNC	Did not converge/compute
DRG	Diagnosis-related group
ECM	Exponential conditional model
EEE	Extended estimation equation
ESAS	Edmonton System Assessment Scale
FAMCARE	FAMily satisfaction with CARE for advanced cancer scale
FMM	Finite mixture model
GGM	Generalised gamma model
GI	Gastrointestinal
GLM	Generalised linear model
HCC	Hepatocellular carcinoma
H-L	Modified Hosmer-Lemeshow test
HRB	Health Research Board (Ireland)
ICU	Intensive care unit
IMRAD	Introduction, methods, results and discussion
IPTW	Inverse-probability treatment weights
IV	Instrumental variable
LOS	Length of stay
MAPE	Mean absolute prediction error
MC	Mount Carmel Health System, Ohio, US
MCW	Froedtert Hospital, Medical College of Wisconsin, Milwaukee, US

MSSM	Icahn School of Medicine at Mount Sinai, New York, US
NCCN	National Comprehensive Cancer Network (United States)
NCI	National Cancer Institute (United States)
NINR	National Institute of Nursing Research (United States)
NLS	Nonlinear least squares
OLS	Ordinary least squares
PC	Palliative care
PCCT	Palliative care consultation team
PC4C	'Palliative Care for Cancer' study
QALY	Quality-adjusted life year
RCT	Randomised controlled trial
RMSE	Root mean squared error
SD	Standard deviation
STROBE	STrengthening the Reporting of OBservational studies in Epidemiology
UB	Universal billing
UC	Usual care
UPMC	University of Pittsburgh Medical Center, Pennsylvania, US
US	United States
USD	United States dollars
VA	Veterans Administration
VCU	Virginia Commonwealth University, Richmond, US
WHO	World Health Organization

Chapter 1: Introduction

Abstract

This thesis focuses on cost analysis of a specialist-led interdisciplinary palliative care consultation team (PCCT) intervention for hospital inpatients with advanced cancer in the United States.

In this chapter I lay out the rationale for my PhD enquiry, detailing the following contextual issues:

- The well-established public health imperative to identify affordable and effective strategies for treating chronic diseases including cancer;
- The rapid growth of palliative care interventions for patients with serious illness in the United States and elsewhere over the last 25 years;
- The current landscape in evidence-based research of palliative care, including economic evaluation;
- A brief overview of my study data and thesis aim.

Established public health priority

Improving access to and quality of treatment for patients living and dying with chronic disease and functional impairment is an established public health priority in most parts of the world (Centers for Disease Prevention and Control, 2013, Busse et al., 2010, Global Burden of Disease Study Collaborators, 2015).

Care for chronic conditions, which are most prevalent among the elderly, accounts for large proportions of healthcare budgets but often results in inappropriate treatment, fragmented pathways, poor outcomes and unmet need (Meier, 2011, Lehnert et al., 2011). And the economic burden of chronic disease is projected to grow substantially in the first half of this century due to ageing populations and shifting patterns of disease, and the limited capacity of health systems originally designed to provide acute and episodic care (Schoen et al., 2009, Lehnert et al., 2011).

In the United States health system these challenges are compounded by long-standing use of high-intensity treatments and technologies. Studies on care for seriously-ill patients report evidence of high-cost futile care (Huynh et al., 2013); potentially inappropriate prescribing (Lindsay et al., 2014); perverse physician incentives towards high-intensity treatments (Malin et al., 2013); and mismatches between patients' preferences and clinicians' perceptions, particularly when patients do not want treatment (Downey et al., 2013). End-of-life care has become more aggressive and resource-intensive since the turn of the century (Teno et al., 2013). Increases in national healthcare expenditure are driven more by rising prices, particularly for drugs, medical devices and hospital care, than by growing demand due to population ageing (Moses et al., 2013).

Cancer specifically is the second most common cause of death in the United States (National Center for Health Statistics, 2015) and accounts for the highest levels of per-capita spending of major chronic conditions among adults (Soni, 2011). Cost growth for cancer treatment has been considerable and persistent, from $72 billion in 2004 to $125 billion in 2010 and projected to be $173 billion by 2020 (Mariotto et al., 2011), driven primarily by increases in spending on pharmaceuticals and imaging (Kelly and Smith, 2014). These trends have not translated into value with many new high-cost treatments resulting in modest increases in survival (Sorenson, 2012), consistent with a wider trend in the United States of high expenditure for poor health outcomes against international comparators (Organisation for Economic Co-operation and Development, 2014).

Over the next two decades, demographic ageing generally observable in high-income countries as well as these US-specific factors and widening insurance eligibility through the Affordable Care Act, will place growing financial pressure on the American healthcare system (Sisko et al., 2014). Strategies are urgently sought to curb cost growth while increasing access to appropriate treatment for patients with advanced cancer and other life-limiting illness (Smith and Hillner, 2011, Levit et al., 2013, Siu et al., 2009, Anderson, 2010).

The growth of palliative care

The World Health Organization (WHO) defines palliative care as follows:

> *"Palliative care is an approach that improves the quality of life of patients and their families facing the problems associated with life-threatening illness, through the prevention and relief of suffering by means of early identification and impeccable*

assessment and treatment of pain and other problems, physical, psychosocial and spiritual. Palliative care:

- *provides relief from pain and other distressing symptoms;*
- *affirms life and regards dying as a normal process;*
- *intends neither to hasten or postpone death;*
- *integrates the psychological and spiritual aspects of patient care;*
- *offers a support system to help patients live as actively as possible until death;*
- *offers a support system to help the family cope during the patients illness and in their own bereavement;*
- *uses a team approach to address the needs of patients and their families, including bereavement counselling, if indicated;*
- *will enhance quality of life, and may also positively influence the course of illness;*
- *is applicable early in the course of illness, in conjunction with other therapies that are intended to prolong life, such as chemotherapy or radiation therapy, and includes those investigations needed to better understand and manage distressing clinical complications."*

- World Health Organization (2015)

The origins of palliative care are in alleviating suffering at the end of life, particularly for cancer patients (Davies and Higginson, 2004a). But in response to rising prevalence of

chronicity and multimorbidity for patients who are not near end of life, care with processes and goals beyond aggressive curative intent are increasingly available for non-malignant diagnoses and/or for patients earlier in the care trajectory (Davies and Higginson, 2004b).

In the United States, as in other high-income countries, palliative care programmes have expanded rapidly over the last 25 years, emphasising pain and symptom management, communication about goals of care and transition management, practical help for family caregivers, and psychological and spiritual support (Hughes and Smith, 2014). Palliative care has also been recommended by the American Society of Clinical Oncology as standard care for any patient with cancer and/or high symptom burden (Smith et al., 2012b).

The intervention evaluated in this thesis is the palliative care consultation team (PCCT), a hospital-based multidisciplinary team comprising a physician, nurse and social worker with access to chaplaincy and psychiatry support. Acting at the invitation of a primary physician, the PCCT assists in the treatment of seriously-ill patients through clarifying diagnosis and treatment options, and helping patients and family members identify and select treatments that match their goals (American Academy of Hospice and Palliative Medicine et al., 2004).[1]

This model is the dominant form of palliative care provision in United States hospitals and has expanded substantially in recent years: in 2000, 24% of hospitals reported having a palliative care team; by 2013 this figure had risen to over two thirds of

[1] For a more detailed description of the intervention in my primary analysis, see **§3.1.4>Intervention**.

hospitals, including over 85% of mid- to large-sized hospitals (Goldsmith et al., 2008, Morrison, 2013a). Improving hospital care to patients with serious illness has long been a recognised priority since most patients prefer not to receive care in an inpatient setting, particularly near end of life (Flory et al., 2004, Gomes et al., 2012), and 35% of direct medical cancer costs in the United States are attributable to inpatient hospital stay (American Cancer Society, 2015).

Evidence and research in palliative care

Where palliative programmes have been evaluated, the evidence suggests they improve outcomes in pain and symptom management, anxiety, communication and decision-making, and patient and family satisfaction for patients with serious illness (Higginson and Evans, 2010, Dy et al., 2008, Dy et al., 2012). These reported improvements in patient experience appear to be derived through a combination of more appropriate decision-making, including expert pain and symptom management; treatment choices and transition plans consistent with patient preferences; and reductions in futile and inappropriate care (Meier, 2011, Zilberberg and Shorr, 2012, Zhang et al., 2009). Multiple studies have also shown that palliative care reduces the average cost of care (Smith et al., 2014).

As palliative care has greater relevance for cohorts not at end of life, earlier interventions are increasingly available 'upstream' in the care trajectory with observable benefits (Parikh et al., 2013). For example, evaluations of concurrent palliative care from point of diagnosis, or use of checklists at hospital admission to identify patients with palliative care needs, have suggested improved outcomes including survival (Temel et al., 2010, Adelson et al., 2013, Zimmermann et al., 2014, Bakitas et al.,

2014). Thus, palliative care has been shown to have an impact not only at end of life but at any time when decisions are made in the treatment of patients with serious illness (Smith et al., 2012b).

However, the size and quality of the evidence base to date is moderate for both patient and family outcomes, and economic analyses (Wheeler et al., 2012, Dy et al., 2012, Smith et al., 2014). This in part reflects low levels of research funding and activity relative to palliative care's role in contemporary healthcare provision (Gelfman et al., 2013, Wheeler et al., 2012, Sleeman et al., 2012). But there are also significant practical challenges in generating high-level evidence for seriously-ill populations. Studying patients at risk of rapid deterioration often leads to low enrolment and high attrition (Fu et al., 2013), and requires the measurement of complex and intangible outcomes (Evans et al., 2013). Randomised controlled trials (RCTs) of palliative care face challenges of feasibility and ethics, often requiring researchers to pursue quasi-experimental designs with observational data (Aldridge Carlson, 2013, Ewing et al., 2004, Penrod et al., 2008, Marcus and Gibbons, 2001).[2]

Continued expansion of palliative care access in the United States is a recognised priority both to address insufficient provision in underserved regions and to develop national capacity as need and awareness of services continues to grow (Morrison, 2013a). However, effective expansion requires further rigorous, evidence-based research to appraise systematically the outcomes of programmes to date, to determine

---

[2] For a more detailed discussion of research design issues for seriously-ill patients, and how I have tried to maximise the quality of evidence generated by my primary analysis, see **§3.1.3** and **§3.2.1**.

the processes and structures underpinning those outcomes, and to identify opportunities to upscale effective models for an ageing population (Goldstein and Morrison, 2005, Kumar et al., 2012, Higginson et al., 2013, Morrison, 2013b).

One essential component of such a research agenda is economic evaluation. Payers and policymakers require evidence to validate current palliative care provision, to explore ways that this could be made more cost-effective, and to assess the case for new programmes (Meier and Beresford, 2009, Harding et al., 2009). Healthcare budget projections for patients living with chronic illness in general, and those with cancer in the United States in particular, emphasise the importance of identifying healthcare responses that can curb cost growth without compromising quality or access (Smith and Hillner, 2011).

Despite the acknowledged significance of economic analyses in evaluating and informing care provision, the economic literature on palliative care programmes is small and disparate, reflecting the complexity of palliative care assessment and the multiplicity of treatment settings, patient diagnoses and levels of severity, and levels of staff specialism (Gomes et al., 2009, Smith et al., 2014).[3]

This PhD thesis

The overall aim of this thesis is to contribute substantively to knowledge on the impact of the PCCT intervention. Specifically, the thesis constitutes a cost analysis[4]: it evaluates how PCCTs are associated with cost of hospitalisation on the assumption that

---

[3] The economic literature on PCCTs is reviewed systematically in **§2**, including a discussion of the limitations of this literature in the context of the current evidence base on the economics of palliative care.

[4] Cost analysis methodology is placed in wider context in **§2.4>Forms of economic evaluation** and subsequent sub-sections.

outcomes for patients who received the intervention were at least as good as for patients who did not receive the intervention. High-quality evidence on the hospital cost effects of the PCCT intervention is essential to the wider research and policy context of understanding the best allocation of resources in providing care to people with serious and life-limiting illness across all settings.

The original research in this thesis forms part of a larger study on the PCCT intervention for patients with advanced cancer in United States hospitals. The study, the 'Palliative Care for Cancer' ('PC4C') study, was conducted to evaluate PCCT effect on patient and family outcomes, goals and processes of care, and in-hospital utilisation.

To identify priorities for my original research I conducted a systematic literature review of previous relevant studies. This work is presented in **§2 Literature Review.** I gained access to the 'PC4C' dataset during an extended period of mentored research at the Icahn School of Medicine at Mount Sinai in New York, consistent with my PhD fellowship funding, which aimed to provide health economists with mentorship and expertise from United States health services researchers (Health Research Board, 2011). For a detailed summary of the 'PC4C' study rationale and methods, and my primary analysis within this study, see **§3 Materials and Methods**.

My original research addresses the objectives identified in **§2** and using methods detailed in **§3** to evaluate PCCT impact on hospital costs, utilisation patterns and length of stay. This research, which has practice and policy implications, as well as highlighting important methodological issues that can strengthen future research in this field, is presented in **§4 Results.** The implications of my findings for practice, policy and

research, including the potential to incorporate my results with other outputs from the 'PC4C' study, are detailed in **§5 Conclusions and recommendations**.

As such this thesis presents original research that contributes evidence on a recognised public health and policy priority worldwide. The study data were collected and analysed using a rigorous and robust design, and so the results constitute a strong contribution to the evidence base. My results will have greatest relevance for the United States healthcare system and setting, but they may also be applicable to other countries. Similarly the analyses are generalisable to patients with cancer diagnoses, but they can inform similar studies of interventions for non-malignancy in future.

Chapter 2: Literature Review

Abstract

This chapter presents the findings of a systematic literature review of the economic impact of hospital palliative care consultation teams (PCCTs). An earlier, edited version of this review has been published in the *Journal of Palliative Medicine* (May et al., 2014).

Background: The economic evidence base on palliative care is small and disparate, and where systematic reviews have analysed economic impact these have tended to report across different settings, diagnoses, levels of specialism and national systems, highlighting patterns without focusing on specific programmes or models of care. A review was undertaken to collect systematically the economic evidence on PCCT interventions specifically, to appraise critically this evidence, and to identify approaches and priorities for my PhD enquiry.

Methods: Economic studies of PCCT interventions were identified using two methodological approaches: systematic meta-review and systematic database review. In the context of multiple relevant prior reviews I identified studies through already-published systematic reviews. For time periods not covered by already-published reviews, I undertook a systematic search of the PubMed, CINAHL, EconLit and AMED databases. Studies were eligible for inclusion if they represented a credible economic evaluation of a specialist palliative care consultation team intervention for adult hospital inpatients versus a usual care comparator.

Results: Twelve studies were included in the review; 11 observational cohort studies and one randomised controlled trial, all from the United States. Studies report a

consistent pattern of cost-saving effect from PCCT interventions but these results were derived using a narrow (and in some cases flawed) methodological approach. The underlying source of cost-savings is also not well established.

Conclusion: PCCT interventions appear to reduce hospital costs for patients with serious illness but these results must be considered carefully and in the context of the methods they employ. Research priorities for my primary PhD research from a hospital costs perspective are identified, as well as areas for future studies beyond the hospital costs perspective.

2.1 Introduction

In the previous chapter I introduced this PhD thesis, outlining the context in which the research was undertaken.

Palliative care programmes have expanded rapidly in recent years to treat patients living and dying with serious illness, but there remains evidence of unmet need and poor quality care for this group. Demand for palliative care services will continue to grow in the context of demographic change, and health system pressures are compounded in the United States by persistent cost growth and high-intensity utilisation. Research activity in palliative care is low proportionate to its significance in healthcare provision, amplifying the need for rigorous, evidence-based knowledge on these programmes. Performing research in this area is complicated by ethical and practical considerations that often preclude randomised controlled trial (RCT) design, and result in enrolment and attrition challenges.

In particular the economic evidence base on palliative care is small and disparate, and where systematic reviews have analysed economic impact these have tended to report across different settings, diagnoses, levels of specialism and national systems, highlighting patterns without focusing on specific programmes or models of care.

Consequently there has been no economic review focused on specialist hospital inpatient consultation teams (PCCTs), the dominant model of provision in the United States hospital setting and the intervention on which this PhD research is focused. A review was therefore undertaken to collect systematically the economic evidence on this model specifically, to appraise critically the evidence, and to identify approaches and priorities for my PhD enquiry.

2.2 Methods

Identifying studies for consideration was performed using two methodological approaches: systematic meta-review and systematic database review.

First, studies were identified by systematic meta-review ('a review of existing reviews'); instead of collating studies from databases, I identified studies through already-published reviews. In the context of multiple relevant prior reviews the early economic evidence on palliative care programmes had broadly been assembled. What the prior reviews had not provided was a detailed examination of the evidence on any specific model of care delivery, or a critical assessment of that evidence. In systematically reviewing relevant previous reviews using clear criteria focusing specifically on economic evaluation of the PCCT model, I am able to assemble the published evidence on my area of interest for time periods covered by prior reviews.

For time periods not covered by already-published reviews, a systematic database search was undertaken.

Review selection

Reviews were identified by systematic searches on the PubMed, CINAHL, EconLit and AMED databases.[5] Relevant search terms for palliative care, review and economics (e.g. palliative, hospice; review, systematic; economic*, cost*) were combined to search titles, abstract and subject headings published by 28th February, 2015. See Appendix to §2.2 for details.

---

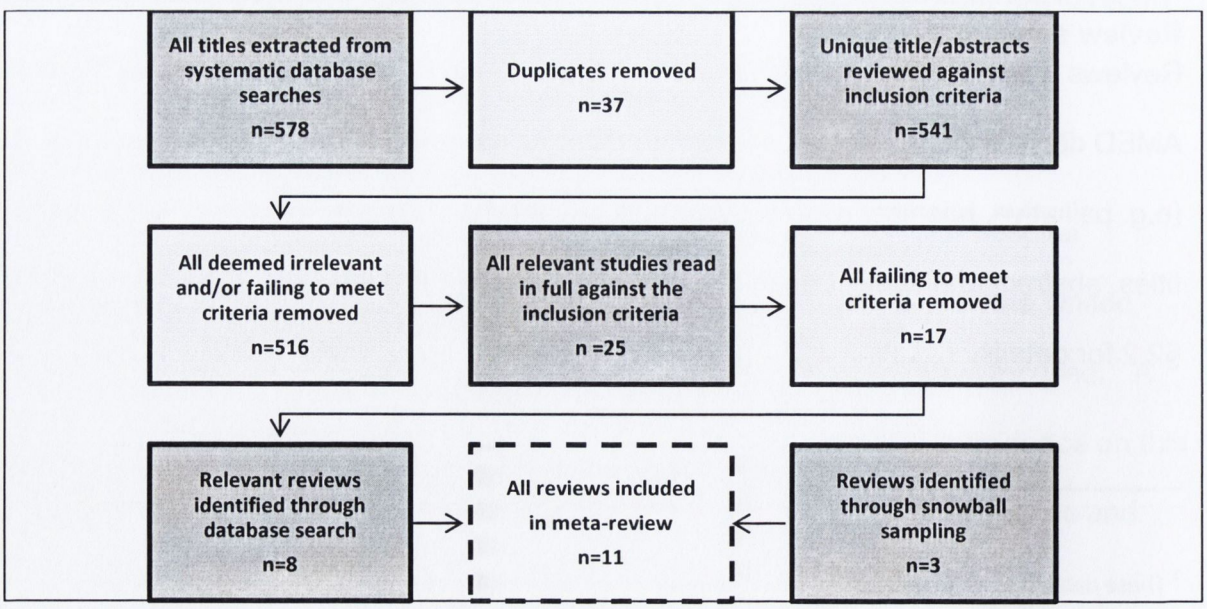
[5] These databases were selected in discussion with my supervisors as covering the key domains of relevant research in palliative care and health economics: medicine, nursing, economics and allied healthcare.

All 578 studies returned by the database search were considered for the meta-review. A review was included in the meta-review only if it was (i) a literature review reporting (ii) a systematic search strategy (iii) examining (but not necessarily limited to) inpatient hospital palliative care programmes (iv) treating adult patients, and (v) identified outcomes of interest as including (but not necessarily limited to) economic analysis. Only English-language journal articles were considered.

I reviewed all unique titles/abstracts against the inclusion criteria; all deemed irrelevant or not meeting the criteria were removed, all others were read in full against the inclusion criteria. Where there was uncertainty about a systematic review's suitability or relevance this was discussed with my supervisors.

All reviews that were included in the meta-review were hand-searched for additional relevant reviews ('snowball sampling'). A full breakdown of the appraisal process is provided in Figure 2.2.1.

Figure 2.2.1 Identification of reviews for meta-review (see also Appendix to §2.2)



The meta-review returned eight published reviews with a relevant focus (Douglas et al., 2003, Thomas et al., 2006, Zimmermann et al., 2008, Smith and Cassel, 2009, Simoens et al., 2010, El-Jawahri et al., 2011, Harris and Murray, 2013, Smith et al., 2014). Three additional reviews were identified by snowball sampling (Higginson et al., 2002, Higginson et al., 2003, Higginson and Evans, 2010), making 11 in total. These are summarised in Table 2.2.1.

Table 2.2.1 Systematic reviews of palliative care programme evaluations

Lead author, year, country	Number of included papers
Higginson, 2002, UK	13
Douglas, 2003, UK	17
Higginson, 2003, UK	15*
Thomas, 2006, CAN	30
Zimmermann, 2008, US	7**
Smith, 2009, US***	21***
Higginson, 2010, UK	59
Simoens, 2010, BEL	15
El-Jawahri, 2011, US	22
Harris, 2013, UK	12
Smith, 2014, IRL	46

* 65 included in total; 15 reported separately in economic analysis.
** 22 included in total; seven reported separately in economic analysis.
*** Not reported as a systematic review. 21 papers discussed in cost analysis.

Of these, 10 systematic reviews variously (a) focused primarily or exclusively on economic factors in palliative care provision (Douglas et al., 2003, Simoens et al., 2010, Harris and Murray, 2013, Smith et al., 2014); (b) reported economic impact as one dependent variable of interest separately alongside clinical and other factors (Higginson

et al., 2003, Thomas et al., 2006, Zimmermann et al., 2008); or (c) evaluated palliative care services without particular emphasis on economic considerations (Higginson et al., 2002, Higginson and Evans, 2010, El-Jawahri et al., 2011). An additional review, not reporting a systematic search strategy but with a highly relevant focus, was included in the meta-review also (Smith and Cassel, 2009).

The reviews had a balance between different systems and perspectives in high-income countries with five written by teams based in the United Kingdom (Higginson et al., 2002, Douglas et al., 2003, Higginson et al., 2003, Higginson and Evans, 2010, Harris and Murray, 2013), three in the United States (Zimmermann et al., 2008, Smith and Cassel, 2009, El-Jawahri et al., 2011), and one each from Belgium (Simoens et al., 2010), Canada (Thomas et al., 2006) and Ireland (Smith et al., 2014).

The timeframe of these 11 reviews provided full coverage of the relevant published literature to the end of 2011. To supplement these findings and identify papers published since 2011, systematic searches were performed on the PubMed, CINAHL, EconLit and AMED databases. Key search terms from the clinical and economic domains (e.g. palliative, hospice; economic*, cost*) were combined to search titles, abstracts and subject headings from 1st January, 2012 to 28th February, 2015. See Appendix to §2.3 for details.

Study selection

All 257 studies included in any of the 11 identified reviews and all 606 studies returned by systematic database search were considered for inclusion.

I reviewed all unique titles/abstracts against the inclusion criteria; all deemed irrelevant or not meeting the criteria were removed, all others were read in full against the inclusion criteria. Where there was uncertainty about a paper's suitability for inclusion this was discussed with my supervisors.

A study was included in the review only if it contained (i) a credible economic evaluation of (ii) a specialist-led multidisciplinary palliative care consultation team to (iii) adult patients in (iv) the hospital inpatient setting, (v) measuring and comparing the costs and/or cost-effectiveness of this intervention (vi) against a usual care (UC) comparator. Studies focused on patients with dementia were excluded since this patient group was not eligible for the study in my own primary research.[6] Only English-language journal articles were considered.

The rationale for these criteria were adapted from the 'gold standard' guidelines for health economic evaluation (Drummond and Jefferson, 1996). Drummond and Jefferson's full checklist for economic evaluations is far greater; these components were identified as constituting a fair 'bare minimum' threshold in a field where economic evaluation is at an early stage.

---

[6] For full details on the eligibility criteria for my primary analysis, see **§3.1.5**.

2.3 Results

Summary

A full breakdown of the appraisal processes are provided in Figure 2.3.1.

Figure 2.3.1 Identification of studies for literature review

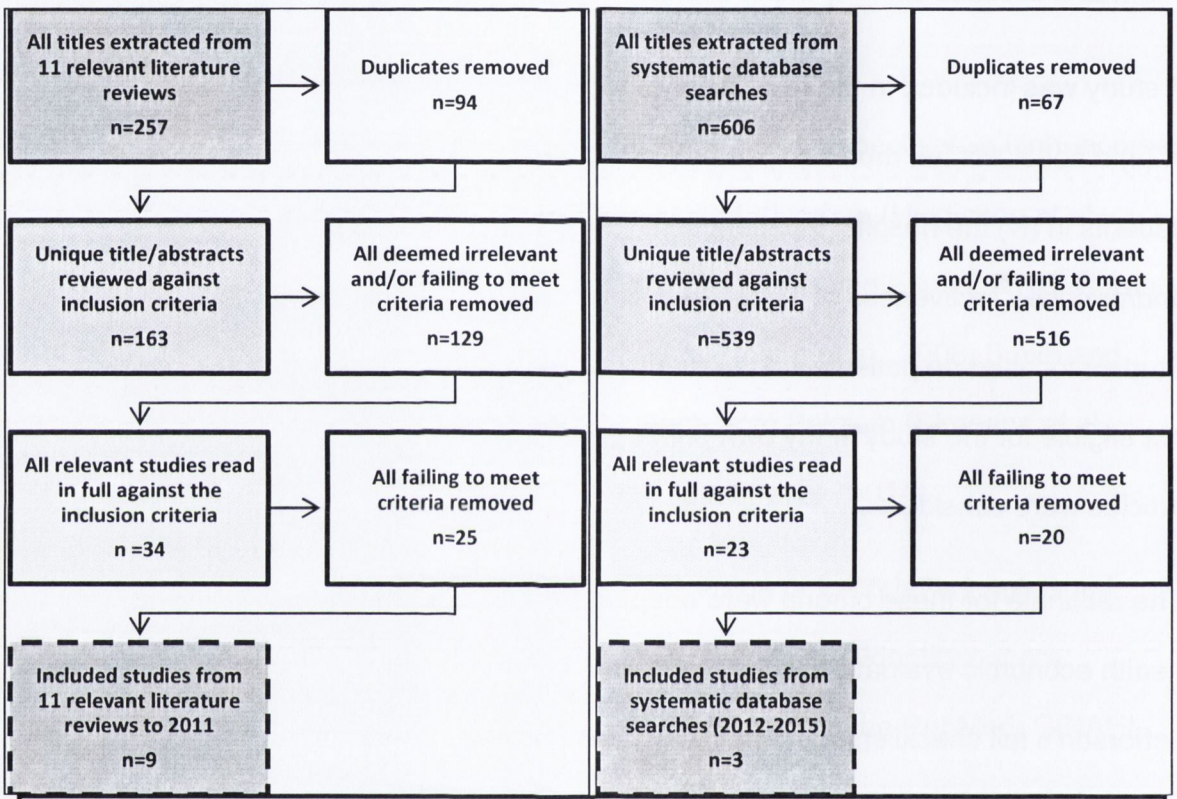


Fig. 2.3.1a Meta-review: to 31/12/2011; see also Table 2.2.1.

Fig.2.3.1b Database search: 1/1/2012 to 28/02/2015; see also Appendix to §2.3.

Nine evaluations identified through the meta-review met the inclusion criteria (Cowan, 2004, Penrod et al., 2006, Ciemins et al., 2007, Bendaly et al., 2008, Gade et al., 2008, Hanson et al., 2008, Morrison et al., 2008, Penrod et al., 2010, Morrison et al., 2011a). Three evaluations returned by the database search met the inclusion criteria (Starks et al., 2013, Whitford et al., 2014, McCarthy et al., 2015). Snowball sampling was performed on all 12 eligible studies for further papers but no additional eligible papers were identified. The 12 studies are summarised in Table 2.3.1.

Table 2.3.1 Economic evaluations of palliative care consultation team interventions for hospital inpatients

Study	Hospital(s) --- Team Label	Design *Matching strategy*	Sample size[8]	Principal Diagnosis[9]	Primary dependent variable[10]	Key findings
Cowan, 2004 **(#1)**	Hospital not categorised by authors --- Advanced Illness Assistance Team (AIA)	Cohort *Matching by DRG*	164 PC 152 UC	Cancer 27% Neurologic 18% Pulmonary 17% CV 12% Organ failure 7% GI 7% Chronic pain 6% Infection 4% Other 2%	Hospital charges Comparison of mean charges for PC group & matched UC comparators	· Lower (~7%) mean daily charges for PC than UC (p=0.006) · For patients with LOS>7 days, PC reduces LOS · Lower LOS for some PC groups but not statistically quantified.
Penrod, 2006 **(#2)**	Two Veterans Administration (VA) facilities --- Palliative Care Consultation Team (PCCT)	Cohort *Propensity scores*	82 PC 232 UC	Cancer 50% Infectious disease 10% CV 7% Pulmonary 10% GI 7% GUR 4% Other 12%	Direct daily cost Comparison of mean costs for weighted PC & UC groups from GLM (gamma, log) regressions	· PC patients 42% less likely to be admitted to ICU · Lower (~22%) daily direct costs for PC than UC (p<0.0001) · Laboratory & radiology also lower; no diff. for pharmacy · LOS reported as descriptive statistic; no sig. diff. between PC & UC (p=0.44)

[8] PC=number of palliative care patients (intervention group); UC=number of usual care patients (comparison group).
[9] For PC group at time of consultation.
[10] Primary dependent variable that is also of interest for this literature review. In some cases studies examined additional outcomes/research questions not relevant to this review which are not reported here.

Ciemins, 2007 **(#3)**	Large, private, not-for-profit medical centre --- Palliative Care Consultation Service (PCCS)	Cohort *Matching by DRG*	27 PC 128 UC	Cancer 100%	Total daily cost Comparison of mean costs for PC group & matched UC comparators	· Lower (~13%[11]) mean daily costs for PC than UC ($p<0.01$) · Lower (~16%) mean total costs for PC than UC ($p<0.0001$) · Lower LOS but not statistically quantified.
Bendaly, 2008 **(#4)**	Public hospital --- Palliative Care consultation (PCc)	Cohort *None reported*	61 PC 55 UC	Pulmonary disorders and/or MV 30% CV 23% Neoplasms 16% Infections with/ w/o sepsis 15% Other 16%	Total charges Comparison of median total charges for PC & UC groups	· Lower (~16%) median total charges for PC than UC ($p<0.001$) · LOS reported as descriptive statistic; no sig. diff. between PC & UC ($p=0.57$)
Gade, 2008 **(#5)**	Three Managed Care Organization hospitals --- Interdisciplinary Palliative Care Service (IPCS)	RCT	275 PC 237 UC	Cancer 27% CHF 9% MI 1% Other HD 3% COPD 13% Other PD 1% Renal 4% Organ failure 12% Stroke 9% Dementia 3%	Total health service costs for 6 months post-discharge Comparison of means for PC & UC groups	· Lower (~32%) total mean health costs for PC than UC ($p<0.001$) · Lower (~23%) total mean health costs for PC than UC once IPCS staffing accounted for · No diff. in physical, emotional symptoms · Improved satisfaction

[11] There is inconsistency in reporting of cost difference as a percentage. Ciemins et al. use the cost of palliative care as the base cost in calculations, i.e. [$\%\Delta C = ((C_{PC} - C_{UC})/C_{PC}) \times 100$] while others (e.g. Hanson et al.) use the cost of usual care, i.e. [$\%\Delta C = ((C_{PC} - C_{UC})/C_{UC}) \times 100$]. In this table and throughout the text all $\%\Delta C$ have been calculated using the latter method.

Hanson, 2008 **(#6)**	Tertiary academic medical centre --- Palliative Care Consultation Service (PCCS)	Cohort *Matching by DRG*	104 PC 1,813 UC	Cancer 61% Cardiopulmonary diseases 11% Neurologic diseases 5% Hepatic/renal failure 4% Acute infections 14%[12]	Direct daily cost Comparison of mean costs for PC group & matched UC comparators	· No difference in total variable costs (p=0.78) · Lower (~10%) daily variable costs for PC than UC (p=0.03) · Larger proportional cost savings per day for PC where LOS is greater · No sig. diff. in LOS (p=0.11)
Morrison, 2008 **(#7)**	Five community hospitals & three academic medical centres --- Palliative Care Consultation Team (PCCT)	Cohort *Propensity scores*	*Live discharge* 2,630 PC 18,427 UC *Hospital deaths* 2,278 PC 2,124 UC	*Live discharge* Cancer 29% Infection 4% CV 19% Pulmonary 15% GI 7% GUR 4% Other 22% *Hospital death* Cancer 19% Infection 11% CV 24% Pulmonary 18% GI 9% GUR 4% Other 14%	Total direct cost Comparison of mean costs for weighted PC & UC groups from GLM (gamma, log) regressions	*Live discharges* · Lower total costs (~14%; p=0.02), total costs per day(~19%; p<0.001), total direct costs(~15%; p=0.004), direct costs per day(~21%; p<0.001) for PC than UC · No sig. diff. in LOS (p=0.12) *Hospital deaths* · Lower total costs (~18%; p=0.001), total costs per day(~22%; p<0.001), total direct costs(~22%; p=0.003), direct costs per day(~25%; p<0.001) for PC than UC · No sig. diff. in LOS (p=0.40)

[12] Primary diagnoses for all PC patients in overall sample (n=304); authors do not report corresponding figures for sub-sample in economic analysis (n=104).

Penrod, 2010 **(#8)**	Five Veterans Administration (VA) facilities --- Palliative Care Consultation Team (PCCT)	Cohort *Propensity scores plus instrumental variable (IV)*	606 PC 2,715 UC	Cancer 62%[13] COPD 36% CHF 28% HIV/AIDS 3%	Daily direct cost Cost effects estimated using gamma IV regression to compare weighted PC & UC groups	· PC patients 44% less likely to be admitted to ICU · Lower[14] daily direct costs for PC than UC ($p<0.0001$) · Pharmacy, laboratory & nursing also lower; no diff. for radiology
Morrison, 2011a **(#9)**	A community hospital, two academic medical centres, and a safety-net hospital --- Palliative Care Consultation Team (PCCT)	Cohort *Propensity scores*	*Live discharge* 290 PC 1,427 UC *Hospital deaths* 185 PC 149 UC	Cancer 58% AIDS 2% CHF 12% COPD 2% Advanced liver disease 19% Prolonged ICU stay 6.9%	Total direct cost Cost effects estimated using GLM (gamma, log) regression of weighted samples See also: Morrison et al. (2011b)	*Live discharges* · Lower total costs (~11%; $p<0.05$), total costs per day(~18%; $p<0.001$) for PC than UC · Lower laboratory (~16%) & imaging (~13%) costs though not stat. sig. · Slightly higher ICU LOS but significantly lower (~42%; $p<0.001$) ICU cost per admission for PC than UC *Hospital deaths* · Lower total costs (~11%; $p<0.05$), total costs per day(~9%; $p<0.01$) for PC than UC · Lower pharmacy (~21%; $p<0.001$) for PC than UC ; no difference for laboratory or imaging · Lower ICU LOS (10.2 days against 13.8; $p<0.01$) for PC than UC; no significant cost difference

[13] Patients could have more than one advanced disease diagnosis; therefore does not add up to 100%.
[14] No UC cost given so not possible to calculate proportional saving.

Starks, 2013 **(#10)**	Two academic medical centres --- Palliative Care Consultation	Cohort *Propensity scores*	*Live discharge* 1,115 PC 1,111UC *Hospital deaths* 679 PC 700UC	*All PC*[15] Cancer 38% COPD 16% Cerebrovasc. 16% CHF 15% Renal 10% Acute myocardial 6% AIDS 3%	Total costs Cost effects for all PC patients using GLM (gamma, log) compared to matched comparators	*Live discharges* · Lower total costs (~19%; p<0.05) for PC than UC for patients with LOS<=7; no difference in either direction for longer LOS. *Hospital deaths* · No difference in either direction. Note: See **§2.4>Sample definition by LOS for** important commentary on these methods and results.
Whitford, 2014 **(#11)**	Integrated medical centre comprising two hospitals --- Palliative Care Consult Service (PCCS)	Cohort *Propensity scores*	*Live discharge* 1,177 PC 3,531 UC *Hospital deaths* 300 PC 900 UC	*Live discharge* Infectious 8% Neoplasm 21% Endocrine 2% Nervous 3% Circulatory 30% Respiratory 16% Digestive 8% Muscloskel. 9% Other 3% *Hospital death* Infectious 4% Neoplasm 32% Endocrine 3% Nervous 4% Circulatory 23% Respiratory 10% Digestive 7% Muscloskel. 3% Other 14%	Total direct costs Cost effects estimated using GLM (gamma, log) regression of weighted sample	*Live discharges* · Lower total costs (~5%; p<0.05) for PC than UC · Lower procedure costs; higher evaluation, imaging, pharmacy costs for PC than UC (no % or p value given) *Hospital deaths* · Lower total costs (~31%; p<0.05) for PC than UC · Lower procedure, evaluation, imaging, laboratory, pharmacy costs for PC than UC (no % or p value given) Note: See **§2.4>Sample definition by LOS** for important commentary on these methods and results.

[15] Data reporting precludes distinguishing disease profiles by discharge status. Frequencies reported sum to 103%.

McCarthy, 2015 **(#12)**	Four community hospitals and one academic hospital --- Palliative Care Consultation	Cohort *Propensity scores*	*Live discharge* 1816PC 33574UC *Hospital deaths* 572PC 1246UC	*Live discharge* Circulatory 1% Cancer 19% CV 15% Endocrine 3% GI 8% GUR 6% Infection 20% Injury 7% Muscloskel. 2% Nervous syst 3% Other 3% *Hospital death* Circulatory 1% Cancer 17% CV 14% Endocrine 1% GI 8% GUR 2% Infection 27% Injury 11% Muscloskel. 1% Nervous syst 2% Other 3%	Total direct costs Cost effects estimated using GLM (gamma, log) regression of weighted sample	*Live discharges* · Lower total costs (~12%; p<0.05) for PC than UC if consult is received within 10 days of admission. · Lower total costs (~14%; p<0.05) for PC than UC for patients with cancer. · Variation in cost effect by hospital site. *Hospital deaths* · Lower total costs (~10%; p<0.05) for PC than UC · Lower total costs (~20%; p<0.05) for PC than UC for patients with cancer. · Variation in cost effect by hospital site.

LOS: Length of stay.
CV: Cardiovascular; MV, mechanical ventilation; COPD: Chronic obstructive pulmonary disease; CHF: Chronic heart failure; HD: Heart disease; MI: Myocardial infarction; GI: Gastrointestinal; GUR: Genitourinary.

Design

All 12 studies are from the United States. Observational designs dominate with one RCT (#5) and 11 cohort studies.[16]

Among observational studies there is a wide variation in size with five having intervention groups of between 27 and 164 patients (#1, #2, #3, #4, #6), and one study 4,908 intervention patients (#7). There is a clear association between study size and publication date; the earliest studies are the smallest, more recent studies are larger.

Five studies treat patients discharged alive and patients who died during hospitalisation as fundamentally distinct patient groups, dividing their samples accordingly (#7, #9, #10, #11, #12).

Bias

Of the 11 studies to use observational data, the earliest studies report small sample sizes and rudimentary matching techniques (#1, #3, #6) or, in one case, no matching technique (#4).

More recent studies are generally larger and use the more sophisticated propensity scores technique (#2, #7, #8, #9, #10, #11, #12).[17]

One study used an instrumental variable to control for unobserved confounding alongside propensity scoring (#8). This can therefore be considered the most robust cohort study in the review for minimising bias.

---

[16] My primary research in this thesis employs data from a cohort study. For a more detailed description of why an observational design is often considered more appropriate for these interventions, see **§3.1.3**.

[17] My primary research employs propensity score matching. For a more detailed description of propensity scoring in theory and practice, see **§3.2.1**.

Setting

The range of settings covered by included studies demonstrates the hospital heterogeneity of the United States healthcare system.[18]

Academic hospitals are the most common, reflecting close relationships between these institutions and research activity and infrastructure (#6, #7, #9, #10, #12). Three of these studies are multi-site including community and academic hospitals side by side (#7, #9, #12). Two studies examine Veterans Administration (VA) facilities (#2, #8). Others variously examine a large, private, not-for-profit medical centre (#3), a public hospital (#4), three Managed Care Organization hospitals (#5), and an integrated medical centre comprising two hospitals (#11). One hospital is not categorised by authors (#1).

Intervention

Eligibility criteria for the review specified an intervention comprising a team led by a specialist physician. Clinical practice guidelines specify a core team of physician, nurse and social worker, with additional support from chaplaincy, psychiatry and other professions as required (American Academy of Hospice and Palliative Medicine et al., 2004).[19] The composition of teams included in the review is broadly consistent with that definition but with some differences.

Six teams were described as comprising at least a physician, a nurse, a social worker and a chaplain (#1 to #6); in some cases these were also described as including a psychologist and/or an oncology nurse specialist and/or nursing assistants.

---

[18] My primary research in this thesis also draws on a diversity of hospital settings. A more detailed description of settings in my primary research is provided in **§3.1.4**.

[19] The PCCT intervention in my primary research is consistent with these guidelines. See **§3.1.4**.

Of the other six, two were multi-site studies where all teams included a physician and nurse but not all included a social worker and chaplain (#7, #9). Two newly-implemented services altered composition during the respective study, one initially comprising a physician and a chaplain before incorporating a nurse (#10); another adding a chaplain to supplement a physician/nurse team (#11). The specific composition of the team was not described in one study (#8), but is indicated to be consistent with a prior related study (#2).

The most inconsistent (both within the study and in comparison to other studies) team compositions were evident in the five-site Study #12; three teams were led by physicians and two by nurse practitioners with variable social worker and chaplain input. Consequently, some sites in their study are not strictly eligible for this review but it was included since some sites in the study are relevant and site-specific results are reported (with no consistent pattern of difference between hospitals by team composition).

A further source of potential variability between interventions is the process and nature of referral. All consultation teams saw patients following referral from another team in the hospital but it is not possible to ascertain how comparable these processes were.

Population

All studies addressed programmes that treated a range of diagnoses, although following matching for economic evaluation one study was restricted to patients with cancer (#3).[20] All study populations have serious illness but not all are at end of life;

---

[20] In my primary research, the patient sample is restricted to cancer patients only since 'Palliative Care for Patients with Cancer' is the focus of the study. For more details of the eligibility for 'PC4C', see **§3.1.5**.

the survival rate during the study period for the intervention group varies between 0% and 80%, with a median of 70%.

In more recent studies, there has developed a tendency to define the sample by length of stay (LOS). Studies #7 and #12 analyse only patients who died or were discharged within seven to 30 days of admission; Studies #9 and #11 set equivalent limits of 6-44 days and 4-90 days respectively. Study #10 stratifies (but does not exclude) their patients by LOS.[21]

Patients who stayed less than 48 hours are routinely excluded from analysis since there is insufficient time for an intervention to have a measurable effect.

Perspective

The observational studies restrict their perspective to the hospital and do not evaluate post-discharge costs, or patient or caregiver outcomes.[22] The RCT (#5) analyses total healthcare costs for six months post-discharge as well as some patient outcome measures, without quantifying the relationship between the two.

Dependent variable

Different studies specify costs as a dependent variable in different ways.

Ten studies examined direct and/or total costs. Direct costs are those attributable to a specific utilisation during hospital stay: services, procedures and relevant staffing costs; indirect costs are the fixed costs of running the hospital, proportionately calculated for each patient.

---

[21] For a critical assessment of these methods, see **§2.4>Sample definition by LOS**.

[22] Studies #3 and #6 evaluated clinical outcomes in separate analyses. My primary research is also concerned with costs from the hospital perspective. See below **2.4 Discussion>Hospital costs perspective** and **§3.1.6**.

Direct costs are therefore not equivalent to variable costs. Within direct costs there are variable and fixed costs. Variable direct costs are those that are wholly dictated by treatment of the specific patient, such as medical supplies, pharmaceuticals, and imaging and laboratory expenditures. Fixed direct costs are those that do not vary with a specific patient's utilisation but can nonetheless be identified with the treatment of that patient; examples therefore include staff salaries, and departmental equipment expenses proportionately calculated for each patient. Indirect costs are those that can neither be identified with a specific patient nor a specific part of the hospital costs that are necessarily incurred in the running of the hospital and borne by all patients and departments. Examples include the hospital organisation's personnel and information technology costs, and general building and maintenance expenses (Taheri et al., 2000).

Total costs are the sum of direct and indirect costs. Those costs can then be reported *in toto* (costs for the entirety of hospital stay) and *per diem* (overall hospital costs divided by LOS).[23]

Five studies primarily examine costs *in toto*, either direct (#11, #12), total (#9, #10), or both (#7). Studies #7 and #9 additionally consider *per diem* costs.

Four studies examine only *per diem* costs, direct (#2, #6, #8) and *in toto* (#3). Two studies look not at costs but at charges from the hospital perspective (#1, #4). One examines total health service costs for 6 months post-discharge (#5).

---

[23] For additional consideration of direct and indirect costs, see **§2.4 Discussion>Total versus direct cost as a dependent variable**. For my primary research I identified direct costs as the primary dependent variable; for details see **§3.1.6**.

Cost-effect estimation

Six studies use a generalised linear model (GLM) with a gamma distribution and a log link to estimate mean incremental effect of the PCCT on costs (#2, #7, #9, #10, #11, #12), while another employed a related approach with an instrumental variable (#8). All used propensity score weights in matching.

Prior to the popularisation of these methods, Studies #1, #3 and #5 all compared the mean effect for each treatment group patient against those of a control matched by diagnosis-related group (DRG). Study #4, without reporting a matching strategy, compared the median effect for the two groups.

Findings

Overall costs

All 12 studies report that palliative care consultations result in lower costs than their usual care comparators for most or all analyses, and there are no significant associations reported between the intervention and higher total costs. Formal meta-analysis is prevented not only by differences in study design, setting, intervention and population, but also differences in approach to expressing costs and calculating cost effects.

Studies that report costs from the hospital perspective mostly find statistically significant savings through palliative care in the 9-25% range (#2, #3, #6, #7, #9). Study #11 reports differences outside of this range with a 5% cost-saving for survivors and 31% for decedents. Of the other four studies to stratify by survivors and decedents, two also report statistically significant cost-saving effects for both groups and greater effects for deaths than live discharges (#7, #9). Study #12 finds a significant effect for patients who died but not for those discharged alive at the overall level, but significant effects for live discharges are observable for specific

sub-samples according to diagnosis and consult timing. Study #10 is distinct in reporting a difference only for those discharged alive within seven days.[24]

Of the two studies examining hospital charges, one reported mean daily charges around 7% lower for PC (#1) and the other median total charges around 16% lower for PC (#4). The only study to take a post-discharge health costs perspective finds costs for palliative patients 32% lower than those for usual care patients over six months (#5). One study's reporting method precluded calculating a proportional difference but daily costs are found to be lower for the intervention group (#8).

Ancillary costs

Where ancillary costs are reported separately, the results are inconsistent. Where statistically significant differences have been identified, costs are typically lower for palliative care interventions, but differences are not always identified. Study #2 reports ancillary (laboratory and radiology) costs 43% lower and no difference in pharmacy; a larger follow-up study found differences in laboratory and pharmacy but not in imaging (#8). Study #7 found differences in pharmacy but not imaging; a second study by the same lead author found no difference in pharmacy (#9). Study #11 found palliative care to be less costly across ancillary categories among patients who died, and different effects by category for patients discharged alive.

ICU costs

Of the six studies to report ICU use as a dependent variable, the results have a clear pattern towards lower use among palliative patients (#2, #8, #5, #7, #11). One found no significant difference, possibly due to lack of power (#6).

---

[24] There are concerns about the validity of results report in Studies #10 and #11. See **§2.4>Sample definition by LOS**.

LOS impact

Of the four studies to report LOS use as a dependent variable, none reports a statistically significant different between treatment and comparison groups (#1, #3, #6, #7). Of the two studies to report LOS as descriptive data for matched samples, neither reports a statistically significant difference between groups (#2, #4).

2.4 Discussion

The review suggests that PCCTs are consistently found to be less costly than usual care comparators in the range 9-25% for hospital costs, and these differences are typically statistically significant. Overall hospital cost-savings appear larger for those who died during hospitalisation than for those discharged alive. The observed effects appear to accrue from reduced intensiveness during hospital stay, with different studies finding significant effects for different ancillary utilisation categories. There is no evidence of reduced LOS in the review. The only study to examine post-discharge costs observed a 32% saving for the intervention group over six months, also statistically significant.

These results represent a clear pattern of cost-saving effect from the PCCT intervention but must be considered carefully and in context of the methodologies they employ.

Broadly, the observational studies can be divided into two classifications: 'earlier' studies, which employ simple analytical techniques (#1, #3, #4, #6), and 'later' studies, which employ more advanced approaches (#2, #7 to #12). Study #5 is distinct not only as an RCT but in taking a different cost perspective, and cannot be compared to the other studies (or inform my own primary analysis) in any detail.

'Earlier' studies employed small sample sizes and one-dimensional matching techniques, raising concerns about bias (Starks et al., 2009). These less complex

early studies also raise concerns in their estimation of incremental effect on cost. Cost data pose well-known challenges in analysis.[25] In particular they typically suffer right-hand skew, resulting from a relatively small number of complex patients utilising high proportions of resources (Jones, 2010). This can be problematic for evaluation since linear statistical models are typically based on an assumption of normally distributed data. The standard response to skewed data is transformation, e.g. the natural log of the dependent variable, which can help to normalise the distribution, but poses a problem in retransformation: mean $ effects are not straightforward to estimate from log-transformed data, and mean $ effects are a primary concern for cost analysis (Manning, 1998).

These 'earlier' studies do not report addressing issues of distribution, raising concerns about the accuracy of their cost-effect estimates. The exception is Study #4, which estimates the median effect rather than the mean. This is an attempt to limit the effect of high-cost outliers but in doing so it adopts a dependent variable that is fundamentally not useful: mean, not median, costs are those that constitute the more relevant data for any payer in any system.

'Later' observational studies are larger and use propensity score matching, a more sophisticated technique that matches treatment and comparison groups on observed covariates but cannot control for unobserved confounding (Rubin, 2007, Stuart, 2010). Instrumental variable approaches can help to minimise unobserved confounding (Angrist et al., 1996) but only one study in the review did so, possibly in part because a valid instrument is difficult to identify (Murray, 2006).

---

[25] For a more detailed description of these challenges, see **§3.2.2.** For my approach to addressing them in my primary research, see **§4.1** and **§4.2**

These studies also demonstrate an awareness of difficulties in analysing cost data, employing a nonlinear regression model (GLM with a gamma distribution and a log link). This is the dominant model in the recent health economics literature and is generally considered appropriate for the characteristics expected of hospital cost data as well as providing results on the raw cost scale, avoiding retransformation problems (Jones, 2010, Manning, 1998).[26]

The 'later' studies can therefore be seen to have addressed the fundamental methodological issues with the 'earlier' studies, and these advances in rigour give increased confidence to the cost-saving pattern of results observed. However, a number of outstanding methodological issues remain. It is particularly important these issues are afforded full consideration given the established pattern in more recent studies of imitating similar or identical methodological procedures.

The remainder of this section is given over to considering these methodological issues in greater detail. The discussion mostly focuses on the methods, results and limitations of the 'later' studies, since these represent the state of the science in economic evaluation of PCCT interventions.

Model selection

The largest, most robust economic evaluations of palliative care have consistently employed one regression model (GLM, gamma, log) to estimate PCCT association with hospital costs. This model has well-established strengths for analysing health cost data and is widely used in the health economics literature as a consequence (Manning et al., 2005, Blough et al., 1999, Basu, 2005).

---

[26] My primary research broadly follows the methods of the 'later' studies, employing propensity score matching and non-linear modelling approaches. See **§3.2, §4.1** and **§4.2.**

However, model selection for cost data is rarely straightforward (Manning and Mullahy, 2001). Multiple modelling approaches are available to analysts, different models often deliver different results for the same data, and comparative performance evaluation does not typically yield a single dominant model (Jones, 2010, Jones et al., 2013a, Jones et al., 2013b). Rather, model selection is performed through trade-offs and is best informed by comparative analysis, clarifying different approaches' strengths and weaknesses, and so qualifying and informing interpretation of their results (Garrido et al., 2012).

No paper included in the review reports model comparison and evaluation prior to analysis. The three most recent studies (#10, #11, #12) explicitly highlighted that their methods follow those of Morrison and colleagues (#7, #9) rather than considering alternative approaches that may be more suitable to their own data. This has the advantage of simplifying comparison of results, but there is no evidential basis on which to assume that a model used in one study is optimal for the data in another.

While GLM (gamma, log) may be the most appropriate model for these studies, issues of model selection and comparison would ideally receive more attention in future studies prior to analysis.[27]

Regression using LOS as a covariate

Two studies report using LOS as a covariate in regression (#8, #11).

