A National Evaluation of the Counselling in Primary Care Service (CIPC)

Charles Brand

Supervised by: Dr. Ladislav Timulak

This dissertation is submitted to the University of Dublin in fulfillment of the requirements for the award of a Doctorate in Philosophy, School of Medicine (2020).

Declaration

20/05/2020

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

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

I consent to the examiner retaining a copy of the thesis beyond the examining period, should they so wish (EU GDPR May 2018).

Charles Brand

Wednesday, 20 May 2020.

Executive Summary

Until this study, there were no national data for the effectiveness of counselling or psychotherapy provided in the primary care context in the Republic of Ireland. Clinical service evaluations are crucial for providing empirical evidence for the effectiveness of counselling in primary care settings. While the efficacy of psychological treatments is well established in randomised controlled trials (RCTs), it is also vital that the public and other important stakeholders can be shown that the efficacy of psychotherapy under highly controlled conditions is converted into effectiveness in routine care environments. Evaluations have been conducted in some parts of the United Kingdom, Sweden, Norway, Australia and the United States and contribute greatly to understanding the role of primary care counselling and psychological interventions in improving public health.

The Irish Health Service Executive's (HSE) national Counselling in Primary Care (CIPC) service was fully implemented in July 2013 to serve people experiencing mild to moderate mental difficulties. The aim of this study was to provide the international psychotherapy research community, mental health policy makers, potential service users and other stakeholders with effectiveness data for counselling and psychotherapy delivered through a nationally distributed primary care service in Ireland for the first time.

This study of initial data involved a large sample of 2,781 CIPC clients attending therapy at locations spread across the Republic of Ireland. The main outcome measure was the CORE OM 34 with additional data provided by way of the Health Related Quality of Life scale. Data collection lasted for three months in three HSE areas and 13 months in the remaining six HSE areas where six month follow up data were also collected. Overall effectiveness of the CIPC service was determined using pre/post therapy paired-sample t tests, calculating standardised effect sizes (Cohen's *d*) and examining the proportions of clients achieving reliable and clinically significant change (RCSC). Overall, study participants experienced large reductions in levels of distress. Almost half of all participants (46%) achieved RCSC with a combined reliable improvement rate of 64% (i.e. those showing RCSC plus those showing reliable improvement). A very large standardised effect size of 1.16 (Cohen's *d*) was also observed (n = 1,397).

The study also examined the effect of the duration of therapy on client outcomes using binary logistic regression in order to determine the likelihood of clients achieving RCSC based on attending more or less therapy sessions and their pre therapy CORE OM scores. The vast majority of participants (88.5%) attended 10 or less sessions and results showed that clients attending fewer sessions had a slightly better chance of achieving RCSC than those attending a higher overall total number of therapy sessions. Further, clients with lower levels of distress at the assessment stage were more likely to achieve RCSC.

The role of the initial severity of clients problems and duration of therapy was examined in terms of their effect on longer term outcomes. A multi-level model was generated and additional CORE OM data were examined from 243 study participants six months after therapy ended. The model indicated a significant effect of time on CORE OM scores between pre therapy, post therapy and 6 months after therapy had ended. The overall model indicated that both pre therapy assessment and the overall number of sessions attended were predictive of CORE OM outcomes. Participants with whom higher numbers of sessions were associated and displayed slightly more severe symptoms at pre therapy appeared to have marginally better outcomes six months after therapy had ended. Sessional data are required in future studies to better understand the role of these factors in client outcomes.

There were no differences in outcomes for clients who received a single therapy approach and those receiving more than one approach and overall no one single approach performed statistically significantly better than others in terms of outcomes.

The study also examined outcomes from the Health Related Quality of Life questionnaire (HRQOL). This measure is used extensively in the United States by the Centre for Disease Control (CDC) as both an outcome measure and population health tracker. Paired t tests showed that between pre and post therapy there was a significant increase in the overall number of days clients felt their general health was excellent and very good, along with significant decreases in the number of days their general health was fair or poor. Significant decreases were also reported by clients in the number of physically and mentally unhealthy days they experienced, as well as the number of days their activity was limited due to the problem with which they had attended the CIPC service. Further, correlational analyses showed a good level of congruence between the outcomes of both the HRQOL scale and the CORE OM and HRQOL items were significant predictors of pre therapy CORE OM levels. Hence, attendance at therapy provided by CIPC had a the effect of improving study participants' perceptions of both their general and psychological health. Results showed they were spending less time suffering the day to day negative consequences of the problem from which they sought relief.

Results of the study also highlighted some important characteristics of the overall national profile of the CIPC service. First, they showed that the CIPC service provided counselling and psychological therapies that were tailored to the needs of each individual client in terms of the different types of approaches currently available and length of the therapy duration. Second, outcome effect sizes and the proportions of clients achieving RCSC, reliable change or deteriorating were on a par with comparable services in other jurisdictions, e.g. the Increasing Access to Psychological Therapies (IAPT) service in the UK, primary care services in Sweden, Norway and the U.S.A.. Also, according to the CORE OM, the CIPC service treated a relatively high proportion of clients presenting in the severe range of distress at the pre therapy stage compared to other services.

The study highlighted several service related areas where improvements could be made. Specifically in the area of counseller/therapist training in the use of the CORENET centralised client information database. This system was purchased from and is developed by a company from the UK and this is reflected in certain elements of the language and clinical emphasis implemented in the system. Updating some aspects of the CORENET system may better reflect the Irish counselling and psychotherapy experience and improve metrics supplied by the system. Further, the HRQOL scale showed some promise as a reliable screening tool to provide pre-assessment information regarding the status of potential clients of the CIPC service.

As an initial study and the first of its kind anywhere in terms of its complete national coverage, this practice based study provides an empirical basis for the effectiveness of the CIPC service in the Republic of Ireland. Further, it also provides a set of benchmark outcomes against which other large, publicy available primary care counselling and psychotherapy services internationally can compare.

Acknowledgements

First and foremost I would like to thank Dr. Ladislav Timulak for his unrelenting support and guidance over the last four years. Were it not for your genuine curiosity and openness to listen, this study would not have been realised in the manner which it was. For helping me pull together the vastness of the psychotherapy research literature and imbuing in me a thirst for a greater understanding of it for the benefit of clients, I sincerely thank you. My gratitude at TCD also extends to Dr. Frédérique Vallieres and Dr. David Hevey who also helped me keep focus on what I was trying to accomplish and how to achieve it.

The endeavour to transform what could have been a relatively small, localised government report into a broad and national empirical research study was made possible by the support of my employment mentor Ms. Fiona Ward. Your encyclopaedic knowledge of primary care mental health services and policy along with a shared understanding of the importance of an empirical basis for claims of effective psychotherapy in CIPC were fundamental to the success of both the Irish Research Council funding application and the overall shape of the study. You supported me in my pursuit to attain as much learning as possible in what was our shot at achieving something not before achieved.

There were a vast number of people within the CIPC service to whom I owe a huge amount of gratitude: the co-researchers in this study, i.e. over 153 CIPC counsellors and therapists who invited over 5,000 clients and successfully recruited 2,781 participants to be part of this study. Over 60 administrative staff and area Coordinators who incorporated new and difficult tasks into their everyday workflow in order to accommodate many research-only related outputs. The directors of CIPC and the NCS service who were open to the idea of completing this study and listened with open hearts and minds to the benefits this research could bring to future clients of the service. Also, the members of the CIPC Research Group who provided additional guidance and support as this study was built from the ground up.

The personal, constant and unwavering support of my family were the main reasons I was able to complete the research. From my mum Helen and step dad Richard, to my aunt Chris and uncle Colm, dad in-law Ned and mum in-law Helen, my sister in-law Annest and step-sisters Sally and Jessie who all took shifts to pitch in at various times while I was “in the cave” and offer practical support, kind words and encouragement.

Finally, I thank my children Leila and Donovan, who in their own way supported me unconditionally, and my wife Aja. Not only were you my partner in life for the last 18 years, but also my main teacher in matters academic. You are instrumental in helping me to understand my place in the world in the context of my learning. You took over the reins of parenting for large periods of time when I could not cope with the pressures of work, writing and the (highly organised) chaos of our young family. You are a monolith of love, knowledge, support, clarity and strength without whom I would never have embarked on my quest to leave the cave and understand the shadows. You taught me to believe that I had this in me, and for that I am forever grateful. We did it!

Table of Contents

Executive Summary **1**

Table of Tables 8
Table of Figures 9

A National Evaluation of the Counselling in Primary Care service (CIPC) **10**

Chapter 1: Introduction **10**

1.1 Background to the study 10
1.1.1 The origins of CIPC 11
1.2 CIPC – a service profile 0
1.2.1 Rationale for the national evaluation 2

Chapter 2: Literature Review **4**

2.1 What are the important factors influencing psychotherapy outcomes? 4
2.2 Assessing psychotherapy outcomes 5
2.3 Research into counselling in primary care - efficacy, pragmatism and effectiveness. . 7
2.3.1 Randomised control trials in psychotherapy research 9
2.3.2 Meta-analyses of psychotherapy research 13
2.3.3 A Pragmatic approach in primary care psychotherapy research 21
2.4 Practice based research – an overview 25
2.4.1 Practice based evidence gathering - a conceptual framework 26
2.4.2 Evaluation Research – the practice based connection 29
2.5 Practice based research literature – closing in on the important factors 31
2.6 IAPT – CIPC's close relative 38
2.7 Issues important for service delivery 43
2.7.1 Overall effectiveness 43
2.7.2 Number of therapy sessions to provide 44
2.7.3 Type of therapeutic approach 52
2.7.4 Health related quality of life 56
2.7.5 Initial symptom severity 57
2.7.6 Employment status and gender 62
2.7.7 Do clients maintain the gains made in therapy? 64
2.8 Conclusion 67

2.9 Research Aims 69
2.10 Research Objectives 69
2.11 Research Questions 70
2.11.1 What are the outcomes of short term counselling conducted in primary care in Ireland:
2.11.1.1 In terms of symptom reduction? 70
2.11.1.2 In terms of treatment duration? 70
2.11.2 Does the initial severity of participants' symptoms predict CORE OM outcomes? 70
2.11.3 Are outcomes sustained six months after therapy has ended? 70
2.11.4 Are there differences in outcomes between therapeutic approaches? 70
2.11.5 What is correlation between CORE OM and Health Related Quality of Life? 70
2.11.6 Does employment status or gender have an effect on client outcomes? 70

Chapter 3: Method 71

3.1 Overview 71
3.2 Philosophical and theoretical foundations 71
3.3 Epistemological position 71
3.4 Measurement of mental attributes 73
3.5 Study Design 74
3.5.1 Background 75
3.5.2 Design rationale 75
3.5.3 Ethical considerations 75
3.5.4 Informed consent 76
3.5.5 Sampling technique 76
3.6 Instrumentation 81
3.6.1 CORE Therapy Assessment form (TAF) 81
3.6.2 CORE End of Therapy form (EOT) 82
3.6.3 The CORE OM 34 83
3.6.4 The Health Related Quality of Life Scale 83
3.7 Data analysis approach 85
3.7.1 Missing data imputation 85
3.8 Pre post outcomes, RCSC and effect size analysis approach 86
3.8.1 Reliable and clinically significant change (RCSC) 89
3.9 Describing the relationship between the number of sessions attended and client outcomes. 93
3.10 Examining the role of Initial severity 94

3.11 Examining the number of sessions and initial severity affect outcomes 6 months after therapy has ended 95
3.11.1 Multi-Level Model Specification 97
3.11.2 Model generation approach 98
3.12 Examining differences in outcomes between and therapists' self-declared therapeutic approaches 100
3.13 Examining the relationship between changes in CORE OM scores and changes in health related quality of life (HRQOL) 101
3.14 Examining employment status or gender in relation to client outcomes 103

4 Chapter 4: Results 104

4.1 Descriptive statistics 104
4.1.1 Presenting problems 107
4.1.2 CORE OM reliability result 109
4.1.3 Missing data handling 109
4.1.4 Attrition 109
4.1.5 Questionnaire completion rates 111
4.2 Pre post outcomes - statistical analyses, RCSC and effect size results 111
4.2.1 Calculating the Reliable Change Index 112
4.3 Pre post therapy results - What are the outcomes in terms of symptom reduction? 113
4.4 What are the outcomes in terms of treatment duration? 114
4.5 Does the initial severity of participants' symptoms predict CORE OM outcomes? .. 118
4.6 Are outcomes sustained six months after therapy has ended? 120
4.6.1 Preliminary exploration of the data 120
4.6.2 Main MLM results 125
4.6.3 Mean CORE OM scores across pre, post and six month follow up points 128
4.7 Are there differences in outcomes between therapeutic approaches? 130
4.8 What is correlation between CORE OM and Health Related Quality of Life? 136
4.9 Does gender or employment status have an effect on client outcomes? 143

5 Chapter 5: Discussion 145

5.1 Introduction 145
5.2 The CIPC Service – client demographic and clinical profile 145
5.3 Pre and post therapy outcomes – overall effectiveness and comparison with other services. 150

5.4 Number of sessions - overall treatment duration..... 156
5.5 Initial severity of clients' problems and longer term effectiveness..... 158
5.6 Does the number of sessions and initial severity affect outcomes 6 months after therapy has ended? 161
5.7 Are there differences in outcomes between and therapist Self-Declared Therapeutic Approaches?..... 165
5.8 HRQOL (Health Related Quality of Life) – was it useful? 171
5.9 Does employment status or gender have an effect on client outcomes? 175
5.10 Limitations..... 175
5.11 Recommendations and future research 177
5.12 Conclusion..... 179

References..... 183

Appendices..... 222

Table of Tables

Table 1: CIPC CHO regions and areas covered 0
Table 2: Demographic information for All Participants and Form Returners sample 105
Table 3: Presenting problems - All Participants sample 108
Table 4: Presenting problems - Form Returners sample 109
Table 5: Types of therapy endings as indicated by counsellor/therapists 110
Table 6: Completion rates for CORE and HRQOL questionnaires 111
Table 7: Pre post outcomes and effect sizes for Form Returner sample 113
Table 8: Reliable and clinically significant improvement rates as a function of total number of sessions attended. 115
Table 9: Logistic regression table showing session number as predictor of RCSC 118
Table 10: Regression table for pre therapy severity depression and anxiety levels predicting CORE OM score 120
Table 11: Mean therapy scores at pre, post and six months after therapy (n = 240) 121
Table 12: Multi level regression table estimating association between total number of sessions attended, pre therapy severity levels, CORE OM scores at pre therapy, post therapy and at six month follow up 125
Table 13: CORE OM scores at pre post and six months after therapy by base line severity levels 130
Table 14: Breakdown of therapy approaches 131
Table 15: Breakdown of therapy approaches by frequency. 132
Table 16: Pre post effect sizes for single versus more than one therapeutic approach. 133
Table 17: Breakdown of rates of RCSI for single versus more than one therapeutic approach. 136
Table 18: Pre post therapy scores for HRQOL items 2,3 and 4. 139
Table 19: Correlations, mean score and standard deviations for CORE OM and HRQOL items 2,3,and 4. 139
Table 20: Logistic regression results for HRQOL items predicting CORE OM score. 140
Table 21: Correlations, means and standard deviations for pre post therapy score differences for CORE OM and HRQOL items 2,3,and 4. 141
Table 22: Logistic regression results for HRQOL items 2,3 and 4, sex, medication status and unemployment status predicting RCSC achievement. 142
Table 23: Comparison of effects sizes and RCSC rates between the current study and the main practice based studies. 152
Table 24: Comparison of the current study single versus multiple approaches with Stiles et al. (2008). 167
Table 25: Table showing corrected effect size calculations for Stiles et al., (2008). 168

Table of Figures

Figure 1: Methodological framework - the five stage model showing the processes employed to employ practice based research. Adapted from Barkham, Stiles, Lambert and Mellor-Clark (2010). 27
Figure 2: Study participation flow diagram showing sample groups and criteria used to exclude participants from each successive sample. 79
Figure 3: Graph showing reliable and significant improvement rates for all participants above clinical cut off level as a function of the total number of sessions attended (n = 1,087). 117
Figure 4: Mean CORE OM responses at pre therapy, post therapy, 6 month follow up and 12 month follow up (n = 240). 121
Figure 5: Plot diagram showing linear fit lines for 17 study participants at pre, post and six months after therapy. 122
Figure 6: Plot diagram showing quadratic fit line for 17 study participants of pre post and six month 123
Figure 7: Box-plots showing CORE OM pre, post and six month follow up as a function of three severity levels; mild (< 1), mildly, moderately, or moderate-to-severely distressed (>1 to <2.5), and severe (> 2.5) (n = 240). 129
Figure 8: Pre post therapy mean outcomes for single versus more than one therapy approach 133
Figure 9: Figure showing interaction of single versus more than one therapy approach on CORE OM scores 135
Figure 10. Pre post change in HRQOL Item 1 - General Health 137
Figure 11. Pre post change in HRQOL Items 2, 3 and 4 - Physically unhealthy days, mentally unhealthy days and activity limitation days. 138
Figure 12: Mean Pre, post and six month follow up outcomes in the current study (n = 240) 162
Figure 13: Baker et al. (2002) - Severity of depression - Outcomes at pre post and 24 month follow up 162
Figure 14: Ray-Sannerud et al (2012) Global Mental Health (GMH) scores at assessment, post therapy and two years after therapy 163

A National Evaluation of the Counselling in Primary Care service (CIPC)

Chapter 1: Introduction

1.1 Background to the study

This study emerged from the need to empirically validate the effectiveness of the therapeutic services provided by the Counselling in Primary Service (CIPC). The researcher was initially hired by CIPC to perform a short, small scale service evaluation of clinical outcomes. Subsequently, Irish Research Council (IRC) Employment Based PhD Scholarship funding was applied for and secured by the researcher in order to perform a full evaluation on a national scale in conjunction with the CIPC service (under the auspices of the National Counselling Service) and The School of Psychology, Trinity College Dublin. That study is described here and was nested within the broader, longer term evaluation and conducted alongside the researcher's normal contractual duties. What follows is an examination and report of initial data, the primary objective being to determine the usefulness of the service. A practice-based research approach was consonant with CIPC service research policies, another rationale for conducting this study in the first instance.

Butler (2002) states when outlining the purpose of evaluation in the Irish health sector "Evaluation has a vital role to play in enabling health service planners and managers to attain the highest standards of effectiveness, efficiency, equity, quality and value for money in the services that they provide, and to demonstrate that attainment for accountability purposes." (2002, p. 6). Bearing this in mind the overarching goal of this study is two-fold: 1) to provide a set of clinical benchmark data on the first national Irish primary care counselling service and

2), to contribute to international scholarly literature regarding the behaviour of variables which are of importance in large scale publicly funded psychological therapy services in other jurisdictions.

1.1.1 *The origins of CIPC*

By 2001 a significant body of international evidence for the benefits of an integrated approach to the provision of psychological and counselling services within the context of primary care had been established (Blount, 2003; Bor & McCann, 1999; Harvey et al., 1998; Hemmings, 2000; Rowland et al., 2000). This facilitated the Irish government in making a commitment to the development of primary care services "so that they become the cornerstone of care and preventive services for communities across the country" as it was recognised that this was "consistent with best international practice." (Department of Health & Children, 2006, p. 17). What followed was the publication of the mental health policy framework document - Vision for Change (2006), which stated that "The popularity of these [talk therapies] approaches and the evidence for their effectiveness has been growing in recent years. The emerging consensus is that they should be regarded as fundamental to basic mental health services" (Department of Health, 2006, p. 246) and concluded that:

> *"Counselling is often the preferred term for a brief intervention with an individual who presents with a circumscribed difficulty, precipitated by some specific challenging event in their life, such as a career crisis, a sudden bereavement or a medical illness."*(Department of Health, 2006, p. 246).

Further, public demand within Ireland for psychological approaches to address emotional and psychological difficulties was coupled with a growing dissatisfaction for the hitherto medication-only approach to such problems prevalent at the time (Batt et al., 2002). The general public perceived psychological therapy services positively, but also reported high levels of

dissatisfaction with the lack of availability, access and choice of services available (Batt et al., 2002; HSRC, 2003).

Of the 11 regional health boards formed in accordance with the 1970 Health Act (Health Act, 1971), some functional areas Ireland had initiated psychological and counselling service provision at a primary care level. The National Counselling Service (NCS) in the Dublin North East (DNE) region conducted a survey of GPs in 2003 to investigate their views on the level of need for counselling, referral practices and their opinions regarding current models of counselling service provision. Almost half of GPs in the region complained of limited referral options due to barriers to accessing counselling. They also cited a lack of available public services, long waiting lists and prohibitive cost (Ward, 2012) which prevented them from referring to secondary services and that there was an urgent need for counselling services at a primary care level to meet many of their patients' psychological difficulties. There was limited counselling available in the DNE region, but what was available was not accessible for many patients due to cost and location.

This service developed a counselling service in primary care called Primary Care Counselling Provision (PCCP), which was initially rolled out on a pilot basis and provided time limited counselling for clients presenting with non-complex psychological difficulties. Between 2007 and 2011 this service received in excess of 5,000 referrals from more than 50 GP practices and several primary care units across the DNE region (Ward, 2012). Similar small local initiatives were reported by Martin, Hawkins, Hicks, & O'Flynn (2008) regarding three services across Dublin in operation between 2005 and 2008 in the North Inner City of Dublin, South Inner City of Dublin and Ballymun, North Wicklow and in the aforementioned Dublin North East region. Finally, Bourke and Byrne (2012) also produced an evaluation of a stepped care model of service delivered in Co. Roscommon which incorporated counselling and psychotherapy. Adopted by

the Improving Access to Psychological Therapies (IAPT) initiative in the UK, this model and its relationship to the CIPC service is fully described later in this review.

Outcomes from these small scale studies were positive and included a report that 89% of clients in the Dublin North Inner City study "…showed a decrease in surgery attendances since completion of therapy…" (Martin et al., 2008, p. 12). Clients' satisfaction levels with the service provided were as high as 97% in terms of their counsellor's ability to listen to and understand their problems and 90% of clients believed that counselling had been effective in addressing their difficulties (Ward, 2012). Bourke and Byrne (2012) also indicated reliable improvement for 75% of those who took part in the stepped care service provision model study in 2012.

Thus, the rationale for the need for greater access to a national service to provide short-term counselling which focussed on people with mild to moderate psychological problems had been established. On foot of these studies, in particular that by Ward (2012), international research findings supporting the positive impact of counselling delivered in primary care contexts and a growing service user movement advocating its provision, increasing access to counselling in primary care was set out as an objective in the 2011 Programme for Government .

1.2 CIPC – a service profile

During the following number of years and in conjunction with the incumbent national healthcare provider the Health Service Executive (HSE), the Irish government commissioned, implemented and rolled out the primary care infrastructure required to deliver evidence-based therapies to the population. This infrastructure comprises the national network of CIPC counsellors, coordinators, directors, support and administration staff and counselling locations involved in hosting and delivering the CIPC service. Part of the rationale for the introduction of the service was a commitment to the adoption of an evidence-based approach and the continuous evaluation of its effectiveness. Further to this commitment was an explicit requirement for service policy improvement based on the provision of clinical outcome

evidence as a result of service evaluation activities. The HSE provided for the roll out of a national CIPC service in its 2012 service plan which was subsequently launched in June 2013. From its initial roll out, the CIPC service was a nation-wide service distributed over 10 Community Health Organisations (CHOs) See Table 1 for details of CHO areas. At the time of writing CIPC employed over 221 counsellor/therapists delivering therapy in 245 locations across 26 counties in the Republic of Ireland. The service is overseen by a board of national directors and coordinated by CHO area service coordinators who administrate over the affairs of all counsellor/therapists (CTs) delivering therapeutic services in their respective CHOs.

Table 1

CIPC CHO regions and areas covered in the study

HSE REGION	AREA COVERED
HSE CHO Area 1	Donegal, Sligo & Leitrim
HSE CHO Area 2	Galway, Mayo & Roscommon
HSE CHO Area 3	Limerick, Clare & North Tipperary
HSE CHO Area 4	Cork & Kerry
HSE CHO Area 6	South Dublin, South East Dublin & East Wicklow
HSE CHO Area 7	South West Dublin, Kildare & West Wicklow
HSE CHO Area 8	Laois, Offaly, Longford & Westmeath
HSE CHO Area 1/8	Meath, Louth, Cavan & Monaghan
HSE CHO Area 9	Dublin North & Dublin North City

Important to note at this stage is that in the UK “counselling” can sometimes have a slightly different connotation from a research and service perspective. Counselling (as a therapy) has been discussed as situated within the context of (but distinct from) other psychological therapeutic approaches (see Eatock, 2000). Moreover, in some of the meta analyses examined later in this literature review, the term “non-directive supportive counselling” in a primary care setting is used to denote “…non-directive supportive therapy [which] is an unstructured therapy ...” (Pybis, Saxon, Hill, & Barkham, 2017). This is because, in some primary studies counselling has sometimes presented as a therapy approach as opposed to forms of therapy such as CBT. Notably, CBT has been compared in terms of outcomes with Person Centred Therapy (PCT) (Gibbard & Hanley, 2008), or in pre post studies as a form of non-specific

generic counselling (Mellor-Clark et al., 2001). CBT has sometimes been included in descriptions of therapies providing "integrative psychological counselling" (Davis et al., 2008, p. 81). Hence, in terms of the studies examining the effectiveness of therapy in primary care in the UK, a heterogeneity exists in descriptions of the types of therapeutic approaches being examined, whereby in different studies counselling can denote "unstructured" approaches, be they stand-alone therapeutic approaches, or be used to describe a collection of approaches involved in a study which include "structured" and "unstructured" approaches. In terms of service provision; all IAPT therapy providers are required to attend IAPT endorsed training with an emphasis on CBT (Clark, 2018b).