This strategy is employed to address unobserved heterogeneity in analysis of observational data. Incorporating (a nonlinear transformation of) LOS as a covariate,

---

[27] In my primary research I performed comparative model evaluation prior to evaluating the intervention. For details see **§3.2.2** and **§4.2**.

or sample selection by LOS, can act as a proxy for otherwise unobserved baseline factors that analysts would like to control for but on which treatment and comparison groups may not be fully matched, e.g. clinical complexity, patient goals (Carey and Burgess, 1999).

However, using LOS as a covariate in regression exacerbates endogeneity concerns: LOS is both a predictor of treatment (long LOS is an indicator of palliative care need) and a determinant of cost (which increases with LOS) (Amporfu, 2010). At the very least sensitivity analyses must be performed with and without the covariate to check the impact on results.[28]

Investigators in Study #8 performed a sensitivity analysis with and without log-transformed LOS, and did not find any difference (personal communication, J. D. Penrod, September 2013), although it is notable that they employed an instrumental variable and were therefore already attempting to control for unobserved heterogeneity. Study #11 authors do not specify what transformation (if any) was used, do not report performing a sensitivity analysis, and did not respond to multiple attempts at personal communication.

Sample definition by LOS

Trimming the sample at seven- and 30-day LOS was first performed in this field in Study #7 to exclude a relatively small number of outliers for whom matching was problematic and unobserved heterogeneity was a particular concern (personal communication, R. S. Morrison, June 2013). Variations were subsequently repeated

---

[28] For an overview of how I addressed 'heterogeneity versus endogeneity' arising from LOS as a covariate in my primary research, see **§4.3**.

by the same lead author (#9) and adopted by studies explicitly imitating Morrison et al. (#11, #12).

As with using LOS as a covariate, the principal justification for trimming the sample by LOS is to control for unobserved baseline heterogeneity. However, this approach constitutes selection of the sample by outcome. It is not consistent with the most rigorous research design, which would enrol and include patients only on the basis of baseline factors. In particular it is antithetical to the philosophy underpinning propensity score analysis (used by all studies which define their sample by LOS), which aims to imitate randomisation by matching on the basis of observed baseline covariates prior to consideration of outcomes (Austin, 2011). Defining a sample by a variable that potentially lies on the causal pathway between treatment and outcome may obscure treatment effects and bias estimates (Garrido, 2014b).

The same criticism could be levelled at dividing samples by discharge status. The justification in this case is that for observational data drawn from short periods of study there will likely be unobserved clinical heterogeneity between patients discharged alive and those who died during hospitalisation; that different treatment decisions and preferences are likely involved; and that there are distinct implications of discharge (the point at which costs are no longer accrued) (Cassel et al., 2010). While sensitivity analysis comparing the impact of dividing by discharge status is optimal, the strategy has a plausible basis and a clearly defined set of rules (discharge status being binary).

The strategy to trim by LOS is more arbitrary. Analysis of possible trimming points for LOS outliers in modelling population cost data has been published (Marazzi et al., 1998) but not for this intervention or patient group. Sensitivity analyses examining

the impact of stratifying by discharge status or LOS are not reported by #7, #9, #11 or #12 but it seems plausible that arbitrary selection by outcome could affect results.[29]

In Studies #10 and #11 concerns from introducing an outcome (LOS) into baseline analysis (propensity scoring) are particularly serious. In #10, the authors stratify their sample by LOS prior to matching as well as including unmatched patients and irrelevant variables; in #11 the authors report using LOS as a covariate on which to match patients (as well as using LOS as a covariate in regression). All of these strategies fundamentally undermine the integrity of research employing propensity score methods (Garrido, 2014a, Garrido, 2014b). Consequently Studies #10 and #11 analyse and report cost effects using frameworks with fundamental endogeneity problems, and these results have limited weight as scientific evidence.

Results according to LOS-defined sub-samples have two further limitations. First, evidence from such methods is of limited value in making treatment decisions because LOS is not known upon admission. Second, there is a long-term trend towards reducing hospital LOS across most health systems, reflecting patient preferences and high institutional costs. Therefore in the future a sub-sample excluding patients who stayed fewer than six or seven days will become a very specific sub-set of patients with complex needs, excluding the majority of shorter-stay patients.[30]

---

[29] Study #10 does report sensitivity analyses but the results are difficult to interpret given (a) the fundamental methodological problems detailed in the subsequent paragraph and (b) there is no indication that propensity scores were recalculated when the sample was redefined by LOS or discharge status (see **§3.2.1>Additional propensity scores** for an overview of the importance of this issue and how I handled it in my primary analysis).

[30] For an overview of how I addressed 'heterogeneity versus endogeneity' in truncation of LOS outliers in my primary research, see **§4.3**.

Daily cost as a dependent variable

Selecting *per diem* cost instead of *in toto* cost typically reduces skew while incorporating LOS indirectly, therefore minimising a major challenge of cost data analysis.

However, it is necessarily limited as a dependent variable. First, daily cost is not synonymous with total cost and must be interpreted carefully. *Per diem* ratios are highest at admission and reduce systematically as LOS grows, so a reduction in LOS from five days to four days does not correspond to a 20% reduction in total costs (Ishak et al., 2012). Additionally, if LOS is systematically different between treatment and comparison groups then daily cost (the ratio of total cost to LOS) is a fundamentally different variable to total cost.

Second, while it may speak to narrow sectional interests in a given incentive structure, generally speaking daily cost is of less interest than total cost. Daily costs are ultimately not the cost that is paid for healthcare, they are only a ratio of that figure to LOS. An intervention that reduces daily cost by 10% but extends LOS by 50% (thereby increasing total cost), will be evaluated positively where daily cost is the primary dependent variable while potentially being viewed negatively by the patient, the payer and the societal impact analysis, all of whom generally favour a shorter stay for lower total cost.[31]

There are circumstances in which daily costs may be a relevant dependent variable. For example, there are insurance programmes in the United States health system

---

[31] Shorter hospital stay is generally favoured by patients, *ceteris paribus*, but early discharge can only be considered a desirable outcome given certain assumptions about the patient and family outcomes that follow the intervention, and also the post-discharge cost implications. For a discussion of the importance of patient and family outcomes, and of post-discharge costs, see below **§2.4>Forms of economic evaluation** and subsequent sub-sections.

that reimburse a fixed fee for each day of hospital stay (Quinn, 2008), meaning that incremental effect on average daily cost is a pertinent result for specific audiences. Nevertheless such analyses should be conducted with the clear understanding that efficient practice is best understood and incentivised by focusing on overall costs, with daily cost at most a secondary variable of interest.[32]

Total versus direct cost as a dependent variable

Different studies in the review take different approaches and the case can be made for either standpoint.

Pure health economic theory points to incorporating proportional levels of indirect costs, thus analysing total costs, as the means to understanding fully the impact of an intervention (Drummond et al., 2005). This approach is widely visible in 'European-style' single-payer systems.

In the United States health system, however, there are important differences. The development of palliative care has often been piecemeal, with programmes expanding much more quickly in hospital settings than in the community (Morrison, 2013a). One important factor in this piecemeal development has been the incentives inherent to the US system's fragmented finance structures (Cassel, 2013). It is common for American providers to be reimbursed for activity and/or numbers of patients, and there are multiple potential payers (Pfuntner et al., 2013). Consequently US hospitals have often taken a 'business case' approach to establishing funding streams for palliative care programmes (Cassel, 2014, Pantilat et al., 2007, Brousseau et al., 2012). Senior managers at individual institutions

---

[32] Daily cost is established as a secondary dependent variable for my primary analysis in **§3.1.6**. The impact of using daily cost instead of total cost as a dependent variable is evaluated in **§4.3**. Incremental effect on daily cost is reported in secondary analyses in **§4.4** and **§4.5**.

assess whether a new programme is in their interest, something that is dependent on which costs are reduced by the programme and how hospitals are reimbursed (Cassel, 2013, Smith et al., 2012a). The development of palliative care services in the US has therefore occurred at least in part due to a preference for cost-containment at large institutions than rather than due to a generalised national policy decision to expand this type of care to appropriate patient groups.

In research, this context has led some analysts to exclude from their projections indirect costs, which in practice are generated by a fixed formula after a patient's direct costs are known, amplifying variations and the impact of interventions, when by definition they still exist for hospitals to pay at the end of the fiscal year irrespective of what care is provided to what patients (personal communication, J. Brian Cassel, March 2013). In restricting their perspective to variable and fixed direct costs, analysts focus on those utilisation categories that the treatment can practicably impact hospital expenditure, and are therefore of most interest to managers for whom cost-containment is an active consideration.

The appropriate approach to total and direct costs rests principally on the time horizon of the analysis. For example, studies of cost effects from a hospital perspective may be more robust if focused on direct costs since those are the costs that can be realistically impacted in the foreseeable future (Weinstein et al., 1996). But a model of how palliative care costs may develop over the next 20 years would

be better to include total costs, since it is not unreasonable that hospital set-up and fixed costs will alter over longer periods of time in line with clinical practice.[33]

Mechanism of observed cost-savings

Cost-savings associated with PCCT interventions can in principle be attributed to two underlying drivers: reduced intensiveness of hospital stay and/or reduced length of hospital stay.

While the overall pattern of cost-saving in this literature review is clear, the findings on the mechanism(s) by which cost-savings are achieved are not. Different studies find different categories of ancillary costs are reduced by treatment, and there is no reported impact on LOS, which is itself a vexed issue in palliative care research (Cassel et al., 2010). Where ICU utilisation is identified as a dependent variable this appears to be reduced systematically, but it is not clear if reduced ICU stay corresponds to reduced LOS overall (Khandelwal et al., 2015).

More evidence is needed to understand and explain the visible pattern of cost-savings through PCCT interventions.[34]

Assumption of homogenous cost effect

Studies #10 and #12 are the only two studies to relax the assumption of cost-effect homogeneity. The other studies analyse mean costs for their samples (albeit sometimes split by discharge status) without consideration for how this effect may vary within the sample(s).

---

[33] For my primary research I identified total direct costs as the primary dependent variable; for details see **§3.1.6**.

[34] The underlying mechanism of savings is examined in my primary analyses in **§4.4** and **§4.5**, and reviewed in **§5.1**. The scope for future research to investigate this variability is discussed in **§5.3.3>Further evidence on how a PCCT operates would extend my results**.

The mean incremental effect on total cost is the obvious *de facto* primary interest in cost analysis of PCCTs. However, the effect is likely not homogenous but instead will vary according to individual and service-level factors (Groeneveld et al., 2013). For example, patient-level variables are strong predictors of utilisation (Tibi-Levy et al., 2006, Simoens et al., 2010, Kelley et al., 2011) while clinical research has shown that palliative interventions earlier in the care trajectory can deliver benefits (Temel et al., 2010, Bakitas et al., 2014, Zimmermann et al., 2014). Investigating cost-effect heterogeneity, examining which interventions for which patients optimise effect, will have obvious benefits for informing efficient provision of care.

Study #12 makes two interesting contributions in this regard, reporting that cancer patients are particularly associated with cost-savings (compared to non-malignant diagnoses) and that there appears to be an association between intervention timing and cost effect, although this relationship is not linear (effect is greatest for a consult after 5-6 days of admission for patients d/c alive and after 3-4 for patients who died). Study #10 reports an associated idea that shorter stay patients are associated with larger cost-saving effect.

However, in both cases their results are limited as evidence due to methodological concerns in defining their samples by LOS (outlined above). Additionally, Study #12 derives its associations between diagnosis or timing and cost effect by 'interaction terms', e.g. fixed-effects dummy variables in the regression to denote a patient's diagnosis. But interpreting the coefficient of a covariate in a propensity score-matched analysis requires that this covariate is balanced between treatment and comparison arms. McCarthy et al. do not report balancing the arms on the interaction terms for diagnosis, and the assumption of balance may not be valid.

Forms of economic evaluation

Health economic evaluation entails "comparative analysis of alternative courses of action in terms of both their costs and consequences" (Drummond et al., 2005). There are multiple forms of evaluation, each a variation on this cost-consequence framework and differentiated by the specific approach to measuring and incorporating consequences (i.e. outcomes) (McPake et al., 2013).

Drummond et al. (2005)[35] differentiate these forms as follows:

- *Cost-benefit analysis*, where outcomes are measured in terms of money, e.g. $, €, £;
- *Cost-utility analysis*, where outcomes are measured using an index that can be employed across different health care services, e.g. the quality-adjusted life year (QALY);
- *Cost-effectiveness analysis*, where outcomes are measured using natural units, e.g. number of life-years gained;
- *Cost analysis*, where outcomes are assumed to be equal, and only the costs of the alternatives are compared.

---

[35] See Table 1.1, page 2, Drummond et al. (2005). Other taxonomies, also based on the cost-consequence framework but with small variations in terminology and definitions, have also been published. E.g. Gold et al. (1996) and Health Information and Quality Authority (2014) refer to 'cost-minimization analysis' (CMA) where Drummond et al. use 'cost analysis'. For consistency with Drummond et al., I use 'cost analysis' throughout this thesis. Within this thesis, this term can be considered synonymous with cost-minimisation analysis.

All 12 studies in Table 2.3.1 are cost analyses: they evaluate the impact of the PCCT intervention on costs without also measuring the effect on outcomes and quantifying the two in a cost-consequence framework.[36]

Evidence of intervention efficacy

Therefore the cost-savings consistently reported in Table 2.3.1 imply that PCCT care is preferable from an economic perspective to usual care only on a 'non-inferiority' assumption, i.e. that outcomes for intervention group patients are at least no worse than those for comparison group patients.

The evidence supporting this assumption has grown in size and sophistication as the PCCT model of care has expanded.[37] Early descriptive studies reported strong patient outcomes, provider satisfaction and caregiver satisfaction, while acknowledging the need for higher levels of evidence (Manfredi et al., 2000, Elsayem et al., 2004, O'Mahony et al., 2005). Subsequently three studies which evaluated patient outcomes separately to costs and are included in Table 2.3.1 variously reported improved pain, dyspnea, and secretions scores (Ciemins et al., 2007), improved care experience and communication (Gade et al., 2008), and improved pain and symptom management (Hanson et al., 2008). Other matched studies of these interventions have reported increased family satisfaction (Casarett et al., 2008, Casarett et al., 2011). And two recent RCTs[38] have reported that

---

[36] The primary analysis in this thesis is also a cost analysis. The context for this methodological approach is reviewed in **§3.1.2>Objectives** and **§3.1.6**. The limitation of this approach in my primary research is reviewed in **§5.2.3**. In future phases of the 'PC4C' project patient and family outcome data will be available to broaden the scope of analysis, as discussed in **§5.3.3**.

[37] For an overview of the growth in access to and research on PCCT services over the last 15 years, see **§1>Evidence and research in palliative care**.

[38] As the literature review shows, RCT designs are relatively unusual in this field. For an overview of why an observational design may be considered more appropriate for this intervention and population, see **§3.1.3**. For steps taken to maximise strength of evidence in my primary analysis, which employs observational data, see **§3.2.1**.

hospital-based PCCTs improve patient quality of life, spiritual well-being and symptom burden (Zimmermann et al., 2014), and patient quality of life, depression and survival (Temel et al., 2010).

Two systematic reviews of specialist palliative care (including but not limited to PCCTs) found improved patient outcomes in the domains of pain and symptom control, and anxiety (Higginson and Evans, 2010), and in quality of life and satisfaction with care (El-Jawahri et al., 2011). And a systematic review of palliative care efficacy in all settings concluded that hospital inpatient palliative care consultations teams have been shown to improve symptom control, quality of life, emotional burden, and caregiver and patient satisfaction (Luckett et al., 2014).

The 'non-inferiority' assumption in this field therefore appears reasonable: there is no evidence of worse outcomes for PCCT patients in any review in Table 2.2.1, in any study in Table 2.3.1, or elsewhere in the literature discussed in this thesis. Moreover, patient outcomes following PCCTs not only appear 'no worse', but often better. While the domain and magnitude of benefits accruing from PCCTs varies between studies across patient and family outcomes, a consistent pattern of positive consequences is discernible with no such evidence of negative consequences.[39]

Outcome measurement in economic evaluation of palliative care

A lack of outcome measurement in economic evaluation is not particular to PCCTs. Smith et al. (2014) found that only two of 46 economic evaluations of palliative care

---

[39] I therefore retain the 'non-inferiority' assumption in my original research. The use of cost analysis methods for my primary analyses is re-iterated in **§3.1.2>Objectives** and **§3.1.6**. For consideration of how I aim to validate the assumption as more data become available, see **§5.3.3>Further research with the 'PC4C' data will extend and strengthen my results**. One issue that the current literature on intervention efficacy has not addressed in detail is *how* PCCTs bring about the reported effects. For further discussion of PCCT treatment goals, mechanisms and causal effects, see **§5.3.3> Further evidence on how a PCCT operates would extend my results**.

in all settings had attempted some form of cost-consequence analysis (Raftery et al., 1996, Higginson et al., 2009).

There are multiple reasons for this limitation in the evidence base, including debate over appropriate outcomes, sometimes known as the 'QALY problem' (Hughes, 2005). While QALYs are known as the 'gold standard' in health economic outcome measurement (Health Information and Quality Authority, 2014) and recommended by some for use in palliative care (Round, 2012), others have questioned their validity for this patient group (Normand, 2009, Chochinov, 2011, Normand, 2012). Best-practice guidelines for research in end-of-life and palliative care failed to reach a consensus on an optimal approach to outcome measurement in the context of this debate (Preston et al., 2012).

Consequently the state of the science of palliative care economics is perhaps best understood through McPake et al.'s framing of economic evaluation:

> *"It is a mistake to think of the different methods of economic evaluation as different approaches. Rather, they should be seen as variations in what is desirable and feasible [...] it is better to think of the problem as being: 'How constrained are investigators in their measurement of outcomes?'"* – p.91-92.

The answer to that question in palliative care research is that substantial constraints are often a prominent challenge for economic and non-economic studies alike (Evans et al., 2013). And this is reflected in the literature, where theoretical discussion of QALYs for palliative care programmes has not been matched by their empirical use, at least in part because of the profound complications in collecting, measuring and synthesising data on complex and intangible outcomes.

A critical assessment of the literature in Table 2.3.1 must therefore acknowledge the limitations of evidence where no quantitative cost-consequence analysis has been performed, while recognising that studies to date do substantively contribute knowledge to the impact of these interventions, that these limitations reflect real-world challenges, that this knowledge can be supplemented by consideration of the non-economic literature, and that a future priority for this field is cost-consequence analysis (Preston et al., 2012, Smith et al., 2014, May et al., 2014).[40]

Hospital costs perspective

With regard to costs, guidelines recommend taking the societal perspective: that is, including in analysis not only costs to the healthcare system but also those to patients, caregivers, wider systems and society (Drummond et al., 2005, Gold et al., 1996, Preston et al., 2012).

Among studies in Table 2.3.1, the focus has been on the hospital 'silo'. Nine examine only costs to the hospital providing care (#2, #3, #6 to #12) while two use only hospital charges (#1, #4), generally considered a poor approximation of hospital costs. One study examines formal health system costs for six months following discharge but no informal costs (#5).

As such, methodology to date has implications for understanding of the role that palliative care plays in enhancing value in health care. Specifically, questions remain as to whether reductions in hospital costs are authentic cost-savings and not simply cost-shifting. If hospital costs are reduced but ultimately passed on to other

---

[40] For an overview of how I aim to extend my thesis results with cost-consequence analysis in future, see **§5.3.3>Further research with the 'PC4C' data will extend and strengthen my results**.

care settings or to family or informal caregivers, the expenditures have not been meaningfully reduced.[41]

Conclusion

The published evidence shows a clear pattern of PCCT interventions reducing hospital costs.

These results must be considered carefully in the context of their methods. Observational designs are typical in this field due to ethical and practical limitations. Earlier studies are small and with rudimentary matching procedures, raising concerns about bias. As methods have become more robust, increasing in size and statistical sophistication, a specific and narrow approach has become established as the norm.

One model (GLM, gamma, log) has been consistently used with little reported comparison and evaluation informing selection. Employing LOS as a covariate and/or sample parameter and/or as the denominator of the dependent variable (i.e. daily costs) has been used to control for unobserved heterogeneity, introducing endogeneity into analysis and limiting the usefulness of results. Most studies have implicitly assumed homogenous PCCT effect on total costs, which is a valid primary analysis but may disguise valuable additional information. There is limited, mixed evidence on the underlying mechanism of observed cost-savings through PCCTs, with differences variously attributable to reductions in different utilisation categories. Understanding of PCCT impact on LOS is limited. All but one study to date has

---

[41] My primary research is also a hospital cost analysis. Restricting the focus to formal hospital costs is part of the 'PC4C' project research design, detailed in **§3.1**. The limitation of this perspective is reviewed in **§5.2.3**.

analysed only the hospital cost perspective with no reported cost-consequence analyses quantifying the relationship between cost and clinical effects.

The stated principal aim of this PhD thesis is to contribute substantively to knowledge on the impact of the PCCT intervention by performing cost analysis on a dataset taken from the hospital perspective. The cost analysis methodology will be employed on a 'non-inferiority' assumption, an approach that appears justified by the well-established pattern of as good (and often better) outcomes for patients who receive a PCCT intervention. From this literature review I identify a number of priorities for my enquiry.

Prior to my main analysis I must perform (i) careful description of the cost data to establish their characteristics and consider appropriate modelling approaches, and (ii) comparative evaluation of different available approaches to assess which models perform best with the study data.

Having done this preliminary work, my primary analysis should focus on (iii) consideration of the impact that controlling for LOS in analysis has on cost-effect estimates, (iv) estimating the impact of PCCT interventions on cost using appropriate methods, including variations in effect according to service- or patient-level factors, (v) the underlying mechanism of any observed cost differences following treatment, and (vi) factors other than treatment that are driving hospital costs for patients with serious illness.

Addressing these priorities will expand the evidence base on the economic impact of PCCT interventions, and so inform practice and policy, while also addressing important methodological issues, and so strengthen future evaluations in this field.

Where new research is designed from scratch, expanding the cost perspective beyond hospital discharge and combining this evidence with clinical outcomes in robust cost-consequence analysis will strengthen understanding of the economic impact of palliative care.

Chapter 3: Materials and Methods

Abstract

This chapter is split into three parts and structured according to items 2-12 in the STROBE guidelines (von Elm et al., 2007).[42]

§3.1 Materials details the study design, patient recruitment and data collection by colleagues in assembling the 'Palliative Care for Cancer' study dataset prior to my research, as well as my objectives and handling of dependent and independent variables (STROBE items 2-8, 10, 11).

§3.2 Methods details my methodological analysis of the dataset, including approaches to bias and statistical methods (STROBE items 9, 12).

§3.3 Summary provides an overview of the key elements of this chapter.

---

[42] For a detailed overview of this PhD thesis according to the STROBE checklist, see Appendix to §3.

3.1 Materials: the 'PC4C' study

3.1.1 Background

The data used in this thesis were collected as part of the 'Palliative Care for Cancer' ('PC4C') study, funded by the National Cancer Institute (NCI) and National Institute of Nursing Research (NINR) in the United States, and led by investigators at the Icahn School of Medicine at Mount Sinai, New York (National Institutes of Health, 2006).

The purpose of the 'PC4C' study was to evaluate the impact of palliative care consultation team (PCCT) interventions for patients admitted to hospital with advanced cancer.[43] Colleagues conducted a multi-site, prospective cohort study employing propensity scores to minimise confounding. The aim was to examine the effect of PCCTs on patient and family outcomes (e.g. symptom control, satisfaction with care), processes of care (e.g. prescribing, transition management, advance planning), and utilisation (e.g. hospital costs) as compared to usual care at five United States hospitals from 2007 to 2011 inclusive.

I accessed the data during an extended period of mentored research (2012-2014) at Mount Sinai under the supervision of R. Sean Morrison, MD, per the terms of my fellowship funding (Health Research Board, 2011). I received a cleaned dataset at the end of the data collection period.

---

[43] Admission to hospital and access to palliative care at all five sites was not affected by diagnosis. PCCTs saw patients with non-malignant conditions. The eligibility criteria reflect not the intervention but the purpose of the study, which was intended specifically to address evidence gaps on the impact of PCCTs for patients with advanced cancer. The implications of this sample specification for results and generalisability are discussed in **§5**.

3.1.2 Objectives

The 'PC4C' study

The aim of the 'PC4C' project was the evaluation of PCCT impact for hospital inpatients with an advanced cancer diagnosis.

This aim was to be achieved through three specific objectives:

1) **Patient and family outcomes**: to evaluate PCCT effect on pain and other physical and psychological symptoms for patients, and on family satisfaction with care for patients (compared to patients who did not receive palliative care).

2) **Processes of care**: to evaluate PCCT use of evidence-based algorithms to communicate information; to establish goals of care and treatment preferences; and to implement discharge plans consistent with goals and preferences (compared to patients who did not receive palliative care).

3) **Utilisation**: to evaluate PCCT impact on resource use during hospitalisation (hospital costs, length of stay (LOS), intensive care unit (ICU) utilisation) in providing care to patients (compared to patients who did not receive palliative care).

This PhD thesis

Of the 'PC4C' study objectives, my PhD thesis is concerned only with (3) Utilisation.

My aim is to contribute substantively to knowledge on the impact of the PCCT intervention through rigorous cost analysis. Specifically I will evaluate the impact of PCCT treatment on selected economic variables of interest for patients admitted to

hospital with an advanced cancer diagnosis, controlling for bias between treatment and comparison groups through propensity score matching.

Consistent with prior economic evaluations of PCCTs, my primary analysis will constitute a cost analysis in which the patient and family outcomes following the intervention are assumed to be at least as good as those following the usual care only comparison. This assumption is made in the context of the published literature on the impact of PCCTs on patient and family outcomes.[44]

The conclusion to the previous chapter highlighted the importance of (i) considering the characteristics of cost data prior to analysis, (ii) comparing performance of alternative models prior to estimating cost effects, (iii) establishing the impact on effect estimates of using LOS to control for unobserved baseline heterogeneity, (iv) estimating the impact of PCCT intervention on hospital cost, and analysing patient- and/or system-level factors that are associated with magnitude of cost effect, (v) examining the underlying drivers of any observed cost differences, and (vi) identifying patient- and/or system-level factors that are associated with hospital utilisation other than palliative care consultation.

---

[44] For a review of this evidence and the broader context of economic evaluation in palliative care, see **§2.4>Forms of economic evaluation** and subsequent sub-sections. The limitations of cost analysis methods are reviewed in **§5.2.3**. My intention to expand the scope of analyses with these data in post-doctoral research is discussed in **§5.3.3**.

My aim will therefore be achieved through the following objectives:

Prior to primary analysis:

(i) **Data overview:** to describe the 'PC4C' cost data and assess its characteristics for the purposes of statistical modelling and cost-effect estimation;

(ii) **Model evaluation and selection:** to compare the performance of alternative modelling approaches with the 'PC4C' data, and so identify the most appropriate model for these data.

In primary analysis:

(iii) **Comparative evaluation of methods using LOS to control for unobserved heterogeneity:** to establish the impact of LOS-control strategies in estimating PCCT cost effect with the 'PC4C' data, and so identify the most appropriate approach to controlling for unobserved baseline heterogeneity while minimising endogeneity concerns;

(iv) **Cost-effect estimation:** to estimate effect of PCCT intervention on hospital costs, to analyse patient- and system-level factors associated with this effect, and to consider the implications of these findings for both provision of care to patients with serious illness and future research of interventions to this patient group;

(v) **Source of any observable cost effects:** to examine the underlying source of any observable PCCT impact on cost, in particular the extent to which cost

differences are attributable to reduced intensiveness of hospital stay, reduced length of hospital stay or a combination of these;

(vi) **Other determinants of utilisation:** to identify patient- and system-level factors associated with hospital cost other than the PCCT intervention, and to consider the implications of these findings for both provision of care to patients with serious illness and future research of interventions to this patient group.

3.1.3 Study design

Using a prospective, observational design, data were collected for patients admitted to one of five United States hospitals from 2007 to 2011.

The rationales for the study design employed are as follows:

Rationale against RCT

A randomised controlled trial (RCT) is the strongest design to test treatment efficacy against a placebo under ideal conditions, but where such conditions are not met an RCT may not be optimal (McKee et al., 1999).

The PCCT is often considered an intervention for which RCT design is neither practical nor useful. Recruiting patients with a high chance of rapid disease progression poses major challenges (Higginson et al., 2013, Aldridge Carlson, 2013, Steinhauser et al., 2006); the process of hospital admission may not provide a sufficient timeframe to recruit and randomise sufficient numbers of patients. Further, consultative interventions (in palliative care and elsewhere) raise issues of ethics, protocol adherence and external validity in an RCT context:

- *Ethical considerations:* Many physicians refuse to take part in a study that restricts the care pathways of their patients (Ewing et al., 2004). Some believe that palliative care is superior to usual care, and so will not enrol their patients in an RCT since they do not want them obligated to remain in the control group. In other cases physicians do not consider palliative care appropriate for their patients and so refuse to participate. As well as raising ethical concerns about an RCT design for palliative care, such physician attitudes lead to enrolment bias.

- *Protocol adherence:* Mandatory consultations are thought to result in lower compliance with treatment recommendations than requested consultations, biasing an RCT towards a null finding (Penrod et al., 2008). Additionally, there is a high risk of contamination through cross-over: if physicians or family members have a positive experience of palliative care during the trial they may request that a patient receiving usual care be transferred to palliative care, biasing the treatment effect (Marcus and Gibbons, 2001).

- *External validity:* For the reasons outlined above, both the physicians and patients who are enrolled in an RCT for palliative care consultation, where the intervention is either mandatory or unavailable, would be atypical in the context of general practice (Black, 1996). Therefore results would not be generalisable to the model of requested consultations as practiced in United States hospitals and other settings.

Thus, while RCTs maximise internal validity, and examples of this design for evaluating PCCTs can be cited (Temel et al., 2010, Zimmermann et al., 2014), many

investigators including those leading the 'PC4C' project prefer to pursue observational studies.

Rationale in favour of observational design

For research questions where an RCT will lead to systematic bias, poor adherence and weak external validity, a rigorous, carefully designed observational study may be considered superior (Rubin, 2007, Higginson et al., 2013).

The most significant weakness in observational design is that, since assignment to treatment and comparison groups is not under analyst control, differences in outcome may result from the intervention, from observed confounders and from unobserved confounders.

Propensity scores, first developed by Rosenbaum and Rubin in 1983, are an increasingly popular means of controlling for bias in observational studies in a variety of disciplines (Rosenbaum and Rubin, 1983, Stuart, 2010). Studies employing propensity scores are structured consistent with an RCT with inclusion criteria and intervention both clearly defined beforehand. Eligible subjects are enrolled prospectively but not randomised; exposure to the intervention is decided by the natural course of medical care. Before analysis of the intervention's effect on outcomes of interest, propensity scores for each subject are calculated. The propensity score represents the probability that a given subject was exposed to the intervention on the basis of observed confounders.

In summary, the 'PC4C' study was conducted to draw valid comparisons between two groups for whom observational data could be most appropriately collected. The study design is intended to preserve important elements of an RCT (other than randomisation itself) by enrolling patients prospectively according to clear eligibility

criteria; by evaluating an intervention that is well-defined, structured and generalisable; and by balancing the treatment and comparison groups on observed confounders. Additionally it avoids some of the weaknesses that make a traditional RCT inappropriate for the evaluation of PCCT interventions: physicians and patients are free to follow the natural course of medical care; potential bias arising from treatment cross-over is eliminated; usual practice patterns and representative physician and patient populations are preserved to ensure external validity.[45]

3.1.4 Setting

The lead centre for the study was the Icahn School of Medicine at Mount Sinai (MSSM), NY. The other participating centres were: Virginia Commonwealth University-Massey Cancer Center, Richmond, VA (VCU); University of Pittsburgh Medical Center, Pittsburgh, PA (UPMC); Mount Carmel Health System, Columbus Ohio, OH (MC), and Froedtert Hospital of the Medical College of Wisconsin, Milwaukee, WI (MCW).

All five participating hospitals were high-volume tertiary-care medical centers with established PCCTs using existing practice guidelines for pain and symptom management, and communication. Specifically, MSSM, UPMC and MCW classify themselves as academic medical centres; MC is a community hospital and VCU a cancer center. These sites were selected for their high patient numbers (conducive to large study sample size), well-established PCCTs (giving an intervention that is generalisable best practice of care), and substantial research capacity (ensuring proficient collection and analysis of study data).

---

[45] For a detailed description of the propensity score methods employed in this thesis, see **§3.2.1**.

The earliest admission for a patient recruited to the project occurred in August, 2007. The latest admission and latest discharge for patients recruited to the project occurred in June, 2011. Hospital data collection occurred between these dates.[46]

Ethics approval for research design

Ethics approval was provided by the institutional review boards of all five participating hospitals prior to initiation of the study. These approvals included acceptance of detailed measures in place for obtaining patient consent, protecting patient and caregiver confidentiality, and preventing risk or harm to subjects.

Trinity College Dublin gave ethical approval for this PhD research specifically.

Intervention

A palliative care consultation was always initiated at the request of the primary treating physician when it was judged that a patient would benefit from the team's expertise. The most common reasons for initiating a consultation are clarification of care goals, assistance with transition planning, and expert pain and symptom control (Kamal et al., 2011).[47] Growing evidence on the clinical efficacy of the consultation intervention, as well as familiarity with and access to palliative care among primary physicians (e.g. oncologists), has led to increased referral rates (Kamal et al., 2011, Hughes and Smith, 2014).

The intervention consisted of three components, reflecting the core elements of palliative care in clinical practice guidelines (American Academy of Hospice and Palliative Medicine et al., 2004):

---

[46] Follow-up via telephone interview after discharge continued for up to six months but post-hospitalisation data form no part of this thesis.

[47] Formal analysis of the reasons for referral with the 'PC4C' data had not been published at the time of thesis submission but provisional analysis suggests consistency with Kamal et al.: treatment of pain and other symptoms, clarifying treatment options, establishing goals of care and advance plans, and assistance with transition planning (personal communication, R.S. Morrison, August 2015).

(1) Pain & symptom assessment and treatment;

(2) Establishment of goals of care and advance care plans;

(3) Transition management based on goals of care.

Consistent with these practice guidelines, at all sites the PCCTs comprised a core team of physician, nurse and social worker, with additional support available from other professionals including psychiatry and chaplaincy as appropriate to patients' needs.

The initial consultation focused on components (1) and (2):

1. Pain and symptom management

- Assess pain and symptom using validated tools;

- Recommend treatment per National Comprehensive Cancer Network (NCCN) protocols communicated to primary physician (National Comprehensive Cancer Network, 2005).

2. Establishment of goals of care and advance care plans

- Initial meeting with patients, families and care teams to establish goals of care, complete advance directives and communicate bad news (if requested) per NCCN protocols.

Follow-up visits were arranged daily or more frequently if needed, and comprised all three components:

1. Pain and symptom management

- Monitor implementation of treatment recommendations;

- Evaluate results of treatment recommendations;

- Assess new and ongoing symptoms. Identified symptoms assessed daily (more if needed) until controlled. Controlled or absent symptoms reassessed every 72 hours;

- Modify treatment recommendations as needed to meet goals of care.

2. *Establishment of goals of care and advance care plans*

- Discussions about new or changing goals of care, communicating bad news (if requested) and associated changes in treatment.

3. *Transition management based on goals of care*

- Meet with patient and family to discuss transition management;

- Work with primary care team to facilitate transition management/discharge plans according to goals of care.

Teams across all five sites were trained in a standardised protocol approach to consultation. Adherence to protocols was monitored by a project Steering Committee (study and site PIs, project managers and study co-ordinator) as well as, for the first six months of data collection, a random monthly audit of 20% of PCCT patient charts by a non-investigator MD. Repeat training was given to all new personnel joining any PCCT.

Usual care

Usual care (UC) comprised routine assessment of pain and other symptoms, function, nutrition, sleep and emotional concerns, as provided by the primary attending physician and their support staff. Symptoms identified at admission and

response to treatment were monitored daily. Chaplaincy and psychiatric support were available at all sites.

Therefore, important differences between PC and UC are that goals-of-care and transition management discussions are not initiated, and pain and symptom management is not provided by palliative care specialists, who have advanced training in this area.

Attending physicians most commonly choose not to involve specialist palliative care in the treatment of a patient where they feel that the team's expertise is not needed, e.g. in cases where pain and symptom management is already good, or in the absence of high symptom burden (Johnson et al., 2008).

Usual care services were also available to PC patients at all five sites.

3.1.5 Participants & recruitment

Eligibility criteria

Patients were considered for enrolment if they were admitted to one of the five participating hospitals, were aged 18 years or older, and had a primary diagnosis of metastatic solid tumour; central nervous system (CNS) malignancies; locally advanced head, neck or pancreas cancers; metastatic melanoma; or transplant-ineligible lymphoma or multiple myeloma.[48]

Subjects were excluded if their attending physician did not give permission to recruit their patients, or if they did not speak English; had a diagnosis of dementia; were unresponsive or nonverbal; had been admitted for routine chemotherapy; died or

[48] The implications of this focus on cancer for findings and generalisability are discussed in **§5.2.4**.

were discharged within 48 hours of admission; or had had a previous palliative care consultation.

Patients had to be enrolled within 48 hours of admission to be eligible.

Enrolment procedure

Prior to initiation, the study was approved by the institutional review board of each participating facility. Approval was sought by each attending physician at each site; where approval was not granted, this physician's patients were not considered for enrolment.[49]

Eligible patients were identified through a daily review of admissions records and administrative databases. For each potential subject identified, the patient's bedside nurse enquired about the patient's willingness to participate in the study. Given patient willingness, a trained clinical interviewer then approached the patient, explained the study and obtained informed consent. With the patient's consent, family members were also approached and enrolled with written informed consent.

3.1.6 Variables

Dependent variables

The primary analysis in this thesis is a cost analysis – a comparison of the costs for an intervention (PCCT) and an alternative (usual care only), on the 'non-inferiority' assumption that patient and family outcomes are at least as good for intervention group patients.[50]

All cost data used in this thesis were collected from the hospital perspective.

---

[49] Over 95% of physicians at all hospitals gave approval.

[50] For an overview of this methodology in the context of health economic theory and palliative care research, see **§2.4>Forms of economic evaluation** and subsequent sub-sections.

Direct and indirect cost data were extracted from hospital databases by colleagues compiling the 'PC4C' dataset. Direct costs are those directly attributable to the provision of medications, procedures, or services, including proportionate attribution of staffing costs, as distinct from indirect costs which represent the overhead costs of operating a hospital (e.g. operations, IT, maintenance) (Taheri et al., 2000).[51]

Cost data were provided for each procedure to each patient, designated by a subject number and a universal billing (UB) code (National Uniform Billing Committee, 2011). For each patient I created *total direct* and *total indirect* costs by summing the respective cost lines, and *total total* costs by summing *total direct* and *total indirect*. Daily cost ratios for each of the three total cost categories were created by dividing by LOS.

Additionally I used the UB codebook to identify major sub-categories of cost: room and board, ICU, pharmacy, laboratory and imaging. These categories were identified as the relevant major sub-categories in the context of similar studies (see previous chapter) and in discussion with R. Sean Morrison, MD.

Date-specific costs were generated for each patient, representing the cost of care received on any given day. These data were then combined with admission, discharge and consult date information to generate day-to-day trajectories of care following admission, prior to discharge and with respect to first day of PCCT treatment. These cost categories are summarised in Table 3.1.1.

---

[51] A more detailed definition of direct and indirect costs was previously provided in **§2.3>Dependent variable**.

Table 3.1.1 Cost variables generated through 'PC4C' study

Cost variable	Notes
Overall	
Total direct	Sum of all direct costs for each patient
Daily direct	(Total direct/LOS)
Total indirect	Sum of all indirect costs for each patient
Daily indirect	(Total indirect/LOS)
Total total	(Total direct + Total indirect)
Daily total	(Total total/LOS)
Sub-categories	
Room and board	Total direct if 100<=ubc<=169 | 170<=ubc<=179 | ubc=206|214 | 230<=ubc<=239 | 550<=ubc<=559 | 660<=ubc<=662 | ubc=700|709 | 760<=ubc<=769
ICU	Total direct if 200<=ubc<=204 | 207<=ubc<=213 | ubc=219
Pharmacy	Total direct if 250<=ubc<=269 | 630<=ubc<=636 | 640<=ubc<=649 | ubc=637 | ubc== 280 | ubc == 289 | ubc == 331 | ubc == 332 | ubc == 335
Laboratory	Total direct if 300<=ubc<=319
Imaging	Total direct if 320<=ubc<=329 | 340<=ubc<=349 | 350<=ubc<=359 | 610<=ubc<=619 | 401<=ubc<=409 | 730<=ubc<=739 | 480<=ubc<=489 | 460<=ubc<=469 | 740<=ubc<=759 | 920<=ubc<=929
Day-by-day	
Since admission	Total direct on *n*th day since admission
Prior to discharge	Total direct on *n*th day prior to discharge
By PC day	Total direct on *n*th day prior to/since first consult

Prior to any analysis I adjusted for regional variation in health care costs using the Medicare Wage Index (Center for Medicare & Medicaid Services, 2011) and standardised all costs to US dollars (USD) in 2011, the final year of data collection, using the Consumer Price Index (Bureau of Labor Statistics, 2014).[52]

The primary dependent variable in this thesis was selected as total direct hospital costs.

[52] Therefore throughout the thesis, all figures reported in $ are for USD adjusted to 2011.

Direct cost data were extracted from hospital accounting databases, and therefore reflect the USD cost to the hospital of providing each medication, procedure or service to each patient. Indirect cost data reflect a hospital administrative formula rather than case-specific costs; indirect costs are approximately 55% of direct costs for patients at all sites. Given the lack of clarity in the generation of indirect costs, as well as the fact that programmes cannot reduce indirect costs in the short run (the costs of running the hospital would still exist), direct costs were considered the more reliable and appropriate (Weinstein et al., 1996). Sensitivity analysis on primary analyses demonstrated absolute consistency between results with *total direct* and *total total* as dependent variables, reflecting the fixed ratio between the two, and as such *indirect costs* and *total total* costs are not considered further in this thesis.[53]

Secondary dependent variables were identified as LOS, daily direct costs, and total direct costs specific to sub-categories: room and board, ICU, pharmacy, imaging and laboratory.

Additional outcomes of interest, not used in this thesis, included: CTM, a survey of patients and family members on their experience of transition from hospital to an alternative setting (Coleman et al., 2005); FAMCARE, which measures family satisfaction with advanced cancer care (Kristjanson, 1993); and – for patients who died in hospital - ABFMI, which evaluates the experience of bereaved family members (Teno et al., 2001).[54]

---

[53] See also **§2.4>Total versus direct cost as a dependent variable** for a discussion of this point in the context of the prior literature.

[54] These data had not been collected and cleaned at the time of thesis analyses. For an overview of how I intend to employ these outcome data in post-doctoral cost-consequence analysis, see **§5.3.3>Further research with the 'PC4C' data will extend and strengthen my results**.

Independent variables

Data were collected to describe the sample at enrolment, to calculate the propensity to be seen by a PCCT during hospital admission, and to model the effects of treatment on dependent variables of interest. These are summarised in Table 3.1.2.

Table 3.1.2 Independent variables collected in 'PC4C' study

	Measure	Source
Study site	Number (1-5)	Medical record
Age	Years	Medical record
Gender	Female, male	Medical record
Race	Asian, Black, Hispanic Black, Hispanic White, Native American, White, Other	Patient interview
Insurance*	Medicare, Medicaid, third payer	Administrative database
Education attained	Elementary school, high school, college graduate	Patient interview
Advance directives	Living will, proxy, none	Patient interview
Primary diagnosis	Tumour type	Medical record
Comorbidities	Total (Elixhauser et al., 1998)	Medical record
Complication	Presence of a hospital complication	Medical record
Symptom burden	ESAS (Bruera et al., 1991) & CMSAS (Chang et al., 2004)	Patient interview
Activity level	ADL & IADL (Katz et al., 1963)	Patient interview
Medications	Equivalent morphine dose (mg) in week prior	Patient interview
Home health aide	Total hours in two weeks prior	Patient interview
Formal homecare	Yes/No in two weeks prior	Patient interview

* Medicare is the federal insurance programme for people over 65 (also covers some younger people with disabilities or specific conditions); Medicaid is the joint federal and state programme for people with limited income (Centers for Medicare & Medicaid Services, 2015). Third-payer is any other health insurance coverage.

Independent variables therefore include demographic descriptives (e.g. age, gender, race), markers of socioeconomic status (e.g. education level, insurance programme),

clinical factors (e.g. diagnosis, comorbidities, functional status) and interaction with formal health services prior to admission (e.g. morphine dose, homecare access).

3.1.7 Data sources/measurement

Initial data collection occurred following enrolment and informed consent. Clinical interviewers met with patients to collect socio-demographic and clinical information.

Daily follow-up interviews were conducted to collect assessment of pain and other symptoms.

Medical record review was performed after hospital discharge; research assistants extracted information from medical records on characteristics, medications and clinical factors.

Cost data were extracted from hospital databases after discharge.

3.1.8 Study size

The 'PC4C' sample size reflects all patients who were eligible for the study during the period of data collection, who gave informed consent to participate, whose physician gave consent for enrolment, and for whom adequate data was subsequently collected for both matching methods (propensity score weights) and statistical analysis (regression of hospital costs on treatment and all observed relevant confounders).[55]

3.1.9 Quantitative variables

For an overview of which variables were collected in the 'PC4C' dataset using which measures, see **§3.1.6**.

---

[55] For a detailed overview of this process and how the final sample was reached, see **§4.1**.

In analyses in **§4 Results**, the principal sample was balanced across treatment and comparison groups on 33 covariates.

Specifically I used binary variables for age (three strata), gender, race (three strata), insurance status (three strata), education level (three strata), advance directive status, lymphoma diagnosis, specific activities of daily living (ADL-6 (Katz et al., 1963)), visiting home services prior to admission, and for levels of pain (three strata) and fatigue (two strata). I used continuous variables for comorbidities (Elixhauser index (Elixhauser et al., 1998)), mean physical and psychological Edmonton System Assessment Scale (ESAS) scores (Bruera et al., 1991), and the Condensed Memorial Symptom Assessment Scale (CMSAS) (Chang et al., 2004)). See Table 3.2.1 in **§3.2.1** for further details.

3.1.10 Statistical software

All analyses in this thesis were performed using Stata version 12 (StataCorp, 2011).

3.2 Methods

3.2.1 Bias & confounding: propensity score methods

Background

The fundamental purpose of the 'PC4C' study is causal inference of the relationship between an intervention (treatment from a PCCT) and multiple outcomes for a sample of patients admitted to hospital with advanced cancer. Specifically this PhD thesis is interested in the causal relationship between PCCT treatment and economic variables, e.g. hospital costs and utilisation patterns.

Identifying and measuring causal relationships is a comparison of potential values:

> *"[T]he causal effect for individual* i *is the comparison of individual* i*'s outcome if individual* i *receives the treatment (the potential outcome under treatment),* $Y_i(1)$*, and individual* i*'s outcome if individual* i *receives the control (the potential outcome under control),* $Y_i(0)$*."* – Stuart (2010), p. 2-3.

Since each individual can only experience either the treatment or the control arm, and so receive one respective outcome, it is not possible to obtain both $Y_i(1)$ and $Y_i(0)$ for any given individual *i*. Rather, causal relationships are inferred by estimating the counterfactual: what would have been the unobserved potential outcome (and what is the difference between this and the observed outcome)? (aka the Rubin Causal Model, Rubin (1974))

Optimal estimation of $Y(0)$ for the treatment group and $Y(1)$ for the control group requires that the two groups do not differ on variables prior to treatment that could impact the outcome of interest. Inherent to RCT design is that treatment and control arms will differ only randomly on both observed and unobserved covariates. The challenge with observational data is that patient characteristics in either group are

potentially associated with both treatment and outcome. This gives rise to different distributions of covariates across treatment and comparison groups, biasing estimation of treatment effects. Minimising differences between groups in the means and variances of baseline covariates is essential to maximise the accuracy of treatment-effect estimates with observational data.

Propensity score methods - overview

Propensity score methods are designed to account for different distributions of covariates in treatment and comparison groups (Rosenbaum and Rubin, 1983). Specifically, a propensity score is a single score representing the probability of receiving treatment given observed covariates. The lack of a treatment assignment mechanism with observational studies is shown to be ignorable contingent on two assumptions. First, treatment assignment is independent of the outcomes of interest $Y_i(1)$ and $Y_i(0)$ given the observed covariates. Second, that all subjects have a non-zero prior probability of receiving both the treatment (1) and the control (0) (Austin, 2011, Stuart, 2010).

Propensity scores have a number of well-established advantages. For research questions where an RCT is inappropriate, propensity scores preserve important advantages of the RCT design. First, observed covariates are balanced between the treatment and comparison groups, reducing concerns about confounding. Second, they allow estimation of marginal/population-average treatment effects, which are particularly useful for cost data given well known challenges in estimating incremental effect on cost (Austin, 2011, Manning, 1998). Third, propensity scores for each subject are calculated before analysis of the outcome of interest (Rubin, 2007).

Additionally they have advantages over previous methods with observational data, such as multivariate regression and matching. For studies such as the 'PC4C' study with a high number of confounders – e.g. demographic, socio-economic, clinical, system – finding exact matches is not realistic. Propensity scores address this difficulty by reflecting all observed confounders in one score. As such, propensity scores can be viewed as an advanced matching technique (Garrido et al., 2014). Propensity scores are complementary to (and optimally used with) regression methods, and allow diagnostic analysis of their performance to validate causal inferences (Stuart, 2010).

The principal limitation of propensity scores is that they do not account for unobserved covariates, and that in the presence of unobserved covariates they may exacerbate the inaccuracy of analyses performed in the presence of such covariates (Brooks and Ohsfeldt, 2013); analysis and reporting must be conducted in the context of this fact.[56]

Propensity score methods - applied

Introduction

For the principal analysis of this thesis, patients were separated by discharge status consistent with previous studies in this area (see **§2 Literature Review**).

Of the 1,023 patients, 969 were discharged alive, 51 died during hospitalisation and three had missing survival data.[57] The 54 patients not confirmed as discharged alive do not constitute a sufficient sub-sample for a propensity score and stand-alone analysis.

---

[56] In my primary analysis I check the robustness of results to propensity score weights by checking that the principal findings are robust without weights in model selection (**§4.2**), estimation of PCCT cost effects (**§4.4** and **§4.5**), and evaluation of patient- and service-level factor impact on costs (**§4.6**).

[57] For a detailed explanation of how this sample size was reached from the start of recruitment, see **§4.1**.

The principal analysis is therefore performed on patients discharged alive (n=969), of whom 256 were seen by a PCCT in hospital (PC group) and 713 received usual care only (UC group).

Creating a propensity score for the principal sample: step-by-step

Project colleagues at the Icahn School of Medicine at Mount Sinai used the 'PC4C' dataset to write a detailed guide with the steps to calculate and evaluate propensity scores for dichotomous treatment (Garrido et al., 2014).

I followed these steps in consultation with colleagues (Melissa M. Garrido, PhD , a statistician experienced in propensity score matching and calculation, and R. Sean Morrison, MD, a palliative care consultant experienced both in clinical practice and applied use of propensity scores in health services research) to create a propensity score for my principal sample as follows:

1. *Choice of variables*

Variables should be included in the propensity score where researchers suspect they are associated with treatment and the dependent variable.

In consultation with Morrison and Garrido I identified the covariates considered most likely to exhibit such an association. For economic evaluation of all patients discharged alive (n=969), I balanced the treatment and control groups for 33 demographic, socio-economic, clinical and service covariates. These are given in Table 3.2.1.

The only covariates that Garrido, Morrison and I identified *a priori* as being important but are not included in the balanced propensity score are specific cancer diagnoses. It was not possible to get propensity scores to balance for eight specific cancer diagnoses using a sub-sample with n=969. Instead we identified lymphoma as the

most distinct in terms of clinical need and financial implication and included this as a propensity score covariate.

Table 3.2.1 Propensity score for principal sample (n=969); variables included with common support (steps 1 & 2, Garrido et al. (2014))

	ESAS Physical: *At admission (mean)*
Age: *55 to75*	**ESAS Psychological:** *At admission (mean)*
Age: *Over 75*	**ESAS Physical:** *At consult/ref day (mean)* [d]
Gender: *Female*	**ESAS Psychol.:** *At consult/ref day (mean)* [d]
Race: *White*	**CMSAS (Number)**[e]: At admission (mean)
Race: *Black*	**CMSAS (Number)**[e]: At consult/ref day (mean)[d]
Living Will: *Yes*	**CMSAS (Severity)**[e]: At admission (mean)
Proxy: *Yes*	**CMSAS (Severity)**[e]: At consult/ref day (mean)[d]
Insurance: *Medicare only*	**Morphine:** *Equivalent dose (mg)* [f ‡]
Insurance: *Medicaid (& Medicare)*	**Pain:** *Somewhat* [g]
Education: *High school* [a]	**Pain:** *Quite a bit* [g]
Education: *College* [a]	**Pain:** *Very much* [g]
Visiting nurse services: *Yes* [b]	**Fatigue:** *A little, somewhat or quite a bit* [g]
Home health aide: *Total hours* [b ‡]	**Fatigue:** *Very much* [g]
Primary diagnosis: *Lymphoma/myeloma*	**ADL:** *Needs partial assistance bathing*
Complication(s): *Yes* [c]	**ADL:** *Needs partial assist transferring from chair*
Comorbidities: *Elixhauser total (mean)*	**ADL:** *Needs complete assistance with 1+ activity*

References cases for binary variables: Age: under 55; Race: Other (neither white nor black); Insurance: third payer (neither Medicare nor Medicaid); Education: Elementary School; Diagnosis: Solid or gynaecological; Pain: A little or none; Fatigue: None.