In Ireland, The Guidance Document on the Provision of Counselling in a Primary Care Setting (Health Service Executive & Irish College of General Practitioners, 2006) which underpinned CIPC's approach to service provision recognised "…the regular confusion of what counselling is, what it is not and how it differs from non-specific approaches which are often appropriately provided by health care professionals who do not necessarily have recognised qualifications in the field." (2006, p. 2). There was acknowledgement that many health professionals provide counselling support to patients across a wide range of disciplines and under various circumstances, However, a delineation was made between those who would be eligible to provide services for CIPC and those who would not. As the document goes on to state "…counselling/clinical psychologists, psychotherapists and counsellors with professional accreditation from recognised professional bodies would be eligible to provide the required level of counselling services within a Primary Care setting." (2006, p. 2). Further, counselling is described as "process", which can involve various different approaches including "psychodynamic, cognitive behavioural humanistic and integrative and systemic and family therapy" (2006, p. 3). This clarified the understanding that the CIPC service considered the qualifications of a potential CIPC therapist as encompassing a broad range of approaches. While individual therapists within the CIPC service may specialise in one (or more) approach

over others, the document is clear that the preferred approach for therapists working for the CIPC service is CBT. However, specific training in CBT or any particular therapeutic approach was not a requirement to be considered for employment with CIPC, once the completion of courses recognised for accreditation as described above were met.

The CIPC service as a whole considers itself a provider of a broad range of psychological therapies rather than the provider of a specific catalogue of CBT based therapy approaches that could be matched with specific presenting problems, as is more the case in the UK, e.g. IAPT. Hence, the basis for the examination of effectiveness in this study is from the perspective of an Irish national service providing a broad range of psychological therapies in order to reduce levels of distress and other symptoms from a range of presenting problems.

1.2.1 *Rationale for the national evaluation*

According to Kusek and Rist (2004), a significant shift occurred in the expectations of an ever increasing array of stakeholders that changed the nature of monitoring and evaluating government provided services internationally. Reporting of routine organisational activities and outputs only up to that time, had been supplanted by a demand for reporting *actual* outcomes (Kusek & Rist, 2004). Subsequently, accountability and service improvement expectations now required answers to more detailed and relevant questions such as whether or not programs and services had achieved the results they predicted, whether they had been executed in accordance with the objectives described during their planning phases, if they included detailed descriptions of problems and resulting rectifications of problems identified during the operational lifetime of a service or whether measures and indicators of success or failure had been readily identifiable (Kusek & Rist, 2004). For this increasing number of stakeholders and important funding decision makers these expectations, specifically in relation to the evaluation of the delivery of counselling and psychotherapy services, emerged as important at this time in

Australia and New Zealand (Jane Pirkis et al., 2005), the U.S.A. (Goodman et al., 2003) and the U.K. (Clark et al., 2009).

It was against this backdrop of transformative change in the approach to evaluating the impact of government programmes internationally, that a working group was set up by the Irish government specifically tasked with the provision of a set of guidelines for the formation of a national counselling service. This group published its recommendations in 2006 which described the future of primary care service delivery with a focus on increasing its commitment to accountability and quality improvement. Contained within the *Guidance Document on the Provision of Counselling in a Primary Care Setting* (Health Service Executive & Irish College of General Practitioners, 2006) were specific references and descriptions of the processes and tools which could be utilised in a process of continuous service evaluation. Further, they highlighted the fact that "Research Evidence is relatively scarce in the Counselling Community in Ireland and it is therefore recommended that evaluation is built into the counselling provision from the outset so as to address this deficit" (Health Service Executive & Irish College of General Practitioners, 2006, p. 10). Hence, from its inception the CIPC service incorporated a continuous evaluation based approach to monitoring outcomes and established a practice based approach to research into the service it provided.

Chapter 2: Literature Review

The first part of the following literature review provides the historical background to the CIPC service, the context from which it emerged and the political and societal rationale for its conception. Next, is an overview of factors emergent in the psychotherapy research literature as consequential for clients as they engage with large scale psychotherapy services such as CIPC. The review also traces the lines of psychotherapy effectiveness research from RCTs, efficacy research to pragmatic studies and meta-analytic approaches and finally to practice based evidence which is the primary research apparatus utilised by this study. As another large nationally distributed and publicly available psychotherapy service geographically closely situated to Ireland, the Increasing Access to Psychological Therapies (IAPT) service in the U.K. will be examined in detail and summarily compared with CIPC before a closer review of the specific variables pertinent to this study are discussed. The review concludes with a specification of the research aims, objectives and questions.

2.1 What are the important factors influencing psychotherapy outcomes?

While the efficacy of therapy for depression and anxiety is no longer disputed, the RCT as a research design approach has been criticised for failing to accurately reflect conditions in actual clinical practice (Leichsenring, 2004; Shean, 2014). Thus, for clinicians working in real world services such as CIPC the generalisability and usefulness of the results from these types of studies, in the absence of additional types of evidence is questionable (Barkham, Hardy, & Mellor-Clark, 2010; Hemmings, 2000; Margison, 2010). Hence there is a growing consensus that the results from RCT and pragmatic style RCTs should be appraised alongside studies that focus on evaluating psychotherapy effectiveness from real world clinical context in order to enhance the overall psychotherapy outcome evidence base (Ammerman, Smith, & Calancie, 2014; Barkham, Stiles, Lambert, & Mellor-Clark, 2010; Castonguay, Barkham, Lutz, & McAleavey, 2013; Leichsenring, 2004; Mellor-Clark, 2000). This sentiment is echoed by Green

(2008) who states “If we want more evidence-based practice, we need more practice-based evidence” (Green, 2008, p. i25).

CIPC is a real world psychotherapy service and this study was conducted under real world conditions. Hence, particular constraints existed while attempting to complete the process of practice based evidence research, as they do in all research endeavours of this type. For example, the varying levels of agreeableness of counsellor/therapists to engage with the research endeavour are influential (see Unsworth et al., 2012), also the decision of research participants to leave the service at the stage at which they feel they no longer need to attend the service (see Bower et al., 2009; Connell, Grant, & Mullin, 2006), thus ending their research participation and effecting levels of statistical power during the analyses. Moreover, it has been noted that the output of practice-based research should be both relevant enough to be of practical use to the particular service under examination (in this case CIPC), but also relevant from the perspective of furthering our understanding of mass delivered psychological interventions in general (Barkham et al., 2010; Castonguay et al., 2013). Hence, this study sought to answer particular questions, as well as address the broader needs of the psychotherapy research community.

2.2 Assessing psychotherapy outcomes

In 1952 the British psychologist H.J. Eysenck’s paper entitled “The Effects of Psychotherapy: An Evaluation” (Eysenck, 1952) was published. In it he rejected the previously reported level of benefit resulting from psychodynamic and eclectic psychotherapy over treatment as usual delivered by GPs or in acute psychiatric settings. In fact according to Eysenck, the more psychotherapy received by the patients in his review, the less they demonstrated recovery. This report has been widely cited as being a critical juncture in the area of psychotherapy research in terms of its disruptive effect on the field (see Bergin, 1971; Horvath, 1988, 1988; Wampold, 2013). Illustrating the vast change which occurred since that time, 61 years later in

2013 the American Psychological Association Council of Representatives overwhelmingly voted to adopt as APA policy a resolution document entitled “Recognition of Psychotherapy Effectiveness” in which it proclaimed its support for the effectiveness of psychotherapy and stated that “…[psychotherapy] should be included in the [U.S.] health care system as an established evidence-based practice” (American Psychological Association, 2013, p. 324).

Subsequent to the claim by Eysenck in his original 1952 paper and prior to the comprehensive endorsement for the benefits of psychotherapy by the APA (2013) a period of unprecedented psychotherapy research activity had occurred (Horvath, 2013; Lambert, 2013; Wampold & Imel, 2015). Much of the early response to Eysenck’s claims sought to challenge the methods adopted by him and how they had led to the outcomes he reported (see De Charms, Levy, & Wertheimer, 1954). There were problems associated with this type of heuristic and qualitative review including a lack of a demonstrable and systemised approach for study selection and exclusion which, it was argued, had resulted in a process based on reviewers’ predisposed positions on the effectiveness of psychotherapy (Wampold, Ahn, & Kim, 2000), and was driven by judgements which appeared to be “…biased or, at least, arbitrary.” (Wampold & Imel, 2015, p. 87). Further, the comparison of treatment groups with non-equivalent control groups was commonplace and evident in Eysenck’s (1952) paper. What the subsequent furore following Eysenck’s report precipitated was a recognition that psychotherapy outcome research was, at that point, lacking a robust and rigorous approach (Wampold, 2013). This motivated psychotherapy researchers to re-evaluate the methods employed to demonstrate the effects of psychotherapy. Hence, there began a movement, albeit a disjointed one (see Moras, 2002), toward the accumulation of empirical data obtained using RCT methodology as opposed to reviews alone (Comer & Kendall, 2013; Wampold et al., 2000).

2.3 Research into counselling in primary care - efficacy, pragmatism and effectiveness.

In the field of medicine, the randomised control trial (RCT) has been described as the gold standard in terms of the level of evidence and the ability to generalise results to particular patient populations (Kabisch et al., 2011). However, it has also been central in defining the understanding of "what works" in the field of psychotherapy and thus providing an evidence base for the efficacy of certain psychological interventions over others (Wampold & Imel, 2015). Further, RCTs have been favoured by both policy decision makers and clinical practitioners seeking clarity in their respective decision-making processes (Petticrew et al., 2012), and these decisions regularly involve considering the reported efficacy of some particular treatments over others, or influence the rationale for making health policy decisions at a government level (Haynes et al., 2012). Hence, RCT results are important from the perspectives of policy and primary research as they can impact on the thoughts of a broad range of stakeholders, from governments to other related policy influencers.

Research designs in medicine adhering to an RCT approach are characterised by three main features: 1) the inclusion of a control group of participants, i.e. participants who receive no treatment, an alternative treatment or some form of "usual" care, 2) the random allocation of study participants to either a treatment group or to a control group (Shean, 2014), and 3) the concealment from the deliverer of treatment whether or not they administered an active treatment rather than a placebo, and/or concealment from the study participants as to whether or not they have been administered the treatment rather than a placebo (Kabisch et al., 2011). This concealment process is better known as blinding and RCTs can be designed as either single blind or double blinded studies. Additional elements may also be part of an RCT research design such as the selection of a strictly defined study population which adheres to predefined set of requirements set out in the study protocol. Also, an intention to treat (ITT) statistical analysis approach may be adopted whereby data belonging to all randomised study

participants are subject to analyses, regardless of whether or not they complete the course of treatment, drop out of the study or deviate from the treatment protocol (Gupta, 2011).

The main aim of an RCT design in medical trials is to determine the degree to which a type of intervention is beneficial under ideal conditions (Scott & Sensky, 2003) and this approach has been adopted in trials examining the efficacy of psychotherapies (Green & Latchford, 2012). Therefore, as many variables as possible thought to effect, or be related to the outcome of interest are as tightly controlled as is practical in order to test hypotheses (Hemmings, 1997; Scott & Sensky, 2003). For example, in order to achieve homogeneity of the final sample, researchers will endeavour to include only participants representative of the target patient population in terms of demographic characteristics and current levels of severity of the target disease during the recruitment phase. In terms of the resulting sample, this should be of an adequate size as to accommodate the planned statistical analytical approach. Finally, the intervention or treatment must be delivered to both control and intervention groups in exactly the same way in terms of dose and preparation in order to deliver a standardised experience to both groups in a study (Watson et al., 2004).

These characteristics and aims of the RCT are all employed in order to minimise bias at all stages of the process and attain the highest level of internal validity, i.e. to minimise or control for the degree to which confounding variables have influenced the outcome results. The purpose of applying these processes is ultimately to strengthen the argument that the cause of an observable change in the outcomes of an experimental group in an RCT is as a result of the effect of the intervention under examination (Shadish et al., 2002). In the case of all RCT trials seeking to compare new and/or existing therapies with usual care or other controls, all of these constraints are designed to prioritise internal over external validity. Further, they are considered necessary in order to determine what is termed the “efficacy” of various treatment

approaches, i.e. the effect of the intervention compared to usual treatment or another comparable treatment under ideal conditions.

2.3.1 *Randomised control trials in psychotherapy research*

A widely cited illustrative example of an RCT research approach which examined the efficacy of treatments for depression was a study conducted by Irine Elkin et al. (1989). A multi-site clinical trial for the National Institute of Mental Health Treatment of Depression Collaborative Research Program (NIMH TDCRP) this was one of the largest, most expensive multi-site clinical RCTs of the effectiveness of psychotherapy conducted up to that time (Wampold & Imel, 2015) and set the standard for hundreds of psychotherapy RCTs which followed. This study is reviewed in detail as it shares features of the methodological employed here, i.e. formation of end point and completer sub-samples for post hoc analysis.

Using a "… randomised, controlled clinical trial research strategy…" (Moras, 2002, p. 529), this study compared the effectiveness of 4 different therapeutic approaches for the treatment of depression (described below). A sample of self-referrals, referrals from psychiatric outpatient and from other mental health facilities comprising 560 patients were selected after a pre-screening process which excluded patients who were "clearly ineligible" to take part in the trial (Elkin et al., 1989, p. 972). Two hundred and fifty patients remained following an interview process by a clinical evaluator to assess suitability for inclusion in the study using the Schedule for Affective Disorders and Schizophrenia interview (Endicott & Spitzer, 1978). Of the remaining 250 patients 239 were randomised across three different physical sites to one of four treatment conditions for a duration of up to 16 weeks: interpersonal psychotherapy (IPT), cognitive behaviour therapy (CBT), imipramine hydrochloride plus clinical management (IMI-CM) or placebo plus clinical management (PLA-CM).

Both psychotherapies were administered by experienced therapists using manualised therapy approaches for IPT (see Klerman, Weissman, Rounsaville, & Chevron, 1984) and CBT (Beck,

1979) and to whom training in their specific treatment approaches was provided. Adherence to their respective treatment approaches was monitored through examination of audio tapes of therapy sessions. Patients were assessed at several time points i.e. baseline (before treatment) at four, eight and 12 weeks, at termination, and at six, 12, and 18-month follow-up periods using the clinician-rated 17-item Hamilton Rating Scale of Depression (HRSD) and Global Assessment Scale (GAS) and the patient-rated Beck Depression Inventory (BDI) and Hopkins Symptom Checklist - 90 Total Score (HSCL-90T). Patients were further categorised by the length of time they remained in treatment; a completer sample of patients who attended at least 12 therapy sessions and 15 weeks of treatments (*n* = 162), an end point sample of patients which included the 162 completers and an additional 42 who received at least 3.5 weeks of treatment before dropping out or being withdrawn (*n* = 204) and a final sample of end point patients consisting of all who entered treatment (*n* =239), i.e. the end point sample plus an additional 35 patients.

Importantly, outcomes for all 4 treatment groups (including the PLA-CM group) improved by statistically significant margins, and the psychotherapies did not differ from each other in their levels of improvement to any substantive degree. Elkin's (1989) study is an illustrative example of a RCT strategy employing all of the essential ingredients of RCT methodology in an attempt to control for as many confounding factors as possible: carefully selected trial participants who adhere to clearly pre-defined inclusion criteria, randomised allocation of those participants to one of the four trial conditions, manualised treatment approaches including adherence checks conducted by independent assessment of audio recordings of sessions, the inclusion of a clinical management component to ensure "standard clinical care" administered in conjunction with the pharmacotherapy medications imipramine (IMI-CM) and placebo condition (PLA-CM), both of which were administered double-blind to participants.

Elkin's study is not without some problems, including the attempt to use a placebo control which is acknowledged by Elkin et al. (1989) and others as problematic in psychotherapy comparison research in general (see Locher et al., 2018). However, the psychological therapeutic approach examined in her study (IPT) and the compiling of completer and two end-point sub-samples makes it highly comparable with outcomes from sub-samples compiled in this study in terms of recovery criteria (although Elkin's recovery criteria omits the clinical significance element introduced by Jacobson and Traux (1991). Moreover, the authors conclude that "All treatment conditions (including PLA-CM) evidenced significant change from pre-treatment to post treatment". (1989, p. 980) and that the initial high severity level of some patient's problems impacted significantly. Both the pre to post differences and their relationship to the initial severity of problems were examined in the current study making the behaviour of these variables of real relevance.

Another large scale pragmatic RCT study of note is that by Roy-Byrne (2010) who also conducted a large trial in 17 primary care clinics in four US cities when they compared the Coordinated Anxiety Learning and Management (CALM) intervention with usual care (UC). Primary healthcare professionals screened and referred patients to the study based on the results of a five question anxiety questionnaire. Eligible patients were those aged 18 to 75 who met Diagnostic and Statistical Manual of Mental Disorders (DSM: Fourth Edition) criteria for 1 or more of panic disorder (PD), generalized anxiety disorder (GAD), social anxiety disorder (SAD), and posttraumatic stress disorder (PTSD) (based on the Mini International Neuropsychiatric Interview; Sheehan et al., 1998) and scored eight or more on the Overall Anxiety Severity and Impairment Scale (OASIS; Campbell-Sills et al., 2009). Following a baseline interview the remaining 501 participants were randomised to receive UC and 503 to receive the CALM intervention (Total *N* = 1004).

Outcome scores on the Brief Symptom Inventory (BSI-12; Derogatis, 1993) for those participants in the CALM intervention group were significantly lower than for those in the UC group at six months, 12 months and 18 months with effect sizes of $d = -0.30$ (95% CI, −0.43 to −0.17), $d = -0.31$ (95% CI, −0.44 to −0.18), and $d = -0.18$ (95% CI, −0.30 to −0.06) respectively. This study is notable for its close adherence to formal RCT requirements as possible, similar to Elkin et al. (1989). For example, eligibility for participation was based on all clients meeting DSM criteria, randomisation of eligible participants to intervention of usual care treatment arms, delivery of a highly prescriptive and manualised treatment program (based on CBT) and the application of blinding on the side of interviewers who administered study questionnaires.

2.3.2 *Meta-analyses of psychotherapy research*

According to Melchert (2016) it was not until the emergence of meta-analysis, more than 25 years after Eyesneck's reproach of the benefit of psychotherapy, that the end of the debate about the general effectiveness of psychotherapy emerged. Meta-analysis provides a method of amalgamating and summarising results from numerous studies in order to compare effects between control and treatment groups or between different therapy types (Wampold et al., 1997). This approach became necessary as a means of analysing what was by the late 1970s a vast and quickly growing body of work comprising many hundreds primary research studies (Wampold & Imel, 2015).

In 1977 Smith and Glass published their first meta-analysis of psychotherapy outcome research on a combination of 375 studies that involved almost 40,000 treated and untreated participants. They used several meta-analytic methods in their analysis which distinguished their methodology from that employed in the previous heuristic or qualitative reviews. First, their analysis included all published and unpublished studies with control groups through 1977, attempting to overcome the problem of study selection bias. Second, they used a well described search strategy to select or exclude studies for analysis, introducing a robust level of replicability. Third, in their review, studies were assigned common characteristics and these features coded in order for their relationship to the effect size of each study to be determined, analysed and compared. This enabled a weighted approach to assessing effects sizes between studies of varying adherence to methodological rigour. The overall combined effect size was *d* =.68 (Smith & Glass, 1977), or put another way the average client in the treatment group was better off than 75% of clients in the untreated group (Wampold et al., 1997). Additionally, in a subsequent larger study, Smith et al. (1980) reviewed 475 studies and calculated an aggregate effect size of *d* = .85 indicating the average treated client was better off than 80% of untreated clients.

There were still some concerns regarding particular aspects of the meta-analytic approach, e.g. the lack of independence of observations due to the use of multiple effect size measures derived from multiple dependent measures (Glass, McGaw, & Smith, 1981; Matt, 1989), but due to the burgeoning quantity of studies being conducted in the area, meta-analytic reviews became a common approach by which to assess the aggregated efficacy and effectiveness of psychotherapy across multiple studies (Wampold et al., 2017). Since the landmark study by Smith and Glass (1977) there has been a very substantial increase in the use of meta-analytic research approaches in the area of psychotherapy research with over 700 meta-analyses of psychotherapy are located in the PubMed database alone (Wampold & Imel, 2015).

Since the first use of meta-analyses to examine the outcomes of psychotherapy, many thousands of primary studies have been conducted in primary care settings using pragmatic and RCT designs. However, Rowland, Bower, Mellor-Clark, Heywood, & Hardy (2000) published the first of what was to become a series of three major systematic reviews of the research evidence of counselling in primary care in the UK over the next 11 years. These reviews included a formalised approach to the study selection process differing from that undertaken by Smith and Glass (1977) who included all studies investigating the efficacy or effectiveness of psychotherapy. Rowland et al., (2000) included detailed reporting of the search strategies employed to select the studies that fit the research objective, explicit reporting of selection process, inclusion and exclusion criteria and justification of decisions, descriptions of assessments undertaken during the review process to identify biasing factors that might influence study outcomes and the documenting of revisions to all of the above as well as the possible impact of any differences to overall outcomes reported in the meta-analyses. This type of methodological approach to performing systematic reviews was driven in part as a response to the recognised weaknesses and biases effecting previous reviews, and partially facilitated by the formation of the Cochrane Collaboration (established in 1993) who

also produced a handbook The Cochrane Collaboration Handbook (see Mulrow & Oxman, 1997).

Beginning with approximately 2,000 abstracts and references which were screened by independent reviewers for their appropriateness for inclusion in the review, a remaining 38 were deemed worthy of more detailed assessment for inclusion in the final review. Importantly from the perspective of the study conducted here, 34 of the studies initially examined did not meet the inclusion criteria because they compared other therapeutic approaches instead of general non-directive counselling or took place in locations other than primary care or those which did not compare against a control group. The process of excluding these studies from the subsequent review more closely aligns the aggregated results with those attained from the results of the study conducted here. Further, recovery data in this review were also analysed twice - for all patients who had complete pre and post treatment scores and also for intention to treat (ITT) data for those patients who dropped out of counselling. Post-intervention outcome data for psychological symptom levels from the included four studies were pooled prior to the resulting meta-analysis.

Analyses of psychological symptoms showed that patients receiving non-directive counselling had significantly better outcomes post intervention than those patients receiving usual GP care. Comparing the Standard Mean Difference (SMD) between the intervention and usual GP care groups revealed a .30 of a standard deviation difference in psychological symptom scores in favour of the intervention group (SMD = -0.30, 95% CI -0.49, -0.11). Further, in terms of recovery rates, patients who had received psychotherapy were significantly more likely to be considered recovered than non-counselled patients at follow up using odds ratio comparisons (OR = 0.54, 95% CI 0.31 - 0.97) (Chi-square = 1.22; df = 1).

In 2006 Bower and Rowland performed updated searches of electronic databases and the Cochrane Controlled Trials register and the Cochrane Collaboration Depression, Anxiety and

Neurosis (CCDAN) trials registers before conducting another review. In addition to the four studies examined in the first review, four further studies were found to meet inclusion criteria and were included in the subsequent meta-analysis. Meta-analysis of pooled outcome data from six of the eight studies included in the review (two studies demonstrated that counselling compared with usual GP care was effective in the short term on a measure of mental health (SMD -0.28, 95% CI -0.43 to -0.13, n = 772). Analysis of the long term effect of counselling was also examined but did not find any statistically significant effect (SMD -0.09, 95% CI -0.27 to 0.10, n = 475). Six years later Bower, Knowles, Coventry, & Rowland (2011) undertook another search of the literature and systematic review to assess the effectiveness and cost effectiveness of counselling for patients with mental health and psychosocial problems in primary care. In total nine studies were included for analyses overall *N* = 1384. When analysis of pooled mental health measures comparing intervention versus usual GP care outcome data were analysed on six of the nine studies, participants receiving counselling had significantly better outcomes 7-12 months after counselling (SMD -0.28, 95% CI -0.43 to -0.13, n = 772). However, the effect of counselling was no different to usual GP care after 12 months (SMD -0.09, 95% CI -0.27 to 0.10, *n* = 475, 4 trials). Further, commenting on the ratings given to the studies which assessed their quality in terms of their adherence to best RCT practices, Bower and Rowland, (2006) stated that key methodological details were not always included by researchers, and that the overall quality ratings would have been improved had they been provided.

Over the course of these meta-analyses, counselling (as understood in the UK context) was shown to be moderately more effective than usual GP care in reducing distress levels in the short term, but generally that difference was not significant after 12 months. However, several important points should be noted. First, in the initial review, Rowland et al. do not provide a specific breakdown of the severity of the problems experienced by participants in the studies they reviewed, except to state: “In the main, patients were described as having emotional

problems and suffering from anxiety, depression and distress" (2000, p. 225). It was acknowledged at the time that a higher level of severity commonly provided a reason for onward referral by GPs to secondary mental health services (see Graham & Graham, 1996). The impact of this practice on the recruitment of participants into studies by GPs is unknown. Further, initial severity had been shown to impact differently on therapy outcomes depending on the type of therapeutic approach undertaken (Elkin et al., 1995). Moreover, in relation to the chronicity of depressive symptoms, Bower et al., (2011) conducted sensitivity analyses which showed that the removal of the study examining a chronic patient group from the overall pool of reviewed studies (i.e. Schroer & Macpherson, 2009), resulted in an increase in the effectiveness of counselling over usual GP care in the overall meta-analysis (SMD -0.36, 95% CI -0.53 to -0.19, n = 611, heterogeneity $\chi^2 = 5.45$, df = 4, $p = 0.24$, $I^2 = 27\%$) , but that stand-alone analysis of the Schroer et al (2009) study showed no significant effect of counselling over usual GP care, indicating at least some impact of the inclusion of those more severe cases on the examination of outcomes in this meta-analysis.