[a] Highest level attained. [b] In two weeks prior to hospitalisation. [c] Presence of a major/minor complication prior to consultation/reference day during the hospitalisation. [d] For PC patients, the reference day was the day of consult; for UC patients, the reference day is the day they had the most similar symptom severity to PC patients. [e] CMSAS (Number): Number of physical symptoms on the CMSAS; CMSAS (Severity): Number of physical symptoms multiplied by the mean severity of physical symptoms on the CMSAS. [f] In week prior to hospitalisation. [g] Taken at consult/reference day with 'None' as the reference category. ‡ Data retransformed as a square root for calculating the propensity score.

2. *Balance across groups*

Stata command –*pscore*– confirms 'common support', i.e. the overlap between subjects in treatment and comparison groups. In the event that common support is inadequate, researchers must adjust the covariates or functional form of covariates included. In the case of the propensity score for the principal sample (n=969), common support was achieved using those variables listed in Table 3.2.1.

3. *Balance of covariates within blocks*

Using Stata output from the –*pscore*- command in Step 2, standardised differences for each covariate within each propensity score block have to be calculated by hand (using the formulae from, for example, Austin (2009), p. 3087).

For continuous variables the standardised difference is calculated as:

$$d = \frac{(\bar{x_{PC}} - \bar{x_{UC}})}{\sqrt{\frac{s_{PC}^2 - s_{UC}^2}{2}}}$$

Where $\bar{x}_{PC}$ and $\bar{x}_{UC}$ are the sample means for each covariate in the palliative care (treatment) and usual care (control) groups respectively, and the s^2 entries are the corresponding variances.

For binary variables this is calculated as:

$$d = \frac{(\hat{p}_{PC} - \hat{p}_{UC})}{\sqrt{\frac{\hat{p}_{PC}(1-\hat{p}_{PC}) + \hat{p}_{UC}(1-\hat{p}_{UC})}{2}}}$$

Where $\hat{p}_{PC}$ and $\hat{p}_{UC}$ are the means for each covariate in the treatment and control groups respectively.

There is no established definition of acceptable difference but some degree of imbalance is expected at this stage; perfect balance is achieved only in RCTs with very large sample sizes (Austin, 2009). An additional consideration in evaluating standardised differences is that greater imbalance is expected in blocks in the tails of the distribution, and is of less concern than persistent differences in the central blocks (Garrido et al., 2014).

In consultation with Garrido and Morrison I used 25% as the guideline limit: if covariates are balanced (d<=0.25) in 75% of cases then this was considered acceptable at this stage.

The *–pscore-* command in Step 2 balanced the principal sample divided into seven blocks. A summary of the balance of the 33 covariates within each of the seven blocks is given in Table 3.2.2. As expected, the imbalance in the extreme tails (particularly blocks 1, 2 and 6) is a little large but the balance in the central blocks is stronger and the overall balance (80% of covariates with d<=0.25) is acceptable.

Table 3.2.2 Propensity score for principal sample (n=969); checking standardised differences within blocks prior to weighting (step 3, Garrido et al. (2014))

Block	1	2	3	4	5	6	7	Total
% variables with d>0.25	33%	27%	15%	0%	21%	30%	12%	20%

4. Choice of matching and weighting strategies

Having achieved a balanced propensity score, analysts must choose how to apply the score in comparison of groups.

The best known comparison strategy is matching, where each subject in the treatment group is matched to a subject in the comparison group with the nearest

propensity score. These matches can be one-to-one or one-to-many as sample size and analyst judgement dictate.

Two recommended alternatives are kernel weighting and inverse-probability treatment weights (IPTW) (Garrido et al., 2014). With kernel weighting each subject in the treatment group is given a weighting of 1 and matched to a subject in the comparison group using a weighted average of the propensity score covariates. With IPTW each subject in the treatment group is given a weight of the inverse of the propensity score, and each subject in the comparison group is given a weight of the inverse of one minus the propensity score.

Selection of a matching and weighting strategy should always be informed by comparing the performance of different approaches in balancing treatment and comparison groups with the data, which occurs in Step 5.

5. *Balance of covariates after weighting*

Using Stata output from the *–pstest-* command, mean standardised differences for the treatment and comparison groups can be compared both overall and for individual covariates.

The overall performance of one-to-one and one-to-three matching, and kernel and IPTW weighting is summarised in Table 3.2.3. All approaches preserve the sample size of 969 and reduce substantially the original sample bias. The best-performing approach is kernel matching, with mean bias of 1.8% after matching.

Table 3.2.3 Propensity score for principal sample (n=969); comparing matching strategies on support and overall bias (steps 4 & 5, Garrido et al. (2014))

Sample type	All (n=)	PC (n=)	UC (n=)	Mean standardised difference in covariates (%)	Median standardised difference in covariates (%)
Original sample	969	256	713	25.4	22.8
Calliper 1:1	969	256	713	5.0	3.9
Calliper 1:3	969	256	713	2.3	1.8
Kernel matching	969	256	713	1.8	1.3
IPTW	969	256	713	4.2	2.9

The performance of all five strategies for matching on individual covariates is given in Table 3.2.4. As in Step 3 there is no standardised definition of acceptable difference but after matching I applied a more stringent guideline for balance than previously: no more than 10% is an acceptable level of difference for each covariate. All instances where a difference exceeds 10% are highlighted in bold. Two approaches have all covariates within 10%: 1:3 matching and kernel weights. Kernel weights is the dominant strategy in Table 3.2.3 and has all variables within 10% in Table 3.2.4, and this is therefore identified as the preferred matching strategy.

Table 3.2.4 Propensity score for principal sample (n=969); comparing matching strategies on standardised differences for each covariate (steps 4 & 5, Garrido et al. (2014))

	Standardised differences				
	Original sample	**1:1 matching**	**1:3 matching**	**Kernel weights**	**IPTW**
Age: *55 to75*	-9.7%	-4.7%	2.1%	3.8%	**-15.5%**
Age: *Over 75*	-1.8%	0.0%	-4.1%	-6.4%	0.5%
Gender: *Female*	-5.2%	0.8%	-0.3%	-3.0%	2.5%
Race: *White*	**-14.7%**	-1.6%	-1.1%	1.1%	**-11.2%**
Race: *Black*	**20.8%**	7.8%	2.9%	-0.6%	-2.9%
Living Will: *Yes*	-9.4%	0.0%	0.8%	0.1%	-5.3%
Proxy: *Yes*	**-22.8%**	-7.1%	-2.1%	-1.9%	4.1%
Insurance: *Medicare only*	3.7%	-4.0%	-1.7%	0.0%	-3.1%
Insurance: *Medicaid*	**28.0%**	6.9%	3.6%	-0.9%	2.4%
Education: *High school*	**32.8%**	7.1%	-1.1%	1.0%	-1.6%
Education: *College*	**-35.5%**	-6.4%	-1.6%	-1.3%	2.9%
Visiting nurse services: *Yes*	7.4%	1.2%	-1.2%	-1.4%	-5.2%
Home health aide: *Total hour*	2.2%	9.0%	2.1%	0.0%	-5.6%
Primary diagnosis: *Lymphoma*	**-10.4%**	-6.0%	-3.5%	-0.6%	**12.4%**
Complication(s): *Yes*	**-15. 0%**	2.1%	0.7%	-1.8%	-4.9%
Comorbidities: *Elixhauser*	**38.1%**	2.2%	3.4%	1.1%	3.7%
ADL: *bathing*	**36.0%**	2.6%	0.9%	-1.0%	-1.0%
ADL: *chair*	**40.7%**	7.1%	3.6%	2.3%	2.6%
ADL: *1+ activity*	**20.6%**	7.4%	2.5%	0.9%	-2.7%
ESAS Physical: *Admission*	**62.6%**	-2.4%	-3.9%	-1.1%	1.4%
ESAS Psychological: *Admission*	**12.4%**	-0.2%	1.1%	2.8%	-6.4%
ESAS Physical: *Consult/ref day*	**65.5%**	**12.7%**	0.6%	0.7%	-5.4%
ESAS Psychol.: *Consult/ref day*	**23.7%**	-2.9%	-7.0%	-4.5%	-0.3%
CMSAS #: *Admission*	**49.7%**	-3.9%	-5.1%	-2.0%	-5.8%
CMSAS #: *Consult/ref day*	**31.9%**	3.6%	-2.3%	-1.3%	**-16.2%**
CMSAS Severity: *Admission*	**56.3%**	-2.4%	-4.5%	-1.3%	2.6%
CMSAS Severity: *Consult/ref*	**52.9%**	**12.7%**	1.6%	1.5%	-5.5%
Morphine: *Equivalent (mg)*	**32.1%**	3.3%	0.6%	4.8%	-0.6%
Pain: *Somewhat*	-8.1%	-3.5%	3.1%	2.5%	-5.4%
Pain: *Quite a bit*	-9.7%	-2.7%	1.5%	-1.9%	0.1%
Pain: *Very much*	-1.8%	**14.8%**	-3.4%	3.1%	-0.4%
Fatigue: *Little,somewhat,q. a bit*	-5.2%	-6.5%	-1.6%	-1.1%	0.2%
Fatigue: *Very much*	**-14.7%**	**10.0%**	1.8%	1.7%	-0.3%

For full description of each covariate see Table 3.2.1.

6. Estimation and interpretation of cost effects

The primary analysis in this PhD thesis seeks to estimate the impact of PCCT interventions on selected economic variables and primarily total direct hospital costs.

For this purpose the average treatment effect (ATE) is the most appropriate measure of inferring a causal relationship between treatment and cost, a.k.a. mean incremental effect. [58] The ATE represents the estimated impact on the dependent variable of moving a patient from the comparison group to the treatment group holding all other values constant[59]; standard errors are adjusted for by bootstrap methods (1000 replications) (Abadie and Imbens, 2008).

Additional propensity scores

In **§4 Results** I perform multiple analyses on the 'PC4C' dataset where the sample is defined differently to the principal sample. For example, I pool patients discharged alive with those who died during hospitalisation, I trim the sample by LOS, I alter the definition of treatment, and I re-define the sample according to patient complexity.

On each occasion the sample was altered from the principal sample profile (n_{ALL}=969; n_{UC}=713, n_{PC}=256) it was necessary to recalculate the propensity score to ensure that any sub-sample was balanced across treatment and comparison groups only according to the baseline characteristics of patients included in that analysis. Balancing all patients included in an analysis (but no other patients included in that

---

[58] In economics it is not typical to refer to 'treatment effect' where cost is the dependent variable, although this is not unusual in the health services research literature. Since this is a PhD thesis in health economics, I follow the economics tradition: in my primary analysis, estimated causal impact of the intervention on costs is reported as an 'estimated incremental effect'. See also '15 Outcome data' in Appendix to §3.

[59] Values held constant are all predictors in the regression other than the treatment variable. Unless otherwise stated, these predictors are all independent variables in Table 3.2.1 and fixed effects for each hospital site. The list of predictors used in regressions is detailed in the respective Methods sections of **§4.2**, **§4.3**, **§4.4**, **§4.5** and **§4.6**.

analysis) is essential to the accurate estimation of a counterfactual that is central to propensity score methodology (Rubin, 2007). Specifically, I calculated a new propensity score weight for each sub-sample (Green and Stuart, 2014).

For each additional propensity score I followed the step-by-step guide by Garrido et al. (2014) as above. I minimised differences between additional scores and the score for the principal sample to ensure comparability of results using different samples.

In steps 1-3 where possible I balanced treatment and comparison groups on the same fixed list of 33 covariates used for the principal sample (Table 3.2.1); where this exact specification did not balance I altered the list of 33 using clear strategies that targeted less important variables and minimised information loss. It is acceptable for different samples to be balanced on different sets of relevant propensity scores but preferable to minimise information loss by transforming variables or removing theoretically less important ones (Green and Stuart, 2014, Garrido et al., 2014).

Specifically, where the sub-sample would not balance according to the list of 33 covariates in Table 3.2.1, I employed the following strategies (in order of preference):

I. **Transformation:** I transformed continuous variables, for example, balancing groups on the square root of raw data rather than the raw data itself;

II. **Re-categorisation:** instead of categorising age as three separate binary variables I employed it as a single continuous variable;

III. **Combination:** I combined variables that were conceptually similar; for example, combining 'living will' and 'proxy' into a single variable 'advance directive', or summing variables that were taken at both admission and reference day (ESAS physical and psychological scores, CMSAS number and severity scores) into an overall variable for each;

IV. **Removal:** I removed the pain and fatigue dummy variables from the original 33 since these are indirectly included in the propensity score already via CMSAS inventories. While controlling for these factors specifically was optimal, Morrison, Garrido and I agreed that removing these variables was acceptable since information loss would be minimised by their inclusion through CMSAS.

In steps 4-5 I used kernel weighting in all cases. First, of the strategies put forward by Garrido et al. this approach consistently performed best in primary analyses.[60] Second, consistency of weighting strategy improves the confidence with which results can be directly compared. In step 6 I use ATE in all cases on the same basis.

Throughout **§4 Results** I signal clearly on each occasion a new sub-sample has been created and indicate where an overview of the propensity score process for sub-samples can be found in a marked appendix.

---

[60] I confirmed that my choice of weighting strategy for the principal sample is robust for other samples by comparing the different alternatives in Table 3.2.3 using the samples in primary analyses in **§4.4** and **§4.5**. Kernel weighting was always the best performing for minimising bias and never suffered greater loss to matching than the second best performing strategy on bias. Kernel weighting is therefore employed as the matching strategy for all analyses in this thesis.

Instrumental variable approach

It is optimal to supplement propensity score methods by using an instrumental variable (IV) to control for unobserved confounding in observational data (Penrod et al., 2009). An instrumental variable is one that is associated with variation in treatment but does not exogenously impact the dependent variable (Angrist et al., 1996).

A common such instrument is identity of physician, provided physicians are randomly allocated to patients (Penrod et al., 2010). However, patients were not randomly assigned to physicians in the 'PC4C' study, and neither I nor any colleague has been able to identify a valid alternative instrument for use with this dataset. This is not unusual - a strong, valid IV is difficult to identify (Murray, 2006), and only Penrod et al. (2010) in previous PCCT studies have employed the IV method (see **§2 Literature Review**).

The analysis in this thesis is therefore performed without an IV approach.

Summary

Propensity score weights are a well-established tool for balancing the distribution of covariates between treatment and comparison groups in observational data analysis. Using guidelines designed by 'PC4C' colleagues I calculated and validated propensity scores for the principal sample for economic evaluation: all patients for whom sufficient data were collected and who were discharged from hospital alive. I repeated this process for all sub-samples of patients used in this thesis. It was not possible to identify a valid instrumental variable for this dataset.

3.2.2 Statistical analysis: modelling costs & estimating cost effects

Background

Healthcare cost distributions typically exhibit features that pose well-known problems for statistical analysis (Jones et al., 2013a):

- *Non-negativity:* cost data are by definition never less than zero.
- *Mass of zero-value observations:* in data drawn from populations (not only recipients of healthcare) a large number of cost data-points will be zero.
- *Positive skew:* a minority of patients may incur a disproportionately high level of costs for the sample, skewing the distribution right.
- *Heteroscedasticity:* variability of costs is unequal across a range of values for important predictors;
- *Leptokurtosis:* clustering of cost observations for a large number of patients with similar care trajectories may result in high 'peaked-ness' of the distribution.

Each of these potentially undermines the assumptions on which routine statistical analyses are based. In the case of the 'PC4C' data, all are potential concerns except for the mass of zero-value observations (there is no-one in the sample who did not receive healthcare and so not accrue costs). Selection of modelling approach to the 'PC4C' data therefore requires careful consideration.

Standard linear regression (Ordinary Least Squares (OLS)) is sensitive to extreme outliers and so is generally not appropriate for healthcare cost data exhibiting skewness, leptokurtosis and heteroscedasticity (Jones, 2010). A rudimentary alternative is to convert skewed cost distributions, typically using log-transformation

but also sometimes square root, to something approaching symmetry and then perform linear regression. However, these often exhibit poor performance, do not address heteroscedasticity and lead to a new problem; the dependent variable (log or square root of costs) no longer has real-world meaning or application, and retransformation of such output is not straightforward:

> *"Although* [log-transformed] *estimates may be more precise and robust* [than using highly skewed distributions of untransformed costs]*, no one is interested in log model results on the log scale* per se*. Congress does not appropriate log dollars. First Bank will not cash a check for log dollars. Instead, the log scale results must be retransformed to the original scale so that one can comment on the average or total response to a covariate* x*. There is a very real danger that the log scale results may provide a very misleading, incomplete, and biased estimate of the impact of covariates on the untransformed scale, which is usually the scale of ultimate interest."* - Manning (1998), p. 285.

These challenges have led health econometricians to develop nonlinear approaches to the analysis of health care cost data and the accurate estimation of $ effects.

Modelling health care costs

The best established model in the health economics literature is the generalised linear model (GLM) (Jones, 2010), and this is also the case specifically in economic studies of PCCTs (see **§2 Literature Review**). Since GLMs estimate effects using the raw cost scale they avoid the transformation problem, and also allow heteroscedasticity to be modelled given certain assumptions (Jones, 2010).

There are two components to the GLM approach, each specified directly by the analyst.

First, a link function is specified describing how mean costs relate to the regressors. Of those provided in statistical software (for example, there are seven GLM link options in Stata version 12), three are typically considered suitable for analysing cost data: log, power and identity (Jones et al., 2013b, Blough et al., 1999, Manning et al., 2005, Manning and Mullahy, 2001).

Second, a distributional family is specified to characterise the relationship between conditional mean and the conditional variance. The most common family distributions are Gaussian (constant variance), Poisson (variance proportional to mean), gamma (variance proportional to square of mean) and inverse Gaussian (variance proportional to cube of the mean) (Jones et al., 2013b, Blough et al., 1999, Manning et al., 2005, Manning and Mullahy, 2001).

Three link and four family options give a total of 12 possible specifications in using a GLM to analyse healthcare costs. Of these, the most common specification in the general health economic literature and the only specification in analyses of PCCTs is GLM (gamma, log) (Jones, 2010).

Non-linear alternatives to GLMs are increasingly visible in the literature (Jones et al., 2013a). These include exponential conditional mean models (ECMs, Jones (2010)); generalised gamma models (GGMs, Manning et al. (2005)); extended estimation equations (EEEs, Basu and Rathouz (2005)); and Finite Mixture Models (FMMs, Conway and Deb (2005)).

These and the better established options for modelling health care costs are summarised in Table 3.2.5.

Table 3.2.5 Summary of principal modelling approaches in analysis of health care costs

Model	Notes
OLS on y	Simple regression model, the most common in orthodox statistics. Typically considered inappropriate for health cost data due to poor performance with skew, leptokurtosis, heteroscedasticity and non-negative values.
OLS on ln(y)	Costs (y) log-transformed to address distributional challenges. Retransforming logged outputs to generate a $-effect estimation is problematic. Alternative: OLS on (y)^0.5.
GLM	Dominant alternative to OLS in recent health economics literature. Avoids retransformation problem and allows heteroscedasticity to be modelled. Analyst specifies one of three link functions (log, identity, power) and four distributional families (Gaussian, Poisson, gamma, inverse Gaussian) to give 12 possible specifications with health cost data.
ECM	Nonlinear specification of the conditional means addresses non-negativity and skewness. Basic approach is nonlinear least squares (NLS). Alternative: Poisson distribution, which is most common for count data.
GGM	Extension of ECM and hazard models commonly used for duration data, also addresses distributional issues. Basic approach assumed homoscedasticity. Alternative: extension allowing for heteroscedasticity.
EEE	Extension of GLM methods that estimates directly appropriate link function and family distribution.
FMM	Alternative approach modelling data assuming the presence of heterogeneous classes. Most commonly used for multi-modal distributions, e.g. with mass of zero-value observations.

I noted in **§2 Literature Review** that while the most robust recent PCCT cost studies all employ GLMs rather than simple linear regression, no study reports evaluating performance of different modelling approaches prior to selecting the GLM. Comparative performance of modelling options suggests that different models perform better according to different criteria and with different datasets, and that

estimation of cost effects should not be performed without first assessing model performance with the relevant dataset (Garrido et al., 2012, Jones et al., 2013b).

Prior to undertaking primary analysis I must therefore compare performance of the modelling options with the 'PC4C' data.[61]

Summary

The distributions of health care cost data typically pose problems in statistical analysis and cost-effect estimation. Various modelling approaches are available to analyse cost data, the most popular of which is GLM (gamma, log). Evaluation of different approaches suggests that no model is dominant and that alternatives should be compared on performance with a specific dataset prior to cost-effect estimation.

[61] For the results of the applied model comparisons with 'PC4C' data, see **§4.2**.

3.3 Summary

The 'PC4C' study was a prospective, cohort study to evaluate the impact of PCCT interventions for patients admitted to hospital with advanced cancer. Clinical and administrative data including cost data from the hospital perspective were collected at five United States hospitals from 2007 to 2011. I accessed the data during an extended period of mentored research at Icahn School of Medicine at Mount Sinai, New York.

This thesis is intended to expand the existing evidence base on economic impact of PCCTs. My primary analysis will evaluate the causal relationship between PCCT treatment and hospital utilisation, based on a 'non-inferiority' assumption that outcomes are at least no worse for intervention group patients and comparison group patients. To address confounding, treatment and comparison groups will be balanced on a large number of baseline covariates using propensity scoring. In the context of cost data distributions which may be problematic for linear statistical approaches, multiple modelling methods will be compared prior to cost-effect estimation.

Specifically in the context of prior studies and my systematic literature review in the previous chapter, I have identified six objectives towards my aim.

Prior to primary analysis I must (i) assess and describe the characteristics of cost data, and (ii) compare model performance with the 'PC4C' cost data so as to maximise the accuracy of cost-effect estimates.

For my primary analysis I intend to examine (iii) controlling for LOS to account for unobserved baseline heterogeneity in cost-effect reporting, (iv) estimation of PCCT impact on cost using appropriate methods, including identification of patient- and/or

system-level factors associated with that impact, (v) examination of the underlying source of any cost differences associated with treatment, and (vi) identification of patient- and/or system-level factors associated with hospital utilisation other than palliative care.

Chapter 4: Results

Abstract

The previous two chapters have provided a detailed contextual basis for my original research.

In **§2 Literature Review** I assembled systematically and assessed critically the economic evidence on palliative care consultations team (PCCT) interventions. Published studies to date demonstrate a clear pattern of cost-saving effect, but these results have been derived using a narrow (and in some cases flawed) methodological framework. An important consideration is the consistent use of length of stay (LOS) to control for unobserved baseline heterogeneity, strategies that may introduce endogeneity into analysis and arbitrarily impact reported results. Also highlighted is the likely existence of cost-effect heterogeneity; most studies to date have focused exclusively on mean overall effects for their sample when there may be important differences of effect within the sample. Identifying factors in *how* and *for whom* treatment has a greater or lesser effect with a sample of patients with serious illness would constitute a substantial extension of the existing evidence base on palliative care programmes.

In **§3 Materials and Methods** I summarised the 'Palliative Care for Cancer' ('PC4C') study and my methodological approach consistent with items 2-12 in the STROBE guidelines (von Elm et al., 2007). I outlined the study design, patient recruitment and data collection by colleagues in assembling the 'PC4C' dataset, and summarised my methods in analysing this dataset, particularly with regard to propensity score matching to minimise bias. I also reviewed the importance of evaluating and selecting a model for cost-effect estimation that is both appropriate to the

characteristics of health utilisation data and performs strongly on evaluative measures with the specific dataset.

In **§3.1.2**, informed by **§2 Literature Review**, I identified six objectives for my own primary research, which takes a cost analysis approach: an assessment of intervention impact on costs, assuming that outcomes for intervention patients are at least no worse than that comparator.

Prior to main analysis:

(i) **Data overview:** careful description of the cost data to establish their characteristics and consider appropriate modelling approaches;

(ii) **Model evaluation and selection:** comparative evaluation of different available modelling approaches to assess which models perform best with the study data.

In main analysis:

(iii) **Comparative evaluation of methods using LOS to control for unobserved heterogeneity:** consideration of the impact that controlling for LOS has on cost-effect estimates;

(iv) **Cost-effect estimation:** estimating PCCT impact on costs and other utilisation, including variations in effect according to system- or patient-level factors;

(v) **Source of any observable cost effects:** evaluating the underlying mechanism of any observed cost differences following treatment;

(vi) **Other determinants of utilisation:** identifying factors other than treatment that are driving hospital costs for patients with serious illness.

This chapter presents the results of my primary research, structured corresponding to these objectives and the STROBE guidelines.[62]

Objective (i): In **§4.1** I provide an overview of the 'PC4C' dataset, summarising the participants, descriptive data and cost data. I outline how the principal sample for economic evaluation (n=969) was reached and demonstrate balance between treatment and comparison groups using propensity score weights. I also show that the cost data exhibit characteristics typical for healthcare utilisation: right-hand skew, heavy tail, leptokurtosis and heteroscedasticity. Therefore prior to estimating cost effects, comparative evaluation of modelling alternatives is essential.

Objective (ii): I perform this comparative evaluation of modelling alternatives in **§4.2**. Following adapted versions of Glick (2014) and Jones et al. (2013c) I compare linear and non-linear modelling approaches to analysing the 'PC4C' cost data. I conclude that the best performing model is a generalised linear model (GLM) with a gamma distribution and a log link, and that sensitivity analyses on primary results should be performed using ordinary least squares (OLS) to confirm the robustness of findings.

Objective (iii): In **§4.3** I apply multiple approaches visible in the prior literature to the 'PC4C' data, controlling for LOS to examine impact on model performance and cost-effect estimates. I find that controlling for LOS has a substantial impact on cost-effect estimates and risks undermining scientific integrity by weakening key elements

---

[62] **§4.1** presents Items 13-15 in the STROBE guidelines. All other sections in this chapter address a specific research question and are structured in the Introduction, Methods, Results and Discussion (IMRAD) format. **§4.2-§4.6** therefore each address STROBE items 16, 17, 18 and 20 for the relevant research question. For a detailed summary of how this thesis is structured consistent with STROBE, see Appendix to §3.

of the methodology and analysis. However, disregarding LOS controls without otherwise controlling for unobserved heterogeneity in the data may result in a failure to detect important associations (i.e. a type II error). A comparison of results suggests that a small number of long-stay patients mask a cost-effect for the majority where LOS is not controlled for in any way. Deriving robust estimates therefore requires identifying this latent class using baseline data only, identifying the important relationships in the data while circumventing the weaknesses inherent to LOS controls.

Objectives (iv and v): Cost-effect estimates and their underlying mechanisms are presented in two parts.

First, I analyse the PCCT's incremental effect on cost, incorporating time-to-consult following admission in **§4.4**. I find that a consultation within six days of hospital admission is associated with a cost-saving effect and that earlier consultation is associated with a larger cost-saving effect, and I identify the source of these cost-savings. A notable secondary finding is that prompt PCCT intervention following admission reduces LOS as well as intensity of hospital stay. No previous study has reported this association, possibly because a failure to specify the intervention accurately (by incorporating timing) has resulted in a failure to detect significant difference (i.e. a type II error). These results are consistent with a growing clinical evidence base that earlier palliative care has observable benefits. In identifying these associations, which are robust to multiple sensitivity analyses, I present results that are both reliable (since they do not employ LOS-control methods) and informative (improving understanding how PCCT impact on cost can be maximised, given assumptions about outcomes). A version of the analyses presented in **§4.4** has been published by the *Journal of Clinical Oncology* (May et al., 2015).

Second, I analyse PCCT impact on cost according to patient multimorbidity in **§4.5**. I find that a consultation within 10 days of admission is significantly cost-saving for all patients with multimorbidity and that this cost-saving effect is larger for patients with higher numbers of comorbidities. Observed cost-savings again accrue primarily through reduced intensity of hospital stay and also (where the intervention is early) through reduced length of stay. The impact of time-to-consult on cost effect is also robust to the multimorbidity association: for any given sample defined by multimorbidity, earlier treatment has a larger effect; for any given definition of the intervention according to timing, the effect is greatest for the most complex patients. These results are robust to multiple sensitivity analyses and demonstrate a relationship between cost effect and multimorbidity for the first time. They have potentially important implications for the organisation of care in a context where prevalence of multimorbidity is growing rapidly, and the care provided to this group consistently yields high costs for poor outcomes.

Objective (vi): Having established methods to derive robust associations with hospital costs using this dataset, I perform a supplementary analysis in **§4.6**, examining which factors other than palliative care treatment are significantly associated with cost of hospital treatment. The majority of associations reflect the expected relationship between patient illness burden and cost of care. Also identified are associations between insurance status and education level, and cost of care, raising equity concerns about a relationship between higher socioeconomic status and higher utilisation during hospital stay.

4.1 Overview of study data

Introduction

This section provides an overview of the 'PC4C' study data, consistent with items 13-15 in the STROBE guidelines (von Elm et al., 2007). [63]

I report details of the study participants (numbers of individuals at each stage of the study) and descriptive data (characteristics of included participants), and summarise primary and secondary cost data.

Additionally I examine the cost data for characteristics such as skew and heteroscedasticity, and consider the implications for estimating mean incremental effect on cost in subsequent sections.

Participants

A total of 5,939 patients were admitted to study sites with an advanced cancer diagnosis during the study period. Of these, 1,562 (26%) refused to participate and 1,159 (20%) did not meet the eligibility criteria, meaning that 3,218 patients were enrolled in the study. There were 266 patients with missing treatment data and a further 1,626 patients for whom insufficient data was collected at baseline (to match for the propensity score) and/or on relevant outcomes (to include in analysis); high levels of attrition and missing data are a familiar challenge in studying populations with serious illness.[64]

Therefore, following data cleaning and matching, a sub-sample of 1,537 patients was identified for whom there were adequate data for matching and analysis (PC n=374, UC n=1163). PC patients were more likely than UC patients to be in the final

---

[63] For a detailed overview of this PhD thesis according to the STROBE checklist, see Appendix to §3.
[64] For a more detailed discussion of challenges in study design with seriously-ill populations, see **§3.1.3**.

analytic sample ($p<0.001$). One site ('Site 4') collected clinical, satisfaction and process data for all other parts of the 'PC4C' project but did not collect cost data. Site 4 patients (n=513) and one additional patient for whom cost data were missing were therefore excluded from economic evaluation papers.

For the principal analysis throughout this thesis, patients from four sites with cost data who died during the hospitalisation (n=54)[65] were excluded. This method has become standard in large cost analysis of palliative care programmes using observational data (see **§2 Literature Review**). The rationales for this approach are that there is likely unobserved clinical heterogeneity between patients by discharge status as well as different treatment decisions and preferences, and distinct implications of discharge (Cassel et al., 2010). Additional sensitivity analysis pooling decedents with those discharged alive is performed in subsequent sections but the relative sub-sample sizes preclude giving those discharged alive (n=969) and those who died during hospitalisation (n=54) equal consideration throughout.[66]

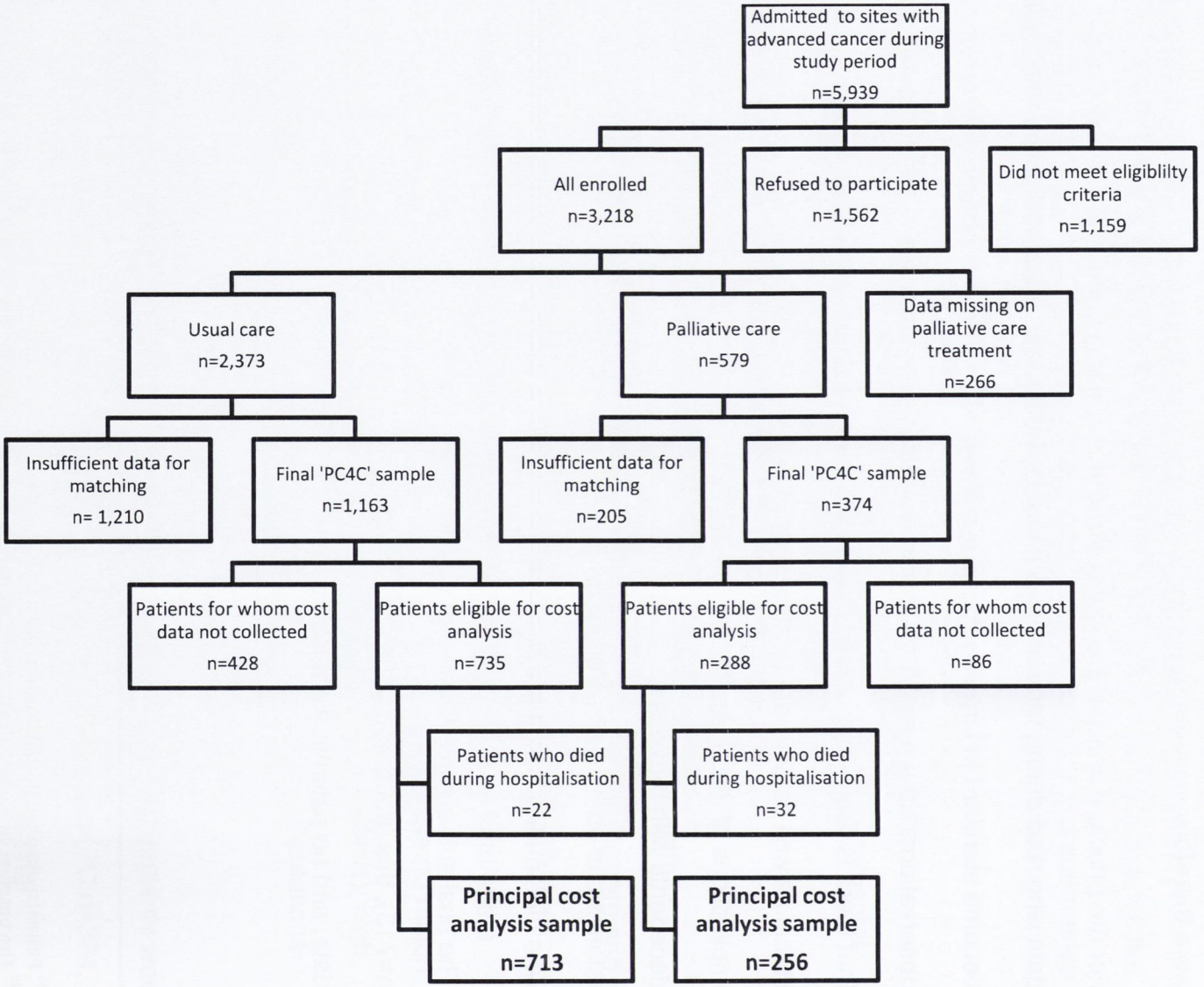
The final sub-sample for economic evaluation was therefore n=969 (PC n=256, UC n=713); this group had no missing data on any propensity score variable or cost data, and no patients were lost to matching. See Figure 4.1.1.

---

[65] Patients who died during hospitalisation, n=51; data missing on survival, n=3.

[66] The baseline characteristics of the decedents compared to those discharged alive are given in Appendix to **§4.1**. There are no significant differences on demographic, socioeconomic or diagnostic factors, but persistent differences on illness burden: the decedents sample had higher mean scores for comorbidities and symptom burden, and lower scores for activities of daily living. It is also worth noting that this study has a high 'survival' rate compared to previous studies: 95% for 'PC4C' patients compared to a median of 70% in **§2**. While this may be somewhat attributable to reduced LOS for hospital patients in recent years, it also reflects that palliative care is increasingly provided 'upstream' to patients who are not approaching end of life (personal communication, R.S. Morrison, February 2015).

Figure 4.1.1 Participants - 'PC4C' study



Descriptive data

Unweighted data

Selected baseline socio-demographic and clinical characteristics are summarised in Table 4.1.1.

The sample population is majority female, white, and aged 55 to 75. A quarter received a palliative care consultation. Across four sites, two hospitals account for 75% of patients.

An important observable difference between usual care (UC) and palliative care (PC) groups prior to matching or weighting is illness burden: palliative care patients have higher numbers of comorbidities, lower activity levels of daily living, and higher physical and psychological symptom assessment scores.

Additionally differences are observable on demographic and socioeconomic factors: palliative care patients are on average younger, with higher proportions of male patients, black patients, and patients on government insurance programmes.

Table 4.1.1 Selected statistics of unweighted principal sample (n=969) at baseline

		Usual Care (n=713)	Palliative Care (n=256)	All Patients (n=969)
Age	*<55*	215 (30.2%)	91 (35.6%)	306 (31.6%)
	55to75	416 (58.3%)	137 (53.5%)	553 (57.1%)
	75<	82 (11.5%)	28 (10.9%)	110 (11.4%)
Gender	*Female*	400 (56.1%)	137 (53.5%)	537 (55.4%)
	Male	313 (43.9%)	119 (46.5%)	432 (44.6%)
Race	*White*	490 (68.7%)	158 (61.7%)	648 (66.9%)
	Black	170 (23.8%)	85 (33.2%)	255 (26.3%)
	Other	53 (7.4%)	13 (5.1%)	66 (6.8%)
Insurance	*Medicare*	129 (18.1%)	50 (19.5%)	179 (18.5%)
	Medicaid	102 (14.3%)	65 (25.4%)	167 (17.2%)
	Other	482 (67.6%)	141 (55.1%)	623 (64.3%)
Advance Directive	*Yes*	429 (60.2%)	130 (50.8%)	559 (57.7%)
Primary Diagnosis	*GI*	159 (22.3%)	52 (20.3%)	211 (21.8%)
	Lymphoma	61 (8.6%)	15 (5.9%)	76 (7.8%)
	CNS	15 (2.1%)	2 (0.7%)	17 (1.8%)
	Breast	67 (9.4%)	29 (11.3%)	96 (9.9%)
	Gynaecological	94 (13.2%)	21 (8.2%)	115 (11.9%)
	HCC	15 (2.1%)	6 (2.3%)	21 (2.2%)
	Lung	84 (11.8%)	53 (20.7%)	137 (14.1%)
	Other	218 (30.6%)	78 (30.5%)	296 (30.5%)
Comorbidities	*Elixhauser Total*	3.22	3.96	3.41
Activity Level	*ADL-6*	10.76	9.8	10.50
Symptoms	*ESAS Physical*	1.45	1.97	1.59
	ESAS Psych.	1.45	1.61	1.49
Site	*1*	95 (13.3%)	27 (10.5%)	122 (12.6%)
	2	76 (10.7%)	30 (11.7%)	106 (10.9%)
	3	293 (41.1%)	38 (14.8%)	331 (34.2%)
	5	249 (34.9%)	161 (62.9%)	410 (42.3%)

Medicare: Patients with Medicare and no other insurance. Medicaid: Patients with Medicaid (and Medicare). This applies in all subsequent tables. For an overview of insurance statuses, see Table 3.2.1.

Weighted data

Propensity score weightings were calculated for the principal sample, balancing the treatment and comparison groups on 33 baseline covariates associated as having an important association with the likelihood to receive palliative care.[67]

The covariates encompass multiple key domains that may be associated with receipt of the intervention: demographic (e.g. age, race), socioeconomic (e.g. insurance status, education level), clinical (e.g. diagnosis, comorbidities) and system (e.g. home services accessed prior to hospital admission). A full list of the 33 included covariates, and the prevalence of each in the treatment and comparison groups following propensity score weighting, are given in Table 4.1.2. There is a demonstrable balance with 32 standardised differences no greater than 5% and all within the 10% rule of thumb for acceptable difference (Austin, 2009).

[67] See **§3.2.1** for a detailed discussion of propensity score weight application in theory and a step-by-step explanation of how the propensity score for the principal sample (n=969) was reached in practice.

Table 4.1.2 Weighted matched sample (n=969) by all 33 covariates in propensity score

	Usual Care	*Palliative Care*	*Standardised difference*
Age: *55 to75*	*51.6%*	*53.5%*	*3.8%*
Age: *Over 75*	*13.0%*	*10.9%*	*-6.4%*
Gender: *Female*	*55.0%*	*53.5%*	*-3.0%*
Race: *White*	*61.2%*	*61.7%*	*1.1%*
Race: *Black*	*33.5%*	*33.2%*	*-0.6%*
Living Will: *Yes*	*40.2%*	*40.2%*	*0.1%*
Proxy: *Yes*	*45.9%*	*44.9%*	*-1.9%*
Insurance: *Medicare*	*19.5%*	*19.5%*	*0.0%*
Insurance: *Medicaid*	*25.8%*	*25.4%*	*-0.9%*
Education: *High school* [a]	*55.0%*	*55.5%*	*1.0%*
Education: *College* [a]	*37.0%*	*36.3%*	*-1.3%*
Visiting nurse services: *Yes* [b]	*13.3%*	*12.9%*	*-1.4%*
Home health aide: *Total hours* [b] ‡	*0.25*	*0.25*	*0.0%*
Primary diagnosis: *Lymphoma/myeloma*	*6.1%*	*5.9%*	*-0.6%*
Complication(s): *Yes* [c]	*2.7%*	*2.3%*	*-1.8%*
Comorbidities: *Elixhauser total (mean)*	*3.94*	*3.96*	*1.1%*
ADL: *Needs partial assistance bathing*	*37.9%*	*37.5%*	*-1.0%*
ADL: *Needs partial assistance transferring from chair*	*35.3%*	*36.3%*	*2.3%*
ADL: *Needs complete assistance with 1+ activity*	*14.6%*	*14.8%*	*0.9%*
ESAS Physical: *At admission (mean)*	*1.98*	*1.97*	*-1.1%*
ESAS Psychological: *At admission (mean)*	*1.57*	*1.61*	*2.8%*
ESAS Physical: *At consult/reference day (mean)* [d]	*1.82*	*1.82*	*0.7%*
ESAS Psychol.: *At consult/reference day (mean)* [d]	*1.43*	*1.37*	*-4.5%*
CMSAS (Number) [e]**:** *At admission (mean)*	*9.08*	*9.02*	*-2.0%*
CMSAS (Number) [e]**:** *At consult/reference day (mean)* [d]	*7.87*	*7.83*	*-1.3%*
CMSAS (Severity) [e]**:** *At admission (mean)*	*16.1*	*16.0*	*-1.3%*
CMSAS (Severity) [e]**:** *At consult/reference day (mean)* [d]	*12.6*	*12.7*	*1.5%*
Morphine: *Equivalent dose (mg)*[f]‡	*3.1*	*3.3*	*4.8%*
Pain: *Somewhat* [g]	*10.5%*	*11.3%*	*2.5%*
Pain: *Quite a bit* [g]	*28.5%*	*27.7%*	*-1.9%*
Pain: *Very much* [g]	*32.7%*	*34.0%*	*3.1%*
Fatigue: *A little, somewhat or quite a bit* [g]	*37.2%*	*36.7%*	*-1.1%*
Fatigue: *Very much* [g]	*29.7%*	*30.5%*	*1.7%*

For more details including reference cases, see Table 3.2.1. [a] Highest level attained. [b] In two weeks prior to hospitalisation. [c] Presence of a major/minor complication. [d] For PC patients, reference day was day of consult; for UC patients, reference day was day they had most similar symptom severity to PC patients. [e] CMSAS (Number): # of physical symptoms; CMSAS (Severity): # of physical symptoms multiplied by the mean severity. [f] In week prior to hospitalisation. [g] Taken at consult/reference day with 'None' as the reference category. ‡ Raw data square-root transformed.

Cost data

Primary dependent variable

The summary statistics of direct cost (USD) of hospital stay, the primary dependent variable throughout this thesis, are presented in Table 4.1.3 and the full distributions for total costs are illustrated in Figure 4.1.2.

The costs exhibit significant positive skew (46.3, as opposed to 0 for normally distributed data) and high kurtosis (5, as opposed to 3 for normally distributed data). Specifically, the costs are skewed right with mean costs 1.35 times the median costs for the whole sample; mean costs are substantially higher than the median for both treatment and comparison groups. Distinct from a dataset drawn from a general population, there are no zero-cost patients in the sample.

Total cost distributions to the 50th percentile are similar in both treatment and comparison groups. In the higher percentiles, costs are higher and more skewed for palliative care than usual care, consistent with the higher illness burden observable on baseline statistics.

Heavily skewed data with only non-negative values present well-known challenges in health cost data analysis[68]; two common approaches to adjustment are log transformation and square-root transformation (Jones, 2010). The distributions of hospital costs for all 969 patients following these transformations are provided in Figure 4.1.3 and Figure 4.1.4. It is clear that log-transformation does bring costs close to a normal distribution without fully accounting for skew (Figure 4.1.3) while square-root transformation does not have the desired effect (Figure 4.1.4).

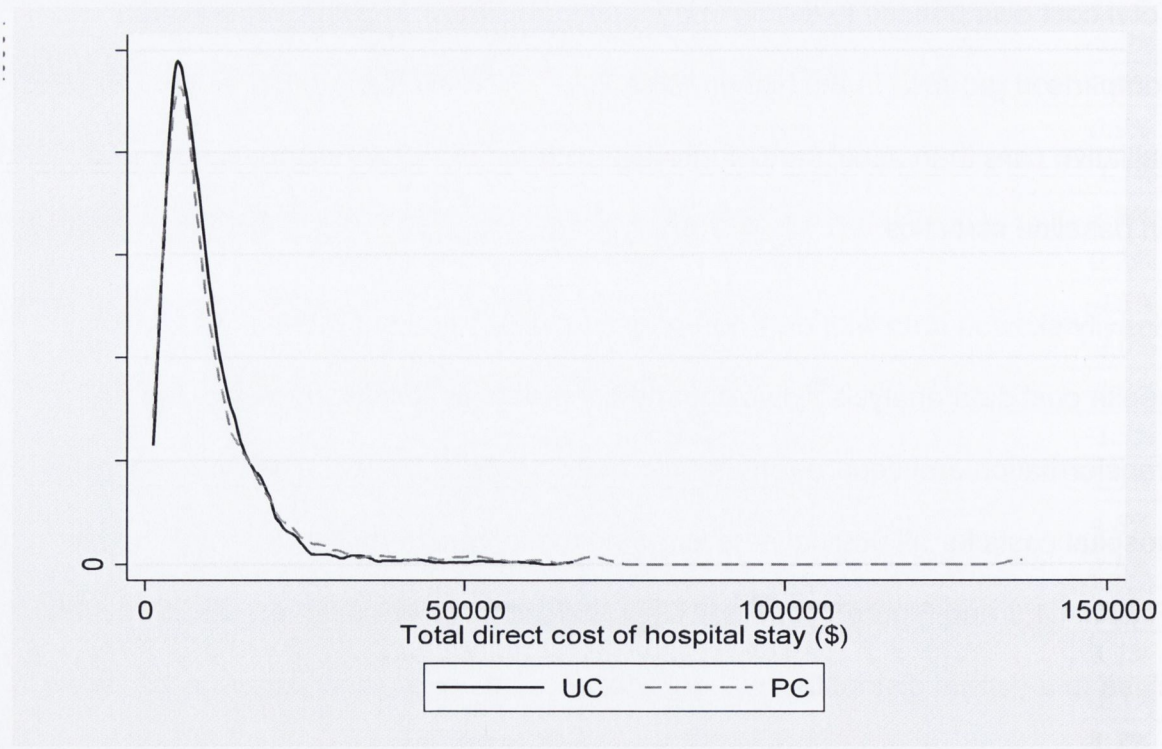
---

[68] For a more detailed discussion of these challenges in theory, see **§3.2.2**.

Table 4.1.3 Summary statistics of total direct hospital costs ($) (n=969)

	Total cost ($)		
	Usual care (n=713)	Palliative care (n=256)	All (n=969)
Mean	9550	11150	9973
Median	7379	7400	7400
Maximum	67075	136930	136930
75th Percentile	11791	12282	11849
25th Percentile	4943	4805	4903
Minimum	1311	1246	1246
Standard deviation	7558	13130	9379
Skewness[69]	3.1	5.0	4.9
Kurtosis	17.1	38.9	46.3

Figure 4.1.2 Distribution of total direct hospital costs ($), by treatment group (n=969)



[69] Skewness and kurtosis statistics are not presented in $. Both statistics are taken from the Stata output from the command *–summarize-*. Skewness is a measure of the lack of symmetry of a distribution (normal distribution has a skewness statistic of 0), and kurtosis is a measure of the 'peaked-ness' of a distribution (normal distribution has a kurtosis statistic of 3). The same presentation is adopted in subsequent tables.

Figure 4.1.3 Distribution of log-transformed total direct hospital costs (n=969)

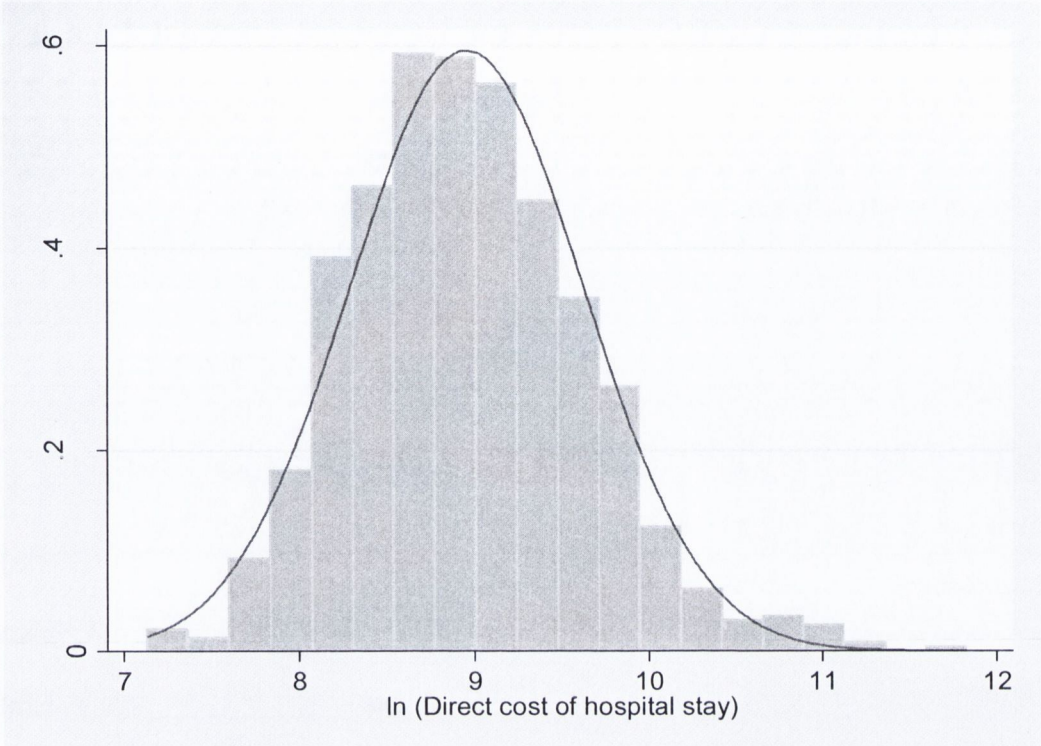
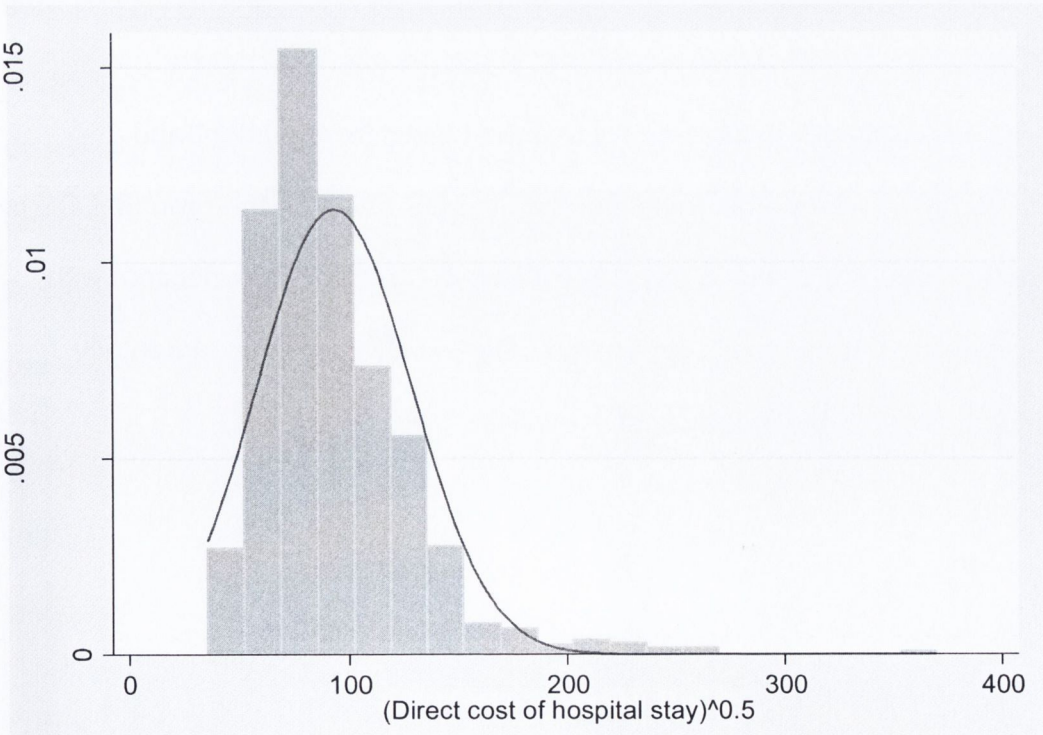


Figure 4.1.4 Distribution of square-rooted total direct hospital costs (n=969)



Secondary dependent variables

Summary statistics of daily direct hospital costs (the ratio of total direct cost to LOS) are provided in Table 4.1.4. Distinct from the total cost summary, mean daily costs are similar across treatment groups at just over $1200 and with lower skew.

Table 4.1.4 Summary statistics of daily direct hospital costs ($) (n=969)

	Daily cost ($)		
	Usual care (n=713)	**Palliative care (n=256)**	**All (n=969)**
Mean	1242	1218	1236
Median	1073	1125	1087
Maximum	6125	5071	6125
75th Percentile	1454	1474	1462
25th Percentile	794	838	816
Minimum	328	311	311
Standard deviation	729	568	690
Skewness	2.9	1.8	2.8
Kurtosis	15.5	10.9	15.6

The apparent inconsistency between daily and total costs for treatment and comparison groups is attributable to a difference in LOS between the groups: LOS is longer in the PC group (9 versus 8 days, see Table 4.1.5). This is consistent with the higher illness burden of the PC group and makes daily cost a fundamentally different dependent variable to total cost in evaluating PCCT impact.

Table 4.1.5 Summary statistics of LOS (days) (n=969)

	Length of Stay (days)		
	Usual care (n=713)	**Palliative care (n=256)**	**All (n=969)**
Mean	8.0	9.0	8.2
Median	6	7	7
25^{th}-75^{th} %ile	5-9	5-10	5-9
Standard deviation	5.0	7.5	5.8

In **§3.1.6** (Table 3.1.1) I detailed how universal billing (UB) codes could be used to isolate and attribute costs to 'major' categories of utilisation. After reviewing the data (and consistent with previous studies; see **§2 Literature Review**), five 'major' categories are identified: room and board, intensive care unit (ICU), pharmacy, laboratory and imaging.

The mean direct hospital costs for each of the five 'major' categories are summarised for the principal sample in Table 4.1.6. Room and board is by far the largest category, accounting for a third of total cost. ICU costs have the most distinct distribution with less than a quarter of patients having non-zero costs while others exceed $12000 (greater than the mean total cost of stay for the whole sample). Pharmacy exhibits the biggest difference between groups, accounted for by one patient with very high costs.[70]

Overall the five 'major' categories account for 69% (UC) and 75% (PC) of total cost of stay. Skewness is consistently observable across categories and treatment groups.

---

[70] Total pharmacy costs= $125999. Total direct hospital costs=$136930, 19 times the median and roughly twice the next most expensive patient.

Table 4.1.6 Summary statistics of total direct hospital costs, by five 'major' sub-categories ($) (n=969)

	Room & Board		ICU		Pharmacy		Labs		Imaging		'Major' Categories	
	UC	PC	UC	PC	UC	PC	UC	PC	UC	PC	UC	PC
Mean	3216	3889	277	288	1973	2917	558	466	594	801	6618	8360
Mean/Total	33.7%	34.9%	2.9%	2.6%	20.7%	26.2%	5.8%	4.2%	6.2%	7.2%	69.3%	75.0%
Median	2544	2811	0	0	1065	1508	393	237	336	492		
Maximum	19372	36126	11927	12829	24517	125999	5236	6460	9581	5920		
75th %ile	3790	4465	0	0	2264	2659	672	574	747	953		
25th %ile	1848	2004	0	0	542	743	203	110	81	185		
Minimum	296	627	0	0	58	36	8	0	0	0		
SD	2286	3512	1019	1282	2861	8297	581	680	865	1001		
Skewness	2.8	4.2	6.0	7.1	4.1	13.0	3.2	4.2	4.1	2.6		

SD: Standard deviation; ICU: Intensive Care Unit; 'Major' Categories sums the costs of the five categories detailed in the table.

Heteroscedasticity

In addition to skewness, a common characteristic of health cost data is heteroscedasticity, where the variability of the dependent variable (in this case, total direct hospital costs) varies across the range of values of predictors. This is a serious problem for analysing health costs as many regression approaches, including the most common ordinary least squares (OLS) assume the absence of heteroscedasticity.

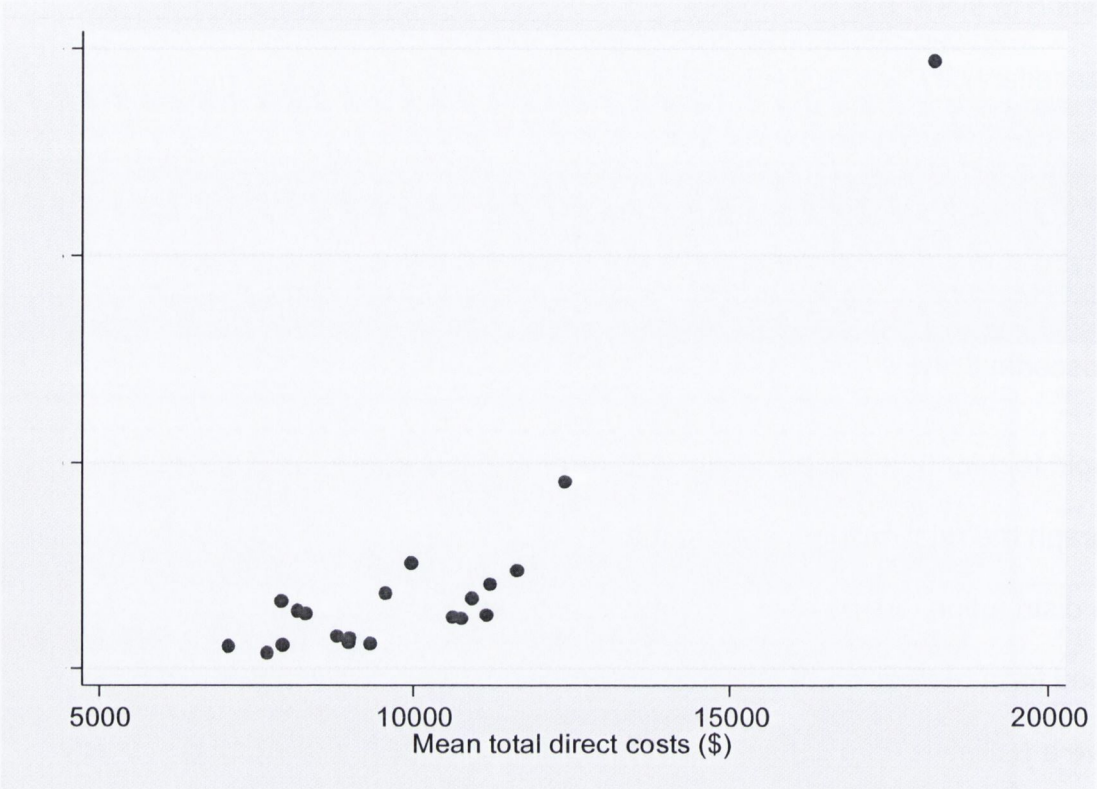
A simple test for heteroscedasticity is to regress costs against relevant covariates, and graph the relationship between mean costs and variance at different percentiles of the distribution (Jones et al. (2013b), p. 39). In Figure 4.1.5 the relationship is graphed for a regression of costs against a binary treatment variable (did patients receive a palliative care consultation during their hospitalisation?), all propensity score covariates (Table 3.2.1) and fixed effects to control for site differences, with the distribution of the conditional mean of costs divided into 20 parts.

The picture is distorted somewhat by one extreme value on the far right of the graph, which reflects very large variance among the highest cost patients. Both with and without this outlier there is a clear pattern of increasing variance as the mean increases - a clear indication that analysis of these cost data must take into account heteroscedasticity.[71] A flexible modelling approach more appropriate to skewed, heteroscedastic data than ordinary OLS approaches, is therefore required to estimate PCCT cost effect with this data.

---

[71] Stata provides formal tests for heteroscedasticity but these are not valid with propensity score weighted data. It was therefore not possible to subject these data to a formal test but the interpretation of Figure 4.1.5 appears straightforward.

Figure 4.1.5 Variance and mean total direct costs ($) for weighted OLS on relevant covariates (n=969, divided into 20 groups)



Summary

This section provides an overview of the 'PC4C' study data.

Following data cleaning and matching, and the removal of 54 patients[72] not discharged alive following hospitalisation, the principal sample for economic evaluation is n=969 (n_{UC}=713, n_{PC}=256). There are clear observable differences between the treatment and comparison groups on baseline data, on both clinical and socio-demographic factors. The groups were balanced on 33 demographic, socioeconomic, clinical and system covariates using propensity score weights.

[72] Died during hospitalisation, n=51; data missing on survival, n=3.

Hospital costs and LOS also both show differences between usual and palliative care patients, consistent with a higher illness burden in the treatment group. Overall patients in the study incurred an average $9973 in total hospital costs with an average stay of 8.2 days.

Most significantly this section establishes that the primary dependent variable – total direct hospital costs - exhibits characteristics typical for health utilisation data: right-hand skew, heavy tail, leptokurtosis and heteroscedasticity.

These raise well-known challenges in modelling costs. Specifically, a flexible modelling approach rather than a simple linear regression is required in modelling costs and estimating cost effects. These challenges are the focus of the next section.

4.2 Model evaluation and selection

Introduction

The previous section summarised the 'PC4C' data, concluding that it exhibits characteristics familiar to health economists: right-hand skew, leptokurtosis and heteroscedasticity. A flexible modelling approach appropriate for data with these characteristics is therefore required.

Selection of the best modelling approach is not straightforward. It was demonstrated in **§2 Literature Review** that the dominant cost model in the economic evaluation of PCCTs is the generalised linear model (GLM) with a gamma distribution and a log link, but that no study to date in this field has reported comparative evaluation of model performance prior to estimating cost effect. In **§3.2.2** I highlighted that GLM (gamma, log) is prominent in the wider literature for addressing health cost data, but also that alternative modelling approaches exist, and that different approaches may offer different strengths and weaknesses rather than one model being dominant. The accuracy and reliability of effect estimates rely on sound model selection.

Alternatives to GLM (gamma, log) include GLMs with different distribution and/or links (Hill and Miller, 2010); exponential conditional mean models (ECMs, Jones (2010)); generalised gamma models (GGMs, Manning et al. (2005)); extended estimation equations (EEEs, Basu and Rathouz (2005)); and Finite Mixture Models (FMMs, Conway and Deb (2005)).[73] All such approaches offer potential solutions to the myriad challenges posed by statistical analysis of health cost data, but only through comparative analysis of performance can a model's suitability to a specific dataset be judged (Jones, 2010, Garrido et al., 2012).

---

[73] An overview of each approach was previously provided in Table 3.2.5.

The purpose of this section is to compare and evaluate the performance of different modelling approaches with the 'PC4C' data, and to use the results to select an appropriate model for estimating incremental effect of the PCCT intervention on hospital costs.

There are multiple tools available to evaluate model performance.

For GLMs specifically, the Modified Park Test for the appropriate family, Pregibon for the appropriate link, and Pearson and Hosmer-Lemeshow for bias are widely proposed (Glick, 2008, Deb et al., 2013). For the comparison of non-GLM approaches, Jones et al. (2013b)[74] propose the three latter tests, R^2 as a goodness-of-fit test, and in-sample indicators of predictive accuracy Root Mean Squared Error (RMSE) and Mean Absolute Prediction Error (MAPE).

Jones et al. (2013b) also suggest out-of-sample testing using v-fold cross-validation. This entails dividing a sample of *n* subjects into groups of equal size *v*. For each group the model is estimated using the sub-sample of *n-v* and predictions are generated using the *v* observations. Repeating the process for each group (i.e. a total of *n/v* times) generates fitted values for all observations, and these can be compared to the observed values. However, the propensity score research design of the 'PC4C' study precludes applying cross-validation methodology because for each occasion that *v* patients are removed from the sample, it is no longer guaranteed that treatment and comparison groups are balanced on propensity score covariates in the remaining *n-v* sub-sample.

---

[74] Specifically, guidance for model evaluation and comparison is provided in 'Chapter 12: Modelling health care costs' in Jones et al. (2013b). Accompanying software and data resources for applied use of the guidance are accessed via Jones et al. (2013c).

This restriction compounds a more general limitation in evaluating performance with a single observational dataset (n=969), namely that there is insufficient sample size to fit the distribution using Monte Carlo methods and out-of-sample testing per large-scale modelling of health care costs (Marazzi et al., 1998, Carey and Burgess, 1999, Burgess et al., 2012, Jones et al., 2013a).

I am therefore restricted to comparison and evaluation using seven in-sample measures, summarised in Table 4.2.1.

Table 4.2.1 Tests for model comparison and evaluation, per Jones et al. (2013b, 2013c)

Measure	**Purpose;** definition
Modified Park Test (Manning and Mullahy, 2001, Park, 1966)	**GLM Family;** constructive diagnostic test, which recommends a GLM family given a specific link. GLM only.
Pregibon link test (Pregibon, 1980)	**Goodness of link;** assesses linearity of response by testing that the square of the predicted values is itself not a significant predictor of y.
Pearson correlation coefficient	**Goodness of fit;** checks for bias by testing correlation between residuals and predicted values.
Modified Hosmer-Lemeshow (Hosmer and Lemeshow, 2000)	**Goodness of fit;** checks for systematic bias by dividing the sample into deciles and regressing residuals against fitted values.
R^2 *In-sample*	**Goodness of fit;** total variance observable in y that is explained by the model. Larger R^2 indicates greater model accuracy, although for nonlinear models, large R^2 values are rare, even for a model that fits the data well.
Root Mean Squared Error (RMSE) *In-sample*	**Goodness of fit;** root of mean squared difference between observed values and the regression line: $\sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}}$ Smaller RMSE indicates greater model accuracy.
Mean Absolute Prediction Error (MAPE) *In-sample*	**Goodness of fit;** absolute difference between observed values and the regression line: $\frac{\sum_{i=1}^{n} abs(y_i - \hat{y}_i)}{n}$ Smaller MAPE indicates greater model accuracy.

y: dependent variable (total direct hospital costs)

Methods

Research question

Which modelling approach is most appropriate for estimating mean incremental PCCT effect on total direct hospital costs with the 'PC4C' data?

Evaluation and comparison

Model evaluation and comparison was performed in three stages.

First, a range of GLMs were subjected to the Modified Park, Pearson, Pregibon and Modified Hosmer-Lemeshow tests. The Modified Park recommends one of four families (Gaussian, Poisson, Gamma, Inverse Gaussian) given a specific link; there are three links applicable to these families (log, identity and power). This gives 12 prospective GLM specifications, which were run with and without propensity score weights to give 24 regressions in total. Tests were run using the *–glmdiag-* Stata programme (Glick, 2014).

Second, the best five performing GLMs were identified and subjected to the additional three tests in Table 4.2.1. Alternative models to GLMs were also run at this stage: ordinary least squares on untransformed, log-transformed and square-rooted costs; ECM – NLS and ECM – Poisson; GGMs with homoscedastic and heteroscedastic errors; EEE; and FMM (gamma). These reflect the range of models proposed by Jones et al. (2013b), minus one (Generalised beta) for which Stata code has not yet been published.[75] The 14 models were subjected to the six tests using a Stata package provided by Jones et al. (2013c), modified[76] to suit my data.

---

[75] Code for this model has not yet been validated for general use (personal communication, AM Jones, December 2014).

[76] Out-of-sample tests excluded per **§4.2>Introduction**. Additionally, in the code published online Glick and Jones take slightly different approaches to the Pregibon test. The former specifies the 'robust' option and

Third, the two best performing models were subjected to sensitivity analyses to ascertain if performance improved for models with a reduced number of covariates, and for samples with long-stay outliers and high-cost outliers removed at the 95th percentile.

Variables

The primary independent variable was a binary treatment variable: was the patient seen by a PCCT during hospital admission? Additional predictors were identified according to propensity score weighting: all covariates (n=33) in the propensity score for the principal cost sample (see Table 3.2.1) were included as predictors. This is the desired specification because associations between the dependent variable and baseline covariates can be drawn from regression results if those covariates are balanced in the propensity score. Fixed site effects to control for hospital-specific differences were also included.

Results

The output of 12 GLMs for four tests, first without propensity score weights and then with, is provided in Table 4.2.2. All instances where a test was passed are highlighted in bold.

The best-performing models are all in the gamma and inverse Gaussian distributions. Of these six, five pass all four tests; that is, a p value ≥0.05 for the corresponding family in the Modified Park test, as well as for the other tests. These are: (gamma, log), (gamma, power^0.5), and the three inverse Gaussian models. One model (gamma, identity) is the only model to pass exactly three tests (it does

---

allows for 30 iterations; the latter specifies the 'difficult' option and allows for 100 iterations. For internal consistency in this thesis, I standardised all Pregibon tests to the 'robust' option.

not pass the Park test for λ=2, the relevant family). None of the Gaussian or Poisson models pass the Park test and none passes more than two additional tests.

Table 4.2.2 Comparing GLM performance, per Glick (2014), (n=969)

Test / GLM	Modified Park	Pregibon	Pearson	Modified H-L
	λ (where p>0.05) *λ=0 (Gaussian)* *λ=1 (Poisson)* *λ=2 (Gamma)* *λ=3 (iGaussian)*	*p value*	*p value*	*p value*
Without weights				
Gaussian, log	None	<0.01	<0.01	<0.01
Gaussian, identity	2 (0.44), 1 (0.05)	<0.01	**0.99**	<0.01
Gaussian, power^0.5	2 (0.15), 1 (0.15)	<0.01	**0.45**	<0.01
Poisson, log	2 (0.50), 3 (0.17)	<0.01	**0.38**	<0.01
Poisson, identity	3 (0.60), 2 (0.15)	0.02	**0.37**	**0.06**
Poisson, power^0.5	3 (0.39), 2 (0.27)	<0.01	**0.40**	0.02
Gamma, log	3 (0.92), **2 (0.10)**	**0.19**	**0.43**	**0.30**
Gamma, identity	3 (0.66)	**0.22**	**0.45**	**0.52**
Gamma, power^0.5	3 (0.90), **2 (0.06)**	**0.27**	**0.51**	**0.75**
iGaussian, log	**3 (0.87),** 3 (0.05)	**0.59**	**0.72**	**0.66**
iGaussian, identity	**3 (0.43)**	**0.88**	**0.95**	**0.92**
iGaussian, power^0.5	**3 (0.69)**	**0.77**	**0.77**	**0.66**
With weights				
Gaussian, log	**0 (0.12)**	<0.01	<0.01	<0.01
Gaussian, identity	1 (0.43)	**0.05**	<0.01	<0.01
Gaussian, power^0.5	1 (0.53)	<0.01	<0.01	<0.01
Poisson, log	2 (0.91)	<0.01	<0.01	<0.01
Poisson, identity	3 (0.19), 2 (0.16)	**0.23**	0.01	<0.01
Poisson, power^0.5	2 (0.27), 3 (0.20)	**0.13**	<0.01	<0.01
Gamma, log	3 (0.90), **2 (0.07)**	**0.97**	<0.01	**0.09**
Gamma, identity	3 (0.46)	**0.56**	<0.01	**0.21**
Gamma, power^0.5	3 (0.76)	**0.84**	<0.01	**0.22**
iGaussian, log	**3 (0.80)**	**0.19**	<0.01	0.02
iGaussian, identity	**3 (0.34)**	**0.38**	<0.01	<0.01
iGaussian, power^0.5	**3 (0.54)**	**0.61**	<0.01	<0.01

Model: principal sample (n=969); total direct costs regressed against treatment, 33 propensity score covariates and fixed-effect site variables.

Once propensity score weights are introduced, there is a noticeable reduction in performance. No model passes all four tests. The best performing is (gamma, log), which passes three tests and fails Pearson. The next best remain (gamma, power^0.5) and the three inverse Gaussian models, all of which pass two out of four tests.

These five GLMs are therefore the best performing both with and without weights and are carried forward to the second stage of analysis.

The results of six tests (all tests except Modified Park, which is relevant only for GLMs) against 11 models are given in Table 4.2.3 (without propensity score weights) and in Table 4.2.4 (with weights). All instances where a test passed are highlighted in bold; the best-performing models for the other evaluative statistics are also highlighted in bold.

In Table 4.2.3 the best performing model is GLM (gamma, log), which passes all three tests as well as delivering comparatively strong results in the goodness-of-fit analyses. This is the dominant GLM; of the alternatives, OLS ln(y), which passes all three tests, and ECM – Poisson are the next strongest.

As in Table 4.2.2, there is a noticeable reduction in performance once propensity score weights are introduced, but Table 4.2.4 otherwise shows a similar pattern to Table 4.2.3.

OLS ln(y) now performs best, passing all three tests and with strong goodness of fit. It is the only model to pass the Pearson test.

Among GLMs, (gamma, log) remains the dominant model. In addition to the results in Table 4.2.2, it has better performance across goodness-of-fit and prediction measures against other GLMs in all cases.

Table 4.2.3 Comparing model performance, per Jones et al. (2013b, 2013c) (no propensity score weights, n=969)

No weights	Link Test p-value	Pearson p-value	H-L p-value	R^2	RMSE	MAPE
					Within sample	*Within sample*
OLS on y	<0.01	-/-	<0.01	**0.07**	9032	5404
OLS on ln(y)	**0.68**	**0.41**	**0.43**	0.05	9121	**5278**
OLS on y^0.5	**0.05**	**0.64**	<0.01	**0.07**	9058	**5289**
ECM – NLS	DNC	<0.01	<0.01	**0.16**	**8750**	5833
ECM – Poisson-ML	**0.07**	**0.38**	<0.01	**0.09**	**8951**	5359
GG – hom	DNC	DNC	DNC	DNC	DNC	DNC
GG – het	DNC	DNC	DNC	DNC	DNC	DNC
GLM gamma, log	**0.19**	**0.43**	**0.30**	**0.07**	**9020**	**5341**
GLM gamma, power^0.5	**0.27**	**0.51**	**0.75**	**0.07**	9055	5345
GLM igaussian, log	**0.59**	**0.72**	**0.66**	0.06	9086	5369
GLM igaussian, identity	**0.88**	**0.95**	**0.92**	0.05	9123	5365
GLM igaussian, power^0.5	**0.77**	**0.77**	**0.67**	0.06	9106	5369
EEE	DNC	DNC	DNC	DNC	DNC	DNC
FMM gamma	-/-	DNC	<0.01	<0.01	125983	11760

Model: principal sample (n=969); total direct costs regressed against treatment, 33 covariates and fixed-effect site variables. DNC: Did not converge/compute.

Table 4.2.4 Comparing model performance, per Jones et al. (2013b, 2013c) (with propensity score weights, n=969)

Weights	Link Test p-value	Pearson p-value	H-L p-value	R^2	RMSE	MAPE
					Within sample	*Within sample*
OLS on y	**0.07**	<0.01	<0.01	0.05	9217	5782
OLS on ln(y)	**0.63**	**0.83**	**0.73**	0.05	**9138**	**5325**
OLS on y^0.5	**0.20**	**0.05**	<0.01	**0.06**	**9117**	**5503**
ECM – NLS	DNC	<0.01	<0.01	**0.17**	9125	6156
ECM – Poisson-ML	**0.67**	<0.01	<0.01	**0.07**	**9075**	5684
GG – hom	DNC	DNC	DNC	DNC	DNC	DNC
GG – het	DNC	DNC	DNC	DNC	DNC	DNC
GLM gamma, log	**0.97**	<0.01	**0.09**	0.05	9191	**5679**
GLM gamma, power^0.5	**0.84**	<0.01	**0.23**	0.05	9217	5700
GLM igaussian, log	**0.19**	<0.01	0.02	0.04	9451	5912
GLM igaussian, identity	**0.38**	<0.01	<0.01	0.03	9384	5941
GLM igaussian, power^0.5	**0.61**	<0.01	<0.01	0.04	9366	5860
EEE	DNC	DNC	DNC	DNC	DNC	DNC
FMM gamma	-/-	DNC	<0.01	<0.01	130657	16661

Model: principal sample (n=969); total direct costs regressed against treatment, 33 covariates and fixed-effect site variables. DNC: Did not converge/compute.

Three models (EEE and the two GGs) fail completely to deliver test outputs, with or without weights.[77] There are difficulties with ECM – NL and FMM, the former failing all tests and the latter failing to compute some scores and delivering very poor performance results in others.

The strongest two models are therefore identified as OLS ln(y) and GLM (gamma, log), although neither exhibits dominant performance in link tests, bias tests or goodness-of-fit measures, and performance demonstrably deteriorates with the addition of weights.

Two common strategies to improve performance are reducing the number of independent variables in a model (to guard against over-fitting) and removing outliers from the model (high-utilisation outliers may distort cost-effect estimates) (Mihaylova et al., 2011). Four new specifications were created according to number of covariates in the model (reduced from 33 to 16|9) and by trimming at the 95th percentile long-stay patients (20 days<=LOS) and high-cost patients ($23,154<y).[78] The results are given in Table 4.2.5.

Altering the number of covariates has little impact for either model. Removing long-stay and high-cost outliers notably improves the goodness of fit for both models, but the latter strategy reduces the performance of GLM (gamma, log) in the bias tests.

---

[77] I also ran GGs and EEEs with a weighted model incorporating a reduced 16 covariates but these also fail to produce an output.

[78] For details of the 'reduced' and 'minimal' models, and samples with 'long stay' and 'high cost' patients removed, see Appendix to §4.2.

Table 4.2.5 Comparing model performance, per Jones et al. (2013b, 2013c), with # of covariates reduced or outliers removed (with propensity score weights)

Weights	Link test p-value	Pearson p-value	H-L p value	R^2	RMSE	MAPE
					Within sample	Within sample
OLS on ln(y)						
'Full' model, per Table 4.2.4	**0.63**	**0.83**	**0.73**	0.05	9138	5325
'Reduced' model	**0.40**	**0.99**	**0.59**	0.04	9186	5363
'Minimal' model	**0.48**	**0.87**	**0.59**	0.03	9213	5396
Long stayers removed	**0.19**	**0.83**	**0.99**	0.09	5487	3959
High-cost removed	0.03	**0.46**	**0.67**	0.08	4567	3617
GLM gamma, log						
'Full' model, per Table 4.2.4	**0.97**	<0.01	**0.09**	0.05	9191	5679
'Reduced' model	**0.80**	<0.01	0.04	0.04	9269	5641
'Minimal' model	**0.39**	<0.01	**0.08**	0.03	9252	5656
Long stayers removed	**0.10**	<0.01	**0.20**	0.09	5538	3983
High-cost removed	<0.01	<0.01	<0.01	0.07	4635	3679

Full model: 969 patients; total direct costs regressed against treatment, 33 covariates and fixed-effects site variables.
'Reduced' model: 969 patients; total direct costs regressed against treatment, 16 covariates and fixed-effects site variables.
'Minimal' model: all 969 patients; total direct costs regressed against treatment, nine covariates and fixed-effects site variables.
Long stayers removed: 924 patients with LOS<=20 days; total direct costs regressed against 33 covariates and fixed-effects site variables.
High-cost removed: 920 patients with total direct costs<23153.96; total direct costs regressed against 33 covariates and fixed-effects site variables.
For details of the 'reduced' and 'minimal' models, and samples with 'long stay' and 'high cost' patients removed, see Appendix to §4.2

Discussion

Detailed comparative evaluation of model performance with the 'PC4C' data re-iterates the findings of previous research: no model is dominant, and selection is necessarily the result of trade-offs.

A comparison of 12 prospective GLMs establishes five with reasonable performance, of which (gamma, log) performs strongest (Table 4.2.2). Evaluation of these five GLMs alongside six other prospective modelling approaches yields a clear conclusion: GLM (gamma, log) and OLS ln(y) are the best performing models (Table 4.2.3, Table 4.2.4).

Despite the comparatively strong performance, doubts about these two models remain. In particular, the introduction of propensity score weights appears to undermine performance and, as noted in **§3.2.1**, in the presence of unobserved confounding propensity score weights can exacerbate problems of uncontrolled heterogeneity (Brooks and Ohsfeldt, 2013). Reducing the number of covariates in the model does little for either GLM (gamma, log) or OLS ln(y); removing long-stay and high-cost outliers improves goodness of fit only (Table 4.2.5). While improved goodness of fit is a desirable quality in a model *per se*, achieving this through removing high-cost patients would require a strong justification, particularly as the removal of such outliers does not appear to improve the models on link, Pearson and Modified Hosmer-Lemeshow tests.[79] Since this is a single observational dataset it is not possible to use simulation to finesse fit to the distribution, a method that has been used in large-scale simulation datasets to justify trimming long-stay outliers (Marazzi et al., 1998).

---

[79] The issue of trimming a sample by LOS outliers is given closer consideration in **§4.3**.

These results highlight the need for judicious trade-offs in model selection. OLS out-performs GLM (gamma, log) once propensity score weights are introduced but log-transformed costs have well-known limitations as a dependent variable (see **§3.2.2**). Moreover, **§4.1** highlighted the presence of heteroscedasticity in the data, which undermines the performance of linear regression approaches. Removing long-stay and high-cost outliers improves model performance in some cases but risks arbitrarily trimming the sample and consequently altering results (see **§4.3**). While propensity score weights demonstrably reduce model performance, they are intrinsic to the research design in minimising confounding.

In summary, OLS on log-transformed costs is the best-performing model on evaluative measures but there are basic concerns about the model's capacity to deliver meaningful $ effect estimates and to cope with heteroscedasticity. GLM (gamma, log) is the best non-linear alternative but this does not pass all diagnostic tests. Checks will be necessary to ensure that derived results are not specific to model selection given that this model is inferior to alternatives on some measures. Additionally checks are necessary to ensure that derived results are not due to the use of propensity score weights, which can exacerbate problems in the presence of unobserved confounding.

Therefore I select GLM (gamma, log) with a treatment variable, 33 covariates and site effects, and propensity score weights, as the primary modelling approach for economic evaluation of the 'PC4C' data. Additionally I will perform sensitivity analyses using alternative approaches:

- Unweighted GLM (gamma, log), ensuring that the addition of propensity score weights are not determining reported results;

- Weighted and unweighted OLS on untransformed costs, ensuring that model selection is not determining reported results.

These sensitivity analyses are applied to the primary analyses in **§4.4** and **§4.5** and presented in marked appendices. Where weighting and modelling alternatives provide substantively similar results to the primary analyses I can be increasingly confident that they are robust findings.

Conclusion

Comparative analysis of model performance for 'PC4C' data shows that the best-performing models are OLS ln(y) and GLM (gamma, log). Given the well-known limitations of linear models with skewed, heteroscedastic cost data, and the difficulty in retransforming log-dollar results into meaningful cost-effect estimates, GLM (gamma, log) was selected as the primary model for all subsequent analysis. Sensitivity analyses with this and other models, altering the specification by use of weights, will be performed where appropriate to check the robustness of primary analyses.

4.3 Controlling for LOS in cost analysis

Introduction

The previous section detailed model evaluation and selection for analysis of the 'PC4C' data. A generalised linear model with a gamma distribution and a log link (GLM (gamma, log)) with propensity score weights was identified as the strongest option given data characteristics and model performance, with sensitivity analysis to be performed to confirm robustness of results in primary analysis.

This section addresses an additional methodological issue, raised in **§2 Literature Review** and observable in the health economics literature more widely, regarding use of length of stay (LOS) in cost analyses. The most robust cost analyses of PCCTs to date have a general methodological consistency, employing propensity scoring to match treatment and comparison groups to minimise bias, and the same regression model (GLM, (gamma, log)) to deal with skew, heteroscedasticity and accurate cost-effect estimation (Morrison et al., 2008, Penrod et al., 2010, Morrison et al., 2011a, Starks et al., 2013, Whitford et al., 2014, McCarthy et al., 2015).

However, **§2.4** also highlighted that there are differences between these studies in approach to using LOS to control for unobserved heterogeneity, with three strategies variously employed:

I. **LOS as a covariate**, where LOS is used as an independent predictor in regression analyses;

II. **LOS as a sample parameter**, where short- and long-stay patients are removed from analysis *ex ante*;

III. **LOS as a denominator in the dependent variable**, where daily cost (the ratio of total cost to LOS) is employed as a dependent variable.

A review of these methods, their potential justifications and potential problems, is provided in Table 4.3.1.

Strategies that control for LOS reflect a tension between heterogeneity and endogeneity in analysis of observational data. Incorporating a nonlinear transformation of LOS as a covariate, or sample selection by LOS, can act as a proxy for otherwise unobserved factors that analysts would like to control for.

But LOS is both a predictor of treatment (long LOS is an indicator of palliative care need) and a determinant of costs (LOS and total cost are closely connected). Regressing or stratifying by LOS in economic evaluation of a programme is therefore at best sub-optimal, while daily cost is a dependent variable of limited practical value. Moreover, daily cost is a fundamentally different variable to total cost if LOS is systematically different between treatment and comparison groups, as is often the case with observational palliative care studies (and is the case with the 'PC4C' data – see Table 4.1.4 and Table 4.1.5).

The purpose of this section is to examine in greater detail the impact of LOS-control methods on 'PC4C' data analysis, and their value in the context of these strengths (incorporating otherwise unobserved heterogeneity) and weaknesses (introducing endogeneity into analysis).

Table 4.3.1 Summary of strategies in prior PCCT studies employing LOS to control for unobserved heterogeneity

Use of LOS	**Definition** – potential justifications	Potential problems	Examples*
I. Covariate	**LOS employed as an independent variable/predictor in regression analysis** Covariate intended to control for: • Unobserved heterogeneity, including clinical complexity; • Uneven accumulation of costs, proportionally more costs being accrued closer to admission.	Use of LOS as a covariate risks introducing endogeneity into analysis, since LOS is not an independent predictor of utilisation. Rather, it is associated with both treatment (long hospital stay suggests clinical complexity) and dependent variables (LOS and other utilisation data are typically closely correlated), thus undermining estimation of the causal relationship of interest (Amporfu, 2010).	Penrod et al. (2010) Whitford et al. (2014)
II. Sample parameter	**Short- and/or long-stay outliers trimmed from sample** Sample parameter(s) intended to control for: • Unobserved heterogeneity, including clinical complexity; • Outliers skewing distribution of utilisation data such as LOS and cost, distorting and disguising effect estimates.	Defining a sample *ex ante* by a factor that is associated with both treatment and the dependent variable risks biasing results and obscuring true effect (Imbens, 2004, Garrido, 2014b). Where propensity score matching is used, *ex ante* trimming is antithetical to the research framework, which aims to estimate a counterfactual using baseline data (Rubin, 2007).	Morrison et al. (2008) Morrison et al. (2011a) Starks et al. (2013) Whitford et al. (2014) McCarthy et al. (2015)
III. Dependent variable denominator	**Daily cost (the ratio of total cost to LOS) employed as primary dependent variable** Dependent variable denominator intended to: • Indirectly limit impact of unobserved heterogeneity by accounting for long LOS; • Reduce skew and leptokurtosis common to healthcare utilisation data distributions; • Address specific stakeholders, e.g. a hospital reimbursed at a fixed daily rate may prioritise daily cost over total cost.	Estimated effect on daily costs is of limited practical value because this is a ratio and not the overall costs accrued: a treatment that reduces daily cost by 10% but increases LOS by 50% is not necessarily delivering desirable outcomes for patients or payers. *Per diem* ratios change systematically with LOS and must be used carefully (Ishak et al., 2012). If LOS differs between treatment and comparison groups then daily cost (total cost/LOS) is a fundamentally different dependent variable to total cost.	Penrod et al. (2006) Ciemins et al. (2007) Hanson et al. (2008) Penrod et al. (2010)

*Examples in the table are from the PCCT literature, as reported in **§2 Literature Review**. Examples of LOS controls in other domains can be cited, e.g. paediatric care (Kuo et al., 2012, Queen et al., 2014) and maternity services (Thompson et al., 2003).

Methods

Research question

What is the effect of controlling for LOS in estimating PCCT impact on hospital costs?

LOS as a covariate

Where health economists have incorporated LOS as a predictor in modelling costs, they have generally done so following non-linear transformation: LOS log-transformed, squared, cubed, etc. (Penrod et al., 2010, Carey and Burgess, 1999).

I therefore estimated five GLMs with a gamma distribution and a log link, differentiated by the use of common nonlinear transformations of LOS as a covariate. I calculated the average change in total direct hospital costs for patients going from usual care to treatment, holding all other covariates at their original values. Bootstrapped standard errors (1000 replications) were calculated consistent with **§3.2.2** (Abadie and Imbens, 2008).

The five models were specified as follows:

(i) glm cost pal_care *varlist*

(ii) glm cost pal_care *varlist* LOS^2

(iii) glm cost pal_care *varlist* LOS^3

(iv) glm cost pal_care *varlist* LOS^4

(v) glm cost pal_care *varlist* ln(LOS)

Where cost is total direct hospital costs, pal_care is a binary variable denoting treatment or comparison, *varlist* is a fixed list of 33 demographic, clinical and system variables included in the principal propensity score (Table 4.1.2) plus fixed effects variables to account for site differences, and propensity score weights are applied to

all regressions.[80] The Akaike information criterion (AIC), a measure of model information loss where a lower score indicates superior performance (Burnham and Anderson, 2004), was calculated in Stata for each regression as an additional indicator of model performance.[81]

LOS as a sample parameter

As summarised in **§2 Literature Review** and Table 4.3.1, previous large studies in this field have often trimmed the sample of short- and long-stay patients. Replicating the most common limits (7<=LOS<=30) is not an option due to the LOS statistics of the 'PC4C' sample (see Table 4.3.2). Only 38% (n=367) of the principal sample for economic evaluation had 7<=LOS<=30.

Table 4.3.2 Distribution of length of stay, principal sample (n=969)

Percentile	1%	5%	10%	25%	50%	75%	90%	95%	99%
LOS (days)	3	4	4	5	7	9	14	20	32

To maximise sample size while investigating the impact of defining the sample by LOS, I trimmed according to LOS percentile. To create sub-samples without longest-stay outliers I removed patients who stayed longer than 25 days (97.5th percentile) and 20 days (95th percentile); to create a sub-sample without shortest-stay outliers I removed patients who stayed less than four days (2.5th/5th percentiles); and to create a sub-sample with both longest- and shortest-stay outliers removed I performed both strategies simultaneously.

---

[80] Model (i) is therefore the same as that used in the principal analyses in **§4.2**.

[81] AIC was not used in **§4.2** because it cannot be compared for different modelling approaches (e.g. GLM versus OLS). In this section, note that AIC values can only be compared between different regressions for the same sample of patients; it is not possible to compare AIC values meaningfully for regressions on different sub-samples.

This generates four sub-samples, summarised in Table 4.3.3. Balance across treatment and comparison groups for each of the four sub-samples was achieved using a sample-specific propensity score (see **§3.2.1** for an overview of propensity score 'stratification'; see Appendix to §4.3 for an overview of the LOS-trimmed sub-samples and the propensity score matching process).

Table 4.3.3 Sub-samples defined by LOS, following propensity score matching

Sample parameter	UC (n=)	PC (n=)	All (n=)	Notes
4<= LOS<=25	717	200	917	Sample trimmed at 2.5th & 97.5th LOS percentile
LOS<=25	700	242	942	Sample trimmed at 97.5th LOS percentile
LOS<=20	686	238	924	Sample trimmed at 95th LOS percentile
4<=LOS	713	229	942	Sample trimmed at 2.5th/5th LOS percentile

For each sub-sample I re-ran GLMs (i) to (v), and estimated incremental effect.

LOS as a dependent variable denominator

Regressions were re-run with daily cost (the ratio of total direct hospital costs and LOS) as the dependent variable for the whole sample and four sub-samples. These were calculated using only Model (i) (no LOS as a covariate) since LOS is already implicitly incorporated in the dependent variable.

Results

LOS as a covariate

The estimated incremental effects for PCCT intervention on costs for the principal sample (n=969) are given in Table 4.3.4.

It is clear that whether (and how) LOS is incorporated into regression as a covariate has an important impact on results. For the model with no LOS as a covariate

(Model i), the estimated effect is negligible: +$153, equivalent to a 1.6% increase in total hospital costs, with a p value of 0.83. Once log-transformed LOS is added to the equation (Model v), the estimated effect is large and significant: -$981, a cost reduction of more than 10%, with a p value of 0.008.

Table 4.3.4 Estimated incremental effect on total direct costs ($) (n=969), by use of LOS as a covariate

Model	Estimated incremental effect		95% Confidence Interval (Standard error)	p value	AIC [Successful replications/1000]
(i)	Mean	+153	-1266 to +1572 (724)	0.83	10554 [1000]
	Median	+147	-1198 to 1493 (687)	0.83	
(ii)	Mean	-10386	-28339 to +7568 (9160)	0.26	10446 [1000]
	Median	-1104	-1781 to -428 (345)	**0.001**	
(iii)	Mean	-121168	-1906604 to +1664269 (910954)	0.89	10487 [1000]
	Median	-814	-1689 to +60 (446)	0.06	
(iv)	Mean	-107998	-7788513 to +7572518 (3918702)	0.98	10518 [1000]
	Median	-480	-1538 to +579 (540)	0.37	
(v)	Mean	-981	-1701 to -261 (367)	**0.008**	10390 [1000]
	Median	-735	-1281 to -188 (279)	**0.008**	

AIC: Akaike information criterion; replications: bootstrap replications/1000 (per **§3.2.1**).

Additionally it is clear that the other three models are of little practical use, producing implausible estimates and very large confidence intervals which imply an absence of predictive power: Model (ii) estimates a mean incremental effect on costs which is higher than mean costs, while (iii) and (iv) deliver estimates more than 10 times larger again.

To estimate mean effect on total costs, this leaves Models (i) [LOS not included as a covariate] and (v) [ln(LOS) included as a covariate]. The results from these two models offer no overlap or consistency; any decision of one over the other has a decisive impact on the results reported.

On the estimates presented, Model (v) is superior to (i). The 95% CI is more efficient (narrower) and it has the superior AIC (information loss) score. However, given that the estimates are mutually incompatible, and that (v) raises endogeneity concerns, it is clear both that use of LOS as a covariate matters and that there is insufficient justification for model selection in Table 4.3.4.

I infer that models (ii), (iii) and (iv) are not delivering reasonable results due to high values in the LOS tail skewing the mean treatment estimation once squared, cubed, etc. The median estimates for (ii), (iii) and (iv) are more plausible, but of little practical use– just as payers do not reimburse for log-transformed means, nor do they for median estimates.

To examine the performance of models (ii), (iii) and (iv) once extreme LOS values have been mitigated, LOS was capped at the 95th percentile (20 days): any LOS over 20 days was recoded as 20 and (ii) to (v) re-run.

The results are given in Table 4.3.5. They reiterate the point that use of LOS as a covariate matters. The results are more consistent than in Table 4.3.4 but still exhibit non-negligible differences. In particular the two best performing models with regard to CI and AIC are Models (ii) and (v), but these offer markedly different incremental-effect estimates and levels of statistical significance. The median results are now more consistent with the mean for each model, with a consistent

pattern of higher mean estimate reflecting the right-hand skew of the cost distribution.

Table 4.3.5 Estimated incremental effect on total direct costs ($) (n=969), by use of LOS as a covariate, LOS recoded so if 20<LOS then LOS=20

Model	Estimated incremental effect		95% Confidence Interval (Standard error)	p value	AIC [Successful replications/1000]
(ii)	Mean	-1597	-2366 to -829 (392)	**<0.001**	10405 [1000]
	Median	-1136	-1691 to -581 (283)	**<0.001**	
(iii)	Mean	-1617	-2438 to -796 (419)	**<0.001**	10418 [1000]
	Median	-1203	-1817 to -588 (313)	**<0.001**	
(iv)	Mean	-1549	-2408 to -691 (438)	**<0.001**	10428 [1000]
	Median	-1187	-1847 to -527 (337)	**<0.001**	
(v)	Mean	-673	-1449 to +103 (396)	0.09	10398 [1000]
	Median	-518	-1125 to +89 (310)	0.09	

LOS as a sample parameter

The estimated mean incremental effects on cost for different sub-samples defined by LOS are given in Table 4.3.6 (4<=LOS<=25), Table 4.3.7 (LOS<=25), Table 4.3.8 (LOS<=20), and Table 4.3.9 (4<=LOS[82]).

A number of clear patterns are discernible.

Defining sub-samples by LOS has an impact on results without LOS as a covariate:

[82] To avoid repetition of the problems caused by long-stay outliers in Models (ii, iii, iv) in Table 4.3.4, LOS was recoded as a covariate for Table 4.3.9 regressions so that it was never greater than 20 (per Table 4.3.5).

- Removing both short- and long-stay outliers increases estimated cost-saving effect somewhat, although this effect is not statistically significant (Model (i) in Table 4.3.6 compared to Model (i) in Table 4.3.4).

- Removing only late-stay outliers increases estimated cost-saving effect, and this effect is statistically significant where LOS<=20 (Model (i) in Table 4.3.7, Table 4.3.8 compared to Model (i) in Table 4.3.4).

- Removing only short-stay outliers decreases estimated cost-saving effect somewhat, although this effect is not statistically significant (Model (i) in Table 4.3.9 compared to Model (i) in Table 4.3.4).

As in Table 4.3.4, the choice to introduce an LOS covariate into the regression generally has a direct impact on the cost-effect estimate and the statistical significance of that estimate:

- Model (i) delivers a markedly different cost-effect estimate to Models (ii) to (v) in Table 4.3.6, Table 4.3.7 and Table 4.3.9. There is consistency between Model (i) and other Models only in Table 4.3.8.

Models using an LOS covariate deliver similar estimates irrespective of specific LOS transformation or sub-sample:

- Models (ii) to (v) deliver consistent results of a statistically significant saving in the $1100-$1600 range in Table 4.3.6, Table 4.3.7 and Table 4.3.8, and (with one quasi-exception - Model iii) in Table 4.3.9.

Model performance is consistent irrespective of sub-sample:

- There is a consistent pattern in terms of which models minimise AIC (information loss) and CI width (efficiency). The descending order of performance is always (v), (ii), (iii), (iv), (i) for AIC; this pattern generally holds for CI width also, with Model (i) consistently the weakest performing, and (v) and (ii) the strongest.

- All models produce 1000 completed replications in the bootstrap process prior to cost-effect estimation. There is no basis on which to favour any model or sub-sample on performance at this stage.

- Excluding Models (ii), (iii) and (iv) in Table 4.3.4 (which are corrected in Table 4.3.5), there is a clear observable consistency in the mean and median cost-effect estimates for each specific model in each table. Therefore with those three exceptions there is no *prima facie* evidence that any Model is delivering cost-effect estimates that are sensitive to outliers.

Table 4.3.6 Estimated incremental effect on total direct costs ($), by use of LOS as a covariate, 4<=LOS<=25 (n=917)

Model	Estimated incremental effect		95% Confidence Interval (Standard error)	p value	AIC [Successful replications/1000
(i)	Mean	-438	-1483 to +607 (533)	0.41	8887 [1000]
	Median	-425	-1437 to +587 (516)	0.41	
(ii)	Mean	-1386	-2081 to -691 (355)	**<0.001**	8812 [1000]
	Median	-1118	-1671 to -564 (282)	**<0.001**	
(iii)	Mean	-1269	-2235 to -302 (493)	**<0.001**	8824 [1000]
	Median	-1063	-1925 to -201 (440)	**<0.001**	
(iv)	Mean	-1136	-1925 to -346 (403)	**0.005**	8835 [1000]
	Median	-974	-1637 to -311 (338)	**0.004**	
(v)	Mean	-1466	-2076 to -855 (312)	**<0.001**	8796 [1000]
	Median	-1173	-1666 to -680 (251)	**<0.001**	

Table 4.3.7 Estimated incremental effect on total direct costs ($), by use of LOS as a covariate, LOS<=25 (n=942)

Model	Estimated incremental effect		95% Confidence Interval (Standard error)	p value	AIC [Successful replications/1000
(i)	Mean	-877	-1957 to +203 (551)	0.11	9868 [1000]
	Median	-863	-1911 to +184 (534)	0.11	
(ii)	Mean	-1517	-2202 to -832 (349)	**<0.001**	9782 [1000]
	Median	-1218	-1764 to -672 (279)	**<0.001**	
(iii)	Mean	-1489	-2173 to -806 (349)	**<0.001**	9797 [1000]
	Median	-1239	-1800 to -679 (286)	**<0.001**	
(iv)	Mean	-1408	-2209 to -608 (408)	**0.001**	9809 [1000]
	Median	-1198	-1868 to -527 (342)	**<0.001**	
(v)	Mean	-1060	-1653 to -467 (303)	**<0.001**	9763 [1000]
	Median	-862	-1347 to -376 (248)	**0.001**	

Table 4.3.8 Estimated incremental effect on total direct costs ($), by use of LOS as a covariate, LOS<=20 (n=924)

Model	Estimated incremental effect		95% Confidence Interval (Standard error)	p value	AIC [Successful replications/1000]
(i)	Mean	-1141	-1983 to -299 (430)	**0.008**	9664 [1000]
	Median	-1127	-1950 to -304 (420)	**0.007**	
(ii)	Mean	-1595	-2227 to -964 (322)	**<0.001**	9599 [1000]
	Median	-1301	-1814 to -789 (262)	**<0.001**	
(iii)	Mean	-1609	-3134 to -84 (778)	**0.04**	9611 [1000]
	Median	-1371	-2641 to -101 (648)	**0.03**	
(iv)	Mean	-1565	-2282 to -849 (366)	**<0.001**	9620 [1000]
	Median	-1379	-1999 to -758 (317)	**<0.001**	
(v)	Mean	-1148	-2228 to -68 (551)	**0.04**	9583 [1000]
	Median	-958	-1793 to -122 (426)	**0.03**	

Table 4.3.9 Estimated incremental effect on total direct costs ($), by use of LOS as a covariate, 4<=LOS (n=942)

Model	Estimated incremental effect		95% Confidence Interval (Standard error)	p value	AIC [Successful replications/1000]
(i)	Mean	+773	-1307 to +2853 (1061)	0.47	9485 [1000]
	Median	+741	-1258 to +2740 (1020)	0.47	
(ii)	Mean	-1423	-2249 to -596 (422)	**0.001**	9354 [1000]
	Median	-1017	-1613 to -420 (304)	**0.001**	
(iii)	Mean	-1358	-2773 to +56 (722)	0.06	9364 [1000]
	Median	-1014	-2141 to +113 (575)	0.08	
(iv)	Mean	-1253	-2159 to -347 (462)	**0.007**	9373 [1000]
	Median	-967	-1657 to -277 (352)	**0.006**	
(v)	Mean	-1193	-2357 to -30 (594)	**0.04**	9348 [1000]
	Median	-917	-1775 to -59 (438)	**0.04**	

LOS as a dependent variable denominator

The estimated incremental effects on daily cost (total direct hospital costs/LOS) using Model (i) (no LOS covariate) only are given for the principal sample and the four sub-samples in Table 4.3.10.

There is a greater consistency, with a statistically significant cost-saving effect irrespective of sample or sub-sample. Nevertheless, the largest estimate of -$174 is over 50% larger than the smallest of -$110, providing further evidence that precise approach to LOS controls directly influences incremental-effect estimates derived. Mean and median estimates are much more similar in magnitude than for total direct costs, reflecting lower skew in the daily cost distribution.

Table 4.3.10 Estimated incremental effect on daily direct costs ($), by LOS-defined sub-sample

Sample	Estimated incremental effect		95% Confidence Interval (Standard error)	p value
Principal (n=969)	Mean	-110	-195 to -24 (43)	**0.01**
	Median	-115	-203 to -27 (45)	**0.01**
4<=LOS<=25 (n=917)	Mean	-174	-249 to -98 (38)	**<0.001**
	Median	-180	-257 to -103 (39)	**<0.001**
LOS<=25 (n=942)	Mean	-138	-215 to -61 (39)	**<0.001**
	Median	-144	-223 to -65 (40)	**<0.001**
LOS<=20 (n=924)	Mean	-155	-227 to -82 (37)	**<0.001**
	Median	-163	-239 to -88 (38)	**<0.001**
4<=LOS (n=942)	Mean	-141	-227 to -54 (44)	**0.001**
	Median	-145	-234 to -56 (45)	**0.001**

Discussion

Previous economic evaluations of in-hospital palliative care consultation have consistently used GLMs (gamma, log) to estimate incremental effect on total cost but

with inconsistent use of LOS as a covariate in regression and/or as a sample parameter. Also commonplace is the consideration of daily cost, i.e. the ratio of total cost to LOS, as a primary dependent variable.

Comparison of incremental-effect estimates derived using these methods confirm the reservations raised in **§2 Literature Review**.

Use of LOS as a covariate and/or as a sample parameter, and use of daily cost as a primary dependent variable, all have a marked effect on both the magnitude and statistical significance of incremental-effect estimates. Simply, the approach chosen in incorporating LOS in analysis has a direct bearing on results. These inconsistencies guard against the arbitrary introduction of LOS-control strategies into analysis. There is also little established guidance on the best approach.

Nonetheless the results do suggest the presence of unobserved heterogeneity within the sample requiring consideration and further analysis.

First, model performance in minimising CI width and AIC information loss is consistently strongest for models that incorporate an LOS transformation as a covariate; on CI and AIC, Model (i - without LOS) is inferior to model (v – ln(LOS)) for the whole sample (Table 4.3.4) and performs worse than all other models for all sub-samples (Table 4.3.6, Table 4.3.7, Table 4.3.8 and Table 4.3.9).

Second, estimated incremental effect on daily cost is consistently statistically significant. Daily costs are a fundamentally different dependent variable to total costs if LOS is systematically different between treatment and comparison groups, as it is with the 'PC4C' data (Table 4.1.4 and Table 4.1.5). Prior to the application of weights, patients in the treatment group exhibit a higher illness burden (Table 4.1.1),

and higher total costs (Table 4.1.3) than comparison patients, but equivalent daily costs (Table 4.1.4). Regressing total and daily costs against treatment and a fixed list of covariates therefore yields different results.