A more recent meta-analyses was conducted which in terms of presenting problems and therapy modality, more closely resembled a sample of service users as that found in the CIPC service. Cape, Whittington, Buszewicz, Wallace, & Underwood (2010) included 34 studies (mainly conducted in the UK, but also several from the U.S.A. and Sweden among others) in a meta-analysis and meta-regression of outcomes from RCTs of adults suffering from anxiety, depression or mixed common mental health problems and comparing treatment of brief psychological therapies (i.e. > 2 sessions but < 10 sessions) with usual GP treatment in primary care. The sample consisted of 3,962 clients, with most studies comparing one of three primary therapy approaches; brief cognitive behaviour therapy (CBT; *n* = 13), counselling (*n* = 8) or problem solving therapy (PST; *n* = 12) with usual GP treatment resulting in a pragmatic research design. Similar to the method of review used by Bower et al. (2006), a quality check on the included studies was carried out using a version of the SIGN quality checklist for the

randomised controlled trials. This included an examination of the randomisation and allocation concealment procedures and attrition (Cape et al., 2010). Effectiveness of brief counselling for depression did not reach a level of significance over that of usual care, but for mixed anxiety and depression was shown to be moderately effective.

In this meta-analysis there were factors however that introduced a level of heterogeneity which may have impacted on the objective of isolating the effects of each of the treatments and subsequent outcomes. For example, while participants in the included studies may have been referred from or recruited in primary care locations, the actual location of their treatment was reported in various studies as occurring in a variety of locations including over the telephone, in participants' homes', both at home and in primary care settings or in primary care settings only, local health centres or simply "in primary care" without further specification. The significance of varying locations of treatment is noted by the authors as an important factor when examining outcomes (Cape et al., 2010). Further, the limitations of meta-regressing outcome data from multiple different questionnaires and rating scales was also acknowledged as possibly introducing systematic biases during comparative analyses (Cape et al., 2010).

Seekles et al. (2013) performed a meta-analysis which focussed on 12 studies that conducted comparisons between experimental (i.e. CBT) and control group outcomes for anxiety disorders only and provided in general practice. Three types of control group were used reported; waiting list (WL), care as usual (CAU) and placebo (PL). They also reviewed only those studies conducted in a primary care setting and included participants demographically and symptomatically similar to that of those seen in the CIPC service, i.e. > 18 years of age. In contrast to Cape et al.(2010) they excluded those studies examining outcomes for participants with a mixed diagnoses of anxiety and depression.

Comparisons were made between the psychological treatments in each of the studies with the relevant control groups. A subsequent overall random effect model showed a medium effect in a pooled analysis of study outcomes (*N* =12), however the fixed-effect model showed a significant moderate to high level of heterogeneity between studies. Certain studies were also grouped together by particular characteristics and coded in order to examine their influence as a priori outcome modifiers. The overall effect size found was Cohen's *d*= 0.57 (95% CI [0.29, 0.84).

Both Cape et al. (2010) and Seekles (2013) attempt to shed light on different aspects of particular factors effecting outcomes of different psychological therapies for depression alone, mixed anxiety and depression and also anxiety disorders alone using meta-analyses of pragmatic RCT studies. Of particular relevance to the current study however was the lack of replication for the significant influence of the service related factor - number of sessions received. For example, Seekles et al. (2013) specifically analysed the number of treatment sessions as a potential modifier of outcomes because Cape et al. (2010) had reported a large affect for brief CBT (i.e. >10 sessions) for anxiety disorders over usual GP care. However, no difference in outcomes between those participants in studies receiving less than eight sessions or those who received eight or more sessions was detected by Seekles et al. (2013).

Cahill, Barkham, & Stiles (2010) conducted a systematic review of practice-based research on psychological therapies in routine clinic settings which included many of the studies relevant to the current study. An overall uncontrolled effect size of Cohen's *d* = 1.29 was calculated from a collection of studies containing almost 11,000 participants (95% CI [1.26, 1.33, *N* = 10,842). They also created benchmark values for effect size and reliable and significant improvement; Cohen's *d* = 1.55 and 54% respectively.

Considering the meta-analytic approach to examining effectiveness of psychotherapy in primary care for depression and anxiety disorders, a recent meta-analytic analysis of studies

determining effectiveness found significant issues in their resulting pool of studies relating to publication bias, the overall methodological quality of the studies and specific problems with the inflated effect on waiting list control groups in comparison to pill placebo or care as usual groups (Cuijpers, Cristea, Karyotaki, Reijnders, & Huibers, 2016). These are issues acknowledged as being problematic within the area of primary research (Goodman et al., 2003; Green & Latchford, 2012; Margison, 2010), but the extent of the issues across the field of research in the area of depression and anxiety is brought to the fore by this meta-analytic perspective of a range of aggregated results and the examination of research processes from large numbers of studies.

Additionally, while these meta-analyses indicate moderate short term effectiveness of psychotherapy delivered in the UK context across a large number of primary studies, in terms of how the pre post results, recovery rates and the role of initial severity relate to the current study, the following is worthy of note: criteria for which potential clients who would be excluded from the CIPC service were included in Bower et al. (2011), i.e. those experiencing obsessive compulsive disorder and substance misuse (i.e. Hemmings, 1997). The inclusion of indictors of initial severity, and thus recovery was not consistent across studies and not uniform in its application to delineate between severity levels (i.e. sometimes, but not always >14 on the BDI). Also, while Cape et al. (2010) did not report significant effect for brief counselling for depression alone, a moderate effect was found for mixed anxiety and depression. This finding is more relevant to the current study since this is the first opportunity to examine mixed anxiety-depressive disorder (MADD) rates in the population attending primary care psychotherapy in Ireland and it has been suggested that half of all cases of common mental disorder could be accounted for by MADD (Das-Munshi et al., 2008). Finally, all of the participants involved in the primary studies and the studies reported on in the meta-analyses examined here (with the exception of Bower & Rowland, 2006; Bower et al., 2002; Rowland et al., 2000), required a formal compliant with the requirements in either the DSM or

ICD-10 for the particular disorder under examination. This is important as no formal diagnosis is required for referral to the CIPC service.

Hence, while some general comparisons were made between variables common to these meta-analyses and primary studies and the current study, the concerns outlined above do illustrate the importance of the provision of an Ireland specific profile of nationally representative data for a large publically funded psychotherapy service and how this study can meaningfully contribute to the international body of research in the field.

2.3.3 *A Pragmatic approach in primary care psychotherapy research*

Thus far, the question of the overall efficacy of psychotherapy for the treatment of depression and anxiety has been examined from the perspective of meta-analytic research concentrating mainly on the results and factors examined in RCTs. The results have shown, in accordance with most other relevant literature in the field that patients who receive psychotherapy tend to have moderately better outcomes than usual GP care or other controls. However, many of these RCT type studies were conducted under highly controlled conditions using highly trained therapists applying manualised treatments to carefully selected study participants. While some RCTs are conducted under strict conditions as close to laboratory conditions as possible, nested within the overall set of meta-analytic results are RCTs carried out in environments more akin to real world services in which psychotherapy is delivered, e.g. secondary care units and primary care centres and involving research participant groups less restricted in their demographic profile and make up of presenting problems – most notably those studies included in meta-analyses by Rowland et al (2000), Bower et al. (2002) and Bower and Rowland (2006). This aligns them more closely with the research conducted here in terms of the overall aim to determine the general effectiveness of psychotherapy in the Irish primary care context.

Pragmatic RCTs involve interventions delivered in more everyday hospital or primary care environments where existing staff administer treatments less adherent to specified therapeutic approach protocols and to less defined groups of patients in terms of criteria for inclusion to participate. These more pragmatic RCTs are now reviewed because they are defined by certain characteristics shared with the current study and have been identified in the literature as important from the perspective of generalising the research results whereby they:

- Compare interventions that are clinically relevant
- Select a diverse population of participants
- Include participants from more than one setting
- Make use of a heterogeneous set of measures to determine and examine change.*

*Adapted from Barkham, Stiles, et al., 2010, p. 36

The aim of a pragmatic RCT conducted in primary care by Ward et al. (2000) was to compare the effectiveness of general practitioner care with two psychological therapies for depressed patients; non-directive counselling and CBT. Patients who met specific inclusion criteria including a score of 14 or above on the BDI (for depression or depression and anxiety) were referred to the trial by GPs from 13 general practices in north London and 11 practices in greater Manchester. The results showed nondirective counselling and CBT were both significantly more effective at reducing depressive symptoms than usual GP care after 4 months according to the BDI ($F = 4.91$, df = 2, $p = 0.008$), but not the BSI ($F = 2.77$, df = 2, $p = 0.065$). This was one of the largest randomised trials of the effectiveness of counselling in the UK at the time (Ward et al., 2000). This particular study can be considered more pragmatic as it included a proportion of patients in their sample who did not meet diagnostic criteria for a depressive disorder, but instead had a diagnosis of some other anxiety state and also those

with no psychiatric diagnosis from the International Diagnostic Classification of Mental Disorders, tenth revision (ICD-10; WHO, 2004).

In this section, we have described studies where psychotherapy researchers have utilised more strict pragmatic RCT designs while demonstrating the effectiveness of psychotherapeutic treatments for depression (i.e. Elkin et al., 1989) and used post-hoc analyses of sub-samples similar to those used in the current study, and examined studies which have shown effectiveness of psychotherapy for multiple different types of anxiety problems simultaneously (i.e. Roy-Byrne et al., 2010). These studies utilised methodological elements similar to those in the current study while attempting to strictly adhere formal RCT design in as many respects as possible, but in many cases, the pre-requisite of exclusion criteria for the purposes of conducting RCTs (pragmatic or otherwise) and many meta-analytic studies, preclude potential patients from participation in the research. This can result in the completion of many examinations of psychotherapy effectiveness for which results are not related to those patients who are subdromal in terms of the severity of their disorder, i.e. they are not in receipt of a formal diagnosis. This can pertain to many users of primary care psychological services with mild to moderate problems in Ireland and in turn, many CIPC users.

We have also examined more pragmatic approaches to conducting psychotherapy research using manualised treatments based on RCT evidence have been adapted in order to conduct studies more sensitive to, as La Bash, Galovski, & Stirman describe "…the complexity of routine clinical treatment settings." (2019, p. 199) (i.e. Bower & Rowland, 2006; Bower et al., 2002; Rowland et al., 2000; Ward et al., 2000). However, there remain fundamental differences between the requirements and objectives of RCTs and pragmatic effectiveness studies and studies gathering practice-based evidence (Cartwright, 2007; Leichsenring, 2004). This situation has led to a paucity of information on whether psychological treatments are generally effective once disseminated through large scale publicly accessible mental health

services in primary care. Studies of this type are required to inform the literature about the comprehensiveness of the effect of therapy in the general public service space when the overall numbers in samples of clients being examined are not overly constrained by factors such as the types of presenting problem or when potential clients are only eligible for assessment by meeting highly specified levels of severity of their presenting problem, age range or the fact they are excluded because of their co-morbidity status or previous diagnoses.

Primary research trials which seek to compare new and/or existing therapies with usual care or other controls prioritise the importance of internal over external validity in an attempt to isolate the effects of the psychological treatment. In so doing they can, as much as possible, compare the efficacy of various treatment approaches and provide an evidence base of a particular type. However, practice based research seeks to determine the effectiveness of therapy as delivered through large scale counselling services such as CIPC in order to provide different outcomes for a different type of evidence base. The requirement for this type of additional outcome information was acknowledged by Rowland et al. who stated "We need information about which patients GPs refer for counselling and individual patient data on which patients improve and which do not. We also need more information on the nature of any improvement; what it means for patients; and how we measure clinical significance." (2000, pp. 228–229). This study will also examine outcomes for a large sample of primary care psychological service participants drawn from the Irish population, whom are previously acknowledged to experience the fourth highest prevalence rate in the OECD for mental health disorders (OECD & European Union, 2018) and will do so across a scope of outcomes recognised as providing a contextualised profile of the results of psychotherapy, i.e. CORE OM and health related quality of life indicators.

2.4 Practice based research – an overview

Practice-based studies focus on routine data collection from clients of a service, the inclusion of all data collected from a large and heterogeneous sample of clients presenting to that service (in both personal characteristics and range of presenting problems) and where the design incorporates routinely collected data but does not rely on the inclusion of control groups or deem necessary a comparative approach during the design or analyses (Castonguay et al., 2013). Importantly, the utilisation of large data sets for the selection of sub-groups for the purposes of analyses can also be a feature. This allows for a more flexible approach to the construction of samples either not readily accessible by researchers of RCT studies in adequate numbers (i.e. under-represented groups in the population), or to the testing of study conditions when randomisation is not possible (Barkham et al., 2010). Additionally, psychological treatments are delivered in usual service locations, are not manualised and practitioners are not subject to additional training in a particular approach for the particular purposes of a study or to checks on adherence to treatment procedures.

As laid out in the Guidance Document on the Provision of Counselling in a Primary Care Setting (Health Service Executive & Irish College of General Practitioners, 2006) which describe the proposed service approaches and structures for the CIPC service, continuous service evaluation of therapy outcomes could be conducted using a standardised outcome measure, e.g. the CORE-OM 34 (Mellor-Clark et al., 1999). Subsequent to the publication of this document, the national CIPC service adopted as standard policy the use of the CORE-OM 34 at pre and post therapy time points for each client with the purpose of providing a measure of therapy effectiveness at the level of both the individual client and the overall service through analysis of individual and aggregated CORE-OM 34 outcome data.

2.4.1 *Practice based evidence gathering - a conceptual framework*

Examples of the use of practice based evidence (also referred to as Practice Oriented Research (POR) (DeFife et al., 2015) can be observed in a diverse range of research areas including prosthodontics (see Dhima et al., 2014), the creative arts (see Clift, 2012), information science (Koufogiannakis, 2012) and medicine (see Seob Kim et al., 2017), and especially in the primary care health sector (Cohen & Davis, 2012; Pirkis, Blashki, Murphy, Hickie, & Ciechomski, 2006; Westfall, Mold, & Fagnan, 2007). Potter et al. (2006) provide a general high level definition of practice based research in the general area of public health as the

> *"…systematic inquiry into the systems, methods, policies, and programmatic applications of public health practice. This definition includes science based inquiry that occurs in practice settings such as field epidemiology, systematic reflection on practice experience, and laboratory analysis to the extent that such inquiry produces generalizable knowledge to improve the outcomes of practice or to inform policy making."* (2006, p. 2).

This understanding emphasises the holistic nature of practice based approaches and how they can be considered in different areas of research, but all with a view to improvement at the individual and service levels. Barkham and Margison (2007a) further define the process of gathering practice based evidence as

> *"…the conscientious, explicit, and judicious use of current evidence drawn from practice settings in making decisions about the care of individual patients. Practice-based evidence means integrating both individual clinical expertise and service-level parameters with the best available evidence drawn from rigorous research activity carried out in routine clinical settings."* (2007a, p. 446).

This also highlights the potential benefit in connecting therapist, client with aggregated outcome data from Large *N* studies conducted in services such as CIPC.

Specifically in areas of medicine and psychotherapy research, the need for the use of more practice based approaches is commonly rationalised by the lack of studies that truly represent the diversity of contexts and situational factors common to the everyday experiences of practitioners (Ammerman et al., 2014; Barkham & Margison, 2007b; Green, 2008). Further, in psychotherapy, practice based evidence can be gathered at the level of the individual client making use of repeated measures in single client case studies to examine within-individual variation, or up to the service level where many thousands of clients attend a service that is physically distributed over a large geographic area and the aggregated data examined in order to determine between-group differences of different types of sub-samples derived from the overall sample (McMillan & Morley, 2010), a point highlighted in the above definition by Barkham and Margison (2007a).

There are methodological frameworks specific to practice based research endeavours outlined in the literature.

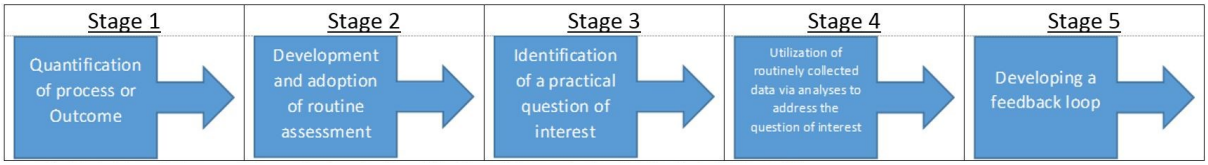


Figure 1: Methodological framework - the five stage model showing the processes employed to employ practice based research. Adapted from Barkham, Stiles, Lambert and Mellor-Clark (2010).

Barkham, Stiles, Lambert, & Mellor-Clark (2010, p. 44) present a five stage methodological framework model to illustrate the processes necessary to employ one type of practice based research. This study addresses each of these stages in the following ways:

I. Stage 1 quantification of process or outcome – As briefly described above (and in detail below), the CIPC service has adopted the CORE OM as its main outcome

measure. In addition, this study employed two further measures in order to broaden the scope and generalisability of the eventual findings.

II. Stage 2 development and adoption of routine assessment – Routine assessment is also built into the existing CIPC intake policy with the national implementation of an electronic centralised recording system for all client data, i.e. CORENET©.

III. Stage 3 identification of a practical question of interest – The overriding issue for management, staff and users of the CIPC service is the question of the overall performance of the CIPC service – as a stand-alone service and as a service comparable to other widely distributed psychology/counselling services. This study addresses the main questions of importance to the CIPC service by assigning them as research questions answered during the course of the study.

IV. Stage 4 utilization of routinely collected data via analyses to address the question of interest – As discussed by Barkham et al. (2010), the analyses and procedures which are used to address questions in practice based studies tend to be broader in terms of their aims than those of basic trials. This is due to the primary focus of practice based research being "…the exploration of relationships between the inputs, processes and outputs of the service delivery system with the aim of improving practice" (Barkham, Stiles, et al., 2010, p. 42), rather than identifying specific components of particular therapies or other minutiae within the therapeutic process which may influence outcomes for clients. Hence, this study employs methodological procedures such as all-inclusive participant criteria and the use of statistical models that vary at more than one level in order to capture and examine data in a manner suitable to meet the needs of the research questions herein.

V. Stage 5 developing a feedback loop – This study is the first national effort by a counselling service in Ireland to evaluate its provision of therapies and outcomes. Hence, information regarding the research aims and objectives provide the first

opportunity for benchmarking service performance. Further subsequent data collection and examination of these data in relation to those produced by this study will begin the process of developing a feedback loop for the benefit of policy and service change.

In addition to the above stage-based framework which contextualises each distinct methodological phase of this study, a major research paradigm relevant to practice based studies is evaluative research. Potter et al. (2006) describe evaluation research as that which "…seeks to assess processes and outcomes of the treatment applied to a problem or the outcome of prevailing practices." (2006, p. 6). How this study's methodological approach supports this description along a justification of specific methodological choices follow.

2.4.2 *Evaluation Research – the practice based connection*

The scientific method underscores the overall ontological and epistemological of evaluation research as pointed out by Speer (1998) who states that "It should be kept in mind program evaluation has its roots in the scientific approach, and that the leading thinkers in the area were scientists first." (Speer, 1998, p. 12). Smith and Glass (1980) derive an understanding of gathering different types of evidence through processes they term "elucidatory" and "evaluative" inquiry (1980, p. 25). Evaluative inquiry is recognised as a different, but valid form of scientific enquiry. Where pure scientific research seeks to provide explanations as to the nature of and mechanisms of change, evaluative enquiry seeks to measure the worth of a program or service (Glass, 1972; Haig, 2018). It is the process of obtaining information for judging the worth of an educational program, product, procedure, or educational objective" (Glass & Worthen, 1971, pp. 150–151). There is no unified theory which comprehensively describes evaluation activities (Alkin & Ellett, 1979). Demarteau defines programme evaluation as "...an information-collection process leading to a value judgement, which will be used in decision-making, leading to action." (2002, p. 547) Further, evaluation research "…collects

evidence systematically from a representative sample of the units of concern. It usually translates the evidence into quantitative terms" (Weiss, 1972, p. 2).

As previously mentioned, RCT type research designs are problematic to implement in large scale, real world counselling services due to the divergent nature of their aims versus the realities and aims of normal service operations (Ammerman et al., 2014), particularly when evaluating psychotherapeutic intervention outcomes (Speer, 1998). Alternatively, instead of the implementation of research as an additional element to service workflow, practice based research designs can be "overlaid" on existing data collected during routine service operations (M. Barkham, Stiles, et al., 2010, p. 42) in a manner that is cognizant of day to day service requirements (Horn & Gassaway, 2007). This approach to integrating research into existing services has several distinct advantages; first, the operational and administrative challenges facing RCTs and pragmatic trials can be incorporated as seamlessly as possible e.g. all existing therapeutic approaches can be considered as relevant, negating the need for additional training of clinical staff to deliver protocol driven therapeutic approaches. Second, sub samples of clients difficult or impossible to recruit using traditional methods can be compiled from an overall service client sample in order to examine questions relevant to the particular service or group of clients with particular characteristics, e.g. examining large numbers of participants adhering to similar demographics such as age, gender and location while also exhibiting the same baseline levels of problem severity. Third, the use of a shared outcome measure by more than one service allows for the leveraging of data from multiple services to a single source for use by all. This allows for the formation of very large datasets for analyses (Barkham et al., 2010). These advantages, all of which were embodied in the research methodology for this study and described in detail in the Method section, position the aims of practice based research in accordance with those acknowledged as pertinent to evaluative and policy improvement endeavours, while also appropriately targeting clients centred areas of service improvement.

2.5 Practice based research literature – closing in on the important factors

What follows is an examination of practice-based studies highly related to the current study in terms of the method of enquiry, questions of interest and type of service provision. In the UK, prior to the introduction of the national implementation of the IAPT initiative individual primary care trusts (PCTs) (in this case the Central Lancashire PCT) conducted practice based evaluations of counselling services. This objectives of this study were closely aligned with the aims and its method (e.g. using the same primary outcome measure – CORE OM) and analytic approach comparable to those of the current study. Moreover, as discussed later, the primary impetus for conducting the evaluation was ultimately to provide initial benchmark data as a basis for future service policy decisions.

Gibbard and Hanley (2008) conducted an evaluation of a primary care counselling service offering only person centred counselling (PCT) over a 5 year period (between April 2002 and March 2007 (*n* = 1098) in order to determine its long term effectiveness. When the service commenced in 2001, six counsellors whose therapeutic approach was person centred therapy (PCT) offered referred patients between six and 12 sessions. However, over the course of the five-year study this number increased to 12 counsellors ranging in experience from considerably experienced and also including some newly- accredited members of the British Association for Counselling and Psychotherapy (BACP). Therapy was also provided by 17 students.

Therapy outcomes were measured by comparing pre and post scores of The CORE Outcome Measure (CORE-OM: CORE System Group, 1998). Effect size for a wait-list comparator group was also calculated (*n* =382). The CORE-OM scores between being referred to the service and beginning counselling were compared to the counselling outcomes of those clients with pre and post counselling CORE-OM scores. Clinical cut-off rates were reported using the clinical cut off levels of 1.19 and 1.29 for men and women, respectively; as recommended by Evans et

al. (2002). Sixty eight percent of clients with valid pre and post CORE-OM scores (*n* = 697) showed reliable improvement, i.e. a change in CORE-OM scores greater than .05 (on the CORE OM). Thirty one percent showed no change and 1.4% reliable deterioration. The overall pre-post Cohen's *d* effect size was 1.2 (*n* =698) and the wait-list pre-post Cohen's *d* effect size was 0.24 (*n* =382).

In order to gauge the performance and effectiveness of this particular service in relation to other services utilising the same outcome measure, i.e. the CORE-OM, a comparison of the percentage of clients who reliably improved according to the CORE-OM (68%) was made against the average improvement rate of clients from the CORE National Research Database (73%) (Bewick, Trusler, et al., 2006). The authors suggest that the discrepancy may be explained by the much higher rate of completed pre and post therapy CORE questionnaires achieved in this study than in those services used to calculate the benchmarks for the national database, i.e. 63% versus 39% respectively. There was also a striking similarity between the percentages of clients who reported feeling better or much better according to exit questionnaires and the percentage of clients who reliably improved according to the CORE-OM – 69% and 68% respectively.

However, in order for post therapy change in each client's CORE-OM score to be considered as clinically significant, their pre therapy CORE-OM score must meet or exceeded the clinical cut-off threshold, in this case 1.19 and 1.29 for men and women, respectively. Gibbard and Hanley do report that over five years of the evaluation period 81% of clients (i.e. men and women combined) scored above clinical cut off at intake, but do not report the combined cut-off value used to calculate the percentage of client presenting above that value, nor do they report a clinically significant change calculation result. This omission restricts slightly the ability for the current study to directly compare reliable and clinically significant change statistics with those in Gibbard and Hanley (2008) later.

Nonetheless the current study and the research carried out by Gibbard and Hanley (2008) share a common central goal; i.e. to provide practice-based evidence sought by a variety of stakeholders (e.g. service managers funding commissioners, health policy makers) in order to justify further support for a service, and a common context; i.e. to conduct the study while being nested within a functioning day to day large-scale service. Gibbard and Hanley (2008) conducted their research in response to a need for the service to demonstrate adherence to clinical governance requirements which was satisfied by assessing clinical outcomes as part of a routine service evaluation. However, after commencement of the initial evaluation, a further requirement emerged for the service to demonstrate not only the general effectiveness of the counselling provided, but also the specific effectiveness of PCT when compared to CBT. The authors refer to the publication of NICE guidelines during this time, with its focus on CBT which triggered an additional need for the service to demonstrate the effectiveness of PCT relative to CBT. It was argued by the authors that while CBT had been demonstrated in the wider literature as being effective (hence its adoption into the NICE guidelines), the need to offer non-CBT approaches was considered a way to diversify the therapeutic offering and thus better meet the needs of the population in their particular area. An important target audience for the study were NHS funders of their local service as ultimately this study was used as part of an effort to increase funding to the local service based on the effectiveness of its broader therapeutic offering, i.e. PCT in addition to CBT (Gibbard & Hanley, 2008). This study exemplifies when practice based research is utilised for the purposes of addressing what Barkham et al. identifies as "…a practical question of interest: A key element […] of interest to practitioners and/or service managers, or an issue that has been identified as problematic" (2010, p. 44).