There are two principal explanations for this trend in the data. Most plausibly, palliative care patients' higher illness burden is resulting in longer LOS than equivalent comparison patients, and this difference is not being fully picked up by the propensity scores; this is given added weight by the consideration that long LOS can be a trigger for palliative care. Alternatively it is possible that the difference is explained by palliative care extending LOS, although there is little evidence for this in the literature and the idea runs counter to the focus of palliative care on patient preferences and transition management. If the former explanation is accepted then unobserved heterogeneity is prevalent in the analysis of total costs.

Third, a comparison of Model (i) results (i.e. incremental-effect estimates not using LOS as a covariate) highlights demonstrable difference. The comparison in Table 4.3.11 shows that for all patients there is very little association between treatment and cost. If the sample is trimmed of both short- and long-stay outliers there is a jump to an estimated cost-saving effect, although this is not statistically significant. If 5% of longest-stay outliers are removed, a large and significant incremental effect is observable.

Table 4.3.11 Estimated incremental effect on total direct costs ($), by LOS-defined sub-sample, Model (i) (no LOS as covariate)

Sample	Mean estimated incremental effect (95% Confidence Interval)	p value
Principal (n=969) [per Table 4.3.4]	+153 (-1266 to +1572)	0.83
4<=LOS<=25 (n=917) [per Table 4.3.6]	-438 (-1483 to +607)	0.41
LOS<=25 (n=942) [per Table 4.3.7]	-877 (-1957 to +203)	0.11
LOS<=20 (n=924) [per Table 4.3.8]	-1141 (-1983 to -299)	**0.008**

Two important conclusions can be drawn from the results in Table 4.3.11.

First, the specific points at which the sample is trimmed by LOS has a direct impact on the incremental effects estimated as well as the statistical significance of those effects. There is no effect estimate emerging using LOS as a sample parameter that could be confidently reported as robust to sensitivity analysis.

Second, there is a significant incremental effect on total cost for a large majority of patients but that this effect is being masked by a small latent class of long-stay patients. Propensity score weights are not fully reflecting the heterogeneity of patients but in Tables 4.3.4 to 4.3.9 this is picked up by the introduction into the regression analysis of LOS as a covariate.

This leaves the question of the best approach to estimating incremental effect on costs using the 'PC4C' data unanswered.

The various strategies summarised in this section remain fundamentally sub-optimal. Use of LOS as a covariate introduces unquantifiable endogeneity into the analysis. Use of LOS as a sample parameter arbitrarily removes patients, losing information

and weakening power. This approach also undermines the fundamentals of the project research design (which emphasises calculation of a counter-factual on the basis of observed covariates) and has limited practical value for informing practice (since clinicians cannot predict LOS at admission). Use of LOS as a dependent variable denominator gives fundamentally different results and for a variable that is of limited practicable value.

However, the results in this section represent strong evidence of unobserved heterogeneity in the data. The approach that minimises endogeneity concerns would be to run Model (i) (no LOS covariate) on the principal sample (n=969) to estimate effect on total cost. But the result from this calculation (Table 4.3.4) suggests a negligible association between treatment and cost, while the results from almost all other approaches suggest the presence of a strong association. Accepting the conclusion that there is a negligible association would therefore miss important relationships evident in the data and potentially increase the risk of a type II error. In particular it appears that a small number of long-stay patients are masking a cost effect for the majority: this effect becomes visible if LOS is controlled for as a covariate, if long-stay outliers are removed or if daily costs are used as a dependent variable.

It raises the question of how the latent heterogeneity suggested by results in this section can be meaningfully identified using input data rather than output. In particular, what defines long-stay patients (other than LOS) and can this be meaningfully identified using input/baseline data? Identifying the latent classes in this way would help to estimate effects on cost that are more robust (identifying heterogeneity of effect) while circumventing the methodological concerns inherent to LOS controls (avoiding the endogeneity problem and selection by outcome).

Conclusion

If and how LOS is controlled for in analysis has a substantial effect on results.

Use of LOS-control strategies in estimating incremental effect on costs risks undermining scientific integrity by weakening key elements of the methodology and analysis. Employment of these strategies is inherently arbitrary and little clear guidance exists to inform selection.

Nonetheless a comparison of results using different strategies strongly suggests unobserved heterogeneity in the sample. In particular a small number of long-stay patients appear to be masking an effect for the majority where no LOS-control strategy is employed.

These results pose a challenge in estimating cost effects. Ideally analysis would identify the latent class(es) using input data rather than output, therefore generating meaningful and robust cost-effect estimates that pick up heterogeneity of effect while using baseline data only, thus circumventing the weaknesses inherent in LOS-control methods.

4.4 PCCT impact on hospital costs by time-to-consult

Introduction

The previous section compared different approaches to incorporating length of stay (LOS) in estimating palliative care impact on cost of hospitalisation.

Regression not employing any such strategy suggested no association between treatment and cost in the 'PC4C' dataset. Once LOS is introduced as a covariate, a sample parameter or a dependent variable denominator then a statistically-significant cost-saving association is generally visible.

These LOS-control strategies are necessarily sub-optimal since LOS is associated with treatment (long LOS is an indicator of palliative care need) and the dependent variable (LOS and total cost of hospital are very closely correlated). However, not employing LOS as a covariate or sample parameter suggests no association between treatment and cost, and accepting this result would be to miss important relationships observable in the data. In particular, a small, latent class of long-stay patients appears to be masking an effect for the majority.

The clear question arising from this analysis is: what other factors define long-stay, high-cost patients, and can these be meaningfully identified using baseline/input data? If it is possible to pick up what is distinct about this latent class without relying on utilisation data then it will be possible to estimate cost effects that are robust and reflect sample heterogeneity without relying on sub-optimal LOS controls.

One potential factor is time-to-consult following admission. By definition, longer-stay patients have a greater scope to receive their first consultation after more days in hospital. This has particular implications in economic evaluations since cost is a cumulative variable: each day that a patient is in hospital prior to consult they are

accruing costs that will appear in the primary dependent variable while not being amenable to treatment.

Palliative care is increasingly available earlier in the care trajectory and studies in the clinical literature report observable benefits (Temel et al., 2010, Adelson et al., 2013, Bakitas et al., 2014, Zimmermann et al., 2014), but there is little evidence on whether earlier treatment has economic benefits. This section therefore examines the association between time-to-consult following hospital admission and PCCT impact on cost.

Methods

Research question

What is the impact of the PCCT intervention on hospital costs, incorporating time-to-consult following admission?

Primary analysis

The sample of patients was the principal sample for 'PC4C' economic evaluation (n=969) summarised in **§4.1**. Generalised linear models (GLMs) with a gamma distribution and a log link were performed consistent with **§4.2**. The primary association of interest was mean incremental PCCT effect on hospital costs (the mean estimated USD effect of moving a patient from the usual care group to palliative care group holding all other values constant) using bootstrapped standard errors (1000 replications), per **§3.2.1**, **§3.2.2** and **§4.3**.

Primary independent variable was a binary treatment variable: was the patient seen by a PCCT within a specified time period? The summary statistics of time-to-consult for the intervention group are given in Table 4.4.1. Time-to-consult is clustered within two days of admission: 77% of PC patients were seen by a PCCT by the end of their second day in hospital.

Table 4.4.1 Time-to-consult following admission (days) for PC patients discharged alive (n=256)

	Min.	25th %ile	Median	75th %ile	90th %ile	95th %ile	Max.	Mean
Days to first PCCT	0	0	1	2	6	10	31	2.3

Discussions with R. Sean Morrison, MD and a provisional literature review suggested no firm clinical basis on which to define treatment according to timeliness. Instead I performed five separate analyses differentiated by timing-sensitive definitions of treatment according to the distribution in Table 4.4.1: being seen by a PCCT (i) at any time during hospitalisation; (ii) within 20 days of admission (97.5% of all PC patients); (iii) within 10 days of admission (95%), (iv) within six days of admission (90%) and (v) within two days of admission (75%). These five definitions are not clinically significant but they preserve the sample size and so ensure efficient use of the data.

Patients who received PCCT after the timing cut-off for each definition ($97.5^{th}/95^{th}/90^{th}/75^{th}$) were dropped from the respective analysis. This gives five over-lapping treatment groups defined by timeliness of the patients' first consults. There is one comparison group sub-sample for all four analyses (n=713). See Table 4.4.2.

Table 4.4.2 Summary of sub-samples, by treatment definition, patients discharged alive

Treatment defined as within _____ days of hospital admission (percentile)	UC (n=)	PC (n=)	All (n=)	Proportion seen by PCCT	Mean total direct cost ($)	Median total direct cost ($)	Mean LOS (days)
Any time (100%)	713	256	969	26.4%	9973	7400	8.2
20 (97.5%)	713	249	962	25.9%	9721	7336	8.0
10 (95%)	713	244	957	25.5%	9603	7300	7.9
6 (90%)	713	231	944	24.4%	9509	7230	7.8
2 (75%)	713	197	910	21.7%	9406	7196	7.7

Balance across treated and comparison groups for each of the five definitions of treatment was achieved with sub-sample-specific propensity scores (see **§3.2.1>Additional propensity scores** for an overview of propensity score 'stratification'; see Appendix to §4.4 (Appendix 4.4a) for an overview of the time-to-consult sub-samples and weighting processes).

This overlapping treatment sub-sample approach was taken for three reasons. First, stratifying the treatment group by specific treatment days creates smaller sub-samples; balance between treatment and comparison arms within these smaller samples is generally weaker.

Second, I am interested less in the point-estimate effect for a consultation on a specific day than in the general association between treatment timing and effect on cost; by creating overlapping sub-samples I leverage a greater proportion of the data to examine this association and avoid over-specification of treatment.

Third, the most prominent methodological alternative - interaction terms, used by McCarthy et al. (2015) and discussed in **§2 Literature Review** – is sub-optimal

because it requires stratification of the sample by LOS (which has an endogeneity problem, per **§4.3**) and assumes balance of propensity score covariates between the comparison group and all treatment groups defined by consult timing, an assumption which may not be valid.

In examining several different definitions of treatment according to timing and re-calculating the propensity score for each sub-sample, samples are always matched on baseline characteristics for a clearly defined intervention. All patients had the opportunity to receive treatment in all analyses because I am defining it as *within* x days of admission, not as specifically *on* the xth day.

Additional predictors were the 33 covariates included in the principal sample propensity score and summarised in Table 3.2.1, and fixed site effects for each hospital. Where propensity scores were balanced using a slightly different set of covariates, regressions were still performed with the list of 33 covariates to ensure direct comparability of effect estimates.[83] No LOS covariate was included in any regression.

Secondary analyses

I re-ran the primary analysis with secondary dependent variables: (i) length of stay; (ii) daily direct hospital costs (total cost/LOS), and (iii) total direct costs for major utilisation categories.

Additionally I analysed the descriptive cost data of intervention group patients according to time-to-consult in two ways: (iv) I calculated the costs accumulated in proportional and absolute terms by patients prior to their first consultation (i.e. those

---

[83] A sensitivity analysis where only sample-specific covariates were included in regressions produced very similar results. A fixed list of covariates was employed to maximise comparability of results.

costs that are in the dependent variable but by definition not amenable to treatment), and (v) I compared the day-to-day cost trajectories of patients according to time-to-consult.

Confirmatory analyses[84]

To check the robustness of my primary analysis I performed confirmatory analyses on results, re-running analyses with (a) late-consult outliers moved to the control group (rather than being excluded from the analysis), (b) long-stay outliers removed from both treatment and comparison groups, and (c) patients who died during hospitalisation pooled with those discharged alive.

Finally, consistent with the conclusions of **§4.2**, I performed sensitivity on the model specification by re-running GLM regressions without propensity score weights, and with an OLS specification.

Results

Primary analysis

Estimates of incremental PCCT effect on total cost taking into account time-to-consult are given in Table 4.4.3.

The results demonstrate a clear pattern: earlier treatment is associated with larger cost-saving effect. PCCT interventions within at least 6 days of admission had an estimated mean effect of -$1312 (95% CI: -2568 to -56; p=0.04) compared to no intervention and within two days of -$2280 (95% CI: -3438 to -1122; p<0.001). The magnitude of cost-saving implied is 14% of total hospital cost for a consult within six days and 24% for a consult within two days.

---

[84] Note: As an additional analysis for time-to-consult I also attempted to model treatment as a 'dose effect' continuous variable. Due to the distribution of consult timing (Table 4.4.1) it was not possible to model or retransform the timing to achieve a normal distribution. The analysis therefore produced no output and is not reported further.

The estimated effect of treatment at any time is negligible. Removing late-consult outliers from the sample precipitates a jump to a non-negligible cost-saving estimate; a small number of high-cost, late-consult patients are masking a cost effect for the majority in the 'any time' result.

Table 4.4.3 Estimated incremental effect on total direct costs ($), by treatment timing, patients discharged alive

Treatment defined as within _____ days of hospital admission (percentile)	UC (n=)	PC (n=)	All (n=)	Estimated incremental effect (95% CI)	P value	Implied saving[85]
Any time (100%)	713	256	969	+153 (-1266 to +1572)	0.83	(2%)
20 (97.5%)	713	249	962	-706 (-2007 to +596)	0.29	7%
10 (95%)	713	244	957	-927 (-2284 to +429)	0.18	10%
6 (90%)	713	231	944	-1312 (-2568 to -56)	**0.04**	14%
2 (75%)	713	197	910	-2280 (-3438 to -1122)	**<0.001**	24%

Secondary analyses

First, I re-ran the primary analysis with LOS as the dependent variable. The results are in Table 4.4.4. The same pattern is visible but the association between the intervention and the dependent variable is weaker in this case. The any-time treatment estimate suggests a small, non-significant increase in LOS; removing late-consult outliers changes the coefficient to negative and as treatment definition is tightened the estimated effect grows. Only for patients who were seen by a PCCT within two days is the effect statistically significant.

[85] In hospital costs from receiving treatment compared to receiving usual care only. This is true for all subsequent tables of incremental effect on cost.

Table 4.4.4 Estimated incremental effect on length of stay (days), by treatment timing, patients discharged alive

Treatment defined as within _____ days of hospital admission (percentile)	UC (n=)	PC (n=)	All (n=)	Estimated incremental effect (95% CI)	P value	Implied reduction[86]
Any time (100%)	713	256	969	+0.6 (-0.3 to +1.5)	0.18	(8%)
20 (97.5%)	713	249	962	-0.1 (-0.9 to +0.7)	0.81	1%
10 (95%)	713	244	957	-0.3 (-1.0 to +0.4)	0.41	4%
6 (90%)	713	231	944	-0.5 (-1.3 to +0.2)	0.14	6%
2 (75%)	713	197	910	-1.0 (-1.7 to -0.2)	**0.009**	13%

Second, I re-ran the primary analysis with daily direct cost as the dependent variable. The results are given in Table 4.4.5.

Table 4.4.5 Estimated incremental effect on daily direct costs ($), by treatment timing, patients discharged alive

Treatment defined as within _____ days of hospital admission (percentile)	UC (n=)	PC (n=)	All (n=)	Estimated incremental effect (95% CI)	P value	Implied saving
Any time (100%)	713	256	969	-110 (-195 to -24)	**0.01**	9%
20 (97.5%)	713	249	962	-117 (-200 to -33)	**0.006**	9%
10 (95%)	713	244	957	-117 (-209 to -26)	**0.01**	9%
6 (90%)	713	231	944	-122 (-207 to -37)	**0.005**	10%
2 (75%)	713	197	910	-161 (-248 to -75)	**<0.001**	13%

Compared to Table 4.4.3 the estimated effect is consistently statistically significant irrespective of time-to-consult definition. Consultation at any time during

---

[86] In total length of hospital stay from receiving treatment compared to receiving usual care only. This is true for all subsequent tables of incremental effect on LOS.

hospitalisation is estimated to reduce daily hospital costs by 9% ($p<0.01$) and estimates continue to show a statistically-significant 9-13% reduction with the largest saving for the earliest treatment.

Third, I analysed the underlying utilisation categories from which observed cost-savings are being drawn. In **§2 Literature Review** and **§4.1** I highlighted that the five 'major' utilisation categories in previous studies and in the 'PC4C' data are room and board, ICU costs, pharmacy, laboratory and imaging. For room and board, pharmacy, laboratory and imaging I re-ran the primary analysis with direct costs attributable to each specific category as a dependent variable.

For ICU costs a different approach was required. ICU costs have a particular distribution with 89% of patients having a zero value (see Table 4.1.6). This results in the standard weighted GLM for all patients against ICU costs failing to converge and deliver a treatment estimate, and the sub-sample of patients with non-zero ICU costs is too small to calculate a new propensity score (n=111).

As a substitute I ran an unweighted GLM on this sub-sample with ICU costs as the dependent variable. The ICU estimates are therefore drawn from treatment and comparison groups that are not matched on observed covariates. Sensitivity analyses of the primary analysis show that unweighted regressions typically estimate a smaller cost-saving effect than weighted regressions because the weights compensate for the higher illness burden among palliative care patients.[87] The ICU sub-sample (n=111) demonstrates similar traits with PC patients having significantly longer length of stay and a trend towards higher symptom burden, though with few

---

[87] For comparisons of weighted and unweighted regressions, see Appendices 4.4c and 4.5c.

significant associations in a small sample.[88] All else being equal, therefore, the ICU results are potentially under-estimates of the cost-saving effect, but it is not possible to verify with a weighted analysis.

The results by utilisation category are given in Table 4.4.6. Laboratory costs are found to be significantly reduced for all definitions of treatment timing, and the magnitude of this effect is greater for earlier treatment (from 26% for the intervention at any time to 51% for an intervention within two days of admission). For other categories, statistically significant results are observable for pharmacy and ICU when the first consult was performed within two days of admission. Moreover for all categories coefficients and p values decrease with time-to-consult, i.e. earlier consult appears to be associated systematically with lower costs in all major cost categories.

There is one anomalous result in the top-left box: palliative care is associated with higher room and board costs when the intervention is defined as at any time. This result is not robust to sensitivity analysis (there is no significant association once timing-sensitive definitions of treatment are employed, and would not persist if high-utilization outliers were removed). I therefore infer that it is attributable to the inadequacy of the model where treatment is not timing-sensitive and LOS is not controlled for in any way.

---

[88] For details, see Appendix 4.4b.

Table 4.4.6 Estimated incremental effect on total direct costs ($), 'major' utilisation categories, by treatment timing, patients discharged alive

Treatment defined as within _____ days of admission (%ile)	Room and board	ICU[89]	Pharmacy	Laboratory	Imaging
All (100%)	**+536 (+141 to +932)** **0.008**	-255 (-1495 to +984) 0.69	+180 (-259 to +619) 0.42	**-149 (-228 to -71)** **<0.001**	-3 (-146 to +140) 0.96
20 (97.5%)	+352 (-23 to +726) 0.07	-696 (-1757 to +365) 0.20	+56 (-400 to +513) 0.81	**-185 (-260 to -110)** **<0.001**	-33 (-174 to +109) 0.65
10 (95%)	+169 (-190 to +528) 0.36	-625 (-1588 to +339) 0.20	-162 (-538 to +214) 0.40	**-207 (-433 to -19)** **<0.001**	-37 (-147 to +74) 0.52
6 (90%)	+42 (-286 to +370) 0.80	-842 (-1878 to +193) 0.11	-253 (-1313 to +806) 0.64	**-228 (-303 to -154)** **<0.001**	-73 (-206 to +61) 0.28
2 (75%)	-144 (-1113 to +825) 0.77	**-1162 (-2133 to -191)** **0.02**	**-332 (-651 to -13)** **0.04**	**-285 (-525 to -45)** **0.02**	-136 (-283 to +12) 0.07
Legend:	PC effect in $ (95%CI) p value				

[89] ICU regression results generated using an unweighted, bootstrapped GLM. All other category results generated using a weighted, bootstrapped GLM consistent with the primary analysis. See text for details.

Fourth, I examined unadjusted cost data to further inform the findings from the regression. In particular, a later intervention by definition has diminished capacity to affect the dependent variable because my primary dependent variable is costs incurred during the hospitalisation. That is, those patients who wait longer before seeing a PCCT than the rest of the intervention group accrue a higher level of costs prior to treatment. Additionally, by definition later consults are associated with longer-stay patients, who have higher *total* costs than the rest of the group. I compared the absolute and proportional costs being incurred, by time-to-consult strata.

The results are given in Table 4.4.7. Costs prior to consult are greater when time-to-consult is greater in both absolute and proportional terms. In particular, the earliest consult patients (within two days) notably incur a minority of their costs (28%) prior to seeing a PCCT whereas all other strata incur a majority of their costs prior to a treatment, with the highest proportion (82%) attributable to the latest consults. There is also a clear association between time-to-consult and length of stay.

Table 4.4.7 Utilisation summary, by time-to-consult, PC patients discharged alive (n=226[90])

t days to first PCCT	PC (n=)	Mean LOS (days)	Mean total direct costs ($)	Mean total direct costs prior to first PCCT ($)	Proportion of costs incurred prior to first PCCT
20<*t*	5	36	55497	44814	82%
10<*t*<=20	4	33	33194	15831	60%
6<*t*<=10	9	15	17983	11775	67%
2<*t*<=6	31	10	13260	6755	59%
***t*<=2**	177	7	8692	2072	29%

Fifth, it is possible that later consults are qualitatively different to earlier consults, leading to diminished impact on cost. For example, an earlier consult may focus on goals-of-care discussions and symptom management, but a later consult may focus more on discharge planning, limiting its influence on subsequent in-hospital treatment decisions.

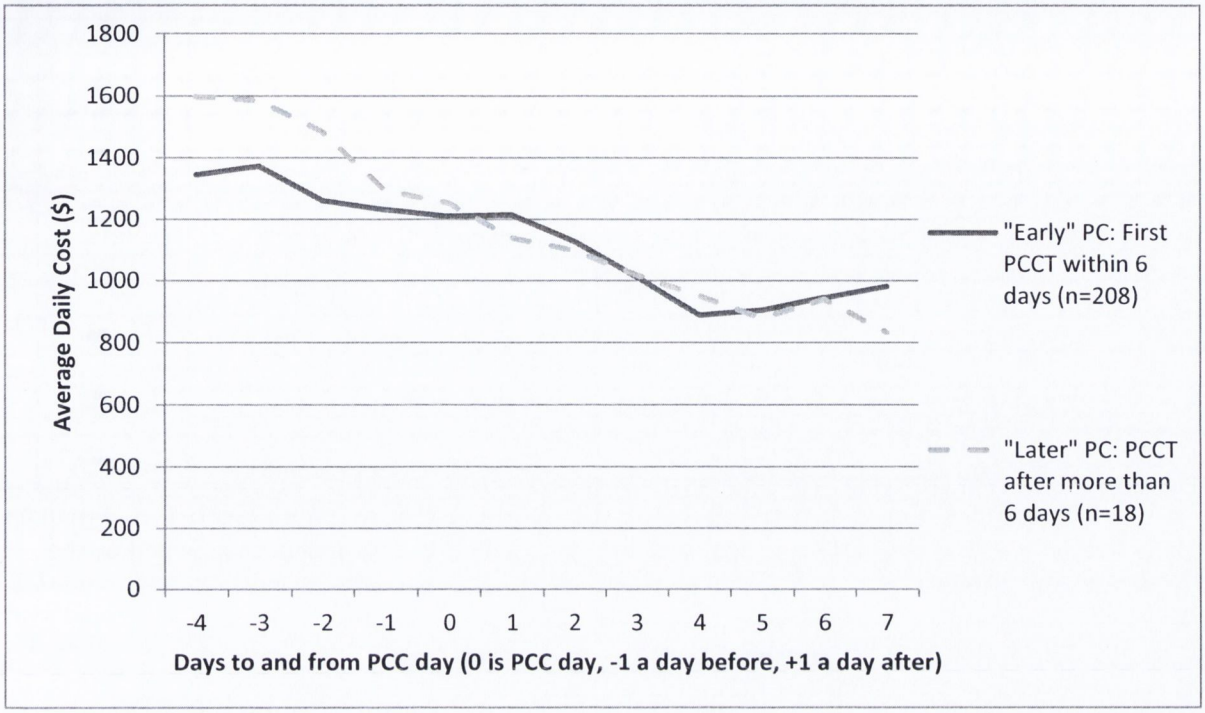
Thus, I examined the average day-by-day cost trajectory for intervention group patients discharged alive, separated into 'early' treatment (within six days, where there is a significant association between treatment and total cost) and 'later' (after six days), in order to examine whether there were differences in cost trajectories across the groups. See Figure 4.4.1.

The strata follow a similar trajectory: in the five days following consult, costs fall by 22% in the 'earlier' group and 18% in the 'later' group.

---

[90] Day-specific cost data available only for three sites (site 2 did not collect data in this way). Day-to-day analysis in Table 4.4.7 and Figure 4.4.1 therefore use n=226, all PC patients for sites 1, 3 and 5.

Figure 4.4.1 Day-by-day direct cost trajectories of intervention patients, by time-to-consult



Costs are also falling prior to palliative care treatment, as is typical over the course of length of stay. There is no evidence in the trajectories that later consults have less of an impact on cost once the consult has been performed.

Confirmatory analyses

I performed a series of additional analyses to confirm the robustness of the primary analysis.

First, it is possible that the late-consult patients excluded from the analysis as definition of treatment was tightened differ systematically from earlier-consult patients. If so, removing these patients may have a substantive impact on results. To address this possibility, I re-ran the primary analysis with late-consult outliers re-categorised as controls. The results are given in Table 4.4.8.

Table 4.4.8 Estimated incremental effect on total direct costs ($), by treatment timing, patients discharged alive; late-consult outliers moved to control group

Treatment defined as within _____ days of hospital admission (percentile)	UC (n=)	PC (n=)	All (n=)	Estimated incremental effect (95% CI)	P value	Implied saving
Any time (100%)	713	256	969	+153 (-1266 to +1572)	0.83	(2%)
20 (97.5%)	719	249	968	-641 (-2001 to +718)	0.36	6%
10 (95%)	725	242	967	-1263 (-2599 to +73)	0.06	13%
6 (90%)	737	232	969	-1993 (-3309 to -678)	**0.003**	20%
2 (75%)	772	197	969	-3262 (-4479 to -2045)	**<0.001**	32%

For two sub-samples one or two patients were not supported by the propensity score (see Appendix 4.4a/Table 2); incremental-effect estimates were calculated using only patients supported by the relevant propensity score.

They are substantively similar to those in Table 4.4.3 with a clear association between earlier consult and cost-saving effect, and again this effect is found to be statistically significant where the first consult was within six days. The implied proportional saving with this method is larger for the earliest consults as a small number of high-cost outliers are being moved to the control group for these calculations instead of excluded.

Second, my method in trimming by time-to-consult is indirectly trimming the treatment group by LOS since late-consult outliers stay longer in hospital. If there are differences between short- and long-stayers that are not fully reflected in the propensity score weights then in trimming the treatment but not the comparison group by LOS I may be biasing my results towards incorrect rejection of the null hypothesis (i.e. a type I error).

To check the robustness of these results controlling for long LOS I re-ran the primary analysis, trimming the sample so that LOS was never greater than day of first

palliative care consult for each definition of treatment: LOS<=20 where treatment is defined as within 20 days, LOS<=10 where treatment is defined as within 10 days, LOS<=6 where treatment is defined as within six days. The results are given in Table 4.4.9.

Table 4.4.9 Estimated incremental effect on total direct costs ($), by treatment timing, LOS outliers removed from both treatment & comparison arms

Treatment defined as within _____ days of hospital admission[91]	UC (n=)	PC (n=)	All (n=)	Estimated incremental effect (95% CI)	P value	Implied saving
20 days (and LOS<=20)	686	238	924	-1141 (-1983 to -299)	**0.008**	13%
10 days (and LOS<=10)	583	192	775	-1376 (-2047 to -705)	**<0.001**	18%
6 days (and LOS<=6)	360	114	474	-1304 (-1950 to -659)	**<0.001**	20%

These suggest the same relationship between earlier treatment and larger proportional cost-saving, confirming that the primary analysis does not reflect trimming of treatment group patients while keeping the comparison group constant. Indeed, estimated cost effect is larger and more often statistically significant using this method.

Third, a sensitivity analysis was performed to check the robustness of results once patients who died during admission (n=54)[92] are included. A repeat of the primary analysis on all patients irrespective of discharge status (n=1,021 following propensity score matching, see Appendix 4.4a) is given in Table 4.4.10. Again, results did not alter substantively, although any difference between survivors and decedents would likely be masked given the respective size of the sub-samples. The magnitude of

---

[91] Patients who saw a consult within 72 hours of discharge excluded as there was insufficient time for treatment to exert any impact on cost.

[92] Died during hospitalisation, n=51; data missing on survival, n=3.

estimates is larger in this analysis, consistent with previous studies that cost effect is larger for decedents than survivors.

Table 4.4.10 Estimated incremental effect on total direct costs ($), by treatment timing, decedents pooled with patients discharged alive

Treatment defined as within _____ days of hospital admission	UC (n=)	PC (n=)	All (n=)	Estimated incremental effect (95% CI)	P value	Implied saving
Any time	735	286	1021	-56 (-1705 to +1592)	0.95	1%
20	735	279	1014	-816 (-2428 to +796)	0.32	8%
10	735	270	1005	-1301 (-2821 to +218)	0.09	13%
6	735	254	989	-2309 (-3867 to -751)	**0.004**	23%
2	735	210	945	-3567 (-5200 to -1933)	**<0.001**	36%

Finally, I performed sensitivity on the model specification by re-running the primary analysis without propensity score weights, and with an OLS specification (see Appendix 4.4c). The results are substantively similar and the weighted GLM results are therefore accepted as valid.

Discussion

The results in this section demonstrate that PCCT intervention within six days of hospital admission is associated with reduced hospital costs, and that earlier consultation is associated with larger effect on cost (Table 4.4.3). This pattern of results is robust in confirmatory analyses incorporating late-consult outliers as controls (Table 4.4.8), controlling for long LOS outliers (Table 4.4.9) and pooling decedents alongside patients discharged alive (Table 4.4.10), and in sensitivity analyses comparing results with different model selection (Appendix 4.4c).

Secondary analysis shows that the cost-savings are attributable to a combination of reduced LOS and reduced intensity of hospital stay. LOS is estimated to be reduced

by one day (13%) when the intervention was within two days of admission compared to usual care patients (Table 4.4.4). A PCCT intervention significantly reduces laboratory costs for all categories of treatment timing, and this is associated with time-to-consult (Table 4.4.6). Earlier consult also appears to reduce systematically the other major utilisation categories of room and board, ICU, pharmacy and imaging. For ICU and pharmacy the relationship is statistically significant when palliative care consult occurred within two days of admission.

The effect estimates for daily cost, which suggest a significant association between intervention and cost for all categories of timing (with earlier treatment exerting a larger effect) (Table 4.4.5), ought to be interpreted carefully. The fact that PCCT is estimated to significantly reduce daily cost irrespective of timing is due to the fact that LOS is higher among PC patients (Table 4.1.5); this means that daily costs between treatment and comparison groups are very similar (Table 4.1.4) while total costs are quite different (Table 4.1.3). The principal conclusion of this analysis – that earlier intervention has larger effect – is verified by the daily cost results but the differing levels of significance bear out the fact that daily cost is a fundamentally different variable to total cost.

Taking the secondary analyses together I infer that the cost-savings reported in the primary analysis are derived primarily from palliative care reducing intensiveness of treatment during hospitalisation, and also from some reduction in LOS if the intervention is provided early. These reductions are larger for an earlier intervention.

The significance of consult timing does not appear to be qualitative, for example with later consults reflecting different patterns of care or consultation team influence. A comparison of day-to-day cost trajectories for earlier and later consult patients shows

no evidence of different effect on costs once the consult has been administered (Figure 4.4.1). Rather, the difference is structural: patients who receive a later consult have systematically accrued higher costs prior to consult than early-consult patients (Table 4.4.7). These costs are included in the dependent variable yet by definition are not amenable to treatment.

These results have a number of implications for both provision of care to patients with serious illness and for research in this field, contingent on the 'non-inferiority' assumption that outcomes from the intervention are at least as good as those from the comparator.[93]

First, to maximise the cost-saving potential of palliative care and improve management of scarce resources, PCCTs should be involved in the care and decision-making of appropriate patients from an early stage of their hospitalisation. This is consistent with a growing body of clinical research suggesting that earlier palliative care access should be more widely implemented.

Second, there is substantial scope to realise currently untapped value from PCCT interventions. In the 'PC4C' study less than a quarter of patients with an advanced cancer diagnosis admitted to hospitals with well-established palliative care programmes were seen by a PCCT within six days of admission. This is not an issue of PCCT capacity at the study sites; instead it reflects a widespread tendency among primary physicians not to consult with palliative care teams concurrently with curative care (personal communication, R. S. Morrison, June 2013). My results

---

[93] For details of this assumption in the wider context of economic evaluation and palliative care evidence, see **§2.4>Forms of economic evaluation** and subsequent sub-sections.

provide further evidence that benefits would be realised through increased take-up of palliative care consultation services by primary physicians.

Third, analysis of the underlying causes of cost-savings suggests reduced resource-intensity and, for early palliative care, reduced LOS. Previous research has suggested mixed results on PCCT impact on LOS, in part due to inconsistent methodology (Cassel et al., 2010); the application of robust methods using a large primary dataset offer strong evidence that early treatment reduces hospital stay.

The sources of reduced utilisation are similar to those found in **§2 Literature Review**, where different studies found different categories to be significant, possibly reflecting different samples. For the 'PC4C' data incremental effect on major utilisation categories was found to be significant only for laboratory costs, plus pharmacy and ICU for an early intervention, but earlier treatment appears to be associated with lower costs across all major categories. Interpretation of the statistically significant associations therefore implies that the intervention reduces costs by reducing the numbers of tests and drugs to which patients are subjected, as well as minimising time in the ICU. But the systematic association between earlier interventions and lower imaging costs implies that this association might also be found to be significant with a larger sample size. [94] The category for which the intervention has the weakest effect is room and board; even though an early intervention reduces LOS, a high proportion of room and board costs are incurred at admission, meaning that this category is not impacted in the same way.

---

[94] For a discussion of how future research can investigate further how the PCCT intervention realises the causal effects reported in this thesis, see **§5.3.3>Further evidence on how a PCCT operates would extend my results**.

Fourth, it appears a mis-specification to measure PCCT incremental effect on cost assuming homogeneity of effect irrespective of consult timing. As detailed in **§2 Literature Review**, prior studies have not incorporated timing into analysis but have controlled for LOS. My results show not only that incremental effect varies by time-to-consult but that this is a structural issue relating to the cumulative nature of costs as a dependent variable. The assumption of homogenous incremental effect irrespective of timing will therefore typically be false, increasing the risk of a type II error; timing directly influences the proportion of the dependent variable that treatment can affect. Moreover, late-consult patients are systematically higher cost patients, meaning that where homogenous effect is assumed these patients have a *larger* impact on mean cost estimates even though their treatment is *less* likely to have a chance to affect the dependent variable. Patients who receive a consult late in their hospital stay are therefore less representative of the intervention studied and more influential on the incremental effect estimated than other patients. While this can be indirectly controlled for by employing LOS strategies it is both more rigorous and more useful to do so by incorporating timeliness explicitly in analyses.

Fifth, combined with **§4.2** and **§4.3**, this section affirms reservations about research methods incorporating LOS to control for unobserved heterogeneity. Mean incremental estimates for the whole sample using such strategies, and mean incremental effect on daily cost would typically show significant associations between treatment and cost irrespective of timing – conclusions that are not robust to sensitivity analysis. This section shows that by identifying a substantive source of heterogeneity that is connected to LOS but does not truncate the sample by outcome, results are attained that are both more robust (minimising endogeneity concerns and robust to sensitivity analyses) and more informative (revealing for the

first time the clear and important association between intervention timing and cost effects, and thus the limits of the assumption of incremental effect homogeneity). To the extent that earlier consultation is associated with more complex patients, methods that incorporate timing may also be incorporating unobserved clinical complexity.

Conclusion

Timely PCCT intervention following admission reduces hospital costs and earlier intervention is associated with a larger cost-saving effect. These results are robust to sensitivity and confirmatory analyses, and have implications for the organisation of hospital care to patients with serious illness. In particular results obtained using these methods are more robust and reveal new and useful information for organisation of, and research on, these interventions.

4.5 PCCT impact on hospital costs by patient multimorbidity

Introduction

In **§4.3** I used the 'PC4C' data to demonstrate the sensitivity of incremental-effect estimates to length of stay (LOS) controls: the estimated impact of palliative care consultation teams (PCCTs) on hospital costs is substantively influenced by how the regression is specified, which patients are included, and which dependent variable is prioritised.

In **§4.4** I illustrated that these differences are (at least partly) attributable to heterogeneity of incremental effect within the sample. Rather than choosing between the different (often mutually exclusive) mean incremental-effect estimates observable for the whole sample in **§4.3**, and using LOS methods to control for unobserved factors, I argued that it is preferable to identify the latent classes within the sample when treatment does (or does not) have a significant effect using baseline data only. In particular, **§4.4** showed that earlier consultation is strongly associated with a larger cost-saving effect, and (in contrast to estimates controlling for LOS) this result is robust to multiple sensitivity analyses.

Where there is significant incremental-effect heterogeneity within a sample, sub-sample-specific results providing evidence of where treatment is (and is not) effective are valuable for at least two reasons. First, they represent more reliable estimates than whole-sample estimates using LOS controls, which are sensitive to specification and potentially suffer endogeneity problems. Second, they can inform organisation and provision of care by identifying in more detail how scarce resources are best allocated.

The **§4.4** results that time-to-consult is strongly associated with incremental effect on cost represents a service-level factor; the results stem from *the way* that care is

delivered. Another potentially important source of incremental-effect heterogeneity is patient-level factors; cost-effect differences may also be observable according *to whom* they are delivered.

One potentially important patient-level factor is multimorbidity (the presence of more than one chronic condition). The relationship between multimorbidity and healthcare costs is well established; for example, among Medicare beneficiaries, the 36% of patients with four or more serious co-occurring conditions account for 74% of programme spending, while the 14% with six or more conditions account for 47% of the budget (Lochner et al., 2013). Health expenditure on serious chronic illness in high-income countries in general and in the United States in particular is projected to exhibit prodigious growth over the next two decades.[95]

As demonstrated in **§2 Literature Review**, no study has examined PCCT cost effect by patient-level factors (with the quasi-exception of interaction terms in McCarthy et al. (2015)) and no study has focused on the association between treatment, cost and multimorbidity. Evidence on this relationship will therefore inform the organisation and provision of cost-effective care to patients with serious illness.

Prior to analysis there are two competing hypotheses on the direction the relationship will take. Specialist palliative care is by definition a complex intervention conceived primarily for patients with serious illness and functional impairment. Multiple chronic conditions work together synergistically (rather than additively) to increase difficulties with finding appropriate medications and treatment regimens that work for all conditions (Ritchie and Zulman, 2013), and so in the abstract one might

---

[95] For a summary of these trends, see **§1 Introduction**.

expect the intervention to have greatest effect where patient need is most complex. However, the analyses in **§4.3** showed a lower incremental effect on cost for longer-stay patients. While this association has been shown to be substantively attributable to consultation timing in **§4.4**, it may also be attributable to a small number of highly complex patients who due to their needs are persistently high-utilisation irrespective of treatment choices, and for whom a cost effect will therefore never be observable.

Methods

Research question

What is the impact of the PCCT intervention on hospital costs, for samples defined by multimorbidity?

Primary analysis

To examine specifically patients with multimorbidity I extracted the number of baseline comorbidities on the Elixhauser scale (Elixhauser et al., 1998) from the principal sample (n=969). The summary statistics are given in Table 4.5.1.

Table 4.5.1 Summary statistics, number of comorbidities at baseline (n=969)

	Baseline Number of Comorbidities (Elixhauser Total)		
	Usual care (n=713)	**Palliative care (n=256)**	**All (n=969)**
Mean	3.2	4.0	3.4
Median	3	4	3
Maximum	11	10	11
75th Percentile	4	5	5
25th Percentile	2	3	2
Minimum	0	0	0

There are 774 patients with multimorbidity (typically one or more serious conditions in addition to advanced cancer; for details see Appendix 4.5b) and this distribution allows me to create four overlapping sub-samples of sufficient size to support a propensity score: patients with 2<=Elixhauser Total, 3<=Elixhauser Total, 4<=Elixhauser Total, and 5<=Elixhauser Total.

Balance across treatment and comparison groups for each of the four sub-samples was achieved using a sample-specific propensity score (see **§3.2** for an overview of propensity score 'stratification'; see Appendix 4.5a for an overview of the multimorbidity sub-samples and their propensity score weights).

Overlapping sub-samples in which less complex patients are incrementally removed was chosen instead of mutually exclusive sub-samples (Elixhauser Total=2, Elixhauser Total=3, etc) for two reasons. First, the latter method entailed small sample sizes with a poorer balance across treatment and comparison groups. Second, I am interested less in a specific effect for patients with a specific number of comorbidities than in the overall association between effect and complexity. This approach allows me to examine this relationship using more of the data and with sub-samples that are not over-specified.

Regression methods were consistent with those used in previous sections: generalised linear models (GLMs) with a gamma distribution and a log link estimating incremental effect on total direct hospital costs using bootstrapped standard errors (1000 replications).

Primary independent variable was a binary treatment variable: was the patient seen by a PCCT within 10 days of admission? In the context of results in **§4.4** this is a definition of treatment for which there is no statistically significant impact on cost in

the overall sample but from which the 12 PC patients with the latest consults are removed as particularly unrepresentative of the intervention and unhelpful in estimating incremental effect on costs.

Additional predictors were the 33 covariates included in the principal sample propensity score and summarised in Table 3.2.1, and fixed site effects for each hospital. Where propensity scores were balanced using a slightly different set of covariates, regressions were still performed with the list of 33 covariates to ensure direct comparability of cost-effect estimates. No LOS covariate was included in any regression.

Secondary analyses

I re-ran the primary analysis with secondary dependent variables: (i) length of stay, (ii) daily direct hospital costs (total cost/LOS), and (iii) direct total cost for major utilisation categories.

I also checked the robustness of the association between time-to-consult and cost effect (per **§4.4**) in the context of sub-samples defined by multimorbidity.

Confirmatory analyses

To check the robustness of my primary analysis I performed confirmatory analyses on results, re-running analyses with (a) late-consult outliers reclassified as controls (instead of excluded from analysis), and (b) patients who died during hospitalisation included alongside those discharged alive.

Finally, consistent with the findings of **§4.2**, I performed sensitivity analyses on the model specification by re-running regressions without propensity score weights, and with an OLS specification.

Results

Data summary

An overview of propensity score matching for the four sub-samples is given in Appendix 4.5a.

Summary unweighted utilisation data for the matched sub-samples are given in Table 4.5.2. As less complex patients are removed from the sample there is a systematic increase in the proportion of patients seen by a PCCT, in total direct costs and in length of stay (LOS).

Table 4.5.2 Summary of sub-samples, by Elixhauser total at baseline, PC within 10 days, patients discharged alive

Sample	UC (n=)	PC (n=)	All (n=)	Proportion seen by PCCT	Mean total cost ($)	Median total cost ($)	Mean LOS (days)
All [Table 4.4.2]	713	244	957	25.5%	9603	7300	7.9
2+ Comorbidities	553	221	774	28.6%	9688	7471	7.9
3+ Comorbidities	436	187	623	30.0%	9946	7698	8.1
4+ Comorbidities	296	136	432	31.5%	10444	8006	8.4
5+ Comorbidities	174	82	256	32.0%	11022	8448	9.0

For details on which comorbidities are most prevalent by sub-sample, see Appendix 4.5b.

Primary analysis

Estimates of incremental effect according to sub-sample are given in Table 4.5.3. There is a clear observable pattern: PCCT treatment is significantly associated with lower total hospital costs for patients with multimorbidity and the effect size grows larger for sub-samples of patients with more comorbidities.

For patients with two or more comorbidities the estimated effect is -$1273 (p=0.04; 95% CI: -2512 to -34), and as less complex patients are removed from the sample, mean incremental effect increases. For three or more comorbidities -$1546 (p=0.02;

95% CI: -2885 to -208); for four more comorbidities -$2113 (p=0.02; 95% CI: -3870 to -357); and for five or more comorbidities -$3050 (p=0.02; 95% CI: -5679 to -422).

These effects are equivalent to 10%, 13%, 15%, 20% and 27% reductions in cost for usual care patients within each sub-sample.

Table 4.5.3 Estimated incremental effect (PC within 10 days) on total direct costs ($), by Elixhauser total, patients discharged alive

Sample	UC (n=)	PC (n=)	All (n=)	Estimated incremental effect (95% CI)	p value	Implied mean saving
All [Table 4.4.3]	713	244	957	-927 (-2283 to +429)	0.18	10%
2+ Comorbidities	553	221	774	-1273 (-2513 to -34)	**0.04**	13%
3+ Comorbidities	436	187	623	-1546 (-2885 to -208)	**0.02**	15%
4+ Comorbidities	296	136	432	-2214 (-4067 to -361)	**0.02**	20%
5+ Comorbidities	174	82	256	-3050 (-5679 to -422)	**0.02**	27%

Secondary analyses

I re-ran the primary analysis with secondary dependent variables: (i) length of stay, (ii) daily direct hospital costs (total cost/LOS) and (iii) direct total cost for major utilisation categories.

The results for LOS are given in Table 4.5.4. The estimated magnitude of reduction in LOS grows for more complex samples but no result is statistically significant.[96]

[96] LOS is significantly reduced for these sub-samples where the first consult was within six days [data not shown]. For an overview of how multimorbidity and time-to-consult factors interact in PCCT's incremental effect on costs, see Figure 4.5.1.

Table 4.5.4 Estimated incremental effect (PC within 10 days) on length of stay (days), by Elixhauser total, patients discharged alive

Sample	UC (n=)	PC (n=)	All (n=)	Estimated incremental effect (95% CI)	p value	Implied mean saving
All [Table 4.4.4]	713	244	957	-0.3 (-1.0 to +0.4)	0.41	4%
2+ Comorbidities	553	221	774	-0.5 (-1.2 to +0.2)	0.18	6%
3+ Comorbidities	436	187	623	-0.6 (-1.4 to +0.2)	0.14	7%
4+ Comorbidities	296	136	432	-0.6 (-1.7 to +0.5)	0.27	7%
5+ Comorbidities	174	82	256	-0.9 (-2.6 to +0.8)	0.30	10%

The results for daily cost effects are given in Table 4.5.5. There is a clear association between PCCT intervention within 10 days and reduced daily cost. Consistent with the results for total cost, this association is strongest for patients with larger numbers of comorbidities, although there is a small inconsistency for the 3<=Elixhauser Total sub-sample.

Table 4.5.5 Estimated incremental effect (PC within 10 days) on daily direct costs ($), by Elixhauser total, patients discharged alive

Sample	UC (n=)	PC (n=)	All (n=)	Estimated incremental effect (95% CI)	p value	Implied mean saving
All [Table 4.4.5]	713	244	957	-117 (-209 to -26)	**0.01**	9%
2+ Comorbidities	553	221	774	-105 (197 to -14)	**0.02**	8%
3+ Comorbidities	436	187	623	-89 (-189 to +9)	0.08	7%
4+ Comorbidities	296	136	432	-149 (-266 to -32)	**0.01**	12%
5+ Comorbidities	174	82	256	-222 (-351 to -93)	**0.001**	18%

The results for major utilisation categories are given in Table 4.5.6. For pharmacy, laboratory and imaging the results follow a clear pattern consistent with **§4.4** (Table 4.4.6): there is a statistically significant association between treatment and reduced

laboratory costs irrespective of multimorbidity sample. This effect and non-significant cost-reducing effects for pharmacy and imaging grow for patients with higher comorbidities, consistent with the primary analysis (Table 4.5.3).

This pattern is not replicated for room and board or ICU costs. In the former case, no result is close to significant with a slight increase in costs associated with treatment. In the latter case, there is a negative coefficient but no significant result and limited inference can be made given the unweighted nature of these regressions.

Table 4.5.6 Estimated incremental effect (PC within 10 days) on total direct costs ($) for 'major' utilisation categories, by Elixhauser total, patients discharged alive

Sample	Room and board	ICU[97]	Pharmacy	Laboratory	Imaging
All **[Table 4.4.6]**	+169 (-190 to +528) 0.36	-625 (-1588 to +339) 0.20	-162 (-538 to +214) 0.40	**-207 (-433 to -19) <0.001**	-37 (-147 to +74) 0.52
2+ Comorbidities	+57 (-268 to +381) 0.73	-742* (-1981 to +496) 0.24	-140 (-532 to +253) 0.49	**-218 (-305 to -132) <0.001**	-47 (-223 to +130) 0.60
3+ Comorbidities	+61 (-295 to +416) 0.74	-967* (-2777 to +844) 0.30	-136 (-522 to +249) 0.49	**-238 (-331 to -146) <0.001**	-137 (-302 to +27) 0.10
4+ Comorbidities	+131 (-393 to +656) 0.62	-1103# (-2261 to +56) 0.06	-332 (-895 to +230) 0.25	**-307 (-426 to -187) <0.001**	**-236 (-470 to -3) 0.04**
5+ Comorbidities	+213 (-616 to +1043) 0.60	-869# (-2330 to +592) 0.24	-351 (-1007 to +304) 0.29	**-257 (-472 to -43) 0.02**	-173 (-579 to +233) 0.40

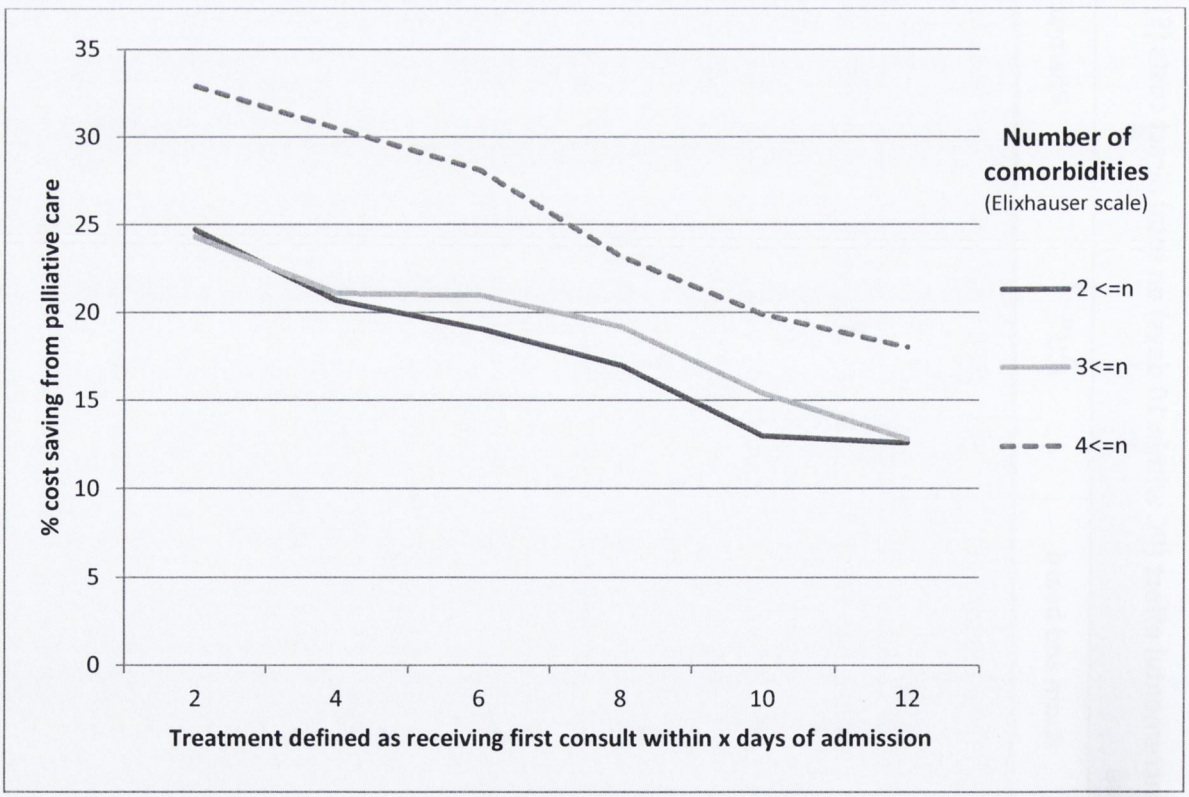
Legend: PC effect in $ (95%CI)
p value

[97] ICU regression results generated using an unweighted GLM (see also Table 4.4.6 and Appendix 4.4c), bootstrapped where possible (denoted *) but bootstrapping not viable for very small sample sizes (n<80; denoted #). All other category results generated using a weighted, bootstrapped GLM consistent with the primary analysis, although (unlike primary analysis) not all 1000 replications are successful in all cases.

Additionally I examined whether the association between time-to-consult and incremental effect observed in **§4.4** is robust for patients with differing levels of multimorbidity. I created overlapping sub-samples defined both by number of comorbidities and by definitions of treatment according to time-to-consult.

The results are illustrated in Figure 4.5.1. They demonstrate that the associations emerging from the analyses in **§4.4** and **§4.5** are both robust. For any given sample defined by complexity, earlier treatment has greater effect; for any given definition of intervention according to timing, the incremental effect is largest for the most complex patients.