Other studies taking a practice based approach have also used similar statistical analysis methods to interrogate data as the current study in order to address issue of the non-independence of data where clients were nested within therapists (or in the case of the

current study, clients nested with different time points). Werbart, Levin, Andersson, & Sandell (2013a) addressed the question of the comparative effectiveness of psychotherapy in a natural clinical setting. Each year for a 3-year period they collected patient outcome data (*N* =1498) and therapist orientation data (*N* =75) across 13 outpatient psychiatric care services in Sweden. They also categorised patients into three groupings i.e. cognitive-behavioural (CBT) and psychodynamic (PDT), but delineated between these and integrative/eclectic psychotherapy (INT) as opposed to person-centred therapy as used in the study by Stiles et al. (2006). A number of different measures were used in order to capture change in different aspects of patients' lives; the Symptom Check List-90 (SCL-90; Derogatis, 1996), The Quality of Life Inventory (QOLI; Frisch, Cornell, Villanueva, & Retzlaff, 1992), The Global Severity Index (GSI) and Self-Rated Health (SRH; Bjorner et al., 1996). From a resulting sample of 180 patients with complete pre post data, positive treatment effects were found for all three therapy types with more patients in the INT group improving across all outcome measures and with least improvements in the QOLI and SRH for patients in the CBT group.

Multilevel regression analyses were conducted in order to investigate differential outcomes due to therapy type. This type of statistical analysis is appropriate when participants naturally cluster into hierarchical groups, in this case patient scores were nested within patients, patients within therapists and therapists within therapy types. At the analysis stage, inter-correlation data between scores for those participants receiving the same therapy type can be examined for significance while taking into account and controlling for the fact that the therapy was delivered by different therapists. The overall effect sizes for the combined approaches on the GSI, QOLI and SRH were *d* = 0.83, 0.77 and 1.20 respectively. No difference between therapy types was found due to the effect of therapy on any of the three outcomes.

Werbart et al. (2013a) reported similar outcomes when compared with a study with similar goals, i.e. Stiles et al. (2008). First, both compared three therapeutic approaches popular in the

jurisdiction of interest and found no differences between them. Second, the same type of participants in both studies (i.e. those receiving CBT) reported fewer presenting problems per patient and fewer interpersonal problems. Third, both studies collected data over a three year period and the comparison of the proportions of each of the study starters who completed and returned pre and post questionnaires was broadly similar – Werbart et al. (2013b) = 180 of 1,294 (13.9%) and Stiles et al. (2008) = 5,613 of 33,587 (16.7%).

There are however several differences between these studies also warrant further mention in the context of their relevance to the current study. Firstly, the mean number of sessions attended by participants varied significantly between the studies. Patients in Werbart et al. attended more psychotherapy sessions (M = 18, 31, and 23 for CBT, PDT, and INT, respectively) than did participants in Stiles et al. (2008) (M = 8 for PDT and 7 for PCT and CBT). This difference may be due to different service parameters for delivering psychotherapy to the population in Sweden and in the U.K. Importantly, despite the large disparity between studies in terms of the mean number of sessions attended by participants, no differences were found in outcomes. This has been attributed to the application of differing treatment lengths based on patient need, a practice shown to eliminate differential outcomes for patients (Watzke et al., 2010). Secondly, there was a difference in clinical effectiveness in terms of reliable and clinically significant improvement as measured by GSI (35%) compared with the CORE-OM (58%). This could be due to differing pre-treatment levels of problems between samples, or measurement differences between the GSI and the CORE-OM or perhaps differential effectiveness levels of therapists between the UK and Sweden (Werbart et al., 2013b).

Outcomes from large scale primary care psychotherapy services in the U.S.A. have also been described recently by Sawchuk et al. (2018). Utilising a large N of 2772 participants, they reported outcomes of a five-site, primary care psychotherapy program using CBT as the

primary therapeutic approach. The severity level of participants' problems (depression, anxiety and adjustment disorders) were measured using the PHQ-9 for depression and the GAD-7 for anxiety related symptoms. In keeping with the practice-based, naturalistic study objectives, they applied very little inclusion or exclusion criteria (with the exception of non-English speaking referrals, due to concerns regarding the lack of translated questionnaires and issues around the fidelity of delivering CBT through interpreters), including the use of cut off scores for involvement in the study. An additional element investigated in this study was mental health utilisation, described as the total number of one- off consultations, total number of CBT treatment sessions, total duration of care and any other mental health services used by participants in the six months prior to the commencement of the study.

In terms of the assessment of outcomes after therapy, t-tests between pre and post therapy scores (for all participants who completed at least two CBT therapy sessions) were used to statistically examine differences between time points and one-way ANOVA to compare all three disorder groups at a single time point. In terms of estimating change, reliable change was set as at least a five point decrease in PHQ-9 or GAD-7 scores between pre and post therapy, cut off criteria for determining intervention treatment response was set as at least a 50% reduction in pre post scores on both the PHQ-9 and GAD-7 and remission was defined as scores on both measures being below 5 at the last measurement.

Post therapy scores across all three diagnostic categories combined were significantly better than those at assessment ($t(1173) = 17.56$, $p < .001$). A little over 53% had registered reliable change, 40.9% were defined as responding to treatment and 22.2% as being in remission a the post therapy stage. Eighty six percent of participants completed 6 or fewer sessions, while those with a diagnosis of adjustment disorder attended significantly fewer numbers of sessions *(*$F(2,1168) = 10.99$, $p < .001$) and had shorter overall treatment durations ($F(2,1168) = 5.53$, $p = .004$) compared to both the groups diagnosed with depression or anxiety. Effect sizes

were also calculated and "…adjusted for paired samples" (Sawchuk et al., 2018, p. 7) resulting in a reported effect size for the full sample of *d* = 0.51 on the PHQ-9 and *d* = 0.57 on the GAD-7.

Unlike Sawchuk et al. (2018), Knapstad et al. (2018) and Gyani et al. (2013a), who both calculated study specific cut off values in accordance with IAPT guidelines (NHS Digital, 2018) which follows the same format as that of CIPC in terms of capturing different levels and types of change in individuals based on the difference between the pre and post therapy points on the respective scales, but also takes into account the possible effect of error inherent in the measurement scales when creating an index of reliability for each client's change amount. By way of contrast, Sawchuk et al. (2018) used different cut-off criteria (and values) to define whether participants were clinical cases or not, which in turn makes reported reliable change, treatment response and remission rates for this study difficult to directly compare with reliable recovery, improvement and reliable deterioration rate with those reported by Knapstad et al. (2018) and Gyani et al. (2013a), a point acknowledged by Sawchuk et al., (2018).

While these studies (Gibbard & Hanley, 2008; Stiles et al., 2006; Stiles et al., 2008; Werbart et al., 2013b) share with the current study an over-all focus on reporting changes in clients' pre to post therapy outcome scores, they also relate to the current study specifically in other ways; 1) how a practice based study can contribute to requirements of clinical governance by providing client and service related data, 2) how the output of a practice based study can include service related data that is readily comparable across other services in other jurisdictions and 3) how large sample sizes can help facilitate modelling the effect of service related factors, i.e. type of therapy delivered.

2.6 IAPT – CIPC's close relative

Of further relevance to this study are practice-based studies conducted in the context of IAPT as this UK based model of service delivery is somewhat similar to that of the CIPC service. IAPT was established on foot of evidence detailed in the guidelines for the treatment of depression and anxiety recommended by the National Institute for Health & Clinical Excellence (NICE). The service was initially trialled in two demonstration sites in Doncaster (Yorkshire) and Newham (inner London) in spring 2006 with the intention to grow to a national service with the remit to provide timely and effective access to psychological services for the treatment of depression and anxiety disorders across the UK (Griffiths & Steen, 2013). The IAPT service was based on a stepped care service model and organized through a collaborative care system. The stepped service model is described as "…a framework of organisation of pathways of care designed to reduce burden to patients while maximising health gain" (NICE, 2011, p. 79). In practice this means that what is considered, subsequent to an assessment, the most effective and least resource intensive treatment option is delivered to patients as a first step. After this, the level of care is considered for 'stepping up' (to a more intensive treatment), to remain the same, or for the current approach to be 'stepped down' (to a less intensive approach) (*Stepped Care*, 2017).

There is much overlap between the national CIPC service and IAPT in terms of the service model employed to deliver psychological therapies to a large population. In common with CIPC, IAPT service policy is centred on continuous evaluation and monitoring of service performance from a bottom up (i.e. client) perspective and was "…considered an integral part of IAPT. The programme stipulated a minimum dataset, which recorded the care provided to each service user and his or her clinical progress." (Gyani et al., 2013a, p. 598). This makes the following studies and the reporting of rates of reliable and clinically significant change and pre post therapy outcome effect sizes highly relevant to this study for the purposes of comparing outcomes from these large, nationally distributed, psychological therapy services. By

calculating the reliable change index (RCI) for the Patient Health Questionnaire (PHQ-9), used throughout the service to measure client levels of depression, they calculated a reliable and clinically significant improvement position (RCSI) for each client in the study – resulting in clients being considered either "recovered" i.e. PHQ score moving from above to below clinical cut-off (calculated as a score of 10 on the PHQ-9 and 8 on the GAD-7) and the difference between pre and post therapy scores being more than six point decrease on the PHQ-9 and 3.53 on the GAD-7, "clinical improvement" indicating a decrease in scores by at least each of the stated change scores, but not changing the overall score from above to below clinical cut-off. "Reliable deterioration" was indicated the relevant increase in scores, while "no reliable change" indicated neither the improvement nor the reliability criteria were met.

In terms of proportions of IAPT clients considered recovered; at the end of the first three years after roll out (i.e. 2012), IAPT reported a recovery rate of 45.6% (Department of Health, 2012), which had increased to an overall recovery rate of 51% (based on data from January to March 2017) reported more recently by Clark (2018b). Gyani et al. (2013a) reported that data collected for a one year period (2008 – 2009) from 24 (of a possible 35) different IAPT sites showed that 40.3% of the 19,395 patients included in the full sample showed reliable recovery. Gyani et al. (2013a) also examined the role of certain predictors influencing outcomes at the patient (*N* =11,535) and service level data. In relation to recovery rates, using logistic regression analyses they found no difference in recovery rates when clients with a diagnosis a depressive episode were treated with CBT or counselling (p = .28), but for a mixed anxiety and depression diagnosis those treated with CBT were 1.7 times more likely to be considered reliably recovered (p = <.001). Other predictors of recovery examined by Gyani et al. (2013) will be discussed in later sections.

Within the body of research carried out within IAPT, there are additional studies that shed light on certain service related factors that can be related to client outcomes. For example,

Pybis, Saxon, Hill, & Barkham, (2017) analysed secondary data from 33,243 patients across 103 IAPT services to investigate whether location (i.e. the site effect) accounted for a significant amount of variability in outcomes while the type of therapy (generic or CBT), would not. They also examined differences in outcomes based on the number of sessions attended.

For clients receiving CBT who entered therapy above the clinical cut-off on the PHQ-9, effect sizes of *d* = 1.59 (95% CI) (1.58 - 1.62) for CBT and *d* = 1.55 (95% CI) (1.52 - 1.59) for counselling were reported. What is important about this study is its ability to propose relevant service-level questions and provide unequivocal answers by way of clinical outcome data pertaining to recovery of clients. These results are generalisable to all IAPT service locations and practitioners in terms of the sample frame used. Additionally, as noted previously in relation to the study by Werbart et al. (2013b) the size and nature of the data set allowed for the employment of multi-level modelling to interrogate the data, providing a layer of complexity and rigour to the analyses which goes some way to addressing some of the aforementioned weaknesses in the practice based research paradigm (see Ammerman et al., 2014; Barkham et al., 2010). The current study also employed such an analytic approach in order to address the question of the influence of certain factors on client outcomes.

Richards and Suckling (2009) conducted an observational study to investigate the effectiveness of psychological therapies on levels of depression and anxiety delivered in IAPT locations. Pre and post therapy data were collected on 2,017 patients who received more than one treatment over a period of 12 months and the main outcome measures used were the PHQ-9 and GAD-7. A pre post effect size of *d* = 1.38 for those patients completing treatment (*N* = 869) for depression and *d* = 1.41 for anxiety was observed along with a combined recovery and remission rate (i.e. patients whose scores reduced by 50% or more in addition to those patients whom scored below the cut-off rate at pre-treatment) of 76%. Richards and Borglin (2011) went on to assess the impact of the IAPT stepped care program over a two year period,

following up patients for a further year. As in Richards and Suckling (2009) the IAPT mandated outcome measures were used for assessing changes, i.e. the PHQ-9 for depression and the GAD-7 for anxiety. Of the 7,859 patients referred to the service during the assessment period the largest effects were observed for those completing treatment (*N* = 2949) for both depression *d* = 1.24 CI 95% (1.01, 1.48) and anxiety *d* = 1.21 CI 95% (1.03, 1.43). Recovery rates were also substantial with 47.3% (1378 of 2906) showing reliable and clinically significant improvement for depression and 47.1% (1367 of 2900) for anxiety.

The IAPT style model of service delivery has also migrated to other jurisdictions, i.e. Norway. Knapstad, Nordgreen, & Smith (2018) report outcomes for the prompt mental health care (PMHC) initiative which is closely based on the IAPT delivery and service model from 12 first pilot sites. An overall recovery rate of 65% was reported for the PHQ-9 and GAD-7 (using the same recovery criteria as that from the previously discussed IAPT studies) along with effect sizes of *d* = .88 95% CI (0.82, 0.94) and .74 95% CI (0.69, 0.79) for the PHQ-9 and *d* = 0.83 (0.78, 0.88) and *d* = 0.71 (0.64, 0.76) for the GAD-7 when analysing clients who attended at least two sessions and the ITT sample respectively using the LOCF method for imputing missing data. Use of sub-sample analyses for clients meeting minimum attendance and ITT criteria were also employed in the current study in order to address 2 reporting issues; 1) so that a relative comparison with results from RCT studies and be made for the sake of comparing outcomes for both research approaches, and 2) so a like for like comparison of outcomes can be made between clients groups from different jurisdictions employing practice based research approaches.

The UK IAPT and Irish CIPC services are similar in many ways: both consider service quality within a framework that includes service evaluation which is based on continuous client progress monitoring and post therapy reporting of outcomes. Both services are large nationally distributed and publicly available, receiving thousands of referrals from a variety of

sources every year. Also, within both jurisdictions the administration by therapists of a standardised set of questionnaires is service policy and these form the dataset for all subsequent analyses conducted in order to establish the general effectiveness of the service. Further, the main criterion used by both IAPT and this evaluation to define the group of clients included in pre post measurement change is the same; completion of an assessment plus at least one therapy session.

However, between the CIPC and IAPT services there are also notable differences. CIPC can be considered less prescriptive when directing therapists of the recommended therapeutic approaches in accordance with service policy. While a preference for CBT as default intervention approach is highlighted in the guidance document on the provision of counselling in a primary care setting (2006), the CIPC service does not have a set of listed disorders which are matched with corresponding recommended interventions in the same way as the IAPT service (see Clark, 2018b). Another difference are the set of measures used by each service to complete the processes of assessment and continuous evaluation. The IAPT service uses a combination of the PHQ-9 for measuring the severity of depressive symptoms, the GAD-7 for symptoms of anxiety and the WSAS as a measure of disability, as well as other disorder specific measures. All measures are administered on a session by session basis. In CIPC the CORE-OM 34 (described in detail in Method section) is a measure of global distress and used for initial assessment and final session scoring for all clients (who have been accepted for therapy), regardless of the type of presenting problems. Sessional data is taken (by some therapists) using the CORE-10 – a shortened form of the CORE-OM 34. In addition to these measures, the current study introduced the Health Related Quality of Life Scale (HRQOL: Centers for Disease Control and Prevention, 2000) and the Work and Social Adjustment Scale (WSAS: Mundt, 2002) which are also discussed in detail in the Method section.

There are more similarities between IAPT and CIPC than there are differences as described above, which means the current study can benchmark CIPC client pre post outcomes against the literature on therapy effectiveness using of other services utilising the same measure (i.e. CORE OM) and more pragmatic (i.e. less controlled) RCTs as reference points.

Information regarding certain other service delivery and client factors are of particular relevance and important to the task of improving large scale therapy services (Richards & Borglin, 2011; Saxon, Firth, & Barkham, 2017; Saxon, Ivey, & Young, 2008). Hence, the current study will seek to address the following questions within the context of the CIPC service as delivered in the Irish primary care health sector.

2.7 Issues important for service delivery

2.7.1 *Overall effectiveness*

The overall effectiveness of counselling has frequently been shown in RCTs addressing questions of the comparative effectiveness of different interventions, where outcomes for clients in general are shown to be significantly better than treatment as usual (Wampold & Imel, 2015). This is also the case in the findings of meta-analyses of the effectiveness of counselling when results from hundreds of studies are aggregated and effect sizes are compared (Bower & Rowland, 2006; Bower et al., 2011; Cape et al., 2010; Rowland et al., 2000; Smith & Glass, 1977). While the current study addresses the question of general effectiveness of the CIPC service, and its comparative effectiveness in relation to other large national services (i.e. IAPT in the UK) is one of the primary aims of this study, related to the overall effectiveness of psychotherapy are questions regarding the specific client and service factors thought to influence this effectiveness data.

For example, the number of sessions required to provide meaningful change in clients' symptoms has been examined by secondary analysis of large datasets (e.g. Howard, Kopta,

Krause, & Orlinsky, 1986; Lambert, Hansen, & Finch, 2001), the impact of the initial severity of clients' problems on therapy outcomes (e.g. Driessen, Cuijpers, Hollon, & Dekker, 2010; Elkin et al., 1995) and the question of the longevity of the effects of therapy (e.g. Des Groseilliers, Marchand, Cordova, Ruzek, & Brunet, 2013; Harder, Koester, Valbak, & Rosenbaum, 2014; Wootton, Bragdon, Steinman, & Tolin, 2015). These factors are, among others, questions that are important to clinicians because they can influence the interaction between therapists and their clients during different phases of therapy including case formulation, monitoring measurement change or making decisions on evaluating methods of tracking individual cases (Margison et al., 2000). Moreover, they are of interest to service funders, managers, directors and other staff involved at different levels of an organisation because of how they may be related to policy decision making. The following sections of the review will focus on research examining these particular components and their relationship to the outcomes of psychotherapy. While certain landmark pragmatic RCTs and meta-analyses will be involved in this discussion, the focus will be on large practice oriented research carried out in a primary care setting.

2.7.2 *Number of therapy sessions to provide*

The treatment dose, or number of sessions provided by all counselling services and subsequent effect on outcomes is important from the perspectives of service cost, client retention and safety (Harnett et al., 2010; Otto, 2002). As Barkham et al. (2006, p. 166) point out these issues "…have particular relevance in a climate in which limited and finite resources tempt administrators and policymakers to impose fixed durations of treatment". Further, the delivery of a number of sessions which adequately treats those experiencing all severity levels of symptoms has direct financial implications on a service and is considered important by funders of services (Harnett et al., 2010). Robust evidence regarding this question can potentially have impact throughout the broader healthcare areas of healthcare policy,

particularly when policy decisions made by U.S. based managed healthcare companies, German private healthcare insurers and Australian Medicare providers involving how much therapy to offer its clients have been viewed as being made on arbitrary grounds and based on little actual empirical data (Harnett et al., 2010). There have been many studies seeking to gain insight into this question (e.g. Baldwin, Berkeljon, Atkins, Olsen, & Nielsen, 2009; Barkham et al., 2006; Howard et al., 1986; Lambert et al., 2001) with two points of view emerging as dominant; the first held that longer therapy engagement (i.e. a higher dose) leads to better outcomes for clients (see Howard et al., 1986), the second proposed that after a certain length of time in therapy individual clients make a decision (in conjunction with their therapist) to focus their therapeutic attention elsewhere or leave therapy based their own summation of the progress that had made (see Barkham et al., 2006).

Each of these positions have been examined using different methodologies. For example Elkin (1989) examined the effect of the treatment dose on study participants using post-hoc analysis of outcome data from during their previously mentioned primary study into the effectiveness of psychotherapy treatment (see page 55). Three different samples of clients were established in order to perform further comparative analyses. An intent-to-treat (ITT) group of 239 patients (End Point 239) compiled all patients who entered treatment, a second group of 204 patients (End Point 204) were those patients who continued their treatment for a minimum of 3.5 weeks and a final group of patients who continued their treatment for at least 15 weeks and attended 12 sessions were termed completers. Subsequent analysis did show differences in the outcomes between the three sample groups from 12 ANCOVAs (3 sites * 4 treatments) across four treatment groups. Significant differences ($P < .10$) were found between the completer sample (n =155) pre and post scores on both of the patient completed questionnaires; the BDI and the HSCL-90T. For the End Point sample (n =239) significant differences were found on both of the clinician-rated questionnaires; the HRSD and the GAS.

There were no significant differences in any of the ANCOVAs at termination across the treatment groups for the End Point 204 sample.

In a widely cited study from roughly the same period, Howard, Mark, Krause, & Orlinsky (1986) attempted to specify the relationship between treatment dosage and the effect of psychotherapy using meta-analysis. Howard et al., (1986) analysed raw data using probit analysis covering a period of over 30 years for 15 studies from 2,431 patients attending an outpatient clinic in Chicago. For the probit analysis data the following sample data from each of the studies were input and analysed, (a) the number of the session at which improvement was assessed, (b) the number of patients assessed at each of these points in treatment and (c) the number of these patients improved at each point. Results showed that, in terms of the dose (i.e. number of sessions) and its effect on outcomes, depressed patients responded most quickly to the lowest dosage of treatment followed by anxious and borderline-psychotic patients respectively. Between 8 and 13 sessions were calculated to be enough for 50% of patients (experiencing depressive and anxious symptoms) to reach a status of improved on both researcher and patient improvement ratings. The resulting Dose-Effect Model (Howard et al., 1986) represented a negatively accelerating relationship between dose (number of sessions) and improvement, i.e. as the number of sessions increased the rate of improvement decreased or the greatest amount of improvement was observed in earlier rather than later stages of therapy.

A study by Kopta, Howard, Lowry, & Beutler (1994) sought to examine the question of differential rates of remission of symptoms during psychotherapy based on the nature of the symptoms at assessment, and in turn explore the amount of therapy required by a typical patient. They used the SCL-90-R to measure symptoms before patients began therapy and at one other point during therapy. Probit analysis was again used to calculate the probability that the effect of psychotherapy (operationalized as the probability that a given symptom score

was derived from a functioning, normal population) differed between three symptom classes; acute distress, chronic distress and characterological distress. These symptom classes were derived from an inspection of the treatment response rates and following this process comparisons of the percentages of clients whom were classed as *recovered* and *not recovered* across doses of therapy, which was used as the dependent variable. Jacobson and Truax's (1991) method for the determination of whether a participant had reached reliable and clinically significant change was used because the data for a functional sample was available for comparison as required. Responses to all 90 symptom items on the SCL-R-90 were examined and those scored as 2 or higher were included in the probit analysis. Each symptom for each patient was classed as *recovered* or *not recovered* and depending on whether or not the item had reached clinical significance. Results showed that each of the symptom classes did respond at different rates to psychotherapy in this study. Interestingly from the perspective of the current study, anxiety and depressive symptoms responded most rapidly to psychotherapy (i.e. during the acute phase) and symptoms associated with more severe psychological disorders, i.e. paranoid, schizoid, schizotypal and borderline personality problems took a greater number of sessions before clinically significant improvement was observed (Kopta et al., 1994).

In the studies discussed thus far, the question of therapy dosage was examined using retrospective or ad hoc designs. However, Barkham et al. (1996) sought to re-examine Howard et al. (1986) and their finding that the dose-effect curve for psychological symptoms was negatively accelerated using and pragmatic RCT design. They randomly assigned 212 depressed clients to receive either eight or 16 sessions of either cognitive-behavioural (CB) or psychodynamic-interpersonal (PI) time-limited psychotherapy. The BDI and IIP-32 total scores were recorded at intake assessment, at mid-treatment of the 16 –session category and at the final session for both the 16 and eight session durations. The BDI was also administered immediately prior to each clients' initial therapy session and two weeks after the eighth

session and two weeks after the conclusion of both the eight and 16 session treatment conditions. Following the completion of collection of data from the BDI, IIP-32 and PQ questionnaires, each participants' PQ item responses were classified into one of four PQ domains: symptoms, mood, self-esteem or relationship. The authors examined changes in the BDI total score and the IIP-32 mean item total score, specific depressive symptoms (as per the BDI items) and the intensity of reported problems as per the PQ which were them classified into symptom categories. These measures were taken at multiple points and categorised as either changed by a clinically significant amount or not. This study sought to investigate the finding by Howard and colleagues that the dose-effect curve for psychological symptoms would be negatively accelerated.