Figure 4.5.1 Implied cost-saving, by consult timing & Elixhauser total, patients d/c alive[98]



Incremental-effect estimates are all statistically significant where palliative care is defined as within 10 days or fewer.

[98] It was not possible to calculate propensity scores for the sub-samples where 5<=Elixhauser Total and PC is defined as within less than 10 days; the sample sizes are too small to support sufficient variables in the score. This figure therefore contains only three multimorbidity sub-samples.

Confirmatory analyses

I performed a series of analyses to confirm the robustness of the primary analysis (Table 4.5.3).

First, I repeated my primary analysis with late consult outliers reclassified as controls instead of excluded from the analysis.

The results are given in Table 4.5.7. The same pattern of increased cost savings for patients with multimorbidity is clearly observable, although the magnitudes of estimated savings are larger.

Table 4.5.7 Estimated incremental effect (PC within 10 days) on total direct costs ($), by Elixhauser total, patients discharged alive; late-consult outliers moved to control group

Sample	UC (n=)	PC (n=)	All (n=)	Estimated incremental effect (95% CI)	p value	Implied mean saving
All [Table 4.4.9]	725	242	967	-1263 (-2599 to +73)	0.06	13%
2+ Comorbidities	564	221	785	-1792 (-3130 to -453)	**0.009**	17%
3+ Comorbidities	447	187	634	-2020 (-3428 to -613)	**0.005**	19%
4+ Comorbidities	306	137	443	-2765 (-4605 to -924)	**0.003**	24%
5+ Comorbidities	182	82	264	-3871 (-6718 to -1024)	**0.008**	30%

Second, a sensitivity analysis was performed to check the robustness of results with patients who died during admission (n=54)[99] included. A repeat of the primary analysis on all patients with multimorbidity irrespective of discharge status is given in Table 4.5.8. Results did not alter substantively, but as noted previously any difference between survivors and decedents would likely be masked given the respective size of the sub-samples.

---

[99] Died during hospitalisation, n=51; data missing on survival, n=3.

Table 4.5.8 Estimated incremental effect (PC within 10 days) on total direct costs ($), by Elixhauser total, decedents pooled with patients discharged alive

Sample	UC (n=)	PC (n=)	All (n=)	Estimated incremental effect (95% CI)	p value	Implied mean saving
All [Table 4.4.10]	735	270	955	-1301 (-2821 to +218)	0.09	13%
2+ Comorbidities	573	251	824	-2048 (-3608 to -489)	**0.01**	20%
3+ Comorbidities	454	212	666	-1958 (-3771 to -144)	**0.03**	19%
4+ Comorbidities	313	162	475	-3008 (-5345 to -671)	**0.01**	28%
5+ Comorbidities	187	99	286	-4392 (-7704 to -1080)	**<0.01**	36%

Finally, I performed sensitivity analysis on the model specification by re-running primary analysis regressions without propensity score weights, and with an OLS specification. The results did not change substantively and the weighted GLM results are therefore accepted as valid (see Appendix 4.5c).

Discussion

The primary analysis in this section demonstrates that palliative care consultation within 10 days of hospital admission is significantly associated with reduced hospital costs for patients with multimorbidity (Table 4.5.3). Moreover the effect on cost is larger for patients with a higher number of comorbidities.

The findings are robust in confirmatory analysis incorporating late-consult outliers as controls (Table 4.5.7) and with patients pooled irrespective of discharge status (Table 4.5.5), and following sensitivity analysis employing alternative model selections (Appendix 4.5c).

Secondary analysis suggests that these savings are predominantly due to reduced resource use during hospital stay: mean effect on daily cost is consistently statistically significant (Table 4.5.5) while mean effect on LOS is negligible in either

direction (Table 4.5.4).[100] The categorical sources of these cost-savings reveal diverse results (Table 4.5.6). Treatment is significantly associated with reducing laboratory costs for all multimorbidity sub-samples and the association is generally larger for more complex patients, consistent with **§4.4**. Incremental effect on pharmacy, imaging and ICU costs is generally larger for more complex patients but results are not statistically significant and the pattern is not systematically consistent as in **§4.4**, possibly due to the much larger variation in sub-sample sizes in **§4.5** than in **§4.4**. There is no significant association between treatment cost and room and board cost, although the direction of the relationship is the opposite of the other findings, possibly reflecting increased complexity (and so LOS) among treatment group patients with high numbers of comorbidities that are not fully controlled for by propensity score weights.

The results reported in this section and in **§4.4** remain clear when the two analytical approaches are combined (Figure 4.5.1). For any given sample defined by complexity, earlier treatment has a greater effect; for any given definition of the intervention according to timing, the incremental effect is largest for the most complex patients.

I am not aware of any previous study to examine this research question, and my results have a number of potentially important implications for provision and research, given the 'non-inferiority' assumption that outcomes from the intervention are at least as good as those from the comparator.[101]

---

[100] Consistent with **§4.4**, earlier consultation does reduce LOS significantly for these sub-samples [data not shown].

[101] For details of this assumption in the wider context of economic evaluation and palliative care evidence, see **§2.4>Models of economic evaluation** and subsequent sub-sections.

First, to maximise the potential financial benefits of palliative care, PCCTs should be involved in the care and decision-making of patients with multimorbidity – and in particular those with high numbers of comorbidities. This is consistent with the prior hypothesis that specialist care is a complex intervention maximally effective for the most complex patients. Early palliative consultations for the sickest patients may help reverse the trend of more aggressive end-of-life care in the United States hospital system. Patterns of improved quality and reduced costs through coordinated, patient-centred palliative care are already evident in the literature (El-Jawahri et al., 2011, Higginson et al., 2003, Smith et al., 2014). My results show for the first time that the cost benefits are larger than previously understood for the most complex patients, and so the case for providing palliative care to this patient group is particularly strong.

Second, there are clear policy implications. The long-term viability of US government health programmes such as Medicaid and the overall national health system depends on reforming care to patients with serious illness in a way that reduces costs without compromising quality and access (Anderson, 2005, Smith and Hillner, 2011). These results further strengthen the case for increased palliative care provision for patients with serious illness through up-scaling the successful PCCT model, and designing of payment and financing mechanisms that reward value for money (Unroe and Meier, 2013).

Third, there are implications for workforce planning. There are demonstrable short- and long-term gaps in the hospice and palliative care staff in the US (Bui, 2012, Lupu, 2010, Quill and Abernethy, 2013). The projected level of future need is such that not all patients will be seen by specialist consultants; rather, specialist staff are and will remain a scarce resource to be allocated in the most effective way. PCCTs

are most impactful where patient needs are most complex, so it is the most complex patients whose treatment by specialist palliative care should be prioritised.

Fourth, and consistent with **§4.4**, further cost-savings could be achieved within current levels of provision by increasing the number of patients seen by a palliative care team. In the 'PC4C' study less than 30% of patients with an advanced cancer diagnosis and multimorbidity admitted to hospitals with well-established palliative care programmes are seen by a PCCT within 10 days of admission. Yet, as already stated in **§4.4**, this is less an issue of limited PCCT capacity than of referral patterns among primary physicians for whom palliative care is an option not taken up. There is demonstrable scope for further cost-savings to be realised if more patients with multimorbidity were referred to a PCCT. Furthermore, these cost-savings are greatest when referral is made early in the hospitalisation.

Fifth, and consistent with **§4.4**, this section confirms the existence of cost-effect heterogeneity in the dataset. The results presented here emphasise the benefits of going beyond mean estimates for the whole sample. In identifying major patient groups for whom treatment cost effects are substantively different, I provide evidence that can further inform organisation and provision of care.

Finally, and consistent with **§4.4**, observed savings accrue through reduced intensity of hospital stay and (where the intervention is delivered early enough) reduced length of stay. My results on resource utilisation category sources are consistent with previous literature: palliative care appears to reduce spending for most major

categories but the statistical significance of relationships varies by (sub-)sample.[102] Interpretation of specific utilisation category effects in Table 4.5.6 is complicated by small sample sizes; there are 256 patients in the '5+ Comorbidities' sub-sample and not all accrued hospital costs in all categories.

Conclusion

PCCT treatment during hospital admission for patients with multimorbidity is associated with a statistically significant cost-saving effect and this association is largest for the most complex patients. These results are robust to sensitivity and confirmatory analyses, and have implications for the organisation of hospital care to patients with serious illness, to policy and finance frameworks, and to research on these programmes.

---

[102] For a discussion of how future research can investigate further how the PCCT intervention realises the causal effects reported in this thesis, see **§5.3.3>Further evidence on how a PCCT operates would extend my results**.

4.6 Patient-level determinants of utilisation

Introduction

The previous results sections have focused on the estimated effect of palliative care consultation teams (PCCTs) on hospital cost for patients with serious illness. Those analyses reflect the primary purpose of the 'PC4C' study and this PhD thesis, namely improving understanding of and evidence on the impact of the PCCT intervention.

Having established robust methodological approaches to cost analysis with the 'PC4C' dataset in those previous sections, this section complements prior analyses by examining which factors other than palliative care treatment impact cost of hospital care for patients admitted to hospital with advanced cancer. Patterns and levels of resource utilisation in providing healthcare to patients with serious illness reflect not only treatment choices but a complex set of relationships between demographic, clinical and system factors (Groeneveld et al., 2013).

In particular patient-level variables may be strong predictors of utilisation (Tibi-Levy et al., 2006, Simoens et al., 2010). Patient-level factors previously identified as potentially significant drivers of resource use among seriously-ill patients include age (Shugarman et al., 2007), gender (Shugarman et al., 2008), ethnicity (Hanchate et al., 2009), socio-economic status (Hanratty et al., 2007), advance directive status (Kelley et al., 2011), insurance status (Kelley et al., 2011), comorbidities (Shugarman et al., 2007), cancer diagnosis (Walker et al., 2011) and functional status (Guerriere et al., 2010).

Information on those factors which are strongly associated with utilisation can inform the provision of care, assisting the efficient allocation of resources and the consideration of equity concerns in patient access and use.

Methods

Research question

Which patient-level factors are significantly associated with hospital costs in the 'PC4C' dataset?

Regression analyses

Analysis in this section was performed in two parts.

First, I examined versions of the regressions used in primary analysis in **§4.4** (Table 4.4.3) and **§4.5** (Table 4.5.3). Specifically, I regressed cost on a binary treatment variable, all covariates included in each sample-specific propensity score as well as fixed effects for each hospital site.[103] This approach varies from that used in previous regressions, which used as independent variables in all cases a set list of 33 covariates included in the primary propensity score (Table 3.2.1) as well as fixed site effects.

The methods used in this section are distinct because only for variables included in each sample-specific propensity score (and so balanced across treatment and comparison groups) can the coefficient be reliably interpreted as its impact on the dependent variable, whereas a fixed list of predictors was preferable in estimating PCCT cost effect since the results are directly comparable.[104]

For each of the regressions in this section I calculated the marginal dy/dx, the discrete change from the base level, for each of the covariates. For binary variables, the marginal dy/dx is the estimated marginal increase in total cost associated with a

---

[103] Where sub-samples defined by time-to-consult or multimorbidity would not balance on the preferred list of 33 covariates, strategies were employed to balance treatment and comparison groups while minimising information loss. See **§3.2.1>Additional propensity scores**, Appendix 4.4a and Appendix 4.5a.

[104] I ran sensitivity analysis on the primary results in **§4.4** and **§4.5** to confirm that incremental-effect estimates are robust whether a fixed list of 33 covariates or only sample-specific propensity score covariates are included as predictors. Estimates exhibit negligible difference.

move from the reference group holding all other covariates to their original values. For continuous variables, it's the estimated cost effect associated with a one-unit increase in the variable. These calculations allow me to examine which patient factors are associated with hospital costs where treatment is controlled for, and there is balance across the treatment and comparison groups.

Second, I ran equivalent unweighted regressions separately on treatment and control groups. Unweighted regressions are useful in this context because they allow analysts to incorporate any variable in the dataset (and not just those in the propensity score). In particular, my methods to date have not incorporated specific cancer diagnoses (with the exception of lymphoma) due to sample size and propensity score balancing issues.[105]

Given the established significance of diagnosis in driving utilisation, this is an important relationship to explore. An additional benefit of the unweighted regressions is that they serve as sensitivity analyses to the propensity score weighting itself, and the extent to which this is influencing results.

In this unweighted analysis I removed one very-high-cost outlier[106] who misleadingly skewed results and divided the remaining sample (n=968) into patients who were seen by a PCCT (n=255) and those who did not (n=713). I repeated the regression methods of the primary analysis in using the same generalised linear model (GLM) with a gamma distribution and a log link, and all covariates from Table 3.2.1 plus six additional diagnosis covariates (gastrointestinal (GI), lymphoma, central nervous

---

[105] See **§3.2.1> Creating a propensity score for the principal sample: step-by-step>Step 1**.

[106] Total direct hospital costs=$136930, 19 times the median and roughly twice the next most expensive patient.

system (CNS), breast, gynaecological, hepatocellular carcinoma (HCC) and lung, with 'other cancer diagnosis' as the reference category).

Therefore in separate analyses of treatment and comparison groups I followed methods used in primary analysis in **§4.4** and **§4.5** except that no propensity score weights were applied and, where previously I controlled for the lymphoma diagnosis only, I instead controlled for seven diagnoses.

Results

The results of the weighted regressions, with treatment and comparison groups pooled and PCCT impact controlled for, are presented in Table 4.6.1 (with sensitivity analyses by timing and multimorbidity). The results of the unweighted regressions are given in Table 4.6.2 (usual care group, with sensitivity analyses by multimorbidity) and Table 4.6.3 (palliative care group, with sensitivity analyses by timing and multimorbidity).

In the weighted regressions (Table 4.6.1) there are five predictors which exhibit a consistent statistical relationship with costs (statistically significant relationship in the principal sample and multiple sensitivity analyses).

- *Number of comorbidities:* higher Elixhauser Total at baseline associated with higher cost;
- *ESAS physical score:* higher symptom distress at admission associated with lower cost;
- *CMSAS score:* higher number of physical symptoms at admission associated higher cost;

- *Education:* College education associated with higher cost, elementary school education (but no further) with lower cost;
- *Insurance:* Medicare and Medicaid associated with lower cost.

In the unweighted regressions among usual care patients (Table 4.6.2) there are seven covariates which exhibit a consistent statistical relationship with costs (statistically significant relationship in the principal sample and more than one sensitivity analysis):

- *HCC diagnosis:* associated with lower costs;
- *Lung diagnosis:* associated with lower costs;
- *Number of comorbidities:* higher Elixhauser Total at baseline associated with higher cost;
- *ESAS physical score:* higher symptom distress at consultation associated with higher cost;
- *ESAS psychological score:* higher psychological distress at consultation associated with lower cost;
- *CMSAS score:* higher number of physical symptoms at consultation associated higher cost;
- *CMSAS severity:* higher severity of physical symptoms at consultation associated lower cost.

In the unweighted regressions among palliative care patients (Table 4.6.3) there are five covariates which exhibit a consistent statistical relationship with costs

(statistically significant relationship in the principal sample and at least one timing sensitivity analysis):

- *Visiting nurses services prior to admission:* associated with lower cost;
- *Gynaecological diagnosis:* associated with higher costs;
- *Lung diagnosis:* associated with higher costs;
- *Number of comorbidities:* higher Elixhauser Total at baseline associated with higher cost;
- *CMSAS score:* higher number of physical symptoms at admission associated higher cost.

Table 4.6.1 Association between patient-level factors and total direct costs ($) - weighted regressions

	Principal sample		Time-to-consult sensitivity analysis								Multimorbidity sensitivity analysis							
	n=969		PC within 20 days n=962		PC within 10 days n=956		PC within 6 days n=944		PC within 2 days n=910		2+ Comorbidities n=744		3+ Comorbidities n=623		4+ Comorbidities n=433		5+ Comorbidities n=256	
	dy/dx	p	dy/dx	p	dy/dx	P	dy/dx	p	dy/dx	p	dy/dx	p	dy/dx	p	dy/dx	p	dy/dx	p
Age: *55 to75*[a]	-1346	0.10	-1153	0.12	-35	0.16	**-1440**	**0.04**	**-1642**	**0.02**	**-1598**	**0.03**	**-1950**	**0.02**	-42	0.33	-27	0.57
Age: *Over 75*[a]	-17	0.99	556	0.62			-149	0.88	-283	0.77	47	0.96	51	0.96				
Gender: *Female*	485	0.47	367	0.56	70	0.91	-96	0.87	119	0.83	18	0.98	27	0.97	-330	0.66	-555	0.57
Race: *White*	-4379	0.20	-4439	0.20	-5006	0.19	-5439	0.18	-8890	0.09	-6986	0.21	-14263	0.18	-16656	0.29	**5655**	**0.01**
Race: *Black*	-4766	0.07	-4715	0.09	-4729	0.11	-5127	0.09	**-7681**	**0.04**	-6317	0.14	-11747	0.12	-12166	0.19	6132	0.06
Living Will: *Yes*[b]	-1947	0.08	-1269	0.21	-1224	0.21	-1373	0.12	**-1644**	**0.02**	-1601	**0.02**	-375	0.71	**-2289**	**0.02**	**-2939**	**0.03**
Proxy: *Yes*[b]	334	0.76	-466	0.63	-340	0.71	-367	0.67					-1253	0.20				
Insurance: *Medicare*	**-2094**	**0.01**	-1389	0.07	-948	0.21	-883	0.22	-468	0.51	-878	0.20	-391	0.63	391	0.73	1127	0.38
Insurance: *Medicaid*	-1186	0.13	-1147	0.10	**-1345**	**0.04**	**-1320**	**0.04**	**-1286**	**0.04**	-637	0.37	-58	0.95	-409	0.71	1307	0.43
Education: *Elementary*	-576	0.63	-1326	0.13	-1445	0.10	-1168	0.18	-1332	0.10	-1312	0.13	-1613	0.09	**-2406**	**0.01**	**-4330**	**<0.01**
Education: *College*	**2253**	**0.01**	**1952**	**0.01**	1237	0.10	**1591**	**0.03**	1270	0.06	942	0.16	987	0.20	747	0.43	102	0.93
Visiting nurse: *Yes*	-808	0.50	-919	0.39	-575	0.60	-169	0.87	-373	0.71	-398	0.69	-153	0.91	236	0.88	-285	0.88
Home aide*: Hours*	-147	0.72	-96	0.82	-192	0.63	-38	0.46	146	0.72	-17	0.76	-51	0.26	-10	0.89	-77	0.07
Diagnosis: *Lymphoma*	2404	0.12	1641	0.22	2007	0.16	2134	0.16	2691	0.11	1937	0.18	1777	0.29	4835	0.07	2220	0.32
Complication(s): *Yes*	-415	0.82	-456	0.79	-187	0.92	-1530	0.23	**-2430**	**0.04**	899	0.67	1072	0.69	1428	0.64	**-3891**	**0.02**

Comorbidity: *Elixh'ser*[c]	**886**	**<0.01**	**650**	**<0.01**	**544**	**<0.01**	**496**	**<0.01**	**578**	**<0.01**	**765**	**<0.01**	**977**	**<0.01**			**1215**	**0.01**
ADL: *Bathing*	-1177	0.22	-745	0.41	-96	0.92	-35	0.97	-459	0.53	-734	0.34	-311	0.72	356	0.76	2421	0.09
ADL: *Chair*	29	0.98	-323	0.72	-443	0.60	-496	0.52	-110	0.88	-37	0.96	-399	0.66	-282	0.82	87	0.95
ADL: *Complete assist 1+*	802	0.47	1359	0.22	913	0.40	-51	0.96	-46	0.96	1338	0.20	1397	0.23	579	0.71	-1172	0.43
ESAS Phy: *Admission*[d]	**-4277**	**0.01**	-439	0.68	**-3177**	**0.02**	**-3529**	**0.01**	**-2525**	**0.04**	**-2898**	**0.03**	**-3020**	**0.04**	-3260	0.10	-1934	0.26
ESAS Psy: *Admission*[d]	-312	0.44	**-517**	**0.01**	-633	0.08	-618	0.08	-329	0.30	-207	0.55	-115	0.77	-310	0.56	-129	0.72
ESAS Phy: *Consult/ref*[d]	2753	0.13			3123	0.06	2857	0.08	2656	0.08	2060	0.19	3163	0.09	915	0.70		
ESAS Psy: *Consult/ref*[d]	-588	0.13			-269	0.45	-288	0.43	-549	0.09	-471	0.17	-617	0.12	-567	0.31		
CMSAS #: *Admission*[d]	**793**	**0.01**	**597**	**0.01**	**877**	**<0.01**	**971**	**<0.01**	**868**	**<0.01**	**662**	**0.01**	548	0.07	236	0.56	15	0.96
CMSAS #: *Consult/ref*[d]	421	0.19			212	0.48	196	0.50	**644**	**0.01**	396	0.11	463	0.12	784	0.06		
Severity: *Admission*[d]	127	0.38	-105	0.36	24	0.84	47	0.68	-45	0.65	51	0.65	100	0.43	122	0.50	113	0.48
Severity: *Consult/ref*[d]	-328	0.11			-285	0.13	-266	0.15	**-435**	**<0.01**	**-313**	**0.04**	**-472**	**0.01**	-278	0.26		
Morphine: *Dose (mg)*	104	0.35	156	0.14	138	0.16	14	0.11	116	0.14	2	0.74	3	0.77	2	0.90	-1	0.95
Pain: *Somewhat*[e]	157	0.87	99	0.92	224	0.81	-21	0.98	359	0.67	-143	0.87	168	0.88				
Pain: *Quite a bit*[e]	371	0.69	331	0.71	-327	0.70	-16	0.98	-114	0.88	142	0.86	318	0.71				
Pain: *Very much*[e]	660	0.52	641	0.47	167	0.86	-62	0.94	717	0.39	543	0.52	502	0.59				
Fatigue: [Not v. much][e]	-29	0.97	342	0.66	-147	0.85	-104	0.89	-120	0.86	494	0.50	813	0.35				
Fatigue: *Very much*[e]	524	0.58	563	0.51	-231	0.78	-22	0.98	-158	0.82	137	0.86	393	0.66				

For a full explanation of covariate selection and definition see Table 3.2.1. Only coefficients for those covariates that are in the sample-specific propensity score can be reliably interpreted as representing the effect on the dependent variable so only these are included in each regression. An overview of how scores were calculated is given in **§3.2.1>Additional propensity scores**. The details of each propensity score are given in Appendix 4.4c and Appendix 4.5c. In this Table: [a] Where 'Age over 75' is missing, 'Age 55 to 75'=Age as a continuous variable. [b] Where Proxy is missing, Living Will=Advanced Directive (Living Will | Proxy). [c] Where Elixhauser is missing, it was not included in the score. [d] Where 'ESAS' and 'CMSAS' scores at ref/consult day are missing, the corresponding 'Admission' entry=the total at baseline and admission/reference day. [e] Where Pain and Fatigue are missing, these were not included in the score.

Table 4.6.2 Association between patient-level factors and total direct costs ($) - unweighted regressions, comparison group

	All		Multimorbidity Sensitivity Analysis							
	n=713		2+ Comorbidities n=553		3+ Comorbidities n=436		4+ Comorbidities n=296		5+ Comorbidities n=174	
	dy/dx	p	dy/dx	p	dy/dx	p	dy/dx	p	dy/dx	p
Age: *55 to75*	400	0.53	728	0.33	578	0.54	-274	0.83	-1292	0.54
Age: *Over 75*	25	0.98	991	0.41	1202	0.43	25	0.99	155	0.96
Gender: *Female*	-836	0.17	**-1536**	**0.03**	-1346	0.11	-1669	0.13	-2534	0.11
Race: *White*	1556	0.14	755	0.57	908	0.58	2921	0.23	2457	0.58
Race: *Black*	1278	0.33	397	0.79	193	0.92	2327	0.43	1433	0.77
Living Will: *Yes*	-565	0.44	-764	0.38	-601	0.57	-1494	0.25	-1785	0.39
Proxy: *Yes*	-1010	0.18	-1186	0.19	-2012	0.07	-2481	0.07	-316	0.88
Insurance: *Medicare*	-640	0.39	-495	0.53	133	0.89	1630	0.21	-73	0.97
Insurance: *Medicaid*	-522	0.52	-114	0.90	150	0.90	37	0.98	-431	0.83
Education: *Elementary* [a]	-1007	0.33	-1337	0.22	-1645	0.17	-2697	0.05	-529	0.82
Education: *College* [a]	154	0.80	205	0.76	121	0.88	897	0.40	1428	0.43
Visiting nurse: *Yes* [b]	521	0.59	225	0.84	715	0.60	10	1.00	-772	0.78
Home aide: *Hours* [b] ‡	-402	0.13	-350	0.20	-319	0.28	-100	0.80	-347	0.49
Diagnosis: *GI*	80	0.92	11	0.99	-102	0.92	-467	0.74	388	0.86
Diagnosis: *Lymphoma*	274	0.80	1097	0.41	1261	0.46	651	0.77	-51	0.99
Diagnosis: *CNS*	**7957**	**0.02**	**8481**	**0.03**	177	0.95	138	0.98	289	0.96
Diagnosis: *Breast*	539	0.63	1851	0.19	814	0.61	-1015	0.58	-2731	0.29
Diagnosis: *Gynaecological*	-783	0.38	-95	0.93	-761	0.55	-1493	0.36	457	0.87
Diagnosis: *HCC*	**-3456**	**0.01**	**-4723**	**<0.01**	**-5379**	**<0.01**	**-5331**	**0.01**	**-7153**	**<0.01**

Diagnosis: *Lung*	**-2155**	**0.01**	**-2156**	**0.01**	**-2460**	**0.01**	**-4136**	**<0.01**	**-4501**	**0.01**
Complication(s): *Yes*[c]	527	0.69	282	0.85	-1035	0.50	-962	0.61	-1257	0.62
Comorbidities: *Elixhauser*	**571**	**<0.01**	**848**	**<0.01**	**1029**	**<0.01**	**1243**	**<0.01**	**1426**	**0.03**
ADL: *Bathing*	-880	0.28	-1003	0.25	-1320	0.20	-1134	0.38	-543	0.78
ADL: *Chair*	756	0.41	834	0.41	1010	0.40	2509	0.13	4511	0.09
ADL: *Complete assist 1+*	1235	0.34	1011	0.45	992	0.51	-1066	0.54	-318	0.91
ESAS Phy: *Admission*	-742	0.55	-2644	0.06	-3110	0.07	**-4678**	**0.03**	**-7404**	**0.02**
ESAS Psy: *Admission*	319	0.32	207	0.57	288	0.53	374	0.52	78	0.93
ESAS Phy: *Consult/ref*[d]	**4013**	**<0.01**	**3636**	**0.02**	**4755**	**0.01**	**5070**	**0.03**	5706	0.11
ESAS Psy: *Consult/ref*[d]	**-951**	**<0.01**	**-677**	**0.04**	**-937**	**0.02**	**-1724**	**<0.01**	**-2377**	**<0.01**
CMSAS: *Admission*	-251	0.27	-287	0.27	-170	0.60	-21	0.96	299	0.64
CMSAS: *Consult/ref*[d]	**369**	**0.01**	284	0.12	**578**	**0.02**	**769**	**0.03**	767	0.12
Severity[e]**:** *Admission*	146	0.19	**348**	**0.01**	**345**	**0.03**	**415**	**0.04**	564	0.05
Severity[e]**:** *Consult/ref*[d]	**-396**	**0.01**	**-359**	**0.03**	**-537**	**0.01**	**-583**	**0.03**	-573	0.13
Morphine: *Dose (mg)*[f]‡	79	0.41	108	0.31	113	0.39	35	0.84	210	0.42
Pain: *Somewhat*[g]	-608	0.43	-642	0.48	207	0.86	-1368	0.31	-67	0.98
Pain: *Quite a bit*[g]	-221	0.77	152	0.86	59	0.96	346	0.80	1102	0.59
Pain: *Very much*[g]	-7	0.99	-284	0.77	-570	0.63	-1348	0.35	856	0.75
Fatigue: [Not v. much][gh]	-198	0.76	238	0.75	665	0.47	95	0.94	-794	0.65
Fatigue: *Very much*[g]	-1168	0.17	-672	0.49	-88	0.94	-398	0.79	-1807	0.41

Table 4.6.3 Association between patient-level factors and total direct costs ($) - unweighted regressions, treatment group

	All		Timing Sensitivity Analysis				M/morbidity Sensitivity Analysis			
	n=256		PC within 10 days n=243		PC within 6 days n=231		Comorbidities<=2 n=234		Comorbidities<=4 n=149	
	dy/dx	p	dy/dx	P	dy/dx	p	dy/dx	p	dy/dx	p
Age: *55 to75*	-611	0.62	-998	0.31	-1489	0.10	-1531	0.21	-1630	0.37
Age: *Over 75*	2118	0.34	2195	0.21	2121	0.20	1674	0.43	-2695	0.25
Gender: *Female*	1405	0.24	755	0.41	163	0.85	1595	0.17	2231	0.18
Race: *White*	-566	0.85	-254	0.91	151	0.94	2646	0.39	3746	0.54
Race: *Black*	-1081	0.72	-236	0.92	119	0.96	2370	0.52	2207	0.77
Living Will: *Yes*	**-3716**	**0.02**	-2294	0.07	-1877	0.11	-2789	0.10	-3633	0.16
Proxy: *Yes*	2570	0.15	1914	0.16	1270	0.32	1029	0.56	908	0.72
Insurance: *Medicare*	-1418	0.32	801	0.52	1131	0.34	-1505	0.26	291	0.89
Insurance: *Medicaid*	-1715	0.21	-860	0.42	-1082	0.27	-1491	0.27	-753	0.72
Education: *Elementary* [a]	197	0.92	-782	0.59	-822	0.53	-158	0.93	485	0.86
Education: *College* [a]	2617	0.06	1030	0.30	961	0.29	2215	0.09	3371	0.09
Visiting nurse: *Yes* [b]	**-4173**	**<0.01**	**-2490**	**0.04**	**-2267**	**0.04**	**-3597**	**0.01**	**-4663**	**0.02**
Home aide: *Hours* [b] ‡	-198	0.78	-326	0.55	-31	0.95	191	0.77	-189	0.85
Diagnosis: *GI*	678	0.68	1009	0.44	685	0.56	136	0.93	-548	0.79
Diagnosis: *Lymphoma*	6987	0.07	5192	0.07	4945	0.08	5311	0.15	3978	0.30
Diagnosis: *CNS*	-3005	0.50	-3446	0.24	-3224	0.22	-2726	0.53	**-9003**	**<0.01**
Diagnosis: *Breast*	464	0.82	490	0.77	497	0.74	648	0.76	2522	0.50
Diagnosis: *Gyn'cological*	**8255**	**0.04**	**5729**	**0.04**	**6912**	**0.02**	5096	0.15	6050	0.26
Diagnosis: *HCC*	-536	0.89	-2235	0.38	-1851	0.43	-2088	0.54	1213	0.89

Diagnosis: *Lung*	**4106**	**0.03**	**3767**	**0.01**	**2869**	**0.03**	3004	0.09	1522	0.50
Complication(s): *Yes*[c]	2562	0.59	5849	0.23	-1983	0.43	3029	0.52	12853	0.27
Comorbidities: *Elixh'ser*	**934**	**<0.01**	342	0.18	214	0.38	**1256**	**<0.01**	**2388**	**<0.01**
ADL: *Bathing*	-2612	0.07	-195	0.87	578	0.60	-2561	0.07	**-4172**	**0.04**
ADL: *Chair*	2524	0.11	1537	0.21	567	0.61	1709	0.27	1826	0.40
ADL: *Complete assist 1+*	2152	0.33	615	0.69	-312	0.82	1731	0.40	593	0.85
ESAS Phy: *Admission*	-4716	0.11	-1124	0.63	-972	0.66	-2998	0.30	-3907	0.34
ESAS Psy: *Admission*	-30	0.97	-397	0.46	-365	0.46	402	0.55	-24	0.98
ESAS Phy: *Consult/ref*[d]	3998	0.29	2384	0.40	2704	0.31	3062	0.40	4456	0.37
ESAS Psy: *Consult/ref*[d]	-893	0.21	349	0.54	269	0.61	-1056	0.13	-2067	0.06
CMSAS: *Admission*	**1183**	**0.02**	**1096**	**0.01**	**1221**	**<0.01**	814	0.08	387	0.57
CMSAS: *Consult/ref*[d]	497	0.40	-384	0.41	-465	0.28	644	0.28	1613	0.09
Severity[e]**:** *Admission*	27	0.92	-196	0.33	-235	0.20	-14	0.96	73	0.83
Severity[e]**:** *Consult/ref*[d]	-430	0.23	-59	0.83	-79	0.75	-449	0.21	-816	0.11
Morphine: *Dose (mg)*[f]‡	209	0.14	**287**	**0.01**	**282**	**0.01**	71	0.62	101	0.63
Pain: *Somewhat*[g]	1690	0.45	2723	0.15	2009	0.24	962	0.64	3666	0.31
Pain: *Quite a bit*[g]	591	0.71	903	0.46	791	0.48	269	0.86	-1172	0.58
Pain: *Very much*[g]	2254	0.20	1543	0.23	993	0.40	2035	0.23	-1671	0.45
Fatigue: [Not v. much][gh]	-706	0.58	-651	0.51	-1060	0.23	-393	0.76	13	0.99
Fatigue: *Very much*[g]	1963	0.23	564	0.64	-57	0.96	1103	0.48	909	0.69

Discussion

In the results of the weighted regressions, two of the five observed associations (Elixhauser total and CMSAS score at admission) reflect a relationship between illness burden and cost of care that is neither unexpected nor obviously informative to future resource allocation. One clinical factor (ESAS physical score is inversely associated with costs) is unexpected.

The associations between Medicare and/or Medicaid and lower cost are consistent with previous studies suggesting that insurance status is a driver of utilisation (Kelley et al., 2011, Shugarman et al., 2008). And the association between a college education and higher cost (and/or an elementary school education and lower cost) suggests that socioeconomic status may be a driver of utilisation since education is a proxy for income and/or health literacy. However, care should be taken not to equate higher utilisation with higher quality of care, since the relationship is not straightforward (Khandelwal et al., 2014).

In the results of the unweighted regressions among usual care patients (Table 4.6.2), four of the associations follow previous research in affirming that utilisation differs by diagnosis: three specific cancer diagnoses exhibit significant relationships with cost. The additional associations are also clinical: number of comorbidities, symptom distress and number of symptoms are all associated with higher costs. Increased psychological distress however is associated with lower costs. And there is one counter-intuitive result: increased severity of symptom burden is associated with lower costs.

In the unweighted regressions among palliative care patients (Table 4.6.3), the number of comorbidities is again significant (the only covariate to be so in all three analyses in this section). Two diagnoses are associated with cost, including lung

cancer, which has the reverse association in the treatment group than in the control group. Number of symptoms at admission has a strong association with cost of hospital stay. Visiting nurse services prior to admission are associated with lower costs, the only association in any table between the health service variables and cost of hospital stay.

These results have the following implications.

The majority of significant associations between patient-level factors and cost are clinical, and in particular covariates related to illness burden exhibit a predictable relationship with hospital costs in all three analyses. The strongest evidence is for a significant positive correlation between numbers of comorbidities on the Elixhauser index and costs; this relationship is visible in all weighted and unweighted regressions. While this result is not unexpected, it further emphasises the findings of the previous section that multimorbidity is a major driver of healthcare costs and that the appropriate treatment of a complex minority will be central to future finance of health care for people with serious illness.[107] It is not clear that the other clinical findings have major implications for future resource allocation, beyond reaffirming the point that sicker patients cost more and so, as the general population increases in illness burden, costs will increase also.

Once specific cancer diagnoses are introduced in the unweighted regressions, these often display significant associations with hospital costs. Planning and organisation of care requires ongoing efforts to understand the detailed resource requirements of

---

[107] For a discussion of how future research can develop these findings further, see **§5.3.3 >Future research can extend my findings on comorbidities, costs and incremental effect on costs.**

different diagnoses. Future studies of this population may also prefer to include specific diagnoses in propensity score matching if methods and sample size allow.

Equity concerns evident elsewhere in the literature – for example as regards age, gender and race – are generally not provided substantive supporting evidence in my results; there are few observable relationships between socioeconomic factors and cost. Two that do raise possible equity concerns are the associations between government insurance and lower cost, and education level and utilisation; both suggest that access to higher-cost care is associated with higher socioeconomic status.

Conclusion

The majority of strong associations between patient-level factors and cost are clinical: patients with a higher illness burden incur higher costs and, in unweighted regressions, some specific diagnoses are associated with utilisation. There are also equity of access concerns raised in respect of a strong association between patient's socioeconomic status and hospital costs, but demographic equity issues such as race and gender visible elsewhere in the literature do not appear a concern in this study.

5. Conclusions and recommendations

Abstract

This chapter is divided into three sections.

§5.1 Summary provides an overview of the principal findings from my PhD research.

§5.2 Limitations qualifies these conclusions in the context of the research's limitations, both in respect of the 'PC4C' study design and my own approach to analysis.

§5.3 Implications details a series of recommendations for practice, policy and research arising from my results.

5.1 Summary of the main conclusions

The aim of this thesis was to expand the existing evidence base on economic impact of palliative care consultation teams (PCCTs) through rigorous analysis of a new observational dataset (the Palliative Care for Cancer ('PC4C') study).

Specifically in the context of prior studies (**§2 Literature Review**) and the 'PC4C' research design (**§3 Materials & Methods**), I identified four key objectives for my primary analysis:

- **Comparative evaluation of methods using length of stay (LOS) to address unobserved heterogeneity:** assessment of approaches which control for LOS in cost analysis of observational data,
- **Cost-effect estimation:** establish PCCT impact on hospital costs, including identification of patient- and/or system-level factors associated with this effect,
- **Source of observable cost effects:** examination of the underlying source of any cost differences associated with the PCCT intervention,
- **Other determinants of utilisation:** identification of patient- and/or system-level factors associated with hospital utilisation other than PCCT interventions.

Each of these objectives was addressed in **§4 Results** and I have drawn the following conclusions in respect of these aims:

5.1.1 Controlling for LOS in cost analysis of observational data undermines the validity of reported results (§4.3)

Using LOS to control for unobserved heterogeneity in cost analysis of observational data has been commonplace in previous economic studies of PCCT interventions. Investigators have variously defined their sample according to length of hospital stay,

employed LOS as a covariate in regression, and/or used daily cost (total cost/LOS) as a primary dependent variable.

Each of these strategies is designed to address challenges in cost analysis of observational data. Including LOS as a covariate in regression, is intended to control for unobserved differences for which LOS may be a good proxy (e.g. clinical complexity, patient preferences). Employing daily cost as a dependent variable simplifies analysis by reducing skew and kurtosis of cost distributions. Removing short- and/or long-stay patients from the sample can in principle address both unobserved heterogeneity in the sample, and highly-skewed and leptokurtic cost distributions, by excluding outliers.

However using LOS in these ways may undermine both the internal and external validity of incremental cost-effect estimates. Defining a sample by a factor that is associated with both treatment (long hospital stay is associated with patient need) and the dependent variable (LOS and hospital costs are by definition very closely correlated) risks biasing results and distorting cost-effect estimates. Use of LOS as a covariate risks introducing endogeneity into analysis, undermining evaluation of the causal relationship of interest. Identifying daily hospital costs as the primary dependent variable focuses evaluation on a measure with little practical application.

These theoretical limitations were born out in empirical evaluation using the 'PC4C' data. Incremental-effect estimates were highly sensitive to different definitions of sample by LOS. For all samples irrespective of LOS, the introduction of LOS as a covariate in regression or as a denominator of total costs precipitated a large increase in the -incremental effect estimate and statistical significance of that estimate.

Use of LOS-control strategies without appropriate sensitivity analyses would have potentially increased the risk of a type I error. Specifically, they would have increased the likelihood of concluding that palliative care consultation *at any time* had a significant cost-saving effect, when my results show that this conclusion is not robust and misses important relationships in the data.

Consequently results using LOS-control strategies are both unreliable and not useful. As discussed in **§3.2.2**, Manning's well-known maxim in modelling costs advises against log-transformation on grounds of practicality: "Congress does not appropriate log dollars. First Bank will not cash a check for log dollars." This principle could be extended to controlling for LOS in estimating the impact of an intervention on costs of hospitalization: healthcare payers do not fund treatments for populations defined *ex ante* by LOS, and in the few instances where payers do so on the basis that treatments reduce the ratio of total cost to LOS during hospital stay, this reflects an inefficient payment system. Overall resource use for a sample defined at baseline is the best way to preserve the rigour and value of these types of analyses.

5.1.2 Earlier palliative care is associated with larger impact on cost (§4.4)

PCCT intervention within six days of hospital admission is shown to be associated with a statistically significant cost-saving effect for patients admitted to hospital with advanced cancer. The observed differences are accrued through a combination of reduced resource-intensity and, for patients who receive a consult within two days of admission, shorter hospital stay.

Moreover, all of these relationships between palliative care intervention and hospital cost are sensitive to time-to-consult following admission: earlier consultation is

associated with larger total cost-savings, larger reductions in cost for major utilisation categories, and larger reductions in LOS.

Comparison of descriptive cost data for earlier and later consults suggests no evidence of difference in impact on costs once a first consult has been administered: proportional reduction in costs in the five days following consultation are very similar for earlier and later consults.

Rather, the differences occur due to the cumulative nature of cost data: patients who were seen by a consultation team later in their hospital stay have accumulated higher costs in proportional and absolute terms prior to receiving the intervention – costs that are included in the dependent variable but by definition are not amenable to treatment.

Pooling all PCCT patients under a single binary any-time treatment variable is therefore a misspecification that potentially increases risk of a type II error. Specifically, this approach (while not controlling for LOS) would have increased the likelihood of concluding that palliative care consultation has no association with cost, when my results show that this conclusion is not robust.

The impact on specific utilisation categories suggests that the intervention reduces the number of tests and pharmaceuticals that patients receive - reductions that if the 'non inferiority' assumption is accepted might be classed as futile care. Where patients were admitted to the ICU these costs were reduced; imaging might also have been shown to be reduced with a larger sample size, and laboratory, pharmaceutical and imaging costs are typically correlated. Therefore it is possible that the intervention is reducing resource intensity in all major categories as a consequence of goals-of-care discussions. Only through more accurate (i.e. timing-

sensitive) definition of the intervention is the systematic relationship between earlier treatment and reduced costs in all categories understood. [108]

The results obtained in reaching this conclusion extend the evidence base on PCCT interventions by revealing new and useful information on the efficacy of these interventions, and using new and robust methods.

5.1.3 Palliative care is associated with larger impact on cost for patients with more comorbidities (§4.5)

PCCT intervention within ten days of hospital admission is shown to be associated with a statistically significant cost-saving effect for patients admitted to hospital with multiple comorbidities. The observed differences are accrued through reduced resource-intensity during hospital stay, and in particular lower laboratory costs. Where treatment is administered within six days, there are also significant reductions in LOS.

Furthermore, the relationship between treatment and hospital costs is sensitive to patient multimorbidity: palliative care is associated with larger cost-savings for patients with higher numbers of comorbidities. This implies that PCCT treatment is a complex intervention most effective for more complex patients.

The association between larger cost effect and patient multimorbidity is robust to time-to-consult (and vice versa): earlier treatment is associated with larger cost-saving effect for any sub-samples defined by number of patient comorbidities; and the estimated effect is largest for the most complex patients irrespective of how treatment is defined by time-to-consult.

---

[108] For a discussion of how future research can investigate further how the PCCT intervention realises the causal effects reported in this thesis, see **§5.3.3>Further evidence on how a PCCT operates would extend my results**.

The results supporting this conclusion extend the evidence base on PCCT interventions by revealing new and useful information on the efficacy of these interventions; no study has previously examined how palliative care effect on cost varies by patient complexity.

5.1.4 Patient-level factors are important drivers of cost irrespective of treatment group (§4.6)

An examination of patient-level factors and their association with hospital costs reveals a number of potentially important relationships.

Most notably, the majority of factors exhibiting a significant association with cost are clinical: patient illness burden is a major driver of in-hospital utilisation irrespective of treatment decisions. And the number of comorbidities a patient has in particular has a large, statistically significant association with cost of hospitalization in all analyses. Given demographic trends and projected disease patterns in the 21st century, this association has potentially great significance for planning and funding of services.

Additionally there are important relationships with socio-economic factors. Both insurance programme and education level appear to be correlated with total hospital costs, raising questions about equity of access.

5.2 Limitations

5.2.1 Unobserved heterogeneity

While measures have been taken to minimise confounding due to selection bias, there is likely unobserved heterogeneity between treated and comparison patients, including reasons contributing to time-to-consult following admission.

Furthermore it is possible for propensity score methods to exacerbate the effect on unobserved confounding in biasing results and disguising important relationships.

Optimally an instrumental variable would be used in conjunction with propensity score matching to control for unobserved heterogeneity but it was not possible to identify a valid instrument within the 'PC4C' dataset.

To confirm the robustness of my findings I have therefore employed multiple sensitivity analyses to check that the main results are robust to different sub-samples, modelling approaches and use of propensity score weights.

One important 'known unknown' of my analyses is patient and physician preferences. Included in the propensity score are binary variables for living will and proxy at baseline. While this is not an explicit inclusion of patient preference, I consider it a useful indicator since people who wish to restrict aggressive life-prolonging treatment are generally more likely to complete a living will and/or nominate a proxy. Additionally, my analyses include rich patient-reported data on many important potential confounders: psychological and physical symptoms, functional status, and formal health care use prior to the hospitalization.

Nevertheless, while my results provide strong evidence for increasing early palliative care access to patients with serious illness, and particularly those with multiple comorbidities, it is not known whether decisions not to pursue palliative care early following admission reflect explicit patient preferences or goals of care. This in turn raises questions about the circumstances under which any decision not to pursue early palliative care is rational given growing evidence that palliative care concurrent with curative care has benefits.

5.2.2 High attrition and loss to follow-up

The 'PC4C' study experienced significant attrition from recruitment to final sample.

Of all 5,939 patients admitted to participating sites with an advanced cancer diagnosis during the study period, 26% refused to participate in the study and 20% failed to meet additional eligibility criteria. Of the 3,218 enrolled in the study, over half did not participate long enough for adequate data to be collected for matching and analysis.

High levels of data attrition and loss to follow-up are commonplace in studies of seriously-ill populations and end-of-life care. The most common reason for incomplete data among enrolled patients was that s/he became too ill to participate. My results can therefore only be generalised to patients admitted to hospital with advanced cancer and who were able to participate in follow-up interviews.

Of the 1,537 in the final 'PC4C' sample, 33% were lost because one site ('Site 4') failed to collect cost data. This error was only noticed at the end of the data collection period when project staff at Mount Sinai including myself analysed the final dataset. This has no obvious qualitative impact on results (there is no reason to think *a priori* that patients differ at Site 4 compared to other sites) but it does limit the sample size and power of my calculations.

5.2.3 Hospital costs perspective

Throughout this thesis, dependent variables are defined from the hospital perspective: cost and length of hospital stay. As summarised in **§2.4**, an economic evaluation in ideal conditions compares a treatment with a comparator on all relevant costs and consequences.

Therefore the hospital costs perspective is restricted in two important ways. First, it does not take account of consequences of treatment (e.g. pain and symptom management, patient and family satisfaction). All reductions in cost and LOS

associated with palliative care that have been presented in this thesis are positive outcomes only on the 'non-inferiority' assumption established in **§2.4** that the consequences of treatment are at least no worse than usual care only.

Second, this thesis does not take account of costs after hospital discharge. One common concern of early hospital discharge is that it is not cost-saving but cost-shifting: patients leave hospital before they are ready, pushing the processes and costs of care onto formal outpatient services or informal networks. All reductions in cost and LOS reported in this thesis are positive outcomes only if reduced intensity and/or length of hospital stay do not impose equivalent (or larger) costs on formal and informal care systems outside of hospital.

Evaluating PCCT impact on both patient and family outcomes, and post-discharge utilisation are objectives of the 'PC4C' study, as presented in **§3.1.2**. Future research by colleagues on the study will therefore address the limitations to the hospital costs perspective I have taken.[109]

In the context of this thesis I consider the assumptions on which these results are reported to be reasonable in the wider research and policy context. As detailed in **§2.4**, there is a substantial and growing evidence base suggesting that palliative care is at least as effective as usual care for these outcomes. And hospital costs represent a large proportion of costs for the United States[110] and all other health systems in high-income countries, and therefore represent a substantial and worthwhile dependent variable in their own right. Eleven of 12 previous economic

---

[109] I also aim to participate in this research, including analyses that quantify the impact of PCCTs on cost, and patients and family outcomes. An overview of priority areas for my post-doctoral research is provided in **§5.3.3**.

[110] 35% of direct medical cancer costs are attributable to inpatient hospital stay; see **§1> The growth of palliative care**.

analyses of PCCTs have taken a hospital perspective and my results make a substantial contribution to this literature.

5.2.4 Generalisability

The data were collected at hospitals with well-established palliative care programmes in the United States. It is not clear how generalisable the results are to new programmes, to patients with diagnoses other than cancer, or to programmes in other health systems.[111]

Restricting eligibility to cancer patients has advantages and disadvantages to be considered in evaluating my results. The principal benefit is homogeneity; the pooling of patients with non-malignant conditions alongside advanced cancer would have introduced substantial new difference into the sample that analysis would have to control for. This has to be weighed against the external validity of the study; as palliative care access for non-cancer patients continues to grow, and non-malignant chronic disease is increasingly prevalent, my results have to be confirmed with that patient group before they can be considered applicable beyond persons with cancer.

The specific points at which treatment is shown to be 'early enough' (within six days, see **§4.4**) to be significant may not be generalisable. Data in **§4.4> Results> Secondary analyses** suggests that the proportion of costs accrued prior to consult is important to timeliness. The definition of how 'early' treatment has to be to be effective is therefore related to length of stay and may vary for other interventions, patient groups and settings. However, the fundamental association – earlier treatment is associated with larger cost-saving effect – should be robust in equivalent analyses.

---

[111] Further discussion of the implications of my results in the Irish context specifically follow in **§5.3.3**.

5.3 Implications for practice, policy and research

5.3.1 Practice

Early palliative care access for patients admitted to hospital with advanced cancer should be more widely implemented

The well-established cost-saving effect of PCCTs appears large (~25%, larger than estimated by most previous studies for patients discharged alive) if the consult occurs early in the hospitalisation (within two days of admission), but this effect may not be significant if the first consult occurs later.

Involvement of a specialist multidisciplinary PCCT in the treatment and transition decisions of this patient group from an early stage of hospitalisation will maximise increasingly scarce health care budgets. A growing literature suggests that early palliative care in general and treatment from PCCTs in particular also bring clinical benefits; there is no evidence that these cost-savings would be achieved through reduced health-related outcomes.[112]

Early palliative care access for patients admitted to hospital with advanced cancer and multimorbidity should be more widely implemented

On the same basis, PCCTs should be involved in the care of patients admitted to hospital with multimorbidity – and in particular patients with high numbers of comorbidities. Specialist palliative care is a complex intervention whose impact is greatest for the most complex patients (and with a larger cost-saving effect than suggested by previous studies which have reported mean effect for a whole sample). This research is the first to show that effect on cost is larger for patients with multimorbidity, and so the evidence for complex patients receiving palliative care during hospitalisation is particularly strong.

---

[112] For an overview of the evidence on palliative care programmes to date, see **§1**. For an overview of how these results may be combined in future with appropriate outcomes from the 'PC4C' study, see **§5.3.3> Further research with the 'PC4C' data**.

The scope for further financial benefits through palliative care is substantial

Taken together, the previous two points suggest substantial scope to reduce the cost of caring for patients with advanced cancer in United States hospitals. In the 'PC4C' principal sample for cost analysis (n=969), 232 patients (24%) were seen by a PCCT within six days of hospital admission; the equivalent figure for matched patients admitted with multimorbidity (n=774), was 214 (28%). Even among very seriously-ill patients with more than five comorbidities in addition to cancer (n=256), the proportion was not dissimilar (n=78; 30%). The potential for further value is also underlined by the association between patient illness burden and cost among usual care patients; while palliative care on average is being received by patients with a higher illness burden, the number of patients with multiple serious conditions who are not receiving palliative care but are accruing large hospital costs remains significant.

The minority take-up of palliative care was observed at large medical and cancer centers with well-established palliative care programmes. It does not predominantly reflect limited treatment capacity at the study hospitals but the treatment patterns of primary physicians for whom palliative care consultation is often not a priority. The full scope of costs and benefits through PCCT interventions will be more fully understood in future once the patient and family outcomes, and goals-of-care domains of the study are completed and reported. Even without those results, accepting the 'non-inferiority' assumption that key patient and family outcomes are at least as good and often better following PCCT interventions, the cost-saving effects identified in this thesis highlight substantial potential benefits in the provision of hospital care to patients with advanced cancer that are yet to be realised.

5.3.2 Policy in the United States[113]

Increase and incentivise access to PCCT interventions for patients with advanced cancer

It is well-established that the sustainability of United States government health programmes Medicaid and Medicare, as well as the overall national health system, depends on controlling costs of care to patients with serious illness without compromising quality and access. Hospital costs constitute a significant proportion of this expenditure.

My analysis affirms and extends the previous literature on the cost-saving effect of PCCTs. Compared to usual care, early palliative care appears to reduce both intensity and length of hospital stay. While PCCTs in US hospitals have expanded substantially over the last two decades, access is not universal. Moreover, even in hospitals with well-established programmes most patients with serious illness do not receive palliative care. There is further scope to upscale the successful PCCT model in US hospitals by implementing payment mechanisms that incentivise and reward clinical and financial benefits in treating seriously-ill patients.