Most relevant from the perspective of the current study was a larger proportion of clients had attained clinically significant change by the end of 16-session treatments (M = 46%; range, 37% to 48%) than by the end of 8-session treatments (M = 30%; range, 28% to 31%). However, clients were more likely to have attained clinically significant change by the end of an eight session therapy period than by the middle session of 16-session treatments doses (M = 21%; range, 15% to 25%). Taking the results in this manner did not support the assertion that as the number of sessions increased, rates of improved slowed. The interpretation of "appropriate dose" was as Barkham et al. (1996) argued problematic because clients in the middle of their therapy course (i.e. session 8 for those participants assigned to 16 sessions in their study) had not received their full dose of treatment. Therefore, taking dose to equate to the total number of sessions "…the effect of the shorter treatment (measured as the proportion of clients achieving clinically significant change) was more than half as large as the effect of the longer treatment, suggesting support for diminishing returns from larger doses (i.e., negative acceleration)" (1996, p. 933).

By way of explanation, Barkham et al. (1996) proposed an alternative hypothesis whereby problems could be assumed to improve in a linear (i.e. steady) fashion until it reaches a good-enough-level (GEL) - a level where both the client and therapist agree to either discontinue therapy or focus the therapeutic efforts in a different area. However, their examination of the rate of improvement based on the characteristics of the problem of the client, led them to assert that the influence of this factor altered the stage at which clients would reach their GEL, thus altering the linear nature of the dose-effect curve to that which was more curvilinear in nature, since different problem types would require increasing or decreasing the overall number of sessions in order for the client to reach GEL. It was also suggested the differential response rates could have been driven by the participants' knowledge that therapy was delivered either on a time-limited or time-unlimited basis (i.e. a characteristic of the treatment), encouraging those receiving time-limited therapy to make progress steadily and until the end of the therapy period (i.e. characteristic of the client). While those allowed unlimited time in therapy may have experienced a deceleration in progress towards the eventual end of their treatment phase (Barkham et al., 1996).

The dose effect phenomenon has also been investigated using hierarchical linear modelling (HLM) technique. Findings from these studies are also useful when assessing the overall trend in this area of research. Using a sample of patients (*n* = 6375) from a mixture of locations; college counselling centres, private mental health centres and four primary care centres, Stulz, Lutz, Kopta, Minami, & Saunders (2013) also asked if individual change on a global clinical symptom measure, i.e. the Global Mental Health Scale (GMH) administered during routine outpatient treatment followed a linear or a negatively accelerated curve. They also asked if the pattern or the rate of change varied according to treatment duration. In terms of dose-effect, they found support for a negatively accelerating relationship between treatment dose and improvement during psychotherapy as proposed by Howard et al. (1986). Interestingly however, Stulz et al. (2013) also found support for the GEL model of change in so far as faster

improvement was associated with shorter treatment times and longer treatment times with slower rates of improvement, which suggested clients may have ended therapy once they felt an acceptable level of improvement. However, there was one quite striking outcome difference between both studies. There appeared to be an inconsistent relationship between the effect of treatment and the number of sessions across clients. They concluded that whether treatment had ended or was ongoing was an important factor in the speed and shape of clients' recovery. However, the resulting level of improvement is usually reached in a log-linear fashion, rather than in a steady even linear patterned manner (Stulz et al., 2013), broadly supporting a more negatively accelerating curve model of the relationship between improvement and treatment dose.

Lambert, Hanson and Finch (2001) used sessional data to track recovery rates. They tracked improvement and recovery (based on Jacobson and Traux (1991) criteria) scores from questionnaires administered on a session by session basis and used survival analysis (based on Cox regression analysis) to determine how patients improved across time. This approach was in contrast to that used by Howard which attempted to predict the position of each patient in relation to a dichotomous variable at a particular point in time, e.g. as either recovered or not recovered at session number X. Howard et al. (1986) had selected specific treatment doses, i.e. 1,2,4,8,13,26,52, and 104 and incorporated the session number at which improvement was assessed and the number of patients assessed at each of the time points into a probit analysis in order to predict the expected percentage of patients to be improved for each of the chosen number of session values. Lambert et al. (2001) deemed this approach to be problematic and sought to address this in their study. Lambert et al. (2001) found that 50% of patients who were assessed as being in the dysfunctional range before therapy achieved clinically significant change after 21 sessions of therapy and using the reliable improvement only standard found 50% of patients improved after seven sessions and 75% after 14. This was in contrast to the

findings of Howard et al (1986) (reported earlier) who found that between only 8 and 13 sessions were enough for 50% of their sample to have reached recovery.

Somewhat aligned with the current study in terms of the method employed and measure used was the study by Barkham et al. (2006) who again examined rates of improvement in psychotherapy as a function of the number of sessions attended this time utilising the CORE-OM 34 and End of Therapy forms. They attempted to fully apply the GEL model to a large dataset of 1,868 clients from 33 NHS sites across the UK using the reliable and clinically significant change indicator (RCSC: Jacobson & Truax, 1991) as a function of number of sessions attended. They did find that the percentage of clients showing improvement stayed constant or declined as a function of the number of sessions – which was consistent with the findings of Howard (1986), Barkham et al. (1996) and Kopta et al. (1994) but stated that for a variety of reasons, this did not prove that clients tended to end therapy once satisfactory levels of gains had been achieved. They again found that rates of improvement were also associated with the nature of a particular problem being addressed by the client and therapist (as in Barkham et al., 1996). Hence, it appeared that more intensive treatment for particular types of problems may require a higher number of sessions.

More recent dose-effect studies (i.e. Barkham et al., 2006; Falkenström et al., 2016; Stulz et al., 2013) tend to provide more support for a GEL explanation of how much therapy is necessary for meaningful clinical change than earlier research in the area (i.e. Howard et al., 1986; Kopta et al., 1994). This could be due to the methods of using sessional data by these later studies, making the exact time when improvement began to occur easier to identify. Studies using session by session data, along with follow-up data (i.e. the modelling of the data longitudinally, rather than cross-sectionally), have presented a challenge to estimated predictions of the number of sessions required for recovery as put forward by Howard et al. (1986) using pre-post measurements. However, some results from the studies examined here

have produced ambiguous results and found support for both the GEL and Dose-Effect models when attempting to determine which model best describes clients' improvement relationship to the number of sessions they attend. Additionally, caution when identifying the appropriate dose of therapy for each client should be observed. As Percevic, Lambert, & Kordy (2006) state "The dose/response curve as defined by Howard et al. is a function not only of the individual recovery course but also of the initial distribution of distress and the cut-off for clinically significant change." (2006, p. 365).

Further, it is acknowledged that comparing improvement percentage rates across these studies is not ideal as criteria for stipulating this end state was not consistent across all studies due to their differing objectives, e.g. Howard et al. (1986) was a meta-analysis and used a combination of subjective and objective ratings as did Barkham et al. (1996) (in their RCT) as their studies sought not only to examine client proportions of change, but also to identify the effects of treatment length on particular components of depressive symptomatology. Alternatively, Barkham (2006) was concerned with model fidelity when applying the GEL model to service data from a large real world psychotherapy service.

The current study analysed pre post therapy outcomes and did not collect sessional data relating to client CORE OM scores. This imposed certain limitations on methods available to explore the relationship between the number of sessions participants' attended, i.e. the duration of their therapy, and their eventual therapy outcomes.

2.7.3 *Type of therapeutic approach*

The IAPT service in the UK was set up as a response to what had been previously limited access to evidence based psychological treatments as recommended by NICE guidelines (see Clark, 2011). This resulted in the adoption of cognitive behavioural therapy (CBT) as the primary standard psychological intervention for IAPT service users experiencing mild to moderate depression and anxiety. While the RCT evidence base for the effectiveness of CBT in

the treatment of depression is substantial (Bortolotti, Menchetti, Bellini, Montaguti, & Berardi, 2008; Cuijpers et al., 2016), the adoption of CBT as the only approach to treating mild to moderate depression to the exclusion of other approaches during the early incarnation of the IAPT service was criticised, not only on the grounds of the huge expense of training new IAPT practitioners in CBT techniques, but also because treatments other than CBT were not considered first line treatments for depression or anxiety even though evidence for their effectiveness was becoming established at the time at the time (Callan & Fry, 2012). By contrast, while the CIPC service was also established in response to long waiting times for access to primary care psychotherapy, the guidance document on the provision of counselling in primary care (2006) states a preference for CBT which is "…typically short-term and uses a treatment model that allows the counsellor to predict how long each client will require service and consequently to manage resources." (2006, p. 3), it goes no further than this and does not impose adherence to the delivery of any one (or more) particular therapeutic approaches on its counsellor/therapists.

Hence, the overall profile of therapeutic approaches delivered in the CIPC service was, before the completion of the current study unreported. Important to bear in mind is that due to the nature of the current study design, the contribution to the debate regarding the superior effectiveness of one form of therapy over another was only ever from a perspective of correlation, not causation. What this study could provide was substantial information on the profile of counsellor/therapists' self-declared therapeutic approaches and the relationship between outcomes and the provision of one or more therapeutic approaches by those counsellor/therapists in the national CIPC service.

This question was addressed specifically by a practice-based research approach in a primary care setting in the NHS in the UK. Importantly, from the perspective of the current study is that Stiles, Barkham, Twigg, Mellor-Clark, & Cooper (2006) set out to address a question of the

comparative outcomes of different therapeutic approaches used as stand-alone approaches and in combination. The CIPC service also used the same method of recording therapeutic approach as Stiles et al. (2006). They used data collected during routine practice in 58 NHS sites in the UK (including primary care, psychology and counselling services, secondary care psychology services, and tertiary care/specialist services). All sites used the CORE suite of assessment, end of therapy and outcome measure forms (CORE System Group, 1998) as part of an ongoing evaluation process as did the current study, which made the resulting data of interest to all services using the same system for recording service related and outcome data.

This was a comparative study of outcomes between CBT, person-centred therapy (PCT) and psychodynamic/psychoanalytic therapy (PDT). In addition to therapists indicating the use of one main approach, they were also asked to indicate one other main approach they used in addition to any of the major approaches. These were signified by CBT+, PCT+ and PDT+ resulting in six classified types of therapy for comparison. Of the 10,351 clients who returned at least one CORE-OM or CORE Assessment form during the three year data collection period 6,789 were excluded due to a variety of reasons; non-completion of pre or post measures, non-completion of a CORE Assessment form, the omission of more than three responses from the CORE-OM form or receiving more than one of the treatments under examination. Of the remaining 3,051 patients who completed pre and post treatment CORE-OM forms, 1,309 met specifications for one of the six groups and. Hence, the valid pre post questionnaire completion rate was 33%.

The mean difference in the effects of therapy between CBT, PCT and PDT was not significant. All six groups showed marked improvement (uncontrolled pre post effect size *d* = 1.39). However, the use of more than a single theoretical orientation (i.e. CBT+1, PCT+1 and PDT+1) resulted in a very small non-statistically significant difference from the single orientation effects sizes (partial η^2=0.003). As in other practice-based studies outcomes were also

assessed by calculating the number of patients who reliably recovered or improved (see Mullin, Barkham, Mothersole, Bewick, & Kinder, 2006). The patients in the '+1' group achieved a statistically significant higher rate of reliable and clinically significant improvement than those in the single therapy orientation group [$\chi^2(1) = 8.38$, $p = 0.004$, $n = 1129$] (66.2% vs. 57.7% respectively).

Stiles, Barkham, Mellor-Clark, & Connell (2008) replicated this comparative study, but with a non-overlapping and substantially larger sample meeting the inclusion criteria (N = 5,613) than that used in the previous Stiles et al. (2006) study. Further, this larger sample contained only clients of primary care mental health services and not from other sections of the mental health services, as was the case in the previous similar study. Using the same selection criteria to classify six treatment approaches as before CBT, PCT and PDT and CBT+1, PCT+1 and PDT+1 an effect size of d = 1.36 was calculated on the CORE-OM. However, in this analysis no significant difference was found between outcomes for participants who received therapy from therapists whom declared they used a single or those implementing multiple therapeutic approaches. While row percentages for each of the single and multiple orientations is stated and rates of RCSI versus non-RCSI across all six treatment groups was reported as not significant [$\chi^2(5) = 4.14$, $p = .530$, $n = 4,954$], no statistical analysis comparing RCSI rates between the groups of three single orientations and ('+1') orientations is reported as in Stiles et al. (2006). However, this comparison is commented on and reported as not having approached significance in Stiles et al. (2008).

When Saxon, Firth, & Barkham (2017) employed multi-level modelling to examine therapist variability and outcomes in relation to treatment modality in the IAPT service, they found differences in outcomes were more dependent on which therapist was seen, the amount of therapy received and whether or not patients finished their therapy, as opposed to the type of therapy received.

2.7.4 *Health related quality of life*

In 1948 the World Health Organisation (WHO) defined health as "…a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity." (WHO, 2014, p. 1) and this has been the first declaration in its constitution document ever since. This emphasis on a more holistic definition of therapeutic change is also evident in Wampold and Imel's (2015) presentation of a contextual model of psychotherapy change, where they posit symptom reduction and a better quality of life for clients as being interdependent. Ronk et al. (2016) also delineates between definitions of recovery presented in clinical and consumer based recovery literature, the latter focussing on additional aspects of change in clients' lives, rather than symptom reduction alone.

The Health Related Quality of Life scale (HRQOL) is a four question sub-set of the Behavioral Risk Factor Surveillance System (BRFSS) that asks respondents for their perception of their own physical and mental health (in the previous 30 day period). The HRQOL has been used in the U.S. as a population health tracker since 1993 and in that time its objective and method have been widely accepted as compatible with the WHO's definition of health (Centers for Disease Control and Prevention, 2000). As such, the HRQOL has been adopted for use in research related to the relationship between HRQOL and a range of chronic physical problems (Slabaugh et al., 2017), including the comparison of those with mental health problems with people suffering from chronic health disorders (Cook & Harman, 2008), and an examination of the relationship between serious mental health problems and HRQOL (Shih & Simon, 2008). Cook and Harman (2008) reported that people with mood disorders reported in excess of twice the number of unhealthy days (i.e. combined mentally and physically unhealthy) as adults in the general population. Shih and Simon (2008) also reported that participants in their study suffering from depression were second only to those suffering serious psychological distress in terms of the number of unhealthy days they reported.

In terms of how this examination of the HRQOL relates to the objectives of this study, additional information related to the duration (and subsequently the severity) of problems experienced by potential clients of the CIPC service has the potential to facilitate improved service provision decisions regarding prioritisation of client allocation to counsellor/therapists. This could be accomplished through a system of flagging high severity clients upon receipt of information before they present for assessment with the CIPC service. The window of opportunity in which to garner this information currently lies between the point referral (usually by a GP) and the assessment session with a counsellor/therapist.

One major advantage of the HRQOL is its brevity which results in a short amount of time required to administer the questionnaire (Slabaugh et al., 2017), hence it could be a candidate as screening tool used to assess potential CIPC clients prior to allocation to a counsellor/therapist. In order to assess its value as such a tool, it was necessary to explore the pre therapy HRQOL's responses alongside pre therapy CORE OM scores in order to determine whether there is an acceptable level of congruence between both measures. It should be noted, that the HRQOL would act as a proxy measure of severity, and not be considered a replacement for the full CORE OM. With this purpose in mind, the study explored the relationship between the CORE OM and the HRQOL in terms of its ability to flag high severity CIPC clients at the assessment stage.

2.7.5 *Initial symptom severity*

According to Wampold and Imel (2015) roughly 50 percent of the variability in psychotherapy outcomes can be attributed to the amount of initial distress reported by clients. The initial severity of clients' problems can be impactful on a large and expansive service. Hence, from this perspective the allocation of resources is of concern to policy makers and funders of a service such as CIPC. Psychotherapy researchers have examined relationships between pre and post therapy levels of distress (i.e. severity). Impactful research has included pragmatic

RCT study designs, meta-analyses and of more relevance to the current study, practice based research incorporating analysis of existing mental health service data, which are now reviewed. Of importance when considering the following body of research in relation to the current study, is the nature of the relationship between clients' level of severity and their chances of achieving change which can be considered clinically meaningful, as this can be helpful to clinicians balancing the priorities within large caseloads (Derek Richards et al., 2016).

Driessen et al. (2010) performed a meta- analysis which included 23 studies totalling 1,365 participants to investigate whether pre-treatment severity moderated psychological treatment outcome relative to controls. Mean pre and post treatment depression scores were used to generate effect sizes for comparison between psychological treatment and control conditions by subtracting the average score of the control condition from the average score of the psychological treatment condition and dividing the result by the pooled standard deviations of both conditions. Depressive symptoms as measured by the Beck Depression Inventory (BDI-I) and the Hamilton Depression Rating Scale (HDRS) were used as the sole outcome measures. Results did not support the hypothesis that post-treatment effect size was moderated by mean pre-treatment depression levels.

Weitz et al., (2015) tested whether baseline depression severity moderated outcomes between CBT and anti-depressant medication (ADM) by performing a systematic literature search and conducting meta-analyses on the aggregated outcome results. Of note in the methodology is the use of individual patient data meta-analysis (IPDMA). This involves collecting raw data from RCTs (requested from and submitted by the authors of the original studies) to be included in the meta-analysis which may allow for a more accurate analysis of the moderators of treatment outcomes (Weitz et al., 2015). All available pre and post therapy

and control scores from 16 studies were included for a total of 1,700 participants - 794 from the CBT condition and 906 from the ADM condition.

In terms of statistical analysis, the interaction between baseline severity and treatment outcomes was added into a multilevel linear regression model. Additionally, an adjusted model was also run where length of treatment, type of medication, demographic variables (age, sex, and marital status), and the risk of bias (sequence generation, allocation of concealment, masking, and ITT analysis) were controlled for. The same analysis was also run on a sample of study completers only (*n* = 1,036). Findings of the IPDMA meta-analysis indicated no moderating effect of baseline depression severity on outcome differences between CBT or ADM on either the clinician-rated HAM-D or the patient-rated BDI on any of the sample profiles created, i.e. the treatment effect (ADM vs CBT) did not differ as a function of severity (Weitz et al., 2015). Hence, those participants with more severe depressive symptoms at baseline measurement were no more likely to require medication than CBT in order to improve than those participants with less severe symptoms according to ratings provided by clinicians or patients.

The preceding studies investigating the role of initial severity were conducted for either the purpose of revealing a moderating effect of initial severity (i.e. Driessen et al. (2010)) or with the purpose of comparing its effect on outcomes using different therapeutic approaches (i.e. Weitz et al., 2015). The results, from a mixture of primary studies and meta-analyses, fail to find any evidence for the moderating effect of clients' level of initial symptom severity on their final therapy outcomes across a range of different measures. While these results are part of the overall discussion regarding how initial severity can impact on client outcomes, they do so mainly from an RCT research perspective. This sets the rationale for the criteria used to construct the included samples and positions the primary focus of the research as being a

comparative exercise, which, as previously discussed, are not the necessary pre-requisites of practice-based research.

Of importance and relevance to the current study was the extent to which clients' pre therapy symptom severity at the client assessment stage in large real world counselling was thought to impact on therapy outcomes. There are have not been enough studies conducted investigating this question in this setting (Kraus et al., 2016; David Saxon & Barkham, 2012). However, it is a factor highlighted as one which should be considered when assessing outcomes in a large managed health care psychotherapy service (Brown et al., 2005). A small number of large, practice based studies of clients attending brief counselling in primary care environments have been conducted in an attempt to identify certain client characteristics which could help predict outcomes including the initial severity of clients' problems the most pertinent of which will be reviewed.

Saxon et al. (2008) sought to identify factors associated with poorer outcomes after counselling and also attempted to produce a predictive model of outcomes based on those factors. Employing a retrospective observational design, CORE System form data from a 2.5 year period involving 1,117 clients of a counselling service in the UK was analysed. Severity levels were recorded by counsellors (*n* =12) during assessment sessions using the CORE-A form. The type of ending of therapy (planned or unplanned) as well as the frequency and number of sessions attended was recorded on the CORE-EOT (end of therapy) forms and used for the study. A number of independent variables were identified a priori as potential covariates for use in building a logistic regression model including having both moderate/ severe levels of anxiety and depression (included as measures of complexity). RCI and clinical cut-off scores were calculated to examine rates of reliable and clinically significant change (RCSC). None of the logistic regression models presented were acceptable in their

discrimination, however having a higher baseline CORE score was associated with not achieving RCSC.

Also within a primary care setting, Bryan et al. (2012) attempted to produce a model showing the typical trajectories for improvement following brief therapy in primary care facilities in the U.S.A. Like that of CIPC, a rapidly expanding model of mental health care in this setting was one of integrated care provided by a behavioural health consultant (BHC). Included in the study were 495 participants referred by their primary care practitioners (PCPs) and attending BHC services integrated and collocated within two large family medicine clinics at Air Force training hospitals in the south-western and south-eastern United States. Participants completed the Behaviour Health Measure-20 (BHM) at assessment and all subsequent therapy sessions. This provided an aggregate score on a Global Mental Health (GMH) scale calculated by taking the average score of all 20 items to provide an index of overall health and was used to distinguish four levels of mental health functioning based on the RCSC criteria as recommended by Jacobson and Truax (1991). Participants all fell within four categories of severity according to their initial BHM scores; 111 (22.5%) in the severe range, 65 (13.2%) were in the moderate range, 137 (27.7%) were in the mild range, and 181 (36.6%) were in the normal range.

Using mixed linear regression modelling, Bryan et al. (2012) attempted to estimate both the frequencies of participants in each of the change categories; Reliable Deterioration, Deterioration, No Change, Improvement, Reliable Improvement and Clinical Improvement and identify changes in GMH score across appointments associated with the effect of appointments and baseline severity of mental health functioning. Of interest to the current study is the performance of the model once patients' initial severity level (i.e., normal, mild, moderate, or severe) was included in Model B as a Level 2 fixed effect to consider differential change patterns over time. Results showed that inclusion of this factor did significantly

improve the overall model fit, i.e. a reduction in the -2 restricted log likelihood (-2LL) between Model A (-2LL = 2,124.492) and Model B (-2LL = 1,184.470).

In the Irish context, there are no studies conducted at a national level within primary care which have sought to determine the role of the severity of clients' presenting problem on outcomes, or the relationship between counsellor/therapist ratings of severity and those determined by different levels of CORE OM scores. The current study will examine these factors and how they are related to therapy outcomes including the impact of the initial level of severity.

2.7.6 *Employment status and gender*

Cross-sectional studies have shown that being unemployed is associated with increased rates of depression compared with people in employment in England and Wales (Jackson & Warr, 1987), Australia (Butterworth et al., 2011) and the U.S.A. (Moorhouse & Caltabiano, 2007). More recently, a large meta-analysis of longitudinal data from Australia, Canada, Croatia, Germany, Israel, Netherlands, Norway, United Kingdom, United States and Ireland found that unemployment and job loss were correlated with poorer mental health (e.g., higher depression and higher psychological distress) (Hergenrather et al., 2015). A number of studies have found a gender imbalance whereby males suffer to a greater extent from the effect of unemployment on their mental health (Paul & Moser, 2009).

Shepherd et al. (2005) conducted a practice based study of a service that utilised a model of delivery closely resembling that of the CIPC service in the current study. This involved a retrospective observational study of data routinely collected from a London borough by counsellors and psychologists across 20 GP practices. Study data were extracted from a database of demographic information, service activity and outcomes from the CORE System (CORE System Group, 1998) set of assessment, screening, outcome and end of therapy forms. A reliable change index (RCI) (Jacobson & Truax, 1991) and clinical cut-off scores were

calculated for the sample in order to form groups for statistical comparison of binary outcomes, i.e. those clients achieving or failing to achieve RCSC. Of particular interest to the researchers were relationships between a number of demographic variables (i.e. age, gender, ethnicity, employment status, and living status), intervention length and outcomes. As the current study seeks to profile the clients of the CIPC service and examine and compare psychotherapy outcomes of a national counselling service for the first time, these relationships are also pertinent to the current study.

Data from a four year period (i.e. 1999-2002) were examined comprising a total of 3,687 referrals. The authors state of those clients who attended assessment (i.e. 62% of referrals), 50% completed an intervention with the service and were discharged and 17% completed an intervention and were referred on for additional specialised therapy (PTSD, eating disorders etc.). Thirty four per cent discontinued treatment prematurely. Client drop-out and subsequent data attrition rates were higher in this study than in RCTs examining psychotherapy outcomes, as is the case with most practice based studies (Connell et al., 2006). Hence, while pre intervention data were available for 64% of clients who attended assessment (*n* = 1,481 of 2,298), just 20% (*n* = 458) of the total number of clients who attended assessment completed both a pre and post treatment CORE OM.

Shepherd et al. (2005) also found that for the sample of all clients who attended assessment and completed pre-treatment CORE-OM questionnaires, the demographic variable gender was significantly related to the number of sessions attended by clients equating to a moderate effect for gender (*n* = 3,682) (slightly more sessions for women) However, once data for only those clients for whom pre and post-treatment scores were available were analysed – no significant relationship effects remained.

Analysis of variance using ANCOVA while controlling for the effect of pre therapy CORE OM scores showed no significant direct effects were found for any of the variables except

employment. Further, there was an interaction effect between gender and employment. What this suggested was while change scores were significantly different for men and women, the direction of the effect showed that men's change scores were effected by unemployment to a much greater extent than change scores for women. Finally, while women achieved a greater reduction in CORE OM scores between pre and post therapy, they also registered higher mean pre therapy scores. Overall Shepherd et al. (2005) found support for the assertion the while women generally came into therapy with higher initial scores than men, they also tended to improve by a greater degree, but being unemployed resulted in much poorer outcomes for men than women.

2.7.7 *Do clients maintain the gains made in therapy?*

Another factor of importance to large scale national mental health therapy services is the likelihood of relapse or a return for further treatment by it users. This can result in additional costs in a service under pressure to justify expenditure on the services it offers (Davis et al., 2008). Hence, implementing evidence based therapeutic approaches which are shown to offer long lasting effects is a priority for service providers. Follow-up periods in the literature range from four weeks (see Bakker, van Balkom, Spinhoven, Blaauw, & van Dyck, 1998) to nine years (see Emmelkamp & Powers, 2010). However, more common follow-up periods tend to be for periods between four and 12 months post therapy.

In terms of RCT studies, Corney and Simpson (2005) examined outcomes a intervals of six, 12 and 36 months for chronically depressed patients in a general practice setting. The overall sample comprised 181 patients screened for inclusion in the trial by scoring between 14 and 40 points on the BDI at assessment. Both the experimental group (n = 92) and the control group (n = 89) were referred to their GP to treat as normal. However, the experimental group participants were also referred onto the resident counsellors in the corresponding practice. BDI and Social Adjustment Scale measures (SAS; Cooper, Osborn, Gath, & Fegetter, 1982) were

administered at assessment, 6 and 12 months and both questionnaires were returned by regular mail at 36 months. No significant differences in BDI or SAS outcome scores were found between experimental and control groups from assessment to six, 12 or 36 months.

Meta analyses of long term effectiveness have also been undertaken including that by Karyotaki et al. (2016). They examined acute phase psychotherapy for depression in adults and its effectiveness compared to control groups in the long term, i.e. >= 6 months. Across Forty-four studies and five companion papers (n =6,096) they found that the psychotherapy group performed better than control groups on all positive outcomes combined (recovery, remission, partial remission, response, and reduction in depression severity) and on quality of life at a follow-up of six months or longer. Additionally, with the exception of non-directive supportive treatment, all psychotherapies maintained treatment gains through six months or longer post-randomisation.

Practice-based studies have also addressed this question. Baker et al. (2002) followed up a pre post therapy effectiveness study reported by Baker, Allen, Gibson, Newth, & Baker (1998) with a longitudinal assessment of outcomes from the same sample of clients. The same questionnaires used in the original study (see Baker et al., 1998) were sent to participants from the original effectiveness study at six months after counselling ended. The wait-list group comprised patients who were sent questionnaires on referral (n =367), and again once they attended their assessment session to begin counselling (n =81). Questionnaire return rates declined as the time since baseline measurement increased, i.e. baseline n =1,724, at three months n =796, at six months n =562. The average time participants waited to begin counselling was 10 weeks. While multiple paired t-tests were used to calculate changes in scores between baseline and subsequent questionnaires, a Bonferroni Correction value of 1.25% was used in an attempt to mitigate the problem of multiplicity and an overall type I error rate of over 5%.

Importantly, the level of comparability of the study by Baker et al. (2002) with the current study undertaken is high for number of reasons; first, over the course of one year all patients considered by their GPs to be suitable for counselling were referred to the service, second, the referred client sample comprised a heterogeneous set of demographic statuses, problem types (once clearly attributable to a definable situation) and excluded chronic and severe problems except mild to moderate anxiety/depression. Third, the same length of time was allowed to pass between follow-up questionnaires being posted to participants, i.e. six months. Fourth, a quality of life scale was utilised (i.e. Watson and Marks Life Impairment Scale (Watson and Marks, 1971)) in conjunction with a measure of symptom severity for depression and anxiety (i.e. shortened version of the Bedford and Foulds Delusion-Symptoms-States Inventory (DSSI; Bedford and Foulds, 1978)) in order to provide a richer, more contextual account of participants' changes in their life circumstances in relation to attending counselling. These methodological similarities make a review of the study by Baker et al. (2002) a pertinent exercise. Three months after counselling anxiety and depression rates for the counselled group had reduced significantly (t = 21.5; df = 720; p < 0.001) and (t = 20.9; df = 701; p < 0.001) respectively. A smaller, but still significant decrease was observed between three and six months for anxiety and depression.

Davis et al. (2008) conducted a long-term evaluation of brief, time-limited counselling with CORE-OM questionnaires being completed by participants (n = 58) during assessment, immediately after and 30 months after counselling. Results showed that mean CORE-OM scores were significantly reduced between the commencement of counselling and immediately after counselling had ended (t [58] =9.03, p<.001, r=.37). Further, there was also a significant statistical difference between CORE-OM scores between the time counselling ended and 30 months later, although the effect was small in real terms (t [58]=1.11, p =.003, r =.14). Clinical cut-off rates of 1.19 for males and 1.29 for females were used (Chris Evans et al., 2002) and 57% showed a clinically significant improvement for all participants above the cut-

off level before counselling. Between post counselling and follow-up there was an increase in those above the clinical cut-off from 29% to 31%.

As mental health services become increasingly expenditure aware, the relationship between psychotherapy outcomes and the length of treatment gains becomes an increasingly important service related question, not least because it has been shown the less impact therapy has on the client the more likely they will seek additional help resulting in higher costs in the long run (Cairns, 2014). With this in mind, follow-up studies become important evidence of a service's overall long term clinical effectiveness as well as cost. The studies reviewed above suggest that regardless of the type of therapy clients receive in the primary care context, most improve significantly and those improvements tend to last for a period of at least three months.

2.8 Conclusion

While some factors discussed have been shown to influence therapy outcomes and are important to both therapists and those involved in commissioning and administration of services, other factors do not appear to have as significant a bearing on the direction of clients' progression through therapy. The current study is concerned with describing the behaviour of the factors examined during the course of the preceding review of the literature as relevant to large scale psychotherapy services such as CIPC. Researchers have addressed questions of relevance using pragmatic RCT research approaches within the context of secondary care, others have sought to answer these questions within the context of primary care, but also using RCT type research approaches. Only more recently have practice-based research approaches been used to demonstrate the effectiveness of psychotherapy and also examined the factors involved in beneficial outcomes for clients in a large scale publicly available service. It has been assumed (by service funders) that if certain psychotherapies were shown to deliver significantly better outcomes than usual care in RCTs, this would also be

the case in large scale services (Pearce & Raman, 2014; Watson et al., 2004). However, this has not always been the case and reasons for this have been suggested as being related to both the perceived limitations and divergent objectives of RCT and practice based evidence gathering (Green & Latchford, 2012; Green, 2008; Hemmings, 2000b; Margison et al., 2000).

As previously discussed, CIPC approaches and service guidelines were the results of an evidence base formed from the results of RCT studies utilising effectiveness and efficacy research approaches. The findings discussed relating to RCTs (both pragmatic and effectiveness) and meta analyses examined above have contributed to, and are essential to the underlying ethos, service formation and service delivery model (i.e. client centred and evidence based) of the CIPC service.

However, notwithstanding the importance of the evidence garnered from RCT studies conducted before the CIPC service was disseminated on a national basis in Ireland, it is practice based research described above which provides more pertinent information when considering the design and implementation of a rigorous evaluation of the CIPC service, as the current study does. Bearing this in mind, the variables included represent some of the most important factors for large scale psychotherapy services such as CIPC to be aware of insofar as they provide the basis from which to interrogate the most important questions. Moreover, they have been implicated in playing a role in outcomes for clients in RCT research. Hence, investigation of these factors becomes relevant from a practice based research perspective in order to deliver information important for the quality of delivery of the service, while the endeavour may also highlight areas important for future investigation using RCT research in order to inform psychotherapy policy. Another influence reflected in the design of this study is that these variables of interest are such because they can be studied in accordance with the realities, constraints and overall objectives of a real world service such as CIPC.

There are many variables of interest in the field of psychotherapy research, but the variables chosen to conduct practice based research should serve two distinct functions: 1) they must be capable of answering questions of relevance to a particular service and 2) also be relevant enough to help answer questions of importance from the perspective of mass delivered psychotherapy services in general. Hence, a practice based research approach was considered as the most appropriate in order to achieve the following aims and objectives and address the subsequent research questions, all of which were considered and formed before data were gathered.

2.9 Research Aims

- To establish the effectiveness of the first large nationally distributed service in Ireland - Counselling in Primary Care.
- To explore the relationships between particular service and therapy related factors considered important to client outcomes by the relevant literature.
- To create a set of national service-level outcome benchmarks for recovery and improvement at service level for counselling as delivered by the CIPC service.

2.10 Research Objectives

- Client outcome and service data will be extracted from the national Counselling in Primary Care database CORENET in order to report the aggregated pre post mean difference and determine the overall effectiveness of the service at the national level.
- Pre and post CORE OM therapy data will also be analysed in order to establish outcome benchmarks based on four categories: recovery, improvement, no reliable change, and reliable deterioration.

- Client outcome and relevant service data will also be analysed and described in order to create a profile of the CIPC service and examine the relationships between client outcomes as measured on the CORE OM and
    - Health related quality of life
    - The number of therapy sessions attended by clients six months after therapy has ended
    - The initial severity of clients' presenting problems.

2.11 Research Questions

2.11.1 What are the outcomes of short term counselling conducted in primary care in Ireland:

2.11.1.1 In terms of symptom reduction?

2.11.1.2 In terms of treatment duration?

2.11.2 Does the initial severity of participants' symptoms predict CORE OM outcomes?

2.11.3 Are outcomes sustained six months after therapy has ended?

2.11.4 Are there differences in outcomes between therapeutic approaches?

2.11.5 What is correlation between CORE OM and Health Related Quality of Life?

2.11.6 Does employment status or gender have an effect on client outcomes?

Chapter 3: Method

3.1 Overview

The philosophical assumptions underpinning the methodology chosen for this study are discussed. Next, the aims and objectives of the research are stated and supporting conceptual frameworks elucidated (evaluative and practice based methods). The methodological framework (use of self-reported measures in quantitative research) is described followed by a description of each of the samples used to answer the relevant research questions. The study recruited over 2,700 participants and the recruitment process, sampling, instrumentation and approach to data analyses are also discussed.

3.2 Philosophical and Theoretical Foundations

The following methodological approach is guided by a positivist research philosophy while employing a practice based quantitative research method. However, the researcher also appreciates certain post-positivist elements in terms of particular ontological and epistemological assumptions when conducting a practice based study such as this. How these assumptions shaped this study are discussed below.

3.3 Epistemological position

A post-positivist research paradigm rests upon certain assumptions, including the understanding that outcomes from scientific research are influenced to varying degrees, by causes which can be measured and analysed in order to address specific and predetermined questions (Creswell, 2003). The overall research philosophy herein is underpinned by a scientific realist perspective and relies upon the understanding that the nature of knowledge is objective (Chilisa & Kawulich, 2012). Further, a scientific realist perspective can take a variety of forms, but most of which "…display a commitment to at least two doctrines: (1) that there is a real world of which we are part and (2) that both the observable and unobservable features

of that world can be known by the proper use of scientific methods." (Haig, 2018). While a post-positive view accepts the expression of absolute truth is not attainable and that all scientific evidence is to some degree imperfect (a position known as critical realism (Creswell, 2003)) and in this respect differs from a strictly positivist perspective, it also advocates that highest form of understanding our world can be achieved through rigorous and robust data collection and subsequent statistical analysis whereby prior claims to knowledge (in the form of hypotheses) can be either supported or rejected based on the most recent set of relevant research findings (Phillips & Burbules, 2000).

Post positivism is associated with the use of quantitative research methodologies and the origins and development of this methodological approach have been dealt with extensively in psychotherapy literature (e.g. Creswell, 2003; Haig, 2018; Luborsky, Auerbach, Chandler, Cohen, & Bachrach, 1971). It has been proposed as possible to render the unobservable (i.e. psychological theories) observable through the development of appropriate tools (i.e. measures and analytical procedures) (Cacioppo et al., 2004). At the heart of the quantitative psychological research approach is an understanding that human behaviour and emotions can be expressed through the collection of (usually numerical) data and subsequently use statistical methods to interrogate those data. Hence, the main variables of interest studied herein (i.e. psychological distress, well-being, symptomatology, risk and health related quality of life) are represented through data collected from self-report measurements, which are considered to serve as a proxy for clients' subjective experiences (Zachar, 2010). Hence, the method employed in this study is reflective of a post positivist epistemological framework in that its methodology facilitates the researcher's attempt to understand the behaviour of variables particular to an existing nationally distributed counselling service.

3.4 Measurement of mental attributes

In order to operationalise the measurement of mental attributes such as depression and quality of life, Guyon et al. (2018) suggest a move away from an empirical (i.e. scientific) realist epistemological position to that of pragmatist realist. In summary, this position rests upon the logical deduction that mental attributes are indicative of neuronal processes in the brain. While these processes are biological in origin and nature they should not be considered as experimentally malleable or to behave in the same way as, for example attributes in physics and therefore justified as belonging in the same framework of measurement as that used in physics. Further, fundamental ontological differences between psychological and physical objects (the former being subjective and the latter objective) in terms of their formal framework of measures, require validation protocols of differing epistemologies; physical objects must adhere to a dominant empirical epistemology, while mental objects may adhere to some alternative. This alternative epistemological position is determined pragmatist realist - whereby mental attributes are "real objects (linked to the brain) but set in a social praxis." (Guyon et al., 2018, p. 2). According to Guyon et al. (2018), adopting this position allows for the operationalisation of the measurement of mental attributes.

However, some criticisms of self-report measures as valid conduits of measuring change in psychotherapy have been made, for example their very administration has been suggested to result in a 'response shift' between pre and post measurement due to tacit learning regarding therapeutic processes and personal reflections on the part of the respondent (McLeod, 2001). Notwithstanding this, the following study adopted a positivist approach whereby its methodology was framed by an understanding that psychological research involves searching for below the surface explanations of empirical phenomena using quantitative data collection and analytic techniques (Glass, 1972).

3.5 Study Design

This study was conducted within the context of a pre-existing national psychotherapy service and employed a practice based research methodology. This research approach will be discussed within the wider context of the overall research effort in order to situate it within the goals and objectives of this study.

3.5.1 *Background*

As laid out in the Guidance Document on the Provision of Counselling in a Primary Care Setting (Health Service Executive & Irish College of General Practitioners, 2006) which described the proposed service approaches and structures for the CIPC service, continuous service evaluation of therapy outcomes could be conducted using a standardised outcome measure, i.e. the Clinical Outcomes in Routine Evaluation (CORE-OM: Mellor-Clark, Barkham, Connell, & Evans, 1999). Subsequent to the publication of this document, the national CIPC service adopted as standard policy the use of the CORE-OM at pre and post therapy time points for each client with the purpose of providing a measure of therapy effectiveness at the level of both the individual client and the overall service through analysis of individual and aggregated CORE-OM outcome data.

3.5.2 *Design rationale*

The overall aim of this research was to conduct an evaluation of clinical outcomes with the primary aim of comparing pre and post therapy outcomes using a practice-based approach. Also under examination were the behaviour of other service and client factors of relevance in a nationally distributed psychotherapy service. The method of investigation was mainly correlational and this study design which was chosen primarily in order to complete the study in spite of the constraints placed on it due to the fact it was nested within day to day CIPC operations.

3.5.3 *Ethical considerations*

All clients who attended assessment during the study period consented to ongoing routine service evaluation and as such the demographic information and baseline characteristics of this group are reported. For the research activities pertaining to data gathering and inferential statistical reporting ethical approval for the study was obtained from the School of Psychology

Research Ethics Committee, Trinity College Dublin. Separate approval was also granted by the respective Research Ethics Committees (RECs) in each of the participating CHOs. Certain areas within some CHOs did not have RECs which covered the entire geographical area under study. Ethical approval to conduct data collection in these areas was granted in writing from the HSE National Director for Mental Health. These approvals combined constituted ethical approval for the completion of the study at a national level in Ireland.

3.5.4 *Informed consent*

Informed consent to participate in the research was sought from all referred clients to the CIPC service at the assessment session during the data collection period. Before arriving for an assessment session all participants were informed by letter or by telephone that the CIPC service was conducting research in addition to the routine evaluation process which was part of normal service provision. During the subsequent assessment session clients discussed the research information letter with their therapists and were informed of the reasons the additional research was being conducted, the potential risks and that participation or non-participation would in no way impact their subsequent therapy in any way. If clients opted to participate in the research they were given copies of the consent forms if requested.

3.5.5 *Sampling technique*

As noted by Barkham et al. (2012) “In practice-based studies, criteria must be applied post hoc to compose plausible treatment groups.” (2012, p. 2). Such was the case in the current study, since in order to address particular research questions samples comprising participants who met certain criteria were required, e.g. attendance of a minimum numbers of sessions or for whom therapy endings were planned. Hence, for the purpose of providing pertinent outcome data not only to the CIPC service, but also to facilitate the comparison of outcomes for the CIPC service with similar primary care services in other jurisdictions, the initial total sample of all clients who consented to participate in this study was further divided into subsamples. The

process by which the sub-samples were formed followed broadly the same process as described by Barkham et al. (2012). These sub-samples were created by the incremental application of increasingly restrictive criteria to the overall number of participants who consented to the research, which reduced the numbers of clients in each resulting sub-sample.

The purpose of this was two-fold: firstly was to recreate comprehensive and realistic representations of the different groups who engage with primary care psychological services based on the defining characteristics of such groups as described in the literature and secondly, to use these groups in analyses to shed light on the behaviour of certain client and service variables recognised as important in large scale services such as CIPC. This is in keeping with the overall objectives of the study.

The initial sampling frame included all referred patients to the CIPC service who attended for an assessment session and consented to take part in the research within the study data collection period (N = 5,070). This sample was the subset of the population of medical card holders in Ireland referred as patients to the CIPC service by their GP or by other primary care health workers. Before further breakdown of this overall sample, all duplicate entries were removed (*n* = 45). Of these, a baseline sample was constructed of all clients that consent to participate in the research. For the remainder of the report this group will be referred to as the *All Participants* sample (N = 2,781).

Next, a set of criteria were incrementally applied to the all participants sample resulting in the formation of four sub-samples. The criteria were applied to each sample in an additive fashion, (i.e. beginning with the first criterion, each subsequent criterion applied was carried forward and applied to the next sample in addition to all previously applied criteria). All criteria are described below as is the order by which they were applied, beginning with the All Participants sample (N = 2,781). These sub-samples were constructed in order to adequately address all aspects of the research questions pertaining to the effectiveness of the CIPC service from a

practice based perspective; which assumes and unrestricted approach to the inclusion of all clients referred to the service and whom provide consent to participate.

The least restricted sample, the *Assessed Sample* comprised all participants from the all participants sample with a set of valid pre therapy CORE OM scores, i.e. with a maximum of one missing item from their pre therapy CORE OM questionnaire (N = 2,507). Second, an *Attender Sample* comprised all research participants from the assessed sample, but was constrained to only those from the assessed sample and who had attended at least one therapy session, i.e. an assessment session plus at least one additional therapy session (N = 2,151). Third, a *Form-returner Sample* comprised all participants meeting inclusion criteria in the attender sample, but was further restricted to those participants with valid post therapy CORE OM results (N = 1,397). Finally, a *Completer Sample* comprised research participants meeting all previously mentioned criteria but placed a further constraint on those participants for whom therapists recorded the context in which therapy ended as a planned ending (N = 1,335). However, the number of additional participants afforded by the formation of the Completer Sample was low (*n*= 62). Hence, except where otherwise stated, the analyses reported in the main text pertain to the Form Returners sample. Results for all of the sample groups are contained in Appendices B and C. See Figure 2 for full breakdown of the numbers of all research participants based on these selection criteria.

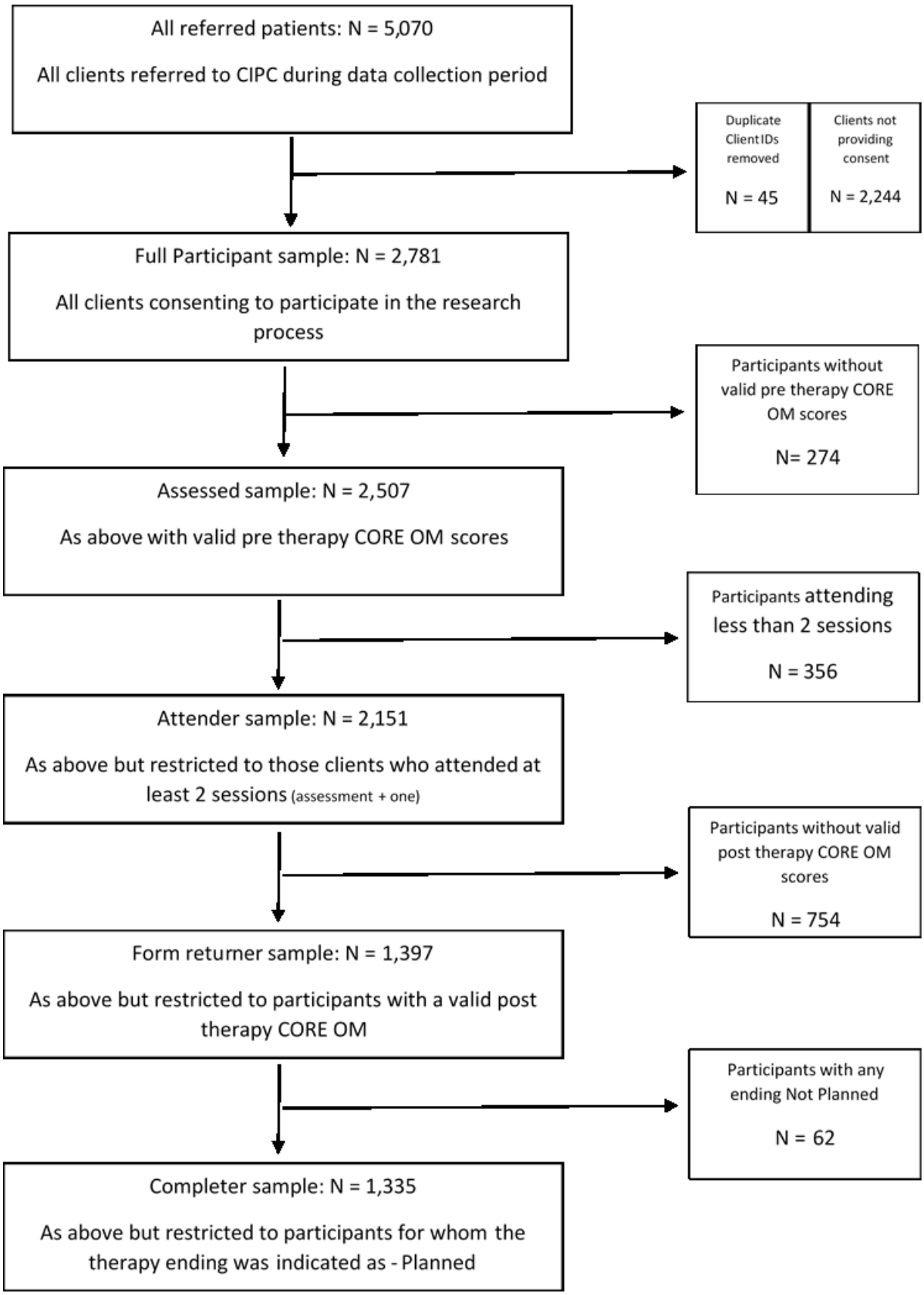


Figure 2: Study participation flow diagram showing sample groups and criteria used to exclude participants from each successive sample.

The primary objective of this study was to profile the national CIPC service and to determine its effectiveness. The Form Returner sub-sample represents the group of major interest to this study as those who attended the service for at least an assessment plus one therapy session and completed post therapy CORE OM questionnaires. The Form Returner sub-sample was analysed for all pre post outcomes analyses, Reliable and Clinically Significant Change (RCSC) and effect size calculations and is reported in the main text of this report. Analyses for the All Participants, Assessed, Attender and Completer sub-samples are included in appendices B and C.

3.6.1 *Procedure*

During the course of this study, the CIPC service migrated from a system whereby all clients' data were separately held and administered on secure computers in each of the CHO areas, to a fully centralised database to accommodate the entry and administration of all CIPC client data. The CORENET database holds all CORE OM questionnaire, assessment and end of therapy for all clients who referred to the CIPC service data. Upon referral to the CIPC service by a GP or other primary care mental health worker details pertaining to the referral are entered into the system. From that point on, all additional service, client and therapy related information is entered into the CORENET system. This use of this CORENET database for the collection of all client data was rolled out and adopted as CIPC policy.

3.6.2 *Data collection*

Normal CIPC service procedures stipulated a process of continuous service evaluation. Hence, all referred patients who attended an assessment session were asked to complete a 34 item CORE OM and completed a CORE Therapy Assessment Form and, when therapy had ceased an End of Therapy Form (described in detail below). In addition, all referred patients who attended for an assessment session during the data collection period were also invited to

participate in this study which required the completion an additional questionnaire, the HRQOL during their assessment session and also the last therapy session. During the assessment session, client questionnaire data were entered into the CORENET system either by the individual client or, because of physical ailments by their counsellor/therapists via electronic means, i.e. tablet, laptop or desktop computer in the same facilities where counselling was provided. In some cases, paper versions of questionnaires were completed and the data entered into the CORENET system at a later time. All questionnaires are included in appendices C, D and E.

3.6 Instrumentation

It has been recognised that when possible, it serves a wider range of prospective decision makers to report on a range of outcome measures including quality of life and to do so longitudinally (Barkham et al., 2010). This study includes such measures and reports the data over a 12 month time period for participants where possible. All participant demographic and service related information was obtained from the CORE Therapy Assessment (TAF) and End of Therapy (EOT) forms uploaded to CORENET by therapists.