Workforce planning

Shortages of specialist palliative care staff in the US health system are well documented.[114] Moreover, projected levels of need over the next three decades are such that not all patients with serious illness will receive specialist-led care in all settings. Specialist palliative care clinicians will remain a scarce resource to be allocated judiciously. PCCTs should therefore be employed where they have the

---

[113] Relevance of the findings beyond the United States are discussed in **§5.3.3> Applying these findings to the Irish context**.

[114] For a summary of these issues, see **§4.5>Discussion**.

greatest impact: on my evidence, patients recently admitted to hospital and/or with multiple comorbidities are those who ought to be prioritised.

5.3.3 Research

Evaluations of PCCT interventions should incorporate time-to-consult following admission in estimating incremental effect on cost

My results show that PCCT cost effect varies according to treatment timing and furthermore that this is systematic: total costs are cumulative, meaning that later consults accrue a higher level of costs prior to consult – costs that are included in the dependent variable but by definition are not amenable to treatment. This is exacerbated by the fact that later-consult patients are comparatively longer-stay patients and therefore comparatively high-cost. If cost effect is estimated using a binary any-time treatment variable for a sample including late-consult patients, those patients will have a disproportionately large impact on the magnitude of effect (due to high costs) despite having received an intervention that is less representative than average (being unable to impact costs to the same extent). Assuming homogenous cost effect irrespective of time-to-consult following hospital admission will therefore increase the risk of a type II error.

This recommendation is equally applicable for evaluating PCCT impact on LOS, as well as for any study of the impact of a binary treatment variable on a cumulative measure of utilisation where treatment can be administered after the utilisation is first accrued.

Incorporating time-to-consult could also contribute to understanding mechanisms of observed cost-savings

In **§2 Literature Review** I noted that previous studies suggested a clear pattern of cost-saving effect from PCCTs but without a consistent source of those savings.

Different studies reported reductions in different ancillary cost categories but none reported any reduction in LOS.

My research is therefore the first to find a significant association between treatment and LOS for this intervention – but only for consult within two days of admission (for all patients; within six days of admission for multimorbidity patients). Contrary to previous results, PCCTs may reduce LOS as well as resource-intensiveness if the intervention is provided early enough. Previous studies may have made a type II error in measuring PCCT impact on LOS without taking into account time-to-consult.

Where future studies incorporate time-to-consult in evaluating PCCTs they will examine more rigorously the extent to which observed cost-savings accrue through reduced intensiveness of hospital stay and through reduced length of hospital stay.

Evaluation of palliative care impact over a longer timeframe would complement my results with new and important evidence

The association between time-to-consult and PCCT impact on cost identified in this thesis is consistent with a growing literature on the benefits of early palliative care.

Clinical studies have increasingly examined palliative interventions earlier in the care trajectory; for example, concurrent palliative care following diagnosis. Many such interventions are therefore concerned with the impact of palliative care on a new patient group: moving 'upstream' from the end-of-life phase where palliative care originated to those who are living with serious illness and functional impairment on a long-term basis.

The 'PC4C' study is somewhat 'upstream' of previous economic evaluations of PCCTs, since 95% of patients in the cost dataset were discharged alive compared to the 12 studies in **§2 Literature Review**, which had a median in-hospital survival rate

of 70%. Even allowing for the trend towards shorter hospital stay, the 'PC4C' patients were not an 'end-of-life' cohort consistent with most previous studies.

Nonetheless the 'PC4C' patients are admitted to hospital with advanced cancer and in the large majority of cases at least one additional chronic condition, i.e. patients who are seriously ill with life-limiting conditions. This is not an 'early' palliative care intervention specifically integrated early in the disease trajectory, for example targeting patients who were recently diagnosed.

The implication of cost-effect heterogeneity by treatment timing is that cost-saving treatment earlier in the care trajectory could potentially impact costs (and other dependent variables) for longer and for a larger overall benefit. Equally, it will be valuable to track patient utilisation following discharge to evaluate if savings from hospital treatment continue to accrue.

The results in this thesis are drawn from a mean timeframe of less than nine days' hospital stay yet the timing of treatment within those nine days has a substantial influence on cost effect. It stands to reason that in the context of long-term chronic illness studies taking a broader time perspective (treatment earlier in the trajectory and/or utilisation following discharge) could result in evidence of even more significant associations between treatment timing and cost.

Evaluations of PCCT interventions should investigate heterogeneity of incremental effect on cost in ways other than timing where possible

In addition to cost-effect heterogeneity by time-to-consult, this thesis identified a larger cost-saving effect for patients with multimorbidity (and larger still for higher numbers of comorbidities).

Unlike the timing recommendation, this is not a question of research robustness. Mean cost-saving effect for all patients (as distinct from all patients with multimorbidity) will be a reasonable and valid primary association of interest in many health care evaluations, assuming the treatment has been correctly specified.

However, additional analyses on how and where treatment is (in)effective can inform organisation and provision of care by identifying in more detail how scarce resources are best allocated. Thus, the results in this thesis demonstrate that for the cost-saving potential of palliative care to be maximised, patients with complex needs should be prioritised.

Further research can explore both this specific association in other populations and settings, and other ways of identifying patients for whom palliative care is particularly (in)effective on cost and other dependent variables.

Future research can extend my findings on comorbidities, costs and PCCT impact on costs

Two key findings of this thesis are that number of comorbidities has a large and significant association with hospital costs, and that the PCCT intervention has a larger impact on costs for patients with higher numbers of comorbidities. The analyses as reported do not offer detailed information on specific comorbidities. These results require further investigation, particularly in the context of health spending in the US and other high-income countries being disproportionately driven by a minority of patients with multimorbidity.

In this thesis multimorbidity was measured using the Elixhauser index, a simple additive index in which a patient is awarded 1 in the presence (and 0 in the absence) of each of 31 comorbidities (Elixhauser et al., 1998). Thus, metastatic cancer and uncomplicated diabetes are afforded the same significance in the Elixhauser index.

It has previously been shown that the Elixhauser index can be modified by weighting each individual comorbidity according to its association with hospital death, delivering a new index that more accurately reflects disease burden in research on morbidity (van Walraven et al., 2009).

It seems plausible that a similar modification could be pursued for research in the economic domain, weighting specific comorbidities according to their association with costs, providing better information on those conditions that are most associated with high costs as well as delivering a more accurate reflection of disease burden, and thus improving accuracy of cost modelling. In principle an indicative index per van Walraven et al. might be generated using the 'PC4C' data. But to follow van Walraven et al. and deliver a robust, generalisable index, it will likely be preferable to employ a large administrative dataset which allows derivation and validation techniques. Having established that cost effect varies by the number of comorbidities, it would also be very informative to identify the comorbidities for which PCCTs have the largest cost-saving effect, but this will require a larger dataset than that collected by the 'PC4C' project.

An important subsequent extension to such work on the association between specific comorbidities, costs and cost effects will be to understand which *combinations* of comorbidities are significant in this context. As noted in **§4.5**, chronic conditions work together synergistically (rather than additively), complicating the identification of appropriate medications and treatment regimens. If it is beneficial to develop a more nuanced approach that distinguishes between metastatic cancer and uncomplicated diabetes, then it may be more beneficial still to distinguish between the combinations and permutations of the 31 diseases in the Elixhauser index.

My findings in **§4.5** and **§4.6** therefore represent only a first step. As demographic change and changing patterns of disease move health services research into an era of multimorbidity, further evidence on these issues will be essential to the planning and provision of cost-effective care to patients with multiple chronic illnesses.

Clinical research should investigate if associations between timing and patient complexity impact patient and family outcomes

The core findings of this thesis contribute new evidence on PCCT impact on cost. It is not clear the extent to which the same associations between incremental effect on cost, and timing and patient complexity are also evident in non-utilisation outcomes.

As described in **§5.2.3** and below, this work is important to confirm the 'non-inferiority' assumption that observed cost-savings are not achieved at the expense of patient or family outcomes. Additionally it is plausible that both associations will also be observable with patient and family outcomes, further strengthening understanding of the impact of PCCT interventions.

These results could also be used in moving beyond hospital perspective in economic evaluation of these interventions and towards analysis incorporating both costs and consequences.

Further evidence on patient-level factors associated with propensity to treat would extend my findings

Patient-level factors are found to be associated with both propensity to treat (patients with a higher illness burden receive palliative care) and with hospital costs (illness burden is positively correlated with utilisation). Yet a majority of patients admitted with advanced cancer to large cancer centers with well-established palliative care teams did not receive palliative care.

Taken together these results suggest that reforming treatment of patients with serious illness to incorporate palliative care requires further evidence on the reasons for low levels of referral. In particular does minority take-up of palliative care reflect patient and physician preferences, and are these preferences rational and consistent with the evidence on palliative care as a concurrent treatment option with curative care?

Additionally, equity concerns highlighted elsewhere in the literature, notably that access to higher-cost care is associated with higher socioeconomic status, are also substantiated by my analyses. Understanding the socioeconomic barriers to care access may in turn improve care to these patients. Such analyses should not assume that utilisation equates to good outcomes (since patient experience may be better where treatment intensity is lower, for example where futile care is reduced). If access is shown to be restricted according to socioeconomic status, it is then worthwhile examining the extent to which this reduced access has negative consequences for patient health and quality of life.

Further evidence on how a PCCT operates would extend my results

In addition to information deficits over referral processes outlined above, there is also scope for improved understanding on another element of clinical practice in this thesis: the PCCT itself.

The formal guidelines for teams and the goals of the intervention are clear.[115] Through a standardised protocol, PCCTs supplement standard care through expert pain and symptom management, goals-of-care discussions, and transition management. The original results in this thesis supplement and strengthen previous

[115] For a more detailed review of these guidelines and goals, see **§3.1.4>Intervention**,

studies reporting hospital cost-savings from PCCTs, and the evidence bases are increasingly well established on both poor outcomes for seriously-ill patients from standard care, and improved patient and family outcomes from palliative care.[116]

However, the mechanisms by which PCCTs realise causal effects have not been the subject of much detailed analysis. Reported improvements in pain and symptom management are inferred to follow the designated expertise of palliative care clinicians. In terms of quality of life and satisfaction, the inference is that patient-centred goals-of-care discussions consider both the advantages and disadvantages of high-intensity treatments, including their likelihood of success, and where patients and their families take a more active role in treatment choices they often prefer to emphasise quality of life over expensive interventions with limited evidence of efficacy. Reduced utilisation is also inferred to result from these choices.

In practice, the work of PCCTs is patient-centred and will vary substantively across diverse domains including expert opioid-prescribing, skilled co-ordination of care, administrative assistance in facilitating discharge, and bereavement counselling. And it was shown in **§2.4** that while the general pattern of PCCT impact is consistent - better or equal patient outcomes, and lower hospital costs - the specific outcomes affected are not so consistent. Some outcome studies have found, for example, improved pain and symptom scores but not improved satisfaction (and vice versa), while cost studies have found, for example, reduced pharmacy costs but not reduced laboratory costs (and vice versa). My original results (Table 4.4.6) show that an early consultation is associated with significantly lower pharmacy and laboratory

---

[116] A detailed review of cost analyses of PCCTs was provided in **§2.** For an overview of the 'high costs, poor outcomes' paradigm in care of people with serious illness, see **§1>Established public health priority**. For a review of the prior evidence on patient and family outcomes, see **§2.4> Evidence of intervention efficacy,**

costs, reaffirming the assumption in the literature that patient-centred treatment choices result in fewer drugs and tests than usual care.

Some variability in the PCCT cost literature more generally may be due to sub-optimal methods employed previously; once intervention timing is incorporated, an earlier intervention appears to reduce systematically costs in all major utilisation categories. It is also plausible that some of the variability in affected categories is due to variability in PCCT practice or patient population: while the guidelines and the goals are standardised, the emphasis and detail of how the intervention is delivered may vary according to which activities are considered by team members to be most appropriate and beneficial for specific patient groups.

The question of 'what is a palliative care consultation?' is therefore one meriting further investigation through primary research.

Investigators may be able to observe incremental effects for complex patients with small sample sizes

Nearly 6,000 patients were admitted to the 'PC4C' sites with an advanced cancer diagnosis during the study period. The principal sample size for economic evaluation includes less than 20% of these patients and reflects all who gave consent and for whom adequate data were collected.

Despite this attrition a number of important relationships are identified with high statistical significance. In particular it is noticeable that associations between treatment within 10 days and cost are discernible for patients with five or more comorbidities, where n=256. For tighter definitions of treatment (within fewer than 10 days) the sample size was typically too small to balance a robust propensity score. Nevertheless, it is both positive and informative to learn that sample sizes as small

as 250 for complex patients can be matched and that significant relationships for a sample of this size can be discerned.

Cost analyses of observational data should avoid controlling for unobserved heterogeneity using LOS

Removing short- and long-stay patients prior to analysis is sub-optimal on a number of counts. First, previous use has been inherently arbitrary, reducing sample size and power without evidence-based justification. Second, results from LOS-defined samples have limited implications for practice and policy since LOS is not known at admission; LOS-dependent evidence cannot be used to inform treatment decisions. Third, it has limited scope as a long-term approach in health services research as hospital LOS is reduced in health systems worldwide, meaning that those patients who do stay (for example) seven to 30 days in hospital will increasingly constitute a distinct minority with complex needs. Fourth, sample selection by outcome prior to matching on baseline characteristics undermines research design with observational data.

The use of LOS as a covariate risks introducing endogeneity into analysis since it is related to both likelihood to receive treatment (in the case of patients with serious illness, long hospital stay suggests palliative care need) and the dependent variable (LOS and total cost are by definition closely correlated). In regressing cost on LOS, analysts undermine the strength of their evaluation of the causal relationship of interest (that of treatment on cost).

Identifying daily hospital costs as primary dependent variable is unwise because it is merely a ratio and does not represent the overall resource use of a hospitalisation. Moreover, if there are differences in LOS between treatment and comparison groups then daily cost is not synonymous with total cost and must be interpreted carefully.

To avoid misleading incremental-effect estimates, LOS controls should be resisted.[117] Where it is considered that controlling for LOS is necessary and worthwhile, any incremental-effect estimates should be checked through multiple sensitivity analyses.

Investigators designing studies using observational methods need new approaches to unobserved heterogeneity

The corollary of my conclusions that LOS-control strategies are weak controls for unobserved heterogeneity and that unobserved heterogeneity exists in established methods in this field, is that future studies must find new ways to address this challenge.

As outlined in **§2.4** and **§3.2.2**, an instrumental variable approach may address many of these concerns where it is possible to identify one. Where a strong, valid instrument cannot be identified, multiple sensitivity analyses must be performed to check the robustness of results.

A priority area for data collection in future is patient and physician preferences. These would address the single biggest missing domain of the 'PC4C' study.

Further research with the 'PC4C' data will extend and strengthen my results

These results represent one group of outputs from the 'PC4C' study. One of my aims in post-doctoral research is to extend these findings by incorporating additional study data as they become available.

---

[117] Additionally, previous studies have matched patients in the propensity score according to LOS. This is more serious than the other concerns raised here, a fundamentally flawed methodology that must be avoided in all circumstances. For details, see **§2.4> Sample definition by LOS**.

My first priority will be to incorporate treatment effects on patient and family outcomes in a cost-consequence analysis.[118] Potential cost-savings highlighted by cost analysis of an intervention only imply potential efficiency gains on a 'non-inferiority' assumption: the outcomes for patients and their families must be at least as good following the intervention than they would have been following the alternative. In this thesis, the assumption was based on a consistent pattern of positive outcomes in evaluations of PCCTs with no evidence of negative outcomes, but it will be possible to validate the assumption using 'PC4C' data in future.

Specifically, discharge data will be made available on the ESAS and CMSAS symptom scales whose baseline data were included in the propensity scoring and regression analyses. These scales are divided into physical (pain, lack of energy, lack of appetite, dry mouth, weight loss, drowsiness, shortness of breath, constipation, difficulty sleeping, difficulty concentrating, nausea) and psychological (worry, sadness, nervousness) sub-scales. Additionally, measures of patient experience with transition management (CTM); family satisfaction for all patients irrespective of discharge status (FAMCARE); and surveys of bereaved relatives for patients who died in hospital (ABFMI) will be made available.[119]

The precise nature of cost-consequence analysis to be employed is yet to be determined. Conceptual debate on outcome measurement in palliative care economics notwithstanding, detailed preference data have not been collected for the 'PC4C' patients, meaning that it will not be possible to perform a weighted synthesis

---

[118] For more detail on the background to these issues, see **§2.4>Forms of economic evaluation** and **§2.4> Outcome measurement in economic evaluation of palliative care**.

[119] For context on how these data fit into the schema of the 'PC4C' project and my primary analyses, see **§3.1.6 Variables**.

of ESAS, CMSAS, CTM, FAMCARE and ABFMI measures into a single QALY-type measure. Rather, a simplified approach, perhaps calculating cost-effectiveness planes for selected patient and family outcome measures, will be pursued.

Additional domains for which data have been collected include processes and goals of care (providing more detail on how and why the PCCT intervention delivers benefits), and post-discharge utilisation data (to track the size and persistence of cost-saving effect following hospitalisation).

Such research will further strengthen the evidence on the PCCT intervention for advanced cancer patients as policy and clinical responses are urgently sought to improve quality and curb costs to this patient group.

Applying these findings to the Irish context

The applicability of these findings to other health systems and settings is unclear.

In-hospital palliative care consultation models are predominantly seen in high-income countries, including Ireland, but there are substantial differences in care provision between these countries.[120] For example, US utilisation patterns in treating patients with serious illness may be higher than in Ireland, meaning lower scope in Ireland for cost-savings through interventions that reduce futile care.

However the relationship between timing and cost should in principle be robust since total hospital costs is a cumulative variable in all settings, and so a cost-saving treatment should have a greater effect if delivered earlier irrespective of local conditions. And the growing prevalence (and associated cost) of multimorbidity is a

---

[120] For an overview of the system-specific pressures in the United States, notably towards high-intensity treatment and pharmaceutical cost growth, see **§1>Established public health priority**. For an overview of how incentives differ between the United States and Europe, and how this influences the development of palliative care programmes, see **§2.4>Total versus direct costs as a dependent variable**.

universal problem in contemporary health care internationally, suggesting scope for complex interventions such as palliative care to exert significant impact on outcomes for complex patients.

The premise of the HRB Economics of Cancer Fellowship which funded my PhD research was that I gain substantive experience of health economics evaluation in the United States with a view to applying this expertise in Ireland in a post-doctoral context. I believe that the analyses presented in this PhD thesis represent a substantial addition to the evidence base and research methodology of this field.

Having accomplished this, my priority is to test the results derived here and associated recommendations in the Irish context. I hope that my future projects will address the impact of timely palliative care for patients with serious illness and multimorbidity in Ireland. In addition to testing my principal conclusions above I will seek to expand the scope of such evaluations to incorporate broader time horizons and patient/family perspectives, as well as analysing the impact of palliative care programmes beyond the hospital at a time when the future of care for people living with serious illness will be increasingly focused on the community setting.

References

ABADIE, A. & IMBENS, G. W. 2008. Notes and comments on the failure of the bootstrap for matching estimators. *Econometrica,* 76**,** 1537–1557.

ADELSON, K., PARIS, J., SMITH, C. B., HORTON, J. & MORRISON, R. S. 2013. Standardized criteria for required palliative care consultation on the solid tumor oncology service. *J Clin Oncol 31, 2013 (suppl 31; abstr 37).*

ALDRIDGE CARLSON, M. D. 2013. Research methods priorities in geriatric palliative medicine. *J Palliat Med,* 16**,** 838-42.

AMERICAN ACADEMY OF HOSPICE AND PALLIATIVE MEDICINE, CENTER TO ADVANCE PALLIATIVE CARE, HOSPICE AND PALLIATIVE NURSES ASSOCIATION, LAST ACTS PARTNERSHIP & NATIONAL HOSPICE AND PALLIATIVE CARE ORGANIZATION 2004. National Consensus Project for Quality Palliative Care: clinical practice guidelines for quality palliative care, executive summary. *J Palliat Med,* 7**,** 611-27.

AMERICAN CANCER SOCIETY. 2015. *Cancer Facts and Figures 2015,* Atlanta, GA.

AMPORFU, E. 2010. Estimating the effect of early discharge policy on readmission rate. An instrumental variable approach. *Health,* 2**,** 504-510.

ANDERSON, G. F. 2005. Medicare and chronic conditions. *N Engl J Med,* 353**,** 305-9.

ANDERSON, G. F. 2010. *Chronic Care: Making the case for ongoing care,* Princeton, NJ: Robert Wood Johnson Foundation.

ANGRIST, J. D., IMBENS, G. W. & RUBIN, D. B. 1996. Identification of causal effects using instrumental variables. *J Am Stat Assoc,* 91**,** 444-455.

AUSTIN, P. C. 2009. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med,* 28**,** 3083-107.

AUSTIN, P. C. 2011. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res,* 46**,** 399-424.

BAKITAS, M., TOSTESON, T., LI, Z., LYONS, K., HULL, J., LI, Z., DIONNE-ODOM, J. N., FROST, J., HEGEL, M., AZUERO, A., AHLES, T., RIGAS, J. R., PIPAS, J. M. & DRAGNEV, K. H. 2014. The ENABLE III randomized controlled trial of concurrent palliative oncology care. ASCO Annual Meeting, 30th May-3rd June, 2014. Chicago, IL.

BASU, A. 2005. Extended generalized linear models: simultaneous estimation of flexible link and variance functions. *Stata J,* 4**,** 501-516.

BASU, A. & RATHOUZ, P. J. 2005. Estimating marginal and incremental effects on health outcomes using flexible link and variance function models. *Biostatistics,* 6**,** 93-109.

BENDALY, E. A., GROVES, J., JULIAR, B. & GRAMELSPACHER, G. P. 2008. Financial impact of palliative care consultation in a public hospital. *J Palliat Med,* 11**,** 1304-8.

BLACK, N. 1996. Why we need observational studies to evaluate the effectiveness of health care. *BMJ,* 312**,** 1215-8.

BLOUGH, D. K., MADDEN, C. W. & HORNBROOK, M. C. 1999. Modeling risk using generalized linear models. *J Health Econ,* 18**,** 153-71.

BROOKS, J. M. & OHSFELDT, R. L. 2013. Squeezing the balloon: propensity scores and unmeasured covariate balance. *Health Serv Res,* 48**,** 1487-507.

BROUSSEAU, R. T., JAMESON, W., KALANJ, B., KERR, K., O'MALLEY, K. & PANTILAT, S. 2012. A multifaceted approach to spreading palliative care consultation services in California public hospital systems. *J Healthc Qual,* 34**,** 77-85.

BRUERA, E., KUEHN, N., MILLER, M. J., SELMSER, P. & MACMILLAN, K. 1991. The Edmonton Symptom Assessment System (ESAS): a simple method for the assessment of palliative care patients. *J Palliat Care,* 7**,** 6-9.

BUI, T. 2012. Effectively training the hospice and palliative medicine physician workforce for improved end-of-life health care in the United States. *Am J Hosp Palliat Care,* 29**,** 417-20.

BUREAU OF LABOR STATISTICS. 2014. *CPI Database (All Urban Consumers)* [Online]. US Department of Labor. Available: http://www.bls.gov/cpi/data.htm [Accessed: 2nd September, 2014.

BURGESS, J. F., DEB, P., GARRIDO, M. M. & PENROD, J. D. 2012. *Evaluating control function estimates of treatment effects in models of healthcare costs,* Working Paper: VA Health System.

BURNHAM, K. P. & ANDERSON, D. R. 2004. Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research,* 33**,** 261-304.

BUSSE, R., BLÜMEL, M., SCHELLER-KREINSEN, D. & ZENTNER, A. 2010. *Tackling chronic disease in Europe : strategies, interventions and challenges,* Copenhagen: European Observatory on Health Systems and Policies.

CAREY, K. & BURGESS, J. F., JR. 1999. On measuring the hospital cost/quality trade-off. *Health Econ,* 8**,** 509-20.

CASARETT, D., JOHNSON, M., SMITH, D. & RICHARDSON, D. 2011. The optimal delivery of palliative care: a national comparison of the outcomes of consultation teams vs inpatient units. *Arch Intern Med,* 171**,** 649-55.

CASARETT, D., PICKARD, A., BAILEY, F. A., RITCHIE, C., FURMAN, C., ROSENFELD, K., SHREVE, S., CHEN, Z. & SHEA, J. A. 2008. Do palliative consultations improve patient outcomes? *J Am Geriatr Soc,* 56**,** 593-9.

CASSEL, J. B. 2013. The importance of following the money in the development and sustainability of palliative care. *Palliat Med,* 27**,** 103-4.

CASSEL, J. B. 29th September 2014 2014. Whose costs are saved when palliative care saves costs? *Health Affairs (Blog)* [Online]. Available from: http://healthaffairs.org/blog/2014/09/29/whose-costs-are-saved-when-palliative-care-saves-costs/.

CASSEL, J. B., KERR, K., PANTILAT, S. & SMITH, T. J. 2010. Palliative care consultation and hospital length of stay. *J Palliat Med,* 13**,** 761-7.

CENTER FOR MEDICARE & MEDICAID SERVICES. 2011. *FY Wage Index (Table 2)* [Online]. US Department of Health and Human Services. Available: http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/Wage-Index-Files-Items/CMS1234173.html [Accessed: 2nd September 2014.

CENTERS FOR DISEASE PREVENTION AND CONTROL. 2013. *The state of aging and health in America 2013,* Atlanta, GA: US Dept of Health and Human Services.

CENTERS FOR MEDICARE & MEDICAID SERVICES. 2015. *Medicare.gov Glossary - M* [Online]. Available: http://www.medicare.gov/glossary/m.html [Accessed: April 22nd, 2015.

CHANG, V. T., HWANG, S. S., KASIMIS, B. & THALER, H. T. 2004. Shorter symptom assessment instruments: the Condensed Memorial Symptom Assessment Scale (CMSAS). *Cancer Invest,* 22**,** 526-36.

CHOCHINOV, H. M. 2011. Death, time and the theory of relativity. *J Pain Symptom Manage,* 42**,** 460-3.

CIEMINS, E. L., BLUM, L., NUNLEY, M., LASHER, A. & NEWMAN, J. M. 2007. The economic and clinical impact of an inpatient palliative care consultation service: a multifaceted approach. *J Palliat Med,* 10**,** 1347-55.

COLEMAN, E. A., MAHONEY, E. & PARRY, C. 2005. Assessing the quality of preparation for posthospital care from the patient's perspective: the care transitions measure. *Med Care,* 43**,** 246-55.

CONWAY, K. S. & DEB, P. 2005. Is prenatal care really ineffective? Or, is the 'devil' in the distribution? *J Health Econ,* 24**,** 489-513.

COWAN, J. D. 2004. Hospital charges for a community inpatient palliative care program. *Am J Hosp Palliat Care,* 21**,** 177-90.

DAVIES, E. & HIGGINSON, I. J. 2004a. *Better palliative care for older people,* Copenhagen: World Health Organization, Regional Office for Europe.

DAVIES, E. & HIGGINSON, I. J. 2004b. *The solid facts: palliative care,* Copenhagen: World Health Organization, Regional Office for Europe.

DEB, P., MANNING, W. G. & NORTON, E. C. 2013. Modeling health care costs and counts. iHEA World Congress, 7th-10th July, 2013. Sydney, Australia. Available: http://harris.uchicago.edu/sites/default/files/iHEA_Sydney_minicourse.pdf.

DOUGLAS, H. R., HALLIDAY, D., NORMAND, C., CORNER, J., BATH, P., BEECH, N., CLARK, D., HUGHES, P., MARPLES, R., SEYMOUR, J., SKILBECK, J. & WEBB, T. 2003. Economic evaluation of specialist cancer and palliative nursing: a literature review. *Int J Palliat Nurs,* 9, 424-8.

DOWNEY, L., AU, D. H., CURTIS, J. R. & ENGELBERG, R. A. 2013. Life-sustaining treatment preferences: matches and mismatches between patients' preferences and clinicians' perceptions. *J Pain Symptom Manage,* 46, 9-19.

DRUMMOND, M. F. & JEFFERSON, T. O. 1996. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ,* 313, 275-83.

DRUMMOND, M. F., SCULPHER, M. J., TORRANCE, G. W., O'BRIEN, B. J. & STODDART, G. L. 2005. *Methods for the Economic Evaluation of Health Care Programmes,* 3rd ed., Oxford, OUP.

DY, S. M., ASLAKSON, R., WILSON, R. F., FAWOLE, O. A., LAU, B. D., MARTINEZ, K. A., VOLLENWEIDER, D., APOSTOL, C. & BASS, E. B. 2012. Closing the quality gap: revisiting the state of the science (vol. 8: improving health care and palliative care for advanced and serious illness). *Evid Rep Technol Assess (Full Rep),* 1-249.

DY, S. M., SHUGARMAN, L. R., LORENZ, K. A., MULARSKI, R. A. & LYNN, J. 2008. A systematic review of satisfaction with care at the end of life. *J Am Geriatr Soc,* 56, 124-9.

EL-JAWAHRI, A., GREER, J. A. & TEMEL, J. S. 2011. Does palliative care improve outcomes for patients with incurable illness? A review of the evidence. *J Support Oncol,* 9, 87-94.

ELIXHAUSER, A., STEINER, C., HARRIS, D. R. & COFFEY, R. M. 1998. Comorbidity measures for use with administrative data. *Medical Care,* 36, 8-27.

ELSAYEM, A., SWINT, K., FISCH, M. J., PALMER, J. L., REDDY, S., WALKER, P., ZHUKOVSKY, D., KNIGHT, P. & BRUERA, E. 2004. Palliative care inpatient service in a comprehensive cancer center: clinical and financial outcomes. *J Clin Oncol,* 22, 2008-14.

EVANS, C. J., HARDING, R. & HIGGINSON, I. J. 2013. 'Best practice' in developing and evaluating palliative and end-of-life care services: a meta-synthesis of research methods for the MORECare project. *Palliat Med,* 27, 885-98.

EWING, G., ROGERS, M., BARCLAY, S., MCCABE, J., MARTIN, A. & TODD, C. 2004. Recruiting patients into a primary care based study of palliative care: why is it so difficult? *Palliat Med,* 18, 452-9.

FLORY, J., YINONG, Y. X., GUROL, I., LEVINSKY, N., ASH, A. & EMANUEL, E. 2004. Place of death: U.S. trends since 1980. *Health Aff (Millwood),* 23, 194-200.

FU, S., MCQUINN, L., NAING, A., WHELER, J. J., JANKU, F., FALCHOOK, G. S., PIHA-PAUL, S. A., TU, D., HOWARD, A., TSIMBERIDOU, A., ZINNER, R., HONG, D. S. & KURZROCK, R. 2013. Barriers to study enrollment in patients with advanced cancer referred to a phase I clinical trials unit. *Oncologist,* 18, 1315-20.

GADE, G., VENOHR, I., CONNER, D., MCGRADY, K., BEANE, J., RICHARDSON, R. H., WILLIAMS, M. P., LIBERSON, M., BLUM, M. & DELLA PENNA, R. 2008. Impact of an inpatient palliative care team: a randomized control trial. *J Palliat Med,* 11, 180-90.

GARRIDO, M. M. 2014a. Propensity scores and palliative care. *J Palliat Med,* 17, 261.

GARRIDO, M. M. 2014b. Propensity scores: a practical method for assessing treatment effects in pain and symptom management research. *J Pain Symptom Manage,* 48, 711-8.

GARRIDO, M. M., DEB, P., BURGESS, J. F., JR. & PENROD, J. D. 2012. Choosing models for health care cost analyses: issues of nonlinearity and endogeneity. *Health Serv Res,* 47, 2377-97.

GARRIDO, M. M., KELLEY, A. S., PARIS, J., ROZA, K., MEIER, D. E., MORRISON, R. S. & ALDRIDGE, M. D. 2014. Methods for constructing and assessing propensity scores. *Health Serv Res,* 49, 1701-1720.

GELFMAN, L. P., DU, Q. & MORRISON, R. S. 2013. An update: NIH research funding for palliative medicine 2006 to 2010. *J Palliat Med,* 16**,** 125-9.

GLICK, H. 2008. Methods for cost-estimation in CEA: the GLM approach. Academy Health Annual Research Meeting, 10th June, 2008. Washington, DC. Available: http://www.uphs.upenn.edu/dgimhsr/documents/acadhlth.glick.061008.pdf.

GLICK, H. 2014. *Stata programs: statistical analysis of costs* [Online]. University of Pennsylvania Health System. Available: http://www.uphs.upenn.edu/dgimhsr/stat-cstanal.htm [Accessed: November 27th, 2014.

GLOBAL BURDEN OF DISEASE STUDY COLLABORATORS 2015. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet,* 385**,** 117-71.

GOLD, M. R., SIEGEL, J. E., RUSSELL, L. B. & WEINSTEIN, M. C. 1996. *Cost-effectiveness in health and medicine*, New York ; Oxford, Oxford University Press.

GOLDSMITH, B., DIETRICH, J., DU, Q. & MORRISON, R. S. 2008. Variability in access to hospital palliative care in the United States. *J Palliat Med,* 11**,** 1094-102.

GOLDSTEIN, N. E. & MORRISON, R. S. 2005. The intersection between geriatrics and palliative care: a call for a new research agenda. *J Am Geriatr Soc,* 53**,** 1593-8.

GOMES, B., HARDING, R., FOLEY, K. M. & HIGGINSON, I. J. 2009. Optimal approaches to the health economics of palliative care: report of an international think tank. *J Pain Symptom Manage,* 38**,** 4-10.

GOMES, B., HIGGINSON, I. J., CALANZANI, N., COHEN, J., DELIENS, L., DAVESON, B. A., BECHINGER-ENGLISH, D., BAUSEWEIN, C., FERREIRA, P. L., TOSCANI, F., MENACA, A., GYSELS, M., CEULEMANS, L., SIMON, S. T., PASMAN, H. R., ALBERS, G., HALL, S., MURTAGH, F. E., HAUGEN, D. F., DOWNING, J., KOFFMAN, J., PETTENATI, F., FINETTI, S., ANTUNES, B. & HARDING, R. 2012. Preferences for place of death if faced with advanced cancer: a population survey in England, Flanders, Germany, Italy, the Netherlands, Portugal and Spain. *Ann Oncol,* 23**,** 2006-15.

GREEN, K. M. & STUART, E. A. 2014. Examining moderation analyses in propensity score methods: application to depression and substance use. *J Consult Clin Psychol,* 82**,** 773-83.

GROENEVELD, I., MURTAGH, F., KALOKI, Y., BAUSEWEIN, C. & HIGGINSON, I. 2013. Determinants of healthcare costs in the last year of life. The Annual Assembly of American Academy of Hospice and Palliative Medicine & Hospice and Palliative Nurses Association, 14th March, 2013. New Orleans, LA.

GUERRIERE, D. N., ZAGORSKI, B., FASSBENDER, K., MASUCCI, L., LIBRACH, L. & COYTE, P. C. 2010. Cost variations in ambulatory and home-based palliative care. *Palliat Med,* 24**,** 523-32.

HANCHATE, A., KRONMAN, A. C., YOUNG-XU, Y., ASH, A. S. & EMANUEL, E. 2009. Racial and ethnic differences in end-of-life costs: why do minorities cost more than whites? *Arch Intern Med,* 169**,** 493-501.

HANRATTY, B., BURSTROM, B., WALANDER, A. & WHITEHEAD, M. 2007. Inequality in the face of death? Public expenditure on health care for different socioeconomic groups in the last year of life. *J Health Serv Res Policy,* 12**,** 90-4.

HANSON, L. C., USHER, B., SPRAGENS, L. & BERNARD, S. 2008. Clinical and economic impact of palliative care consultation. *J Pain Symptom Manage,* 35**,** 340-6.

HARDING, R., GOMES, B., FOLEY, K. M. & HIGGINSON, I. J. 2009. Research priorities in health economics and funding for palliative care: views of an international think tank. *J Pain Symptom Manage,* 38**,** 11-4.

HARRIS, I. & MURRAY, S. A. 2013. Can palliative care reduce futile treatment? A systematic review. *BMJ Support Palliat Care,* 3**,** 389-98.

HEALTH INFORMATION AND QUALITY AUTHORITY 2014. Guidelines for the economic evaluation of health technologies in Ireland. Dublin: HIQA.

HEALTH RESEARCH BOARD. 2011. *Health Economics of Cancer Fellowships 2011* [Online]. Dublin: HRB. Available: http://www.hrb.ie/research-strategy-funding/grants-and-fellowships/hrb-grants-and-fellowships/grant/66/ [Accessed: 31st July 2014.

HIGGINSON, I. J. & EVANS, C. J. 2010. What is the evidence that palliative care teams improve outcomes for cancer patients and their families? *Cancer J,* 16**,** 423-35.

HIGGINSON, I. J., EVANS, C. J., GRANDE, G., PRESTON, N., MORGAN, M., MCCRONE, P., LEWIS, P., FAYERS, P., HARDING, R., HOTOPF, M., MURRAY, S. A., BENALIA, H., GYSELS, M., FARQUHAR, M. & TODD, C. 2013. Evaluating complex interventions in end of life care: the MORECare statement on good practice generated by a synthesis of transparent expert consultations and systematic reviews. *BMC Med,* 11**,** 111.

HIGGINSON, I. J., FINLAY, I., GOODWIN, D. M., COOK, A. M., HOOD, K., EDWARDS, A. G., DOUGLAS, H. R. & NORMAN, C. E. 2002. Do hospital-based palliative teams improve care for patients or families at the end of life? *J Pain Symptom Manage,* 23**,** 96-106.

HIGGINSON, I. J., FINLAY, I. G., GOODWIN, D. M., HOOD, K., EDWARDS, A. G., COOK, A., DOUGLAS, H. R. & NORMAND, C. E. 2003. Is there evidence that palliative care teams alter end-of-life experiences of patients and their caregivers? *J Pain Symptom Manage,* 25**,** 150-68.

HIGGINSON, I. J., MCCRONE, P., HART, S. R., BURMAN, R., SILBER, E. & EDMONDS, P. M. 2009. Is short-term palliative care cost-effective in multiple sclerosis? A randomized phase II trial. *J Pain Symptom Manage,* 38**,** 816-26.

HILL, S. C. & MILLER, G. E. 2010. Health expenditure estimation and functional form: applications of the generalized gamma and extended estimating equations models. *Health Econ,* 19**,** 608-27.

HOSMER, D. W. & LEMESHOW, S. 2000. *Applied Logistic Regression*, New York, Wiley.

HUGHES, J. 2005. Palliative care and the QALY problem. *Health Care Anal,* 13**,** 289-301.

HUGHES, M. T. & SMITH, T. J. 2014. The growth of palliative care in the United States. *Annu Rev Public Health,* 35**,** 459-75.

HUYNH, T. N., KLEERUP, E. C., WILEY, J. F., SAVITSKY, T. D., GUSE, D., GARBER, B. J. & WENGER, N. S. 2013. The frequency and cost of treatment perceived to be futile in critical care. *JAMA Intern Med,* 173**,** 1887-94.

IMBENS, G. W. 2004. Nonparametric estimation of average treatment effects under exogeneity: a review. *Rev Econ Stat,* 86**,** 4-29.

ISHAK, K. J., STOLAR, M., HU, M. Y., ALVAREZ, P., WANG, Y., GETSIOS, D. & WILLIAMS, G. C. 2012. Accounting for the relationship between per diem cost and LOS when estimating hospitalization costs. *BMC Health Serv Res,* 12**,** 439.

JOHNSON, C. E., GIRGIS, A., PAUL, C. L. & CURROW, D. C. 2008. Cancer specialists' palliative care referral practices and perceptions: results of a national survey. *Palliat Med,* 22**,** 51-7.

JONES, A. M. 2010. *Models for health care*, Working Paper 10/01, York: Health Economics and Data Group, University of York.

JONES, A. M., LOMAS, J., MOORE, P. & RICE, N. 2013a. *A quasi-Monte Carlo comparison of developments in parametric and semi-parametric regression methods for heavy tailed and non-normal data: with an application to healthcare costs*, Working Paper 13/30, York: Health Economics and Data Group, University of York.

JONES, A. M., RICE, N., BAGO D'UVA, T. & BALIA, S. 2013b. *Applied Health Economics,* 2nd, Oxford, Routledge.

JONES, A. M., RICE, N., BAGO D'UVA, T. & BALIA, S. 2013c. *Applied Health Economics: Software and Data Resources* [Online]. York: HEDG, University of York. Available: http://www.york.ac.uk/economics/postgrad/herc/hedg/software/ [Accessed: November 27th, 2014.

KAMAL, A. H., SWETZ, K. M., CAREY, E. C., CHEVILLE, A. L., LIU, H., RUEGG, S. R., MOYNIHAN, T. J., SLOAN, J. A. & KAUR, J. S. 2011. Palliative care consultations in patients with cancer: a Mayo Clinic 5-year review. *J Oncol Pract,* 7**,** 48-53.

KATZ, S., FORD, A., MOSKOWITZ, R., JACKSON, B. & JAFFE, M. 1963. The index of ADL: a standardized measure of biological and psychological function. *JAMA,* 185, 914-919.

KELLEY, A. S., ETTNER, S. L., MORRISON, R. S., DU, Q., WENGER, N. S. & SARKISIAN, C. A. 2011. Determinants of medical expenditures in the last 6 months of life. *Ann Intern Med,* 154, 235-42.

KELLY, R. J. & SMITH, T. J. 2014. Delivering maximum clinical benefit at an affordable price: engaging stakeholders in cancer care. *Lancet Oncol,* 15, e112-8.

KHANDELWAL, N., ENGELBERG, R. A., BENKESER, D. C., COE, N. B. & CURTIS, J. R. 2014. End-of-life expenditure in the ICU and perceived quality of dying. *Chest,* 146, 1594-603.

KHANDELWAL, N., KROSS, E. K., ENGELBERG, R. A., COE, N. B., LONG, A. C. & CURTIS, J. R. 2015. Estimating the effect of palliative care interventions and advance care planning on ICU utilization: a systematic review. *Crit Care Med,* 2015 Jan 9. [Epub ahead of print].

KRISTJANSON, L. J. 1993. Validity and reliability testing of the FAMCARE scale: measuring family satisfaction with advanced cancer care. *Soc Sci Med,* 36, 693-701.

KUMAR, P., CASARETT, D., CORCORAN, A., DESAI, K., LI, Q., CHEN, J., LANGER, C. & MAO, J. J. 2012. Utilization of supportive and palliative care services among oncology outpatients at one academic cancer center: determinants of use and barriers to access. *J Palliat Med,* 15, 923-30.

KUO, D. Z., SISTERHEN, L. L., SIGREST, T. E., BIAZO, J. M., AITKEN, M. E. & SMITH, C. E. 2012. Family experiences and pediatric health services use associated with family-centered rounds. *Pediatrics,* 130, 299-305.

LEHNERT, T., HEIDER, D., LEICHT, H., HEINRICH, S., CORRIERI, S., LUPPA, M., RIEDEL-HELLER, S. & KONIG, H. H. 2011. Review: health care utilization and costs of elderly persons with multiple chronic conditions. *Med Care Res Rev,* 68, 387-420.

LEVIT, L., BALOGH, E., NASS, S. & GANZ, P. A. (eds.) 2013. *Delivering High-Quality Cancer Care: Charting a New Course for a System in Crisis,* Washington (DC): Institute of Medicine/The National Academies Press.

LINDSAY, J., DOOLEY, M., MARTIN, J., FAY, M., KEARNEY, A. & BARRAS, M. 2014. Reducing potentially inappropriate medications in palliative cancer patients: evidence to support deprescribing approaches. *Support Care Cancer,* 22, 1113-9.

LOCHNER, K. A., GOODMAN, R. A., POSNER, S. & PAREKH, A. 2013. Multiple chronic conditions among Medicare beneficiaries: state-level variations in prevalence, utilization, and cost, 2011. *Medicare Medicaid Res Rev,* 3.

LUCKETT, T., PHILLIPS, J., AGAR, M., VIRDUN, C., GREEN, A. & DAVIDSON, P. M. 2014. Elements of effective palliative care models: a rapid review. *BMC Health Serv Res,* 14, 136.

LUPU, D. 2010. Estimate of current hospice and palliative medicine physician workforce shortage. *J Pain Symptom Manage,* 40, 899-911.

MALIN, J. L., WEEKS, J. C., POTOSKY, A. L., HORNBROOK, M. C. & KEATING, N. L. 2013. Medical oncologists' perceptions of financial incentives in cancer care. *J Clin Oncol,* 31, 530-5.

MANFREDI, P. L., MORRISON, R. S., MORRIS, J., GOLDHIRSCH, S. L., CARTER, J. M. & MEIER, D. E. 2000. Palliative care consultations: how do they impact the care of hospitalized patients? *J Pain Symptom Manage,* 20, 166-73.

MANNING, W. G. 1998. The logged dependent variable, heteroscedasticity, and the retransformation problem. *J Health Econ,* 17, 283-95.

MANNING, W. G., BASU, A. & MULLAHY, J. 2005. Generalized modeling approaches to risk adjustment of skewed outcomes data. *J Health Econ,* 24, 465-88.

MANNING, W. G. & MULLAHY, J. 2001. Estimating log models: to transform or not to transform? *J Health Econ,* 20, 461-494.

MARAZZI, A., PACCAUD, F., RUFFIEUX, C. & BEGUIN, C. 1998. Fitting the distributions of length of stay by parametric models. *Medical Care,* 36, 915-27.

MARCUS, S. & GIBBONS, R. 2001. Estimating the efficacy of receiving treatment in randomized clinical trials with noncompliance. *Health Serv Outcomes Res Methodol,* 2**,** 247-258.

MARIOTTO, A. B., YABROFF, K. R., SHAO, Y., FEUER, E. J. & BROWN, M. L. 2011. Projections of the cost of cancer care in the United States: 2010-2020. *J Natl Cancer Inst,* 103**,** 117-28.

MAY, P., GARRIDO, M. M., CASSEL, J. B., KELLEY, A. S., MEIER, D. E., NORMAND, C., SMITH, T. J., STEFANIS, L. & MORRISON, R. S. 2015. Prospective cohort study of hospital palliative care teams for inpatients with advanced cancer: earlier consultation is associated with larger cost-saving effect. *J Clin Oncol,* 33**,** 2745-52.

MAY, P., NORMAND, C. & MORRISON, R. S. 2014. Economic impact of hospital inpatient palliative care consultation: review of current evidence and directions for future research. *J Palliat Med,* 17**,** 1054-63.

MCCARTHY, I. M., ROBINSON, C., HUQ, S., PHILASTRE, M. & FINE, R. L. 2015. Cost savings from palliative care teams and guidance for a financially viable palliative care program. *Health Services Research,* 50**,** 217-36.

MCKEE, M., BRITTON, A., BLACK, N., MCPHERSON, K., SANDERSON, C. & BAIN, C. 1999. Methods in health services research. Interpreting the evidence: choosing between randomised and non-randomised studies. *BMJ,* 319**,** 312-5.

MCPAKE, B., NORMAND, C. & SMITH, S. 2013. *Health Economics : An international perspective,* Third edition., Routledge.

MEIER, D. E. 2011. Increased access to palliative care and hospice services: opportunities to improve value in health care. *Milbank Q,* 89**,** 343-80.

MEIER, D. E. & BERESFORD, L. 2009. Palliative care cost research can help other palliative care programs make their case. *J Palliat Med,* 12**,** 15-20.

MIHAYLOVA, B., BRIGGS, A., O'HAGAN, A. & THOMPSON, S. G. 2011. Review of statistical methods for analysing healthcare resources and costs. *Health Econ,* 20**,** 897-916.

MORRISON, R. S. 2013a. Models of palliative care delivery in the United States. *Curr Opin Support Palliat Care,* 7**,** 201-6.

MORRISON, R. S. 2013b. Research priorities in geriatric palliative care: an introduction to a new series. *J Palliat Med,* 16**,** 726-9.

MORRISON, R. S., DIETRICH, J., LADWIG, S., QUILL, T., SACCO, J., TANGEMAN, J. & MEIER, D. E. 2011a. Palliative care consultation teams cut hospital costs for Medicaid beneficiaries. *Health Aff (Millwood),* 30**,** 454-63.

MORRISON, R. S., DIETRICH, J., LADWIG, S., QUILL, T., SACCO, J., TANGEMAN, J. & MEIER, D. E. 2011b. Palliative care consultation teams cut hospital costs for Medicaid beneficiaries (Detailed Methods Appendix). *Health Aff (Millwood),* 30**,** [Online only].

MORRISON, R. S., PENROD, J. D., CASSEL, J. B., CAUST-ELLENBOGEN, M., LITKE, A., SPRAGENS, L. & MEIER, D. E. 2008. Cost savings associated with US hospital palliative care consultation programs. *Arch Intern Med,* 168**,** 1783-90.

MOSES, H., 3RD, MATHESON, D. H., DORSEY, E. R., GEORGE, B. P., SADOFF, D. & YOSHIMURA, S. 2013. The anatomy of health care in the United States. *JAMA,* 310**,** 1947-63.

MURRAY, M. P. 2006. Avoiding invalid instruments and coping with weak instruments. *J Econ Perspect,* 20**,** 111-132.

NATIONAL CENTER FOR HEALTH STATISTICS 2015. Deaths: Final data for 2013. Hyattsville, MD: National Vital Statistics Reports.

NATIONAL COMPREHENSIVE CANCER NETWORK. 2005. *Guidelines for Supportive Care,* Fort Washington, PA.

NATIONAL INSTITUTES OF HEALTH. 2006. *Project information: Palliative care for hospitalized cancer patients* [Online]. Bethesda, MD: US Department of Health & Human Services. Available: http://projectreporter.nih.gov/project_info_description.cfm?projectnumber=5R01CA116227-04 [Accessed: 1st August 2015.

NATIONAL UNIFORM BILLING COMMITTEE. 2011. *Official UB-04 Data Specifications Manual,* Chicago, IL: American Hospitals Association.

NORMAND, C. 2009. Measuring outcomes in palliative care: limitations of QALYs and the road to PalYs. *J Pain Symptom Manage,* 38, 27-31.

NORMAND, C. 2012. Setting priorities in and for end-of-life care: challenges in the application of economic evaluation. *Health Econ Policy Law,* 7, 431-9.

O'MAHONY, S., BLANK, A. E., ZALLMAN, L. & SELWYN, P. A. 2005. The benefits of a hospital-based inpatient palliative care consultation service: preliminary outcome data. *J Palliat Med,* 8, 1033-9.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. 2014. *How does the United States compare? ,* Paris: OECD Health Statistics 2014 - Country Notes.

PANTILAT, S. Z., RABOW, M. W., KERR, K. M. & MARKOWITZ, J. D. 2007. *Palliative care in California: The business case for hospital-based programs,* Oakland, CA: California Healthcare Foundation.

PARIKH, R. B., KIRCH, R. A., SMITH, T. J. & TEMEL, J. S. 2013. Early specialty palliative care - translating data in oncology into practice. *N Engl J Med,* 369, 2347-51.

PARK, R. E. 1966. Estimation with Heteroscedastic Error Terms. *Econometrica,* 34, 888.

PENROD, J., MORRISON, R. S. & MEIER, D. E. 2008. Studying the effectiveness of palliative care. *JAMA,* 300, 1022-3; author reply 1023-4.

PENROD, J. D., DEB, P., DELLENBAUGH, C., BURGESS, J. F., JR., ZHU, C. W., CHRISTIANSEN, C. L., LUHRS, C. A., CORTEZ, T., LIVOTE, E., ALLEN, V. & MORRISON, R. S. 2010. Hospital-based palliative care consultation: effects on hospital cost. *J Palliat Med,* 13, 973-9.

PENROD, J. D., DEB, P., LUHRS, C., DELLENBAUGH, C., ZHU, C. W., HOCHMAN, T., MACIEJEWSKI, M. L., GRANIERI, E. & MORRISON, R. S. 2006. Cost and utilization outcomes of patients receiving hospital-based palliative care consultation. *J Palliat Med,* 9, 855-60.

PENROD, J. D., GOLDSTEIN, N. E. & DEB, P. 2009. When and how to use instrumental variables in palliative care research. *J Palliat Med,* 12, 471-4.

PFUNTNER, A., WIER, L. M. & ELIXHAUSER, A. 2013. *Overview of hospital stays in the United States, 2011,* Rockville, MD.: Agency for Healthcare Research and Quality.

PREGIBON, D. 1980. Goodness of link tests for generalized linear models. *J R Stat Soc. Series C (Applied Statistics),* 29, 15-14.

PRESTON, N. J., SHORT, V., HOLLINGWORTH, W., MCCRONE, P., GRANDE, G., EVANS, C., ANSCOMBE, E., BENALIA, H., HIGGINSON, I. J. & TODD, C. 2012. MORECare research methods guidance development: recommendations for health economic evaluation in palliative and end-of-life care research [abstract]. *Palliat Med,* 26, 541.

QUEEN, M. A., MYERS, A. L., HALL, M., SHAH, S. S., WILLIAMS, D. J., AUGER, K. A., JERARDI, K. E., STATILE, A. M. & TIEDER, J. S. 2014. Comparative effectiveness of empiric antibiotics for community-acquired pneumonia. *Pediatrics,* 133, e23-9.

QUILL, T. E. & ABERNETHY, A. P. 2013. Generalist plus specialist palliative care - creating a more sustainable model. *N Engl J Med,* 368, 1173-5.

QUINN, K. 2008. New directions in Medicaid payment for hospital care. *Health Aff (Millwood),* 27, 269-80.