3.6.1 *CORE Therapy Assessment form (TAF)*

The TAF collects service related information including some demographic information i.e. age, sex, employment status and ethnicity and the coded location of the therapy site, therapists ID, referrer type and date, assessment dates (first and last), the total number of assessments attended by the client, whether they have been a client of the service prior to the current assessment and if so the number (i.e. number) of the current treatment phase. Whether or not the current session is a follow up or review session and their relationship/support status is recorded along with an indication of whether or not they are currently attending other psychological services and if so, what kind and how many. Clinically relevant information is also gathered on the TAF including: a description of the reason for referral, and an indication

of the type and level of severity of the client's presenting problems. Problems are listed in categories, i.e. depression, anxiety, interpersonal etc. and the level of severity selected from one of the following options – *1 Causing minimal difficulty, 1 Causing mild difficulty, 3 Causing moderate difficulty, 4 Causing sever difficulty.* Level of risk as determined by the therapists is indicated by selecting one of the following options: *Mild, Moderate, Severe* from a list of categories that contains Suicide, Self-harm, Harm to others, Legal/Forensic. The final piece of client information asked for is if any actions have been taken on the part of the client to cope with or avoid the aforementioned problems and if so, what they were. Once complete, the therapist indicates the assessment outcome from a list of outcomes including *Accepted for therapy, Unsuitable for therapy at this time, Referred to another service* among others.

3.6.2 *CORE End of Therapy form (EOT)*

The CORE EOT form collects service level and clinical information. Some of this information is duplicated from the TAF such as site location (i.e. Site ID), but the number of sessions planned as decided at the start of the therapy in the contract with each client, as well as the final number actually attended and unattended are recorded as well as the type, mode and frequency of therapy. Therapy ending type (e.g. planned or unplanned) is also recorded here. A review of the presenting problems is undertaken of which the categories available are duplicated with one extra option, that of *0 Causing no difficulty* which is intended to capture the successful alleviation of that particular problem by way of the application of the respective treatment. Risk levels are again addressed here with the same intention of tracking any change from pre to post therapy. Motivation, Working Alliance and Psychological Mindedness are rated as contextual factors on a scale from *Poor*, *Moderate* or *Good*. Any perceived benefit of therapy can be either marked as *Improved – Yes* or *No* or *Not Addressed* from a list that includes Symptoms, Personal Insight/Understanding, and Day to day functioning among others. The client then indicates whether or not any change to their mediation status has

change and if so in what direction – increased/decreased or no change. The therapists then indicates whether they believe any change (if any has occurred) is of benefit to the client. Finally, the therapists indicates if the client has been offered a follow up session and how long (in months) it will be before it occurs.

3.6.3 *The CORE OM*

The CORE-OM (Clinical Outcome in Routine Evaluation; CORE System Group, 1998; Mellor-Clark, 2006) is a 34 item global measure of distress (see Appendix D) and is used extensively in clinical settings across primary and secondary services in Ireland, the UK and Europe. The CORE-OM 34 comprises 34 items addressing domains of subjective well-being (4 items), symptoms (12 items), functioning (12 items) and risk (6 items: 4 'risk to self' items and 2 'risk to others' items). Items are scored on a five-point scale from 0 to 4 ('*Not at all*', '*Only occasionally*', '*Sometimes*', '*Often*', and '*All or most of the time*'). 17 of the 34 items focus on low-intensity problems (e.g. 'I feel anxious/nervous') and 17 focus on high-intensity problems (e.g. 'I feel panic/terror'). Eight items are scored positively.

High completion rates of the CORE-OM in primary care counselling service settings attest to its usefulness as a brief measure in primary care settings (Barkham, 2005) and Evans et al. (2002) reported good levels of discrimination between clinical and non-clinical populations, an internal consistency of .94 (Barkham et al., 2001) and test-retest correlation over a one month period of .88 (Barkham, Mullin, Leach, Stiles, & Lucock, 2007). Reliability analysis results are reported in the result chapter.

3.6.4 *The Health Related Quality of Life Scale*

The Health Related Quality of Life (HRQOL: Moriarty et al., 2003) measure is used to assess perceptions of health status and activity limitation (Hagerty et al., 2001). The HRQOL-4 scale is a component of the Behavioral Risk Factor Surveillance System employed by the Centre for

Disease Control in the U.S. It has been validated as measuring distinct mental and physical health components (Hjermstad et al., 1998) and is shown to be a reliable measure of physical and psychological status (Moum, 1998). Additionally, the construct validity of the HRQOL-4 measure has been demonstrated in primary care populations in Sweden (Alfredsson, 2002), Canada (Ounpuu et al., 2000) and Puerto Rico (National Center for Chronic Disease Prevention and Health Promotion, 2002). The HRQOL-4 was evaluated on 14 criteria for determining the validity and usefulness with 21 other QOL indexes by Hagerty et al. and described as being "…quite satisfactory as a measure of the health domain in QOL," (2001, p. 14). Further, a systematic review of studies using the HRQOL (*n* =110) revealed "strong literature support for using Healthy Days to measure HRQOL among comparative populations, including those with and without health conditions." (Slabaugh et al., 2017, p. 15).

HRQOL data as a contributor to a more holistic profile (at the individual and aggregated levels) is recognised as potentially valuable to both clinicians and policy makers seeking to improve not only intervention outcomes, but also the overall quality of life of citizens (Slabaugh et al., 2017). Moreover, it is argued the HRQOL gives a more personal voice to a multi-dimensional concept of physical, mental emotional and social functioning as directly experienced by respondents which can fail to be captured by health agencies and professionals charged with profiling public health profiles. For example, in the U.S. while poor mental health may is not a diagnosable mental disorder in and of itself, it has a much higher prevalence rate in the general population than those mental disorders which can be labelled with a diagnosis. However, the rate of frequent mental distress (as calculated in the HRQOL) is a valid measure of poor mental health and as such has been utilised to investigate practical problems related to the primary care health system, e.g. the relationship between the cost of healthcare and the reluctance of older people to access primary care GP services (Segev et al., 2012) and other public health issues, e.g. the relationship between rates of frequent mental distress and rates of smoking in cancer survivors (Poghosyan et al., 2016).

The HRQOL-4 measure consists of four questions relating to 1) self-rated general health (physical and mental), 2) client rated number of days when their physical health and 3) mental health were not good, and 4) the number of days when activity was limited as a result of poor physical or mental health. These responses were analysed to generate an estimation of "healthy" and "unhealthy" days as perceived by clients during the 30 day period prior to the questionnaire being administered.

3.7 Data analysis approach

3.7.1 *Missing data imputation*

Missing data in the context of a large distributed service such as CIPC has been identified as causing problems related to the accurate reporting of outcomes and several sources have been identified as being responsible: the client, the practitioner and/or other administrative staff (Connell, Grant, & Mullin, 2006). The rate of missing post therapy outcome data in a service is usually directly related to client drop out as clients can leave the service without completing post therapy questionnaires. For most of the pre post and effect size analyses, the Form Returners sample was used. For samples containing participants in the study for whom post therapy scores were not available, a last observation carried forward (LOCF) approach (i.e. a fixed-value class of imputation) was applied in order to have both measures available for analyses. No data were replaced in the All Participants sample. The LOCF approach to missing data imputation used to replace missing post therapy scores in the Assessed and Attenders samples was the same as the method used by Barkham et al. (2012). LOCF was not required for the Form Returners and Completers sub-samples as no post therapy data were missing. Congruence with the method for imputation of missing data between the current study and that of Barkham et al. (2012) facilitates a more parsimonious comparison of outcomes between both studies. This is important since the current study seeks to both ascertain the

baseline performance of the CIPC service and also compare outcomes with other services of a similar service provision model, i.e. that reported in Barkham et al. (2012).

The sub-dividing of the main All Participants sample described earlier was for the purpose of observing how the constraining criteria affected outcomes for the various sub-samples. Hence, pre post statistical analyses, effect sizes and RCSC are reported for all samples when possible, i.e. using the LOCF technique described above to replace missing post therapy data for Assessed and Attender sub-samples. A full set of results for these samples are contained in the appendices (see appendices B and C) However, for all other analyses, the Form Returners sample was used and reported in the main body of the text. This sub-sample represents a large N (1397) with true pre and post scores (and considerable six month follow up score data (*n* = 240) while its overall *n* was not restricted by the criterion requiring planned endings as recorded on the End of Therapy forms.

3.8 Pre post outcomes, RCSC and effect size analysis approach

Statistical differences between pre and post CORE OM means were calculated using paired-samples t tests. This is a parametric test suited for use when comparing two means from groups containing the same individuals measured at two different time points. The main purpose of the test is to determine whether the mean difference between two measurements from a pair of observations is statistically significantly different from zero.

In addition, this study also utilised the *reliable and clinically significant improvement* (RCSI) index (Jacobson & Truax, 1991) which is discussed in more detail later. This method describes change as it pertains to a minimum of one statistical case in a sample which scores above the cut-off point. Effect sizes and RCSC proportions reported separately for the Form Returners sample in the main body of the text and in appendices B and C for all other sub-samples.

Effect sizes was calculated and reported in order to quantify the mean difference between pre and post measures. Effect size (ES) is a descriptive statistic which calculates a standardised

difference two sample means. In this study effect sizes were calculated and reported because *p* values from have been described as being difficult to interpret and particularly uninformative when reported in isolation (Sullivan & Feinn, 2012) and this statistic is frequently used to report differences between the pre and post measurements of improvement or deterioration in practice based studies of psychotherapy services in primary care (Barkham, 2005; Evans, Margison, & Barkham, 1998). In most primary care practice based psychotherapy research, ESs are calculated to produce a version of Cohen's *d* by subtracting the mean of the post-treatment test scores (*M*2) from the mean of pre-treatment scores (*M*1) and dividing the outcome by the pre therapy group (*M*1) standard deviation (SD) (Lipsey & Wilson, 2000):

$$Cohen's\ d\ =\ \mathrm{M1}\ -\ \mathrm{M2}/\mathrm{SD}_{\mathrm{pre}}$$

This method for calculating ESs (or standardized mean difference) has been used extensively throughout practice based psychotherapy research literature (see Barkham et al., 2012; Knapstad, Nordgreen, & Smith, 2018; Stiles, Barkham, Mellor-Clark, & Connell, 2008). However, this method has been shown to produce positively biased ESs (Durlak, 2009; Lakens, 2013). Other researchers have used a pooled SD value as the denominator in the Cohen's *d* calculation (e.g. Werbart et al., 2013b)

$$\mathrm{ES}\ =\ \mathrm{M1}\ -\ \mathrm{M2}/\mathrm{SD}_{\mathrm{pooled}}$$

Where $\mathrm{SD}_{\mathrm{pooled}}$

$$\sqrt{\frac{(SD_1^2 + SD_2^2)}{2}}$$

The problem with these methods of determining the size of the effect of an intervention is that this method assumes an independence of observations between pre and post groups (Lakens, 2013), which if not accounted for can also lead to an over-estimation of effect size

(Dunlap et al., 1996). Single group pre post research designs however, are based on the understanding of non-independence of observations between measurements for participants in the study. Hence, an alternative method of calculating and subsequent reporting of effect sizes for the samples in this study was considered appropriate.

An alternative approach for calculating effect sizes, which acknowledges the non-independent nature of the pre post data points for participants in single group repeated measure designs was described by Dunlap et al. (1996). This calculation incorporates results from a paired groups t-test statistic and the resulting the correlation (i.e. Pearson's *r*) which produces effect sizes "…consistently more moderate results than raw or severity adjusted methods…" (Seidel et al., 2014, p. 12). This method also demonstrated how ES could be calculated for multiple studies with an overall narrower range of resulting values, i.e. the calculation produced more conservative estimates by reducing bias caused by variability introduced by different choices of denominator in previous calculations (e.g. SD_{pre} above).

As part of the current study a comparison of effect sizes estimated using both these methods was considered prudent as recent literature has suggested appropriate alternative estimations can result in more moderate effects, less biased and more interpretable effect sizes (Seidel et al., 2014). The equations for calculating Cohen's *d* while accounting for the correlation between the same participants' measurements according to Dunlap et al. (1996) are:

$$\mathrm{d}p \;=\; t_c[2(1-r)/n]^{\,1/2}$$

Where *r* = the matched pairs t test correlation coefficient and *n* = matched pre post sample size.

and

$$t_C = M_{diff}/(SD_{diff}/n^{1/2})$$

Where t_C = the t statistic for the correlated observations, M_{diff} = difference between pre and post therapy scores and SD_{diff} = standard deviation of client change scores (Seidel et al., 2014).

Hence, the current study calculated effect sizes and reported results using both methods, i.e. unadjusted effects sizes and Dunlap et al. (1996) as a comparative exercise of interest in two areas: 1) for the purposes of comparing previously reported unadjusted effect sizes for other large scale primary care counselling/psychotherapy services (i.e. Barkham et al., 2001; Connell et al., 2007; Stiles, Barkham, Mellor-Clark, & Connell, 2008), and 2) to examine the range of effect sizes between both estimation methods to establish if the method devised by Dunlap et al. (1996) does produce more conservative estimations overall. Calculations for the single group repeated measures ESs were completed using an online calculator to compute Cohen's *d* from the test statistics of dependent t-tests using the calculations by provided by (Lenhard & Lenhard, 2016) based on calculations by Dunlap et al. (1996) and are denoted as $ES_{RepeatedMeasures}$ in the remainder of the text.

3.8.1 *Reliable and clinically significant change (RCSC)*

Prior to the early 1990s, the language used to deliberate symptomatic change as a result of psychotherapy was considered at best ambiguous (Evans, Margison, & Barkham, 1998). It could be said this was because of the different objectives of clinicians and researchers; the former being more interested in monitoring individual therapy progress, the latter in reporting the factors effecting outcomes at the group mean level. However, a method for assessing and reporting client change information which consolidated these divergent interests emerged from work by Jacobson, Follette, & Revenstorf (1986; 1984), Christensen and Mendoza (1986) and Jacobson and Truax (1991). Reliable and Clinically Significant Change (RCSC) as proposed by Jacobson and Truax (1991) summarised a technique for assessing an individual client's change along any continuous scale in the context of changes observed for an entire sample.

This allows for the comparison of outcomes between individual clients of a service or practitioner, as well as comparisons between services or populations based on proportions of clients meeting various different RCSC criteria. Importantly, calculation of the reliability criterion of RCSC (described below) incorporates both the respective scale's reliability and the study population's mean standard deviation and the ability to use these components, which were particular to the sample under study underscored the individuality of the concept (Schennach et al., 2016). Use of RCSC analyses has increased steadily since its introduction into psychotherapy literature and it is now the most widely used classification method of clinically significant change, although there are others (see Ronk et al., 2016). As this study intended to benchmark the effectiveness of the national CIPC service for the first time, use of RCSC analyses was considered necessary and appropriate in order to add to the knowledge base of outcome profiles of large nationally distributed counselling and psychotherapy services.

Jacobson and Truax's (1991) method of assessing RCSC classifies each individual client's therapy outcome based on the reliability of the pre to post therapy change on the CORE OM score and their end state post therapy, i.e. whether or not they had moved from the clinical to the non-clinical population. This method takes cognizance of the fact that all clients' score contain some amount of measurement error, or standard error (SE) which, along with the 1-month test-retest reliability value of .88 as reported by Barkham et al. (2007) is used in the calculation for the reliable change index (RCI) criterion, i.e. the reliability element of RCSC. Hence, each client is assessed on both the amount of change observed (is the change reliable) and their final end point state post therapy (is the change clinically significant or have they crossed a predetermined point on the mean outcome scale for the particular sample).

For this study, RCI criterion, overall change value and the final end points were used to calculate an RCSC for each client and the proportions of RCSC were reported for the entire sample under examination. The first task of calculating RCSC was establishing whether the

amount of change observed for each client was in excess of that which might have been attributable to measurement error, i.e. was the amount of change observed due to actual change or could the change be the result of variability caused by instrument itself? In order to assess the measurement variability (known as the reliable change index (RCI)) a variation on the standard error of measurement which utilises the difference between two measurements (e.g. pre and post therapy), known as the standard error (SE) of the difference which was represented by the following formula:

$$SE_{diff} = SD1\sqrt{2}\sqrt{1-r}$$

Where *SD*1 is the standard deviation of the participant's pre therapy score and *r* is the Cronbach's alpha (i.e. reliability coefficient) of the instrument. According to Jacobson and Truax (1991) any change value which exceeds 1.96 * SE_{diff} is unlikely is unlikely to have occurred due to the unreliability of the measure in question in more than 5% of cases.

Standard error of the difference (SE_{diff}) which is represented by the following formula:

$$SE_{diff} = SD1\sqrt{2}\sqrt{1-r}$$

Where *SD*1 = standard deviation of the measure and *r* = test-retest reliability of the measure

$$SE_{diff} = 0.59\sqrt{2}\sqrt{1-0.88}$$

$$SE_{diff} = 0.59 * 1.141 * 0.346$$

$$SE_{diff} = 0.240$$

According to Jacobson and Truax (1991) any change value which exceeds 1.96 * SE_{diff} is unlikely to have occurred due to the unreliability of the measure in question in more than 5% of cases.

RCI = 1.96*0.240

RCI = 0.47

Once the reliability of each client's change in pre to post therapy measurement has been determined (in the case of this study a pre post change of 0.47 in either direction on the CORE OM scale), the next consideration in calculating RCSC is assessing whether any improvement is clinically significant. The purpose of what of Jacobson and Traux referred to as *clinical change* was intended as "a precise method for classifying clients as "changed" or "unchanged" on the basis of clinical significance criteria." (Jacobson & Truax, 1991, p. 13). This clinically significant change criterion score was dependent on whether or not a client's post therapy CORE OM score changed, moving them from within the range of the clinical population to that of the non-clinical population. The method of determining whether this is the case for each participant depends on the demarcation point calculated for a clinical sample in relation to a non-clinical sample, which can be used to separate both populations (Evans et al., 1998).

The clinical cut off used in this study is that recommended by Connell et al. (2007) of 1.00 on the CORE OM scale. While it has been commented that the value at which to apply a cut-off point on the CORE OM scale is important and has implications for other indicators of improvement (i.e. RCSI) (Barkham et al., 2012), as a cut off this value of 1.00 (or 10 depending on the scale format used by the study in question) has been widely adopted in many practice based studies using the CORE OM scale (see Cahill et al., 2010, p. 442). Hence, it is of relevance to this service and the use of this cut-off value in this study may also provide useful information to services in other jurisdictions. For these reasons the number of clients scoring above and below the cut-off point is calculated and reported.

Hence, once the reliability index and the cut-off point between the clinical and general population were established using the techniques described above, each client's end point status was determined as residing in either the *clinical* or the *non-clinical* population and their change amount as reliable or not. Both reliability and significance (improvement) criteria must

be met in order for a client to be considered reliably and clinically significantly improved. The proportions of participants meeting each of the following criteria in the *Form Returners* sample were reported in the main text:

- *Reliable and clinically significant improvement (RCSI)* = Both the reliability and clinical significance criteria are satisfied
- *Reliable improvement only (RIO)* = Only the criterion of reliability is satisfied, i.e. pre to post therapy scores improve by at least an order of 0.47*(RCI)
- *No reliable change (NRC)* = Neither criteria satisfied
- *Reliable deterioration (RD)* = Participants whose pre to post therapy scores were worse by at least an order of 0.47*(RCI)

Finally, in order to calculate the RSCI rate for those samples without completed and valid CORE OM forms at the pre and /or post therapy stages the following processes were undertaken:

- *Attender and Assessed samples* – in order to calculate a mean pre to post change score for this sample, pre-treatment CORE OM scores were carried forward for those participants for whom they were missing. This is consistent with the Last observation Carried Forward (LOCF) method as used in Brakham et al., (2012).

Differences between pre and post treatment CORE OM scores were determined using two different methods in order to meaningfully differentiate between statistical and clinically relevant changes, while also facilitating easy comparison between studies and also outcomes from different jurisdictions.

3.9 Describing the relationship between the number of sessions attended and client outcomes

Gyani et al. (2013b) found that those clients in the IAPT service who attended a higher number of treatment sessions had a better chance of reaching recovery. Hence, a logistic regression was used to investigate participant and service level variables that predicted achieving RCSC. Logistic regression analyses are used to predict membership of a category from a number of predictor variables and compute the log odds that a particular outcome will occur. Results are

expressed in the form of the ratio of the probability of an event occurring, or not occurring, which in this case was that participants belonged or did not belong to the group who achieved RCSC based on each of the predictor variables included in the model while controlling for all others.

The following predictor variables were included in the model; the total number of sessions attended by the participant (i.e. the duration of the treatment) and the pre therapy CORE OM score. For independent variables measured on a continuous scale, e.g. the number of session attended, the odds ratios described the amount change in the likelihood of participants achieving RCSC for every one unit change in the total number of sessions attended (from the mean number of sessions). Particular independent variables were chosen for inclusion in the binary logistic model because they were related to indicators of success in primary care counselling services. For example Saxon, Ivey and Young (2008) collated a number of variables from the CORE Assessment and CORE End of Therapy forms along including the Number of Sessions (as indicated by counsellor/therapists) along with CORE scores in an attempt to build a predictive model of poorer outcomes. They conducted tests of difference (both independent t-tests and Chi square) in order to assess the suitability of each of the proposed variables for inclusion in the eventual logistic regression model used to predict poorer outcome. Bearing this in mind, similar analyses were conducted on a number of variables including those from the study by Saxon, Ivey and Young (2008) as the overall profile of their study is similar to that of the current study (i.e. sample make-up, outcome measure, source of the independent variables included in the model and primary care dataset).

3.10 Examining the role of Initial severity

For the examination of the relationship between initial severity and CORE OM outcomes, first an analysis of the correlation between pre and post therapy CORE OM scores was performed in order to statistically determine whether they were related to one another, and if so the

nature of that relationship (i.e. positive or negative). This was followed by regression analysis of severity levels as indicated by counsellor/therapists ratings on CORE OM pre therapy scores in order to observe whether there was a congruence between these determinations of client symptomology. Next, severity categories were created based on three groups: non-distressed (≤1.00), mildly, moderately, or moderate-to-severely distressed (scores between 1 and 2.5), and severely distressed (≥ 2.5). These severity category bands were shown by Barkham et al. (2005) to have good discriminative validity between groups of patients recieving psychological services delivered in primary and secondary care settings.

A 2*3 (Time*Severity) within-subjects ANOVA was conducted (where Time was the between subjects factor and severity was the within subjects factor) in order to determine if there was a differential effect of pre therapy severity on participants' pre to post therapy CORE OM scores.

3.11 Examining the number of sessions and initial severity affect outcomes 6 months after therapy has ended

As recommended by Heck et al. (2014), before constructing the model building process a visual inspection of the pattern of pre, post and six month CORE OM data plots was undertaken in order to determine the general shape of the growth trajectory. This also served to aid the selection of the appropriate polynomial term to be applied to each participant's change pattern and was added into the data set as a variable. Previous meta-analytic research would suggest that the positive effects of short term psychotherapy tend to recede in the months after therapy has ended (Bower et al., 2011). Further, in keeping with the literature on psychotherapy outcome change, the term "growth curve" was replaced with "decay curve" reflecting the fact that change is represented negatively accelerating curve or downward trend. This is in keeping with the CORE OM used in this study, which considers progressively lower scores as correlated with decreasing levels of symptomatology, i.e. improvement. This approach is intended to make interpretation of the results more intuitive.

A decay curve model, also referred to as a multi-level mixed effect model (MLM) with 2 levels (Peugh, 2010) was proposed in order to examine the relationship between changes in CORE OM scores and baseline severity and the total number of sessions attended – pre therapy, post therapy and at six month follow up time-points. Like all statistical models, MLMs make predictions based on the mean of the outcome variable and one or more predictors. However, there are several conceptual reasons this approach was chosen over other statistical approaches.

First, a basic assumption of a linear regression model and other modelling approaches based on the ordinary least squares (OLS) approach is an independence of responses. However, studies such as this one include the results of multiple completions of the same questionnaire items (over different time points), rendering each respondent's set of responses inter-dependent, not independent. Analyses of independent and non-independent data should be approached differently since the overall error term for non-independent data is more complex and composed of several elements (Brauer & Curtin, 2018). MLMs can take into account inter-group and intra-person variation simultaneously and within the same model and determine how much variation in the data can be explained by differences between these. This is in contrast to single fixed-effect approaches to subject-analyses and item-analyses whereby the variation in each of these sources is performed in isolation and averaged across all subjects and items respectively (Winter, 2013).

Second, MLMs also allow for the inclusion of variables as "fixed", i.e. the values of these data are generally the same from study to study and have been gathered from all levels of the variable of interest (Brauer & Curtin, 2018) and "random", i.e. variables included when the intention of the analyses is to suggest inferences about the general population of interest (Dickinson & Basu, 2005). MLMs incorporate these random effects as variation associated with

different baseline values for each participant which can be modelled and estimated as individual random intercepts in the model (Winter, 2013).

Third, MLM also incorporates the ability for researchers to specify level 1 and level 2 covariance structures that best represent the nature of the repeated observations (Heck et al., 2014), i.e. how the covariance matrices are dimensionalised at each of the different levels. Finally, non-linear trends in outcome data can also be explored using MLMs. Adding higher order polynomials to MLMs facilitates the examination of differential rates of individual change to account for the effect of factors other than time alone (Heck et al., 2014).

3.11.1 *Multi-Level Model Specification*

In order to explore the effects of time and baseline severity on COREOM scores a longitudinal multilevel analysis of the relationship between participants' CORE OM scores, pre therapy severity levels and the total number of sessions attended using maximum likelihood estimation was conducted in IBM SPSS© 24.0. Following the process as outlined by Heck Thomas and Tabata (2014) the nature of the relationships the level 1 model was specified as follows:

$$\mathrm{COREOM}_{1j} = \beta_{0j} + \beta_{1j}\mathrm{COREOMSEV}_{ij} + SESSIONS_{ij} + \varepsilon_{ij}$$

Where COREOM_{ij} represents the COREOM score for participant i at time j; β_{0j} represents the continuous participant specific intercept, or baseline COREOM score; β_{1j} represents the participant-specific slope of change in COREOM scores over time; $\mathrm{COREOMSEV}_{ij}$ represents the categorical participant specific severity level at baseline (i.e., non-distressed, mild to moderately distressed, or severe); $SESSIONS_{ij}$ represents the total number of sessions attended by participant i at time j; ε_{ij} is the error term, i.e. the deviation of the observed COREOM scores from the predicted scores of each participant i and each measurement session$j$.