RAFTERY, J. P., ADDINGTON-HALL, J. M., MACDONALD, L. D., ANDERSON, H. R., BLAND, J. M., CHAMBERLAIN, J. & FREELING, P. 1996. A randomized controlled trial of the cost-effectiveness of a district co-ordinating service for terminally ill cancer patients. *Palliat Med,* 10, 151-61.

RITCHIE, C. S. & ZULMAN, D. M. 2013. Research priorities in geriatric palliative care: multimorbidity. *J Palliat Med,* 16, 843-7.

ROSENBAUM, P. R. & RUBIN, D. B. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika,* 70, 41-55.

ROUND, J. 2012. Is a QALY still a QALY at the end of life? *J Health Econ,* 31, 521-7.

RUBIN, D. B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol,* 66**,** 688–701.

RUBIN, D. B. 2007. The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. *Stat Med,* 26**,** 20-36.

SCHOEN, C., OSBORN, R., HOW, S. K., DOTY, M. M. & PEUGH, J. 2009. In chronic condition: experiences of patients with complex health care needs, in eight countries, 2008. *Health Aff (Millwood),* 28**,** w1-16.

SHUGARMAN, L. R., BIRD, C. E., SCHUSTER, C. R. & LYNN, J. 2007. Age and gender differences in Medicare expenditures at the end of life for colorectal cancer decedents. *J Womens Health (Larchmt),* 16**,** 214-27.

SHUGARMAN, L. R., BIRD, C. E., SCHUSTER, C. R. & LYNN, J. 2008. Age and gender differences in medicare expenditures and service utilization at the end of life for lung cancer decedents. *Women Health Iss,* 18**,** 199-209.

SIMOENS, S., KUTTEN, B., KEIRSE, E., BERGHE, P. V., BEGUIN, C., DESMEDT, M., DEVEUGELE, M., LEONARD, C., PAULUS, D. & MENTEN, J. 2010. The costs of treating terminal patients. *J Pain Symptom Manage,* 40**,** 436-48.

SISKO, A. M., KEEHAN, S. P., CUCKLER, G. A., MADISON, A. J., SMITH, S. D., WOLFE, C. J., STONE, D. A., LIZONITZ, J. M. & POISAL, J. A. 2014. National health expenditure projections, 2013-23: faster growth expected with expanded coverage and improving economy. *Health Aff (Millwood),* 33**,** 1841-50.

SIU, A. L., SPRAGENS, L. H., INOUYE, S. K., MORRISON, R. S. & LEFF, B. 2009. The ironic business case for chronic care in the acute care setting. *Health Aff (Millwood),* 28**,** 113-25.

SLEEMAN, K. E., GOMES, B. & HIGGINSON, I. J. 2012. Research into end-of-life cancer care - investment is needed. *Lancet,* 379**,** 519.

SMITH, S., BRICK, A., O'HARA, S. & NORMAND, C. 2014. Evidence on the cost and cost-effectiveness of palliative care: a literature review. *Palliat Med,* 28**,** 130-150.

SMITH, T. J. & CASSEL, J. B. 2009. Cost and non-clinical outcomes of palliative care. *J Pain Symptom Manage,* 38**,** 32-44.

SMITH, T. J., COYNE, P. J. & CASSEL, J. B. 2012a. Practical guidelines for developing new palliative care services: resource management. *Ann Oncol,* 23 Suppl 3**,** 70-5.

SMITH, T. J. & HILLNER, B. E. 2011. Bending the cost curve in cancer care. *N Engl J Med,* 364**,** 2060-5.

SMITH, T. J., TEMIN, S., ALESI, E. R., ABERNETHY, A. P., BALBONI, T. A., BASCH, E. M., FERRELL, B. R., LOSCALZO, M., MEIER, D. E., PAICE, J. A., PEPPERCORN, J. M., SOMERFIELD, M., STOVALL, E. & VON ROENN, J. H. 2012b. American Society of Clinical Oncology provisional clinical opinion: the integration of palliative care into standard oncology care. *J Clin Oncol,* 30**,** 880-7.

SONI, A. 2011. *Top 10 most costly conditions among men and women, 2008: estimates for the U.S. civilian noninstitutionalized adult population, age 18 and older*, #331, Rockville, MD: Agency for Healthcare Research and Quality.

SORENSON, C. 2012. Valuing end-of-life care in the United States: the case of new cancer drugs. *Health Econ Policy Law,* 7**,** 411-30.

STARKS, H., DIEHR, P. & CURTIS, J. R. 2009. The challenge of selection bias and confounding in palliative care research. *J Palliat Med,* 12**,** 181-7.

STARKS, H., WANG, S., FARBER, S., OWENS, D. A. & CURTIS, J. R. 2013. Cost savings vary by length of stay for inpatients receiving palliative care consultation services. *J Palliat Med,* 16**,** 1215-20.

STATACORP 2011. Stata Statistical Software: Release 12. College Station, TX: StataCorp LP.

STEINHAUSER, K. E., CLIPP, E. C., HAYS, J. C., OLSEN, M., ARNOLD, R., CHRISTAKIS, N. A., LINDQUIST, J. H. & TULSKY, J. A. 2006. Identifying, recruiting, and retaining seriously-ill patients and their caregivers in longitudinal research. *Palliat Med,* 20**,** 745-54.

STUART, E. A. 2010. Matching methods for causal inference: a review and a look forward. *Stat Sci,* 25**,** 1-21.

TAHERI, P. A., BUTZ, D., GRIFFES, L. C., MORLOCK, D. R. & GREENFIELD, L. J. 2000. Physician impact on the total cost of care. *Ann Surg,* 231, 432-435.

TEMEL, J. S., GREER, J. A., MUZIKANSKY, A., GALLAGHER, E. R., ADMANE, S., JACKSON, V. A., DAHLIN, C. M., BLINDERMAN, C. D., JACOBSEN, J., PIRL, W. F., BILLINGS, J. A. & LYNCH, T. J. 2010. Early palliative care for patients with metastatic non-small-cell lung cancer. *N Engl J Med,* 363, 733-42.

TENO, J. M., CLARRIDGE, B., CASEY, V., EDGMAN-LEVITAN, S. & FOWLER, J. 2001. Validation of toolkit After-Death Bereaved Family Member Interview. *J Pain Symptom Manage,* 22, 752-8.

TENO, J. M., GOZALO, P. L., BYNUM, J. P., LELAND, N. E., MILLER, S. C., MORDEN, N. E., SCUPP, T., GOODMAN, D. C. & MOR, V. 2013. Change in end-of-life care for Medicare beneficiaries: site of death, place of care, and health care transitions in 2000, 2005, and 2009. *JAMA,* 309, 470-7.

THOMAS, R. E., WILSON, D. & SHEPS, S. 2006. A literature review of randomized controlled trials of the organization of care at the end of life. *Can J Aging,* 25, 271-93.

THOMPSON, A. H., ALIBHAI, A., SAUNDERS, L. D., CUMMING, D. C. & THANIGASALAM, N. 2003. Post-maternity outcomes following health care reform in Alberta: 1992-1996. *Can J Public Health,* 94, 104-8.

TIBI-LEVY, Y., LE VAILLANT, M. & DE POUVOURVILLE, G. 2006. Determinants of resource utilization in four palliative care units. *Palliat Med,* 20, 95-106.

UNROE, K. T. & MEIER, D. E. 2013. Research priorities in geriatric palliative care: policy initiatives. *J Palliat Med,* 16, 1503-8.

VAN WALRAVEN, C., AUSTIN, P. C., JENNINGS, A., QUAN, H. & FORSTER, A. J. 2009. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. *Med Care,* 47, 626-33.

VON ELM, E., ALTMAN, D. G., EGGER, M., POCOCK, S. J., GOTZSCHE, P. C. & VANDENBROUCKE, J. P. 2007. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet,* 370, 1453-7.

WALKER, H., ANDERSON, M., FARAHATI, F., HOWELL, D., LIBRACH, S. L., HUSAIN, A., SUSSMAN, J., VIOLA, R., SUTRADHAR, R. & BARBERA, L. 2011. Resource use and costs of end-of-life/palliative care: Ontario adult cancer patients dying during 2002 and 2003. *J Palliat Care,* 27, 79-88.

WEINSTEIN, M. C., SIEGEL, J. E., GOLD, M. R., KAMLET, M. S. & RUSSELL, L. B. 1996. Recommendations of the panel on cost-effectiveness in health and medicine. *JAMA,* 276, 1253-8.

WHEELER, J. L., GREENE, A., TIEMAN, J. J., ABERNETHY, A. P. & CURROW, D. C. 2012. Key characteristics of palliative care studies reported in the specialized literature. *J Pain Symptom Manage,* 43, 987-92.

WHITFORD, K., SHAH, N. D., MORIARTY, J., BRANDA, M. & THORSTEINSDOTTIR, B. 2014. Impact of a palliative care consult service. *Am J Hosp Palliat Care,* 31, 175-82.

WORLD HEALTH ORGANIZATION. 2015. *WHO Definition of Palliative Care* [Online]. Available: http://www.who.int/cancer/palliative/definition/en/ [Accessed: March 4th, 2015.

ZHANG, B., WRIGHT, A. A., HUSKAMP, H. A., NILSSON, M. E., MACIEJEWSKI, M. L., EARLE, C. C., BLOCK, S. D., MACIEJEWSKI, P. K. & PRIGERSON, H. G. 2009. Health care costs in the last week of life: associations with end-of-life conversations. *Arch Intern Med,* 169, 480-8.

ZILBERBERG, M. D. & SHORR, A. F. 2012. Economics at the end of life: hospital and ICU perspectives. *Semin Respir Crit Care Med,* 33, 362-369.

ZIMMERMANN, C., RIECHELMANN, R., KRZYZANOWSKA, M., RODIN, G. & TANNOCK, I. 2008. Effectiveness of specialized palliative care: a systematic review. *JAMA,* 299, 1698-709.

ZIMMERMANN, C., SWAMI, N., KRZYZANOWSKA, M., HANNON, B., LEIGHL, N., OZA, A., MOORE, M., RYDALL, A., RODIN, G., TANNOCK, I., DONNER, A. & LO, C. 2014. Early palliative care for patients with advanced cancer: a cluster-randomised controlled trial. *Lancet,* 383, 1721-30.

Appendices

Appendix to §2.2: Search terms used and results returned in identifying systematic reviews through systematic database search (Figure 2.2.1)

The purpose of Appendix 2.2 is to detail the search terms used and results returned in identifying systematic reviews for the meta-review in **§2** (Figure 2.2.1). The search terms used to identify relevant reviews of palliative care interventions are detailed in the tables that follow.

The four database searches returned 483, 63, 3 and 29 papers respectively – hence (483+63+3+29=) 578 papers were considered in Figure 2.2.1.

In all tables below * indicates a truncation; where "cost*" is searched, results are returned for any word beginning with "cost": cost, costs, costing, etc.

Appendix 2.2/Table 1 Meta-review search for systematic reviews: PubMed search strategy

Step	Subject heading or search term	Studies returned
1	economic*[Title/Abstract] OR cost*[Title/Abstract] OR financ*[Title/Abstract] OR utilization[Title/Abstract]	669816
2	palliative care[Title/Abstract] OR hospice care[Title/Abstract] OR terminal care[Title/Abstract] OR end of life care[Title/Abstract]	22651
3	2 AND 3	2073
4	"palliative care/economics"[MeSH Terms] OR "terminal care/economics"[MeSH Terms] OR "palliative care/utilization"[MeSH Terms]	1928
5	3 OR 4	3572
6	review[Title] OR "review"[Publication Type]	2069557
7	"1980/01/01"[Date - Publication] : "2015/02/28"[Date - Publication]	19061471
8	English[Language]	20253149
9	5 AND 6 AND 7 AND 8 AND 9	483

Last search: March 25th, 2015.

Appendix 2.2/Table 2 Meta-review search for systematic reviews: CINAHL search strategy

Step	Subject heading or search term	Studies returned
1	MH cost savings OR MH cost benefit analysis OR MH (costs and cost analysis) OR MH Economics OR MH cost control OR MH health care costs OR MH health resource utilization OR MH health facility costs	60482
2	MH hospice OR MH palliative care OR MH terminal care OR MH critical illness	31782
3	1 AND 2	773
4	AB cost* OR AB econom* OR AB finan*	79569
5	AB palliative OR AB terminal OR AB end of life OR AB hospice	19883
6	4 AND 5	1388
7	3 OR 6	2015
8	TI review OR AB review	155600
9	8 AND 7	228
10	9 AND [Limiters: Published before 31/02/2015; English Language; Age Groups: All Adult; Publication type: Journal article]	63

Last search: March 25th, 2015.

Appendix 2.2/Table 3 Meta-review search for systematic reviews: EconLit search strategy

Step	Subject heading or search term	Studies returned
1	AB palliative care OR AB terminal care OR AB end of life care OR AB hospice care	67
2	TI Review or AB Review	72464
3	1 AND 2	4
4	3 AND [Limiters - Published Date: 19800101-20150231; Publication Type: Journal Article]	3

Last search: March 25th, 2015.

Appendix 2.2/Table 4 Meta-review search for systematic reviews: AMED search strategy

Step	Subject heading or search term	Studies returned
1	TI cost* or AB cost*	4932
2	TI palliative care or AB palliative care	6299
3	TI review or AB review	15531
4	1 AND 2 AND 3	29
5	4 AND [Limiter: Published before 31/02/2015]	29

Last search: March 25th, 2015.

Appendix to §2.3: Search terms used and results returned in identifying studies for the literature review through systematic database search (Figure 2.3.1b)

The purpose of Appendix 2.3 is to detail the search terms used and results returned in identifying studies for the literature review through systematic database search in §2 (Figure 2.3.1b). The search terms used to identify relevant reviews of palliative care interventions are detailed in the tables that follow.

The four database searches returned 395, 188, 18 and 5 papers respectively – hence (395+188+18+5=) 606 papers were considered in Figure 2.3.1b.

Appendix 2.3/Table 1 Database search for economic studies of PCCTs (2012-2015): PubMed search strategy

Step	Subject heading or search term	# of studies returned
1	"palliative care"[Title/Abstract] OR "end of life care"[Title/Abstract] OR hospice[Title/Abstract] OR "terminal care"[Title/Abstract]	26821
2	cost*[Title/Abstract] OR economic*[Title/Abstract] OR financial[Title/Abstract]	553059
3	"palliative care"[MeSH Terms] OR "hospice care"[MeSH Terms] OR "terminal care"[MeSH Terms]	75410
4	"cost benefit analysis"[MeSH Terms] OR "cost savings"[MeSH Terms] OR "cost control"[MeSH Terms] OR "economics"[MeSH Terms] OR "health care costs"[MeSH Terms] OR "health resources"[MeSH Terms] OR "health services/economics"[MeSH Terms] OR "hospital costs"[MeSH Terms] OR "costs and cost analysis" [MeSH Terms]	526481
5	1 AND 2	2034
6	1 AND 4	2568
7	3 AND 2	2708
8	3 AND 4	4596
9	"palliative care/economics"[MeSH Terms] OR "palliative care/utilization"[MeSH Terms] OR "terminal care/economics"[MeSH Terms] OR "critical illness/economics"[MeSH Terms]	2067
10	5 OR 6 OR 7 OR 8 OR 9	7903
11	hospital*[Title/Abstract] OR consult*[Title/Abstract]	953061
12	"2012/01/01"[Date - Publication] : "2015/02/28"[Date - Publication] AND "english"[Language] AND "journal article"[Publication Type]	2877442
13	10 AND 11 AND 12	395

Last search: 31st March, 2015.

Appendix 2.3/Table 2 Database search for economic studies of PCCTs (2012-2015): CINAHL search strategy

Step	Subject heading or search term	# of studies returned
1	MH cost savings OR MH cost benefit analysis OR MH (costs and cost analysis) OR MH Economics OR MH cost control OR MH health care costs OR MH health resource utilization OR MH health facility costs	60482
2	MH hospice OR MH palliative care OR MH terminal care OR MH critical illness	31782
3	1 AND 2	773
4	AB cost* OR AB econom* OR AB finan*	79569
5	AB palliative OR AB terminal OR AB end of life OR AB hospice	19883
6	4 AND 5	1388
7	3 OR 6	2015
8	7 AND [Limiters: Published between 01/01/2012 and 31/02/2015; English Language; Age Groups: All Adult; Publication type: Journal article]	188

Last search: 30th March, 2015.

Appendix 2.3/Table 3 Database search for economic studies of PCCTs (2012-2015): EconLit search strategy

Step	Subject heading or search term	# of studies returned
1	AB palliative care OR AB terminal care OR AB end of life care OR AB hospice care	67
2	1 AND [Limiters - Published Date: 20120101-20150231; Publication Type: Journal Article]	18

Last search: 30th March, 2015.

Appendix 2.3/Table 4 Database search for economic studies of PCCTs (2012-2015): AMED search strategy

Step	Subject heading or search term	# of studies returned
1	TI cost* or AB cost*	4932
2	TI palliative care or AB palliative care	6299
3	TI hospital or AB hospital	9124
4	1 AND 2 AND 3	75
5	4 AND [Limiter: Published Date: 20120101-20150231]	5

Last search: 30th March, 2015.

Appendix to §3: STROBE checklist for my primary analysis

Appendix 3/Table 1 STROBE checklist of items for reports of observational studies

	Item No	Notes
Title and abstract	1	Provided at the beginning of the thesis.
Introduction		
Background/rationale	2	Broad policy and research context provided in *§1*. State of the science in economic evaluation of PCCT intervention established in *§2*. Specific context for 'PC4C' study and my primary research given in *§3.1.1*.
Objectives	3	See *§3.1.2*
Methods		
Study design	4	See *§3.1.3*
Setting	5	See *§3.1.4*
Participants	6	(*a*) For eligibility criteria, and the sources and methods of recruitment, see *§3.1.5*. (*b*) For matching process and criteria, see *§3.2.1* and Table 3.2.1. For number of exposed and unexposed, see *§4.1* and Figure 4.1.1.
Variables	7	See *§3.1.6*
Data sources	8	See *§3.1.7*
Bias	9	See *§3.2.1*
Study size	10	See *§3.1.8*
Quantitative variables	11	See *§3.1.9*
Statistical methods	12	(*a*) For those used to control for confounding, see *§3.2.1*. For those used to model costs and estimate PCCT impact on outcomes of interest, see *§3.2.2* and *§4.2*. All analyses are performed within this framework. Additional section-specific context is provided in the respective Methods sub-sections within *§4.2*, *§4.3*, *§4.4*, *§4.5* and *§4.6*.
		(*b*) For an overview of those used to control for confounding in subgroups, see *§3.2.1>Additional propensity scores*. The details of sub-sample-specific propensity scores are provided in Appendices, clearly labelled within *§4.3*, *§4.4*, *§4.5* and *§4.6*. Where interactions between patient-level factors and cost are examined, methods are explained in *§4.6>Methods*.
		(*c*) Where patient was missing cost data or baseline data required to calculate the propensity score, s/he was excluded. Where survival data were missing, patients were categorised as dying during hospitalisation. See *§4.1* and Figure 4.1.1.
		(*d*) Loss to follow-up not applicable in this study, which examines one

		hospitalisation per patient.
		(*e*) Multiple confirmatory/sensitivity analyses are detailed and reported in the respective Methods and Results sub-sections within *§4.3*, *§4.4*, *§4.5* and *§4.6*.
Results		
Participants	13	See *§4.1>Participants* and Figure 4.1.1.
Descriptive data	14	(a) For characteristics of study participants (e.g. demographic, clinical, social) and information on exposures and potential confounders, see *§4.1>Descriptive data*, and Table 4.1.1 and Table 4.1.2. For characteristics of sub-samples, see relevant Appendices.
		(b) In principal sample (n=969) there are no participants with missing data for any variable of interest.
Outcome data[121]	15	See *§4.1>Cost data*, and Table 4.1.3, Table 4.1.4, Table 4.1.5, and Table 4.1.6.
Main results	16	A preliminary analysis establishing appropriate economic modelling techniques is reported in *§4.2*. The main analysis of the thesis is the presented in *§4.3*, *§4.4*, *§4.5* and *§4.6*. Primary analyses are detailed and reported in the respective Results sub-sections within *§4.2*, *§4.3*, *§4.4*, *§4.5* and *§4.6*.
Other analyses	17	Secondary analyses are detailed and reported in the respective Results sub-sections within *§4.3*, *§4.4*, *§4.5* and *§4.6*.
Discussion		
Key results	18	Main results and secondary results are considered in the respective Discussion sub-sections within *§4.3*, *§4.4*, *§4.5* and *§4.6*.
Limitations	19	See *§5.2*
Interpretation	20	Interpretation of the primary analysis results are considered in the respective Discussion sub-sections within *§4.3*, *§4.4*, *§4.5* and *§4.6*. An overall interpretation of primary results is provided in *§5.1*. A summary of implications for future practice, policy and research is provided in *§5.3*.
Generalisability	21	See *§5.2*

[121] Costs are not typically categorised as outcomes in economic evaluation, which distinguishes between costs and consequences (see **§2.4>Forms of economic evaluation**). However, in the health services research literature from which the STROBE guidelines originate, costs are readily understood as an outcome (of interest). Since this is a PhD thesis in health economics, I follow the economics tradition in the main text: costs are referred to as a dependent variable (of interest) rather than as an outcome, and the phrase 'outcome of interest' is used only in general conceptual discussion (e.g. **§3.2.1 Bias & confounding: propensity score methods**). For the purposes of the STROBE checklist (and in papers originating from the thesis but published in non-economics journals), costs may be categorised as an outcome: they are the primary dependent variable on which the causal effect of the intervention is estimated.

Other information		
Funding	22	'PC4C' study funded by NCI & NINR (National Institutes of Health, 2006). Economics of Cancer fellowship funded by HRB and NCI (Health Research Board, 2011). Sponsors had no role in study design; in data collection, analysis and interpretation; or in writing or submission of the thesis. See also *§1* and *§3.1*.

Note: I did consider including a checklist for economic evaluations. However, the equivalent list for health economics is that of Drummond and Jefferson (1996). Their checklist is concerned with a series of issues that are not applicable to my thesis (including measuring benefits and time horizon issues) and does not include issues that are fundamental to my thesis, including baseline heterogeneity and heterogeneity of incremental effect. I therefore concluded that the STROBE checklist was adequate to illustrate the format of my enquiry.

Appendix to §4.1: Overview of patients not discharged alive

In all primary analyses in **§4 Results**, patients are separated by discharge status due to a possible heterogeneity problem.[122] Since there are insufficient patients in the decedent group to support a propensity score, the survivor group (n=969) is the principal sample throughout the thesis with sensitivity analysis performed where patients are pooled irrespective of discharge status (n=1,023).

In **§4.1** I provided an overview of how the principal sample was reached. The purpose of Appendix 4.1 is to provide an equivalent overview of patients who were not discharged alive.[123]

A comparison of samples by discharge status according to selected baseline statistics is provided in Appendix 4.1/Table 1, corresponding to Table 4.1.1 in the main document.

Appendix 4.1/Table 1 shows that there are no significant demographic or socio-economic differences between patients discharged alive and those who died during hospitalisation.

The most substantive differences relate to illness burden: patients who died had more comorbidities, lower activities of daily living and higher physical symptom scores. Additionally there were fewer patients in the decedent group with the primary diagnosis of 'other' cancer.

---

[122] Separating by discharge status is standard practice in this field but sensitivity analyses following pooling irrespective of discharge status have not been standard. See **§2.4**, **§3.2.1>Propensity score methods applied**, and **§4.1**.

[123] Patients coded as dying in hospital, n=51; patients with missing survival data, n=3.

Appendix 4.1/Table 1 Selected statistics of unweighted samples, by discharge status

		Discharged Alive (n=969)	Died in Hospital (n=54*)	t-test (p-value)
Age	*<55*	306 (31.6%)	11 (20.3%)	0.08
	55to75	553 (57.1%)	36 (66.7%)	0.17
	75<	110 (11.4%)	7 (13.0%)	0.72
Gender	*Female*	537 (55.4%)	28 (51.9%)	0.61
	Male	432 (44.6%)	26 (48.1%)	
Race	*White*	648 (66.9%)	32 (59.3%)	0.25
	Black	255 (26.3%)	19 (35.2%)	0.15
	Other	66 (6.8%)	3 (5.6%)	0.72
Insurance	*Medicare*	179 (18.5%)	12 (22.2%)	0.49
	Medicaid	167 (17.2%)	7 (13.0%)	0.42
	Other	623 (64.3%)	35 (64.8%)	0.94
Adv. Directive	*Yes*	559 (57.7%)	27 (50%)	0.27
Diagnosis	*GI*	211 (21.8%)	17 (31.5%)	0.10
	Lymphoma	76 (7.8%)	3 (5.6%)	0.54
	CNS	17 (1.8%)	0 (0.0%)	0.33
	Breast	96 (9.9%)	8 (14.8%)	0.25
	Gynaecological	115 (11.9%)	6 (11.1%)	0.87
	HCC	21 (2.2%)	2 (3.7%)	0.46
	Lung	137 (14.1%)	9 (16.7%)	0.61
	Other	296 (30.5%)	9 (16.7%)	**0.03**
Comorbidities	*Elixhauser Tot.*	3.41	5.00	**<0.01**
Activity Level	*ADL-6*	10.50	8.83	**<0.01**
Symptoms	*ESAS Physical*	1.59	2.00	**<0.01**
	ESAS Psych.	1.49	1.60	0.55
Site	*1*	122 (12.6%)	4 (7.4%)	0.36
	2	106 (10.9%)	8 (14.8%)	0.28
	3	331 (34.2%)	17 (31.5%)	0.69
	5	410 (42.3%)	25 (46.3%)	0.57

* 51 patients were recorded as dying during hospitalisation; three had missing survival data.

Appendix to §4.2: Details of models with covariates reduced &high-utilisation outliers removed

In **§4.2** I compared performance of different modelling approaches in estimating the incremental effect of the PCCT intervention on direct cost of hospital stay for the principal sample (n=969). This was done employing the following independent predictors (in addition to treatment): the 33 covariates included in the principal sample propensity score (Table 3.2.1) and fixed effects for each site.

In Table 4.2.5 I compared the impact on model performance of (i) reducing the number of covariates in the model to those considered most important for estimating incremental effect, and (ii) removing long-stay and high-cost outliers from the sample (since these distributions are heavily right-skewed, complicating analysis of the dependent variable).

The purpose of Appendix 4.2 is to detail these models and samples.

Models with fewer covariates

Details of the predictors employed in the 'reduced' and 'minimal' models in Table 4.2.5 are provided in Appendix 4.2/Table 1.

Appendix 4.2/Table 1 Predictors included in models with reduced # of covariates (per Table 4.2.5)

Predictors included in the 'Reduced' model:		**Predictors included in the 'Minimal' model:**	
Age	Elixhauser Total	Age	Fixed site effects
Gender	ESAS Phys. score†	Gender	
Race: White	ESAS Psych. score†	Race: White	
Race: Black	CMSAS Number †	Race: Black	
Advance directive	CMSAS Severity†	Advance directive	
Insurance: Medicare	Morphine dose	Insurance: Medicare	
Insurance: Medicaid	Home health aide	Insurance: Medicaid	
Diagnosis: Lymphoma	Fixed site effects	Elixhauser Total	
Visiting nurse services		Diagnosis: Cancer type	

For a detailed explanation of these variables, see Table 3.2.1.
† These scores are the sum of scores taken at baseline/admission and at consult/reference day.

Models with high-utilisation outliers removed

For the primary sample (n=969) in this thesis, propensity scores were calculated to balance the treatment and comparison groups on observed confounders (see **§3.2.1>Propensity score methods – applied**). Where the treatment and/or comparison group is redefined - for example, the removal of high-utilisation outliers as in **§4.2** - a new sample-specific propensity score was calculated (see **§3.2.1>Additional propensity scores**).

Overviews of the matched samples without the 5% of longest-stay patients (LOS<=20) and without the 5% of highest cost patients (Total direct costs<23153.96) are provided in Appendix 4.2/Table 2. In all cases balance was checked before matching (step 3, per Table 3.2.2 for the principal sample) and after matching (step 5, per Table 4.1.2 for the principal sample). Following matching, treatment and comparison groups were balanced on all covariates to within the 10% guideline highlighted by Garrido et al. (2014). No patient was lost to matching in either sample.

Appendix 4.2/Table 2 Summary of propensity score matching for sub-samples in Table 4.2.5

Sub-sample *Discharged alive &* _____	**Steps 1 & 2:** Included covariates differ from the 33 used to balance primary sample in Table 3.2.1?	**Step 5:** Patients supported (UC/PC/All) and patients lost to matching (LTM)			
		UC (n=)	**PC (n=)**	**All (n=)**	**LTM (n=)**
LOS<=20	No difference	686	238	924	0
Total direct costs<23153.96	Age re-categorised as a continuous variable; advance directive (=1 if living will=1 | proxy=1) replaces living will and proxy.	683	236	919	0

Appendix to §4.3: Summary of propensity scores

For the primary sample (n=969) in this thesis, propensity scores were calculated to balance the treatment and comparison groups on observed confounders (see **§3.2.1>Propensity score methods – applied**). Where the treatment and/or comparison group is redefined - for example, by patient length of stay as in **§4.3** - a new sample-specific propensity score was calculated (see **§3.2.1>Additional propensity scores**).

Overviews of the matched samples without short- and/or long-stay outliers are provided in Appendix 4.3/Table 1. In all cases balance was checked before matching (step 3, per Table 3.2.2 for the principal sample) and after matching (step 5, per Table 4.1.2 for the principal sample). Following matching, treatment and comparison groups were balanced on all covariates to within the 10% guideline highlighted by Garrido et al. (2014). Two patients were lost to matching in each sample; no alternative matching strategy (steps 4 &5, per Table 3.2.3 for the principal sample) reduced this loss to matching without a marked increase in bias between groups.

Appendix 4.3/Table 1 Summary of propensity score matching for sub-samples in Table 4.3.3

Sub-sample *Discharged alive &* _____	**Steps 1 & 2:** Included covariates differ from the 33 used to balance primary sample in Table 3.2.1?	**Step 5:** Patients supported (UC/PC/All) and patients lost to matching (LTM)			
		UC (n=)	**PC (n=)**	**All (n=)**	**LTM (n=)**
4<=LOS<=25	No difference	700	217	917	2
LOS<=25	Advance directive (=1 if living will=1 | proxy=1) replaces living will and proxy.	700	242	942	2
4<=LOS	Advance directive (=1 if living will=1 | proxy=1) replaces living will and proxy.	713	229	942	2

* Equivalent figures for LOS<=20 sample were previously provided in Appendix 4.2/Table 2.

Appendix to §4.4: Summary of propensity scores; overview of ICU sub-sample; results of sensitivity analyses

Appendix 4.4a

For the primary sample (n=969) in this thesis, propensity scores were calculated to balance the treatment and comparison groups on observed confounders (see **§3.2.1>Propensity score methods – applied**). Where the treatment and/or comparison group is redefined - for example, by intervention timing as in **§4.4** - a new sample-specific propensity score was calculated (see **§3.2.1>Additional propensity scores**).

Overviews of the matched samples for the main and secondary analyses in **§4.4** (Table 4.4.3, Table 4.4.4, Table 4.4.5, Table 4.4.6) are provided in Appendix 4.4a/Table 1. Equivalent tables for sensitivity analyses with late-consult outliers re-categorised as controls (Table 4.4.8), trimming both treatment and comparison arms by LOS (Table 4.4.9) and pooling all patients irrespective of discharge status (Table 4.4.10) are given in subsequent tables.

In all cases balance was checked before matching (step 3, per Table 3.2.2 for the principal sample) and after matching (step 5, per Table 4.1.2 for the principal sample). Following matching, treatment and comparison groups were balanced on all covariates to within the 10% guideline highlighted by Garrido et al. (2014). No more than four patients were lost to matching in any sample; no alternative matching strategy (steps 4 & 5, per Table 3.2.3 for the principal sample) reduced loss to matching without a marked increase in bias between groups.

Appendix 4.4a/Table 1 Summary of propensity score matching for sub-samples in Table 4.4.3, Table 4.4.4, Table 4.4.5, Table 4.4.6

Sub-sample *Discharged alive & intervention defined as within _____*	**Steps 1 & 2:** Included covariates differ from the 33 used to balance primary sample in Table 3.2.1?	**Step 5:** Patients supported (UC/PC/All) and patients lost to matching (LTM)			
		UC (n=)	**PC (n=)**	**All (n=)**	**LTM (n=)**
20 days	ESAS & CMSAS scores summed from admission and reference day creating one overall variable for each.	713	249	962	0
10 days	Binary age variables converted to a single continuous variable.	713	244	957	0
6 days	No difference	713	231	944	1
2 days	Advance directive (=1 if living will=1 | proxy=1) replaces living will and proxy.	713	197	910	0

Appendix 4.4a/Table 2 Summary of propensity score matching for sub-samples in Table 4.4.8

Sub-sample *D/C alive, intervention defined as within _____ & late-consult outliers controls*	**Steps 1 & 2:** Included covariates differ from the 33 used to balance primary sample in Table 3.2.1?	**Step 5:** Patients supported (UC/PC/All) and patients lost to matching (LTM)			
		UC (n=)	**PC (n=)**	**All (n=)**	**LTM (n=)**
20 days	Pain and fatigue variables excluded.	719	249	968	1
10 days	No difference	725	242	967	2
6 days	No difference	737	232	969	0
2 days	No difference	772	197	969	0

Appendix 4.4a/Table 3 Summary of propensity score matching for sub-samples in Table 4.4.9

Sub-sample *D/C alive, intervention defined as within _____ & LOS<_____*	**Steps 1 & 2:** Included covariates differ from the 33 used to balance primary sample in Table 3.2.1?	**Step 5:** Patients supported (UC/PC/All) and patients lost to matching (LTM)			
		UC (n=)	**PC (n=)**	**All (n=)**	**LTM (n=)**
10 days	No difference	583	192	775	4
6 days	No difference	360	114	474	0

* Equivalent figures for LOS<=20 sample were previously provided in Appendix 4.2/Table 2.

Appendix 4.4a/Table 4 Summary of propensity score matching for sub-samples in Table 4.4.10

Sub-sample *All patients irrespective of d/c status & intervention defined as within _____*	**Steps 1 & 2:** Included covariates differ from the 33 used to balance primary sample in Table 3.2.1?	**Step 5:** Patients supported (UC/PC/All) and patients lost to matching (LTM)			
		UC (n=)	**PC (n=)**	**All (n=)**	**LTM (n=)**
Any time	No difference	735	286	1021	2
20 days	No difference	735	279	1014	2
10 days	ESAS & CMSAS scores summed from admission and reference day creating one overall variable for each.	735	270	1005	2
6 days	No difference	735	254	989	1
2 days	ESAS & CMSAS scores summed from admission and reference day creating one overall variable for each.	735	210	945	1

Appendix 4.4b

In **§4.4>Secondary analyses** I examined the underling source of observed cost-savings by utilisation category (see Table 4.4.6). As detailed in the text it was not possible to repeat primary analysis methods with intensive care unit (ICU) costs as a dependent variable due to their unusual distribution: 11% (n=111) had non-zero ICU costs, an insufficient number to calculate a propensity score. Instead, regressions were run on an unmatched sample.

The purpose of Appendix 4.4b is to summarise the characteristics of this unmatched sample, in comparison to the principal sample (n=969) and separated by treatment group. These data are provided in Appendix 4.4b/Table 1.

The only statistically significant difference between ICU and non-ICU patients is that the former had higher comorbidities on the Elixhauser scale.

Among ICU patients only there are differences between palliative and usual care groups on advance directives, insurance programme and race. PC patients stayed significantly longer in hospital compared to UC patients in the ICU. There are no significant differences on illness burden but on balance the PC patients exhibit higher burden with (slightly) higher comorbidities and (noticeably) higher symptom scores. The UC patients have higher activities of daily living, which is not consistent with other samples in this thesis.

Given that PC patients in the ICU stayed significantly longer and had at least as large an illness burden, the estimated effect from an unweighted regression reported in Table 4.4.7 is taken to be indicative of the relationship; all else being equal the estimate might be larger in a weighted regression.

Appendix 4.4b/Table 1 Patients with non-zero ICU costs (n=111) compared to non-ICU patients (n=858) and separated by treatment group

		Non-ICU (n=858)	ICU (n=111)	t-test (p-value)	ICU Usual Care (n=85)	ICU Palliative Care (n=26)	t-test (p-value)
Age	*<55*	31.7%	30.6%	0.82	27.1%	42.3%	0.14
	55to75	57.0%	57.7%	0.89	61.2%	46.2%	0.18
	75<	11.3%	11.7%	0.90	11.8%	11.5%	0.98
Gender	*Female*	56.2%	49.5%	0.19	44.7%	65.4%	0.07
	Male	43.8%	50.5%		55.3%	34.6%	
Race	*White*	66.2%	72.1%	0.22	**77.6%**	**53.8%**	**0.02**
	Black	26.6%	24.3%	0.61	**18.8%**	**42.3%**	**0.01**
	Other	7.2%	3.6%	0.15	3.5%	3.8%	0.94
Insurance	*Medicare*	17.8%	23.4%	0.15	24.7%	19.2%	0.57
	Medicaid	17.5%	15.3%	0.57	**10.6%**	**30.8%**	**0.01**
	Other	64.7%	61.3%	0.48	64.7%	50.0%	0.18
Adv. Directive	*Yes*	57.9%	55.9%	0.68	**61.2%**	**38.5%**	**0.04**
Diagnosis	*Solid*	79.5%	86.5%	0.08	85.9%	88.5%	0.74
	Lymphoma	8.0%	6.3%	0.52	7.1%	3.8%	0.56
	Gynaecological	12.5%	7.2%	0.11	7.1%	7.7%	0.91
Comorbidities	*Elixhauser*	**3.3**	**4.2**	**<0.01**	4.1	4.3	0.68
Activity Level	*ADL-6*	10.5	10.4	0.66	10.6	9.7	0.09
Symptoms	*ESAS Physical*	1.6	1.5	0.18	1.4	1.7	0.08
	ESAS Psych.	1.5	1.5	0.86	1.5	1.5	0.86
	CMSAS #	7.9	7.4	0.18	7.1	8.3	0.11
	CMSAS Sev'rity	12.3	10.8	0.15	10.1	13.1	0.15
LOS	*Days*	**7.8**	**11.8**	**<0.01**	**10.9**	**14.8**	**0.04**

Appendix 4.4c

The purpose of Appendix 4.4c is to detail sensitivity analyses performed on the primary analysis in **§4.4**. The analyses are intended to check the robustness of the main results in two aspects specifically.

First, the model comparison selection process in **§4.2** concluded with the selection of generalised linear model (GLM) with a gamma distribution and a log link for primary analysis of the data on the basis of its comparatively strong performance in evaluation tests and its strength in providing a reliable mean incremental effect estimate in $ form. However, the selection was not straightforward: GLM (gamma, log) was the dominant GLM specification but did not dominate non-GLM models in evaluation; it failed one test (Pearson) with propensity score weights (Table 4.2.4) and other non-GLM models had better in-sample accuracy on some measures. In particular, the ln(OLS) approach performed better in evaluation once weights were added but this method has well known limitations of its own, detailed in **§3.2.2>Background**. Consequently I wanted to check my primary results with a standard OLS; I am conscious that the point estimates for incremental effect are unlikely to be accurate but want to ensure that the patterns observed in my main analysis are not due to model specification only.

Second, in the presence of unobserved confounding, it is possible for the impact of unobserved variables to worsen when using a propensity score approach (Brooks and Ohsfeldt, 2013). Consequently it is advisable to check results without weights for robustness.

With respect to these concerns, Appendix 4.4c/Table 1 presents results that repeat the Table 4.4.3 analysis with OLS on untransformed costs, with and without weights, and GLM (gamma, log) without weights.

These sensitivity analyses confirm the robustness of the results reported in Table 4.4.3: earlier palliative care consult is associated with larger cost-saving effect for a variety of modelling approaches, with and without propensity score weights.

For both models the unweighted estimates are lower than the equivalent result with weighted regressions, which is expected as weights balance the treatment and comparison groups (*inter alia*) on the higher illness burden of palliative care patients (see Table 4.1.1, Table 4.1.2).

The point estimates for effect vary by model, which is also expected. The OLS on untransformed costs has lower estimates than the GLM and much larger, less efficient confidence intervals, affirming the model's unsuitability for estimating cost effects with skewed, leptokurtic data.[124]

Therefore I am confident that the incremental-effect estimates produced in the primary analysis are robust to model selection and propensity score use (the strong observable association between timing and estimated effect on cost is persistent in all cases). Additionally I take my primary results to be more meaningful than those presented in Appendix 4.4c/Table 1 because (a) propensity score weights are important to account for confounding between treatment and comparison groups,

---

[124] Equivalent regressions on log-transformed costs using a smearing factor to re-convert estimates to USD find larger incremental effects than the GLM used in primary analysis, further affirming the unreliability of OLS approaches [data not shown].

and (b) GLM (gamma, log) is a more accurate and efficient estimator of incremental effect on cost than OLS approaches.[125]

[125] See also Appendix 4.5c, where the same patterns were observable and the same conclusions drawn for multimorbidity analyses.

Appendix 4.4c/Table 1 Sensitivity analysis on Table 4.4.3 results, checking robustness to model selection and propensity score weights (see §4.2)

Treatment defined as within _____ days of hospital admission (%ile)	UC (n=)	PC (n=)	OLS on untransformed costs						GLM (gamma, log)		
			Unweighted			Weighted			Unweighted		
			$ Effect (95% CI)	P value	% Saving	$ Effect (95% CI)	P value	% Saving	$ Effect (95% CI)	P value	% Saving
Any time (100%)	713	256	+1201 (-786 to +3188)	0.24	(12%)	+917 (-946 to +2780)	0.34	(10%)	+730 (-789 to +2249)	0.35	(8%)
20 (97.5%)	713	249	+276 (-1705 to +2257)	0.79	(3%)	-124 (-1988 to +1739)	0.90	1%	+69 (-1401 to +1540)	0.93	1%
10 (95%)	713	243	-300 (-2119 to +1519)	0.75	3%	-629 (-2262 to +1004)	0.45	7%	-383 (-1732 to +967)	0.58	4%
6 (90%)	713	231	-641 (-2514 to +1232)	0.50	7%	-746 (-2754 to +1262)	0.47	8%	-748 (-2175 to +679)	0.30	8%
2 (75%)	713	197	-1293 (-3435 to +850)	0.24	14%	-1399 (-3212 to +414)	0.13	15%	-1442 (-2888 to +4)	0.05	15%

$ effect=Mean estimated incremental effect in USD; % saving=Implied mean saving.

Appendix to §4.5: Summary of propensity scores; overview of multimorbidity sub-samples; results of sensitivity analyses

Appendix 4.5a

For the primary sample (n=969) in this thesis, propensity scores were calculated to balance the treatment and comparison groups on observed confounders (see **§3.2.1>Propensity score methods – applied**). Where the treatment and/or comparison group is redefined - for example, by patient multimorbidity as in **§4.5** - a new sample-specific propensity score was calculated (see **§3.2.1>Additional propensity scores**).

Overviews of the matched samples for the main and secondary analyses in **§4.5** (Table 4.5.3, Table 4.5.4, Table 4.5.5, Table 4.5.6) are provided in Appendix 4.5a/Table 1. Equivalent tables for sensitivity analyses with late-consult outliers re-categorised as controls (Table 4.5.7) and pooling all patients irrespective of discharge status (Table 4.5.8) are given in Appendix 4.5a/Table 2 and Appendix 4.5a/Table 3 respectively.

In all cases balance was checked before matching (step 3, per Table 3.2.2 for the principal sample) and after matching (step 5, per Table 4.1.2 for the principal sample). Following matching, treatment and comparison groups were balanced on all covariates to within the 10% guideline highlighted by Garrido et al. (2014). Where there was loss to matching, no alternative matching strategy (steps 4 and 5, per Table 3.2.3 for the principal sample) reduced this loss to matching without a marked increase in bias between groups.

Appendix 4.5a/Table 1 Summary of propensity score matching for sub-samples in Table 4.5.3, Table 4.5.4, Table 4.5.5, Table 4.5.6

Sub-sample *Discharged alive, PC within 10 days &* _____	**Steps 1 & 2:** Included covariates differ from the 33 used to balance primary sample in Table 3.2.1?	**Step 5:** Patients supported (All/UC/PC) and patients lost to matching (LTM)			
		UC (n=)	**PC (n=)**	**All (n=)**	**LTM (n=)**
2+ Comorbidities	Pain and fatigue variables excluded.	553	221	774	3
3+ Comorbidities	No difference	436	187	623	1
4+ Comorbidities	Advance directive (=1 if living will=1 | proxy=1) replaces living will and proxy. Age re-categorised as continuous. No Elixhauser total (considered appropriate for this sample where distribution of Elixhauser total is very narrow). No pain and fatigue dummy variables.	296	136	432	4
5+ Comorbidities	Advance directive (=1 if living will=1 | proxy=1) replaces living will and proxy. Age re-categorised as continuous. ESAS & CMSAS scores summed from admission and reference day creating one overall variable for each. No pain and fatigue dummy variables.	174	82	256	2

Appendix 4.5a/Table 2 Summary of propensity score matching for sub-samples in Table 4.5.7

Sub-sample *Discharged alive, PC within 10 days, late-consult outliers controls &* _____	**Steps 1 & 2:** Included covariates differ from the 33 used to balance primary sample in Table 3.2.1?	**Step 5:** Patients supported (All/UC/PC) and patients lost to matching (LTM)			
		UC (n=)	**PC (n=)**	**All (n=)**	**LTM (n=)**
2+ Comorbidities	Advance directive (=1 if living will=1 | proxy=1) replaces living will and proxy.	564	221	785	3
3+ Comorbidities	No difference	447	187	634	1
4+ Comorbidities	No difference	306	137	443	3
5+ Comorbidities	Advance directive (=1 if living will=1 | proxy=1) replaces living will and proxy. Age re-categorised as continuous. ESAS & CMSAS scores summed from admission and reference day creating one overall variable for each. No pain and fatigue dummy variables.	182	82	264	2

Appendix 4.5a/Table 3 Summary of propensity score matching for sub-samples in Table 4.5.8

Sub-sample *All patients irrespective of discharge status, PC within 10 days &* _____	**Steps 1 & 2:** Included covariates differ from the 33 used to balance primary sample in Table 3.2.1?	**Step 5:** Patients supported (All/UC/PC) and patients lost to matching (LTM)			
		UC (n=)	**PC (n=)**	**All (n=)**	**LTM (n=)**
2+ Comorbidities	ESAS & CMSAS scores summed from admission and reference day creating one overall variable for each. No pain and fatigue dummy variables.	573	251	824	0
3+ Comorbidities	No pain and fatigue dummy variables.	454	212	666	1
4+ Comorbidities	Advance directive (=1 if living will=1 | proxy=1) replaces living will and proxy. No pain and fatigue dummy variables.	313	162	475	2
5+ Comorbidities	Advance directive (=1 if living will=1 | proxy=1) replaces living will and proxy. Age re-categorised as continuous. No Elixhauser total (considered appropriate for this sample where distribution of Elixhauser total is very narrow). ESAS & CMSAS scores summed from admission and reference day creating one overall variable for each. No pain and fatigue dummy variables.	187	99	286	4

Appendix 4.5b

The primary analysis in **§4.5** stratified the sample according to number of patient comorbidities and these sub-samples were presented in Appendix 4.5a.

The prevalence of specific comorbidities within these sub-samples in Appendix 4.5b/Table 1.

In over 95% of cases in the matched multimorbidity group (739 out of 774 patients), at least one recorded comorbidity is a cancer diagnosis (Elixhauser groups 17-19). However, some cancer diagnoses that made patients eligible for the 'PC4C' study are not recorded on the Elixhauser index, e.g. myeloma. Therefore in the multimorbidity sample there are a small number of patients (n=35) who had multiple non-cancer comorbidities and an advanced cancer, but for whom the cancer diagnosis is not reflected in their Elixhauser score.

In addition to cancer, the most prevalent comorbidities are hypertension and fluid/electrolyte disorders, which affect around half of patients, and cardiac arrhythmia, chronic pulmonary disease, weight loss and depression, which affect a quarter to a third of patients.

Conditions for which there is markedly higher prevalence in samples with higher numbers of comorbidities include congestive heart failure, cardiac arrhythmia, pulmonary circulation, hypertension, chronic pulmonary disease, renal failure, coagulopathy, fluid/electrolyte disorders, weight loss and depression.

Appendix 4.5b/Table 1 Prevalence of specific comorbidities within primary analysis sub-samples

	2+ Comorbidities (n=774)	3+ Comorbidities (n=623)	4+ Comorbidities (n=432)	5+ Comorbidities (n=256)
Congestive Heart Failure	61 (8%)	61 (10%)	58 (13%)	53 (21%)
Cardiac Arrhythmia	121 (16%)	117 (19%)	105 (24%)	83 (32%)
Valvular Disease	30 (4%)	30 (5%)	28 (6%)	24 (9%)
Pulmonary Circulation	41 (5%)	41 (7%)	37 (9%)	32 (13%)
Peripheral Vascular Disorders	29 (4%)	27 (4%)	24 (6%)	20 (8%)
Hypertension Uncomplicated	351 (45%)	321 (52%)	230 (53%)	145 (57%)
Hypertension Complicated	51 (7%)	51 (8%)	50 (12%)	41 (16%)
Paralysis	23 (3%)	21 (3%))	17 (4%)	15 (6%)
Other Neurologic Disorders	47 (6%)	42 (7%)	32 (7%)	22 (9%)
Chronic Pulmonary Disease	154 (20%)	149 (24%)	128 (30%)	94 (37%)
Diabetes Uncomplicated	40 (5%)	38 (6%)	32 (7%)	23 (9%)
Diabetes Complicated	13 (2%)	13 (2%)	13 (3%)	11 (4%)
Hypothyroidism	78 (10%)	75 (12%)	62 (14%)	43 (17%)
Renal Failure	65 (8%)	63 (10%)	59 (14%)	47 (18%)
Liver Disease	32 (4%)	30 (5%)	24 (6%)	20 (8%)
Peptic Ulcer Disease	8 (1%)	8 (1%)	7 (2%)	6 (2%)
AIDS/HIV	3 (<1%)	2 (<1%)	1 (<1%)	1 (<1%)
Lymphoma	54 (7%)	45 (7%)	31 (7%)	19 (7%)
Metastatic Cancer	544 (70%)	450 (72%)	324 (75%)	191 (75%)
Solid Tumour without Metastasis	564 (73%)	466 (75%)	337 (78%)	205 (80%)
Rheumatoid Arthritis	14 (2%)	13 (2%)	12 (3%)	8 (3%)
Coagulopathy	45 (6%)	43 (7%)	40 (9%)	32 (13%)
Obesity	21 (3%)	21 (3%)	20 (5%)	15 (6%)
Weight Loss	177 (23%)	169 (27%)	140 (32%)	89 (35%)
Fluid/Electrolyte Disorders	303 (39%)	285 (46%)	233 (54%)	150 (59%)
Blood Loss Anaemia	4 (1%)	4 (1%)	2 (<1%)	1 (<1%)
Deficiency Anaemia	24 (3%)	22 (4%)	21 (5%)	15 (6%)
Alcohol Abuse	23 (3%)	23 (4%)	19 (4%)	19 (7%)
Drug Abuse	22 (3%)	21 (3%)	15 (3%)	11 (4%)
Psychoses	11 (1%)	11 (2%)	10 (2%)	8 (3%)
Depression	154 (20%)	146 (23%)	118 (27%)	83 (32%)

Appendix 4.5c

The purpose of Appendix 4.5c is to detail sensitivity analyses performed on the primary analysis in **§4.5**. The analyses are intended to check the robustness of the main results in two aspects specifically: to model selection, and to propensity score matching (see **§4.2>Discussion** and Appendix 4.4c for a fuller explanation of the rationale for these sensitivity analyses).

Appendix 4.5c/Table 1 presents results that repeat the Table 4.5.3 analysis with OLS on untransformed costs, with and without weights, and GLM (gamma, log) without weights.

These sensitivity analyses confirm the robustness of the results reported in Table 4.5.3: palliative care consult is associated with larger cost-saving effect for patients with the highest number of comorbidities for a variety of modelling approaches, with and without propensity score weights. As in the primary analysis, the distinction between patients with 4+ and 5+ comorbidities and other sub-samples is clearer than for 2+ and 3+ comorbidities, which return similar results. Proportionally fewer results in Appendix 4.5c/Table 1 return statistically significant results than in Table 4.5.3, possibly due to OLS models' lack of responsiveness to smaller sample sizes.

For both models the unweighted estimates are lower than the equivalent result with weighted regressions, which is expected as weights balance the treatment and comparison groups (*inter alia*) on the higher illness burden of palliative care patients (see Table 4.1.1, Table 4.1.2).

The point estimates for incremental effect vary by model, which is also expected. As in Appendix 4.4c, the OLS on untransformed costs has lower estimates than the GLM but much larger, less efficient confidence intervals.

Therefore I am confident that the incremental-effect estimates produced in the primary analysis are robust to model selection and propensity score use (the strong observable association between timing and estimated effect on cost is persistent in all cases). Additionally I take my primary results to be more meaningful than those presented in Appendix 4.5c/Table 1 because (a) propensity score weights are important to account for confounding between treatment and comparison groups, and (b) GLM (gamma, log) is a more accurate and efficient estimator of incremental effect on cost than OLS approaches.[126]

[126] See also Appendix 4.4c, where the same patterns were observable and the same conclusions drawn for time-to-consult analyses.