COREOM was modelled as a function of the level 1 predictor TIME (i.e. pre therapy, post therapy and six months after therapy) and the level 2 predictors baseline severity level (i.e., normal, moderate, or severe) and total number of sessions attended.

The level 2 model was specified as:

$$\beta_{1j} = \gamma_{10} + \gamma_{11}\mathrm{COREOMSV}_{0j} + SESSIONS_{ij} + U_{1j}$$

Where γ_{10} represents the overall slope of change in CORE OM scores for all participants; γ_{11} represents the deviations from γ_{10} accounted for by the fixed variables $\mathrm{COREOMSV}_{0j}$ and $SESSIONS_{ij}$; and U_{1j} represents the unexplained deviation from the participant-specific slopes of CORE OM change. The participant-specific slope of CORE OM change (β_{1j}) is therefore modelled as a function of the overall slope of CORE OM change, plus the deviations from the overall slope, plus the unexplained deviation from the participant-specific slope. Baseline severity level ($\mathrm{COREOMSV}_{0j}$) was chosen as a predictor variable because previous research suggests that baseline severity level can affect the course of improvement (i.e., β_{1j}) among primary care patients (Elkin et al., 1995; Simpson et al., 2000). The total number of sessions attended was chosen as a predictor of CORE OM scores because previous research suggests that up to a certain number, more, as opposed to fewer sessions attended by clients of primary care counselling services is associated with better outcomes (Falkenström et al., 2016; Hansen et al., 2006; Lambert et al., 2001). Also, the initial severity of patients' problems has been implicated in the degree to, and rate at which improvement occurs after therapy (Elkin et al., 1995) and at follow up (Bryan et al., 2012).

3.11.2 *Model generation approach*

Included in this analysis were all participants in the Form Returners sample who provided CORE OM data at three time points: pre therapy, post therapy and six months after therapy had ended (*n* = 240). Pre therapy severity was rated by their counsellor/therapist after

assessment and the total number of sessions attended was reported on each participant's end of therapy form. As demonstrated by Heck et al. (2014) pre, post and follow up related outcome measurement patterns can be best captured by a quadratic polynomial term when the required function is to "…test for the presence of a change in the rate of growth occurring over time." (2014, p. 173). Hence, two transformed time variables were added to the models: one linear and one quadratic in order to represent this aspect of change. This process also reduced potential collinearity problems as outlined in Heck et al. (2014).

In order to best ascertain the influence of the predictors in the models, the analyses was conducted in a number of steps. First a basic unconditional model with no level 2 predictors was constructed in order to estimate a basic model with which to compare subsequent models (Null Model). Generating such a model provided the opportunity to examine the amount of variance in the outcomes partitioned between the within and between-group elements of the sample by calculating the intra-class correlation (ICC) coefficient using the formula $ICC = \tau00 / (\tau00 + \sigma^2)$ (Peugh, 2010), where $\tau00$ represents the variance between time points for CORE OM scores and σ^2 represents the between participant variance of CORE OM scores. This aided the decision to continue exploring further model generation and examination. Next, the linear and quadratic time components were added (Model A) in order to analyse the effect of time only on CORE OM scores and determine is a quadratic term in the model provided a better overall fit to the shape of the decay curve. This was also a method to capture how the relationship between time and CORE OM scores was itself changing. This was followed by the introduction of a random slope for the time variable into the model (Model B) to allow the time variable to vary randomly across all participants in the sample (i.e. at level 2) as well as the predictors; pre therapy severity level (i.e. mild, moderate and severe) and total number of sessions attended. Further, in order to account for covariance between the randomly varying slopes and intercepts for all participants, an unstructured covariance type was applied at level 2 of the model and a diagonal structure at level 1. Comparison between models was performed by

calculating the deviance of each nested model from the null predictor model per differences in degrees of freedom (Heck et al., 2014).

3.12 Examining differences in outcomes between and therapists' self-declared therapeutic approaches

Counsellor/therapists in the CIPC service indicated the type of therapy approach taken with each client on the CORE end of therapy form once they had determined that the current episode had ended. The type of therapy undertaken was selected from a list of options as shown in on the End of Therapy form (Appendix F p. 226). Once inputted into the CORENET web based client information database, this is the only occasion and location this information is recorded as the Therapy Type for each client episode.

As discussed in the literature review, the CIPC service did not perform adherence checks on the therapy conducted by counsellor/therapists. It is possible that CIPC therapists employed either one single approach or a range of therapeutic approaches with clients depending on the requirements and preferences of each individual client who presented to the service for therapy. It could also be the case that counsellor/therapists applied a therapeutic approach they were most comfortable with or for which they had received the most up to date training. The focus of enquiry for this part of the study was not verify the fidelity of adherence to specific approaches, nor to address questions of the superiority of different approaches over others. This questions asks whether outcomes were different for participants who received a single therapeutic approach or a combination of at least two different approaches as declared by their counsellor/therapist once therapy had ended.

In order to examine the effect of single therapy approaches versus more than one therapy approach this study replicated the approach reported in the literature review by Stiles et al. (2006; 2008). Six groups of participants were formed from participants in the Attender sample (n = 2,151) in order to compare their outcomes. First, three groups which were indicated by

their therapist as having received only one single therapy approach; either CBT (included only those participants who received either Cognitive, Behavioural or Cognitive Behavioural only), INT (included only those participants who received Integrative therapy only) or PDT (included those participants who received either Psychodynamic or Psychoanalytic therapy only). Next, three separate groups were formed each comprising those participants who received one of the single approaches plus one other therapy approach which was not one a single approaches already described (i.e. one of the following treatments Supportive therapy, Art therapy, Integrative therapy, Systemic therapy, Structured Brief therapy or Other therapy type).

First, a one-way analysis of variance (ANOVA) compared pre-therapy means across all six groups to determine if any began treatment with significantly different pre-treatment CORE OM scores. This was followed by a repeated-measures (pre-treatment v. post-treatment) ANOVA, with treatment approach (CBT v. PCT v. PDT) and +1 approaches as fixed factors to examine whether there was a pre post interaction effect.

Finally, a Chi-square test was conducted in order to determine whether any significant differences in RCSI rates between all of the treatment conditions was detected. Importantly, before this test was conducted, all participants below the clinical cut-off as determined by the RCI calculation for this study were removed. This was done to remove those participants who could not fulfil the criteria necessary to achieve RCSC, i.e. move from above to below clinical cut-off (Jacobson & Truax, 1991).

3.13 Examining the relationship between changes in CORE OM scores and changes in health related quality of life (HRQOL)

Based on the use of the HRQOL in past literature, most pertinent to this study were four HRQOL outcomes; 1) changes in participants' perception of their overall general health, 2) the total number of perceived physically unhealthy days in the previous 30 day period, 3) the total number of perceived mentally unhealthy days in the previous 30 day period and 4) the

number of days during the same period that participants' felt their normal levels of activity were limited because of their reported problem.

Hence, with the exception of item 1 on the HR QOL which deals with self-perceived general health, these data are numerical and constricted to a maximum of 30, i.e. the maximum number of days in a month. Hence, analysis of these data consists of reporting pre and post comparisons of these numerical data as mean differences using repeated measures t-tests in order to assess any increases or decreases from pre to post therapy in terms of days and if any differences were statistically significant.

Correlational analyses were conducted next in order to examine the relationship between pre therapy CORE OM scores and HRQOL pre therapy scores. The correlations between these variables and the CORE OM pre therapy scores were analysed in order to determine whether these relationships were congruent within the context of the literature review presented earlier. For example, the direction of the associations should be in keeping with the broad objectives of both measures and congruent with each other, e.g. participants with higher levels of symptomatology according to the CORE OM should also display increasing numbers of days when their mental health was indicated as increasingly not good. This step can be viewed as a preliminary assessment of congruent validity between the CORE OM and the HRQOL scales.

Next and following on from the results of the correlational analyses a linear regression analysis was conducted. This analysis used only the variables observed as statistically significantly correlated to the CORE OM in the previous analysis as predictors. The objective was to explore the level to which the pre therapy HRQOL items Q2, Q3 and Q4 could predict which participants would have elevated pre therapy CORE OM scores. These HRQOL items were also used in a further model to assess their predictive value to identify participants who achieved RCSC.

Finally, a binary logistic model was constructed using the HRQOL items Q2, 3 and 4 as predictor variables into the model as Block 1 and adding a set of demographic and service related variables as Block 2 variables (i.e. Sex, being on medication and being unemployed). Both sets of variables were regressed on post therapy RCSC achievement as the dependent outcome variable. The rationale for performing this analysis as a two-step process is as follows. Both of the steps involve different variable group types which add different aspects of client pre therapy information. Step one related only to the HRQOL items Q2 and Q3 and are included for the purposes of assessing their predictive value in a more complex overall model. The second block of independent variables was made up of demographic variables and one service related variable which have been linked to primary care outcomes in some way, but also used in other HRQOL studies as covariates: gender (Cook & Harman, 2008; Ogrodniczuk & Staats, 2002), being on medication (Fournier et al., 2009; Slabaugh et al., 2017), being unemployed (Chin et al., 2015; Shih & Simon, 2008). Hence, the two-step process is intended to deliver predictive information regarding two distinct aspects of participant outcomes: 1) how well the HRQOL items Q2 and Q3 predict achievement of RCSC for participants and 2) how well the model predicted RCSC achievement once the HRQOL variables are included in the overall model.

3.14 Examining employment status or gender in relation to client outcomes

A one-way between-subjects analysis of variance was conducted in order to identify whether being male or female, employed or unemployed contributed significantly to the variance in pre to post CORE OM score differences for the participants in the study.

Chapter 4: Results

Introduction

As emphasised in the previous chapters, this practice based study focussed on providing a detailed profile and answering questions relevant to the national CIPC service counselling service. The results which follow address the questions outlined in the literature review and are methodologically the result of the processes and approaches proposed in the methods chapter.

First, a description of the samples and sub-samples used for the calculations of ES and RCSC is provided in terms of their demographic and characterological breakdown. The results of the ES and RCSI calculations are presented followed by analyses of the relationship between participant outcomes and the number of sessions they attended before leaving the service. The prevalence of differences in outcomes between participants attending different therapy approaches are reported next, followed by the relationship between outcomes and HRQOL data. Next, the question of whether the initial severity of presenting problems of participants was predictive of the number of sessions they ultimately attended is addressed followed by the effect of gender and employment status on outcomes. Finally, a series of multi-level models were constructed and the question of the longer terms effects of therapy (up to six months) and the influence of initial severity and the number of sessions attended on these outcomes was addressed, the results of which are presented at the end of the chapter.

4.1 Descriptive statistics

The first section of the results addresses a question pertinent to the current study and also the wider CIPC service regarding the characteristics of the overall sample from a demographic and clinical presentation perspective, but also from a broader requirement to build a service

profile regarding the nature of the problems clients present with at the assessment stage of service provision.

Table 2

Demographic information for All Participants and Form Returners sample

	All participants (*n* = 2,781)	Form returner sample (*n* = 1,397)
Patient demographics		
Age *M (SD)*	42.6 *(14.6)*	43.6 *(14.3)*
Sex		
Female % (*N*)	74.9 *(2082)*	73.4 *(1026)*
Male % (*N)*	25.1 *(699)*	26.6 *(371)*
Ethnicity % (*N*)		
White Irish	73.2 *(2036)*	74.7 *(1044)*
White Irish Traveller	0.6 *(18)*	0.4 *(6)*
Any other White Background	11.4 *(317)*	0.5 *(7)*
Black, or Black Irish - African	0.5 *(14)*	0.4 *(5)*
Black, or Black Irish - any other Black background	0.5 *(14)*	0.5 *(7)*
Asian or Asian Irish - Chinese	0.3 *(8)*	0.3 *(4)*
Asian or Asian Irish - any other Aisan background	0.3 *(9)*	0.4 *(5)*
Other including mixed background	1 *(27)*	0.9 *13)*
N/A	12.1 *(337)*	10.9 *(152)*
No Data	0.1 *(1)*	0.1 *(1)*
Employment status % (*N*)		
Employed	32.7 (*910*)	30.2 (*422*)
Unemployed and seeking work	12.6 (*351*)	13.7 (*192*)
Student	5.6 (*155*)	6.2 (*151*)
Long Term Sick, Disabled or Benefits	15.5 (*430*)	17 (*237*)
Homemaker not working or actively seeking work	16.9 (*471*)	15.7 (*220*)
Not receiving benefits and not working or actively seeking worl	0.4 (*11*)	0.5 (*7*)
Unpaid voluntary work, not working or actively seeking work	0.8 (*23*)	1.1 (*16*)
Retired	7.3 (*202*)	7.8 (*109*)
N/A	8 (*223*)	7.6 (*108*)
No Data	0.2 (5)	0.1 (2)

Of all 5,025 valid patients referred to the CIPC service during the course of the study, 44.5% (*n* = 2,244) did not provide consent to participate in the study, while 55.6% did select to participate (*n* =2,781). The was an almost perfect 1/4 to 3/4 ratio split between men and women with the average age of all participants over 42 years old (i.e. 42.6 years). The vast majority (84.6%) declared as either White Irish or other White background, with White Irish Travellers forming the largest group of minority ethnicities (0.6%) followed by Black, or Black Irish – African, Black, or Black Irish - any other Black background (1%), Other including mixed background (1%) and Asian or Asian Irish – Chinese, Asian or Asian Irish - any other Asian background (1%). A significant number of participants had a designated ethnicity of Not Applicable (12.1%) and one participant had no data in this regard.

Almost a third of participants declared as employed at the time the study took place (32.7%). Homemakers formed the next largest group (16.9%) followed by those who were Long Term Sick, Disabled or on Benefits (15.5%). Unemployed participants made up 12.6% of all participants and more retirees than students were represented (7.3% and 5.6% respectively). A significant minority (8%) of participants were not recorded alongside any particular employments status and were marked as N/A, while data were missing for this demographic element for 0.2% of participants.

A series of chi-square tests of independence and t-tests were performed in order to compare study participants with non-participants along a number of assessment factors. Inspection of the bar chart showing the distribution of the age variable across all referred patients indicated that age was not normally distributed, hence an independent samples Mann-Whitney U test was conducted and indicated a significant difference in age between study participants; who were younger (*Mdn* = 41 years) than non-participants (*Mdn* = 44 years), *U* = 2642159.000, *p* = <.001.

Participants did not differ significantly from non-participants in terms of gender: χ^2 (1, *N*= 5,025) = 2.15, *p*=.142), but did in terms of employment status: χ^2 (8, *N*= 5,025) = 195.91, *p* <.001) where there were almost twice as many participants as non-participants associated with being Employed (18.1% versus 10.6%). However, the association was not strong Φ = 0.198 and accounted for 3.9% of the overall variance. There were also significant differences in ethnicity between participants and non-participants: χ^2 (9, *N*= 5,025) = 117.49, *p* <.001). Again the association was weak, Φ = 0.153 accounting for 2.3% of the overall variance.

A visual inspection of the distribution of available pre therapy CORE OM (*n* =4532) scores confirmed a normal distribution. A subsequent independent t test showed that mean CORE OM scores at assessment differed to a statistically significantly degree between participants (*M* = 1.57, *SD* = .61) and non-participants (*M* = 1.53, *SD* = .66), *t*(4532) = 2.02, *p* = .043. In real

terms however the clinical significance was negligible and the *p* value was very close to non-significance.

In terms of the severity of symptoms at assessment stage, significant differences were detected between participants and non-participants presenting with problems known to affect the majority of those presenting to primary care psychotherapy services, i.e. depressive and anxiety related symptoms (see Barkham et al., 2001; Connell et al., 2007; Chris Evans et al., 2002; Stiles, Barkham, Mellor-Clark, & Connell, 2008). A chi-square test showed the relationship between study participants and being identified at some level of severity for depressive symptoms was significant: χ^2 (1, *N*= 5,025) = 49.34, *p*<.001) and there were over 10% more study participants with such an indication than non-participants (32.8% versus 22%), the association was weak however Φ = -0.99 accounting for 1% of the variance. This was also the case for those 44.4% of all study participants who presented with levels of anxiety as opposed to 29.8% of non-participants, the chi square results was also significant for these groups: χ^2 (1, *N*= 5,025) = 120.55, *p*<.001), Φ = -0.155 accounting for 2.4% of the variance.

4.1.1 *Presenting problems*

Therapists working for the CIPC service during the study period indicated the type of presenting problem they thought was the most significant contributor to their clients' presenting problems. Multiple presenting problems were frequently identified by therapists for individual participants, resulting in the total percent of cases being in excess of 100% of the actual number of participants in the study. The most frequent determination of severity was in the moderate category, followed in order by mild, severe and minimal.

The breakdown of indicated problems recorded during assessment for the All participant sample (*N* =2,781) is shown in Table 3. As described in the methods chapter, counsellor/therapists indicated the presence of particular problems from a list provided on the therapy assessment form (TAF) (see Appendix D) during, or shortly after the initial assessment

session with each client. The most commonly reported type of problem recorded was anxiety stress with 89.2% of participants recorded as presenting with some level of severity for this problem type. Next, levels of depression were indicated for 67.3% of participants, followed by problems in interpersonal relationships at 51.6% and self-esteem at 44.4%, Bereavement/loss accounted for severity ratings for 38.2% of all participants, while for physical problems and trauma this was 29.3%. The remaining problem types accounted for a range of 8.4% to 25.8% of participants from the cognitive learning to the living/welfare categories respectively. Across all problem types a significant number of participants (4.5%) were indicated as presenting with a level of severity which was indicated as none.

Table 3

Presenting problems - All Participants sample (n = 2,781)

	Level of severity n *(%)*						
Problem type	Mild	Moderate	Severe	No Data	None	Total indications for problem type	Total % indications for problem type
Anxiety stress	618	1134	480	241	8	2481	89.2
Depression	790	685	174	190	34	1873	67.3
Interpersonal relationship	414	603	227	170	20	1434	51.6
Self-esteem	391	501	183	153	6	1234	44.4
Breavement	275	383	208	168	28	1062	38.2
Physical Problems	222	271	177	116	30	816	29.3
Trauma	243	264	147	129	33	816	29.3
Living/Welfare	227	235	99	125	31	717	25.8
Work/Academic	203	149	68	115	32	567	20.4
Addictions	71	39	21	127	57	315	11.3
Eating Disorder	62	36	26	116	60	300	10.8
Psychosis	33	40	13	148	59	293	10.5
Other	12	25	14	134	54	239	8.6
Personality Problems	38	17	5	123	53	236	8.5
Cognitive /Learning	33	15	7	121	57	233	8.4
Total number of indications *(% of total N)*	3632 *(28.8)*	4397 *(34.9)*	1849 *(14.7)*	2176 *(17.2)*	562 *(4.5)*	12616	

Table 4 shows the problem severity breakdown for the Form Returner sample (*n* = 1,397). With the exception of some slight variation of the order of the final five severity categories in the table, it mirrors the numerical order of presenting problems as observed in the All Participants sample. As with the data in Table 3 showing the breakdown of problems and severity for the All Participants sample, the overall number of indications far exceeds the

number of participants because counsellor/therapists regularly indicated multiple problems types and severities against individual participants.

Table 4

Presenting problems - Form Returners sample (n = 1,397)

	Level of severity n *(%)*					indications for problem	indications for problem
Problem type	Mild	Moderate	Severe	No data	None	type	type
Anxiety Stress	276	608	208	-	33	1125	80.5
Depression	312	351	74	-	97	834	59.7
Interpersonal relationship	185	329	103	95	37	749	53.6
Self-esteem	150	283	80	82	26	621	44.5
Bereavement Loss	107	200	95	101	23	526	37.7
Physical Problems	89	144	82	84	36	435	31.1
Trauma Abuse	98	123	68	92	20	401	28.7
Living/Welfare	91	130	36	89	27	373	26.7
Work /Academic	84	77	35	84	29	309	22.1
Addictions	21	15	8	101	16	161	11.5
Other	6	11	4	108	-	129	9.2
Eating Disorder	12	5	3	94	13	127	9.1
Cognitive/Learning	11	1	1		7	118	8.4
Personality Problems	1	-	1	98	6	106	7.6
Psychosis	1	-	-	-	1	2	0.1
Total number of indications *(%)*	1444 *(24)*	2277 *(37.8)*	798 *(13.3)*	1028 *(18.7)*	371 *(6.2)*	6016	

4.1.2 *CORE OM reliability result.*

The internal consistency of the CORE OM scale was calculated on the all participants sample using Cronbach's (α) coefficient (Cronbach, 1951) and was found to be highly reliable 0.95 (*n* =1396)

4.1.3 *Missing data handling*

Missing post therapy data for the Assessed and Attender samples were imputed on a last observation carried forward basis (LOCF). This is in keeping with the method used by Barkham et al. (2012) which is the comparative point of reference for effect size and main outcomes for the current study. The remaining sample sub groups (i.e. Form Returners and Completers) did not require data imputation as the criteria by which they were compiled included the requirement for complete pre and post therapy outcome data.

4.1.4 *Attrition*

The All Participants samples (*n* = 2,781) was used to determine the proportion of participants indicated by counsellor/therapists as terminating their therapy early, unexpectedly or without prior consultation. In accordance with normal CIPC service, the end status of each participant's episode of engagement with CIPC was indicated by therapists as either "Planned" of "Unplanned". According to the CORE System User Manual (CORE System Group, 1998) planned endings are indicated when therapy ends after a particular number of sessions have taken place was agreed between the client and the therapist prior to the commencement of therapy or at a point during therapy when both parties agreed to end. Unplanned endings are indicated when therapy ends due to a crisis on the client side, the client does not want to continue due to some other reason or the client fails to attend ongoing therapy without giving any reason. All planned endings are represented in a single category below and unplanned endings are separated into their various categories in Table 5. Importantly, a significant number of participants from the All Participants sample were indicated as having planned endings while also being indicated with other statuses which were inconsistent with them being part of the Completer sample, i.e. less than 2 sessions attended, insufficient pre or post therapy CORE OM items completed. As described in the Method section, the Completer sample (*n* = 1,335) represented all participants from the All Participants sample meeting the following criteria; age $\geq$ 18 years, valid pre and post therapy CORE OM, attended more than 2 sessions and indicated as having a planned ending.

Table 5

Types of therapy endings as indicated by counsellor/therapists on CORE EOT form

Reason indicated for therapy ending	*n*	%
Planned ending with client	1724	62
Counselling discontinued by client	280	10.1
Client did not attend/cancelled therapy sesions offered	290	10.4
Client declined counselling	39	1.4
Client did not attend initial assessment	1	0
Client uncontactable	54	1.9
Missing data	393	14.1
Total	2781	100

Classification of participant drop out for this study was the same as based on the experiences of other publicly available large distributed psychological service operations based in the primary care sector with a similar client intake procedure as that of the CIPC service (e.g. Connell, Grant, & Mullin, 2006; Werbart, Levin, Andersson, & Sandell, 2013).

The total percentage of participants who were recorded as unplanned endings was 23.8%. If we assume that missing data also represents those participants who ceased attending without informing their therapist, or where therapists did not record the type of ending, the unplanned ending rate increases to 37.8%.

4.1.5 *Questionnaire completion rates*

Completion rates for CORE OM and HR QOL index items are detailed in Table 6 below. Overall rates dropped by an average of 32% across all HRQOL index items between pre and post therapy completion. Of 2,781 clients who consented to participate in the study 2,507 completed valid CORE OM pre therapy questionnaires (90.1%) and 1,493 completed CORE OM post therapy questionnaires (53.7%), this represented 59.5% of the 2,507 valid pre therapy forms completed.

Table 6

Completion rates for CORE and HRQOL questionnaires (n = 2,781)

		Pre therapy	Post therapy
HRQOL Question		n *(%)*	
	CORE OM	2507 *(90.1)*	1493 *(53.7)*
Q1	Q1 General health	868 (*62.1*)	291 *(20.8)*
Q2	Q2 Physical health	867 *(62)*	295 *(21.1)*
Q3	Q3 Mental health	723 *(51.8)*	292 *(20.9)*
Q4	Q4 Activity limitation	521*(37.3)*	276 *(19.75)*

Note. HRQOL data is for Form Returners sample

4.2 Pre post outcomes - statistical analyses, RCSC and effect size results

As discussed in the Method chapter, stratification of the main sample of participants in the study was applied in order to examine changes in CORE OM scores on these particular groups

before and after therapy. The rationale for each of the progressively restrictive criteria applied to the all participants sample is based in the natural formation of these groups in primary care services similar to CIPC. Appendices B and C (located on pages 218 and 219) contain tables showing results of pre post analyses and RCSC proportions for all samples, including the All Participants, Assessed and Attenders samples as well as separated data for those participants above and below clinical cut off. The remainder of this chapter will focus on results of analyses on for the Form Returner sample as this sample is the most representative of the typical cohort of people assisted by the CIPC service insofar as they were more likely to be female, heterogeneous in terms of their demographic characteristics, the type, number and severity of their presenting problems, they mainly attended between six and 9 sessions after most ended their therapy in consultation with their counsellor/therapist.

4.2.1 *Calculating the Reliable Change Index*

The RCI calculation used in this study incorporated the mean baseline CORE OM (0.76) and standard deviation (SD 0.59; 95 %CIs 1.04 – 1.16) for the non-clinical sample as determined by Evans et al (2002) for a non-clinical sample (N = 1084). As advised by Evans (2002) a test-retest reliability was used in the calculation as this ensures the inclusion of a real source of variation – which was part of the real world determination of reliability when it was conducted on the normative sample. This was a value of .88 as reported by Barkham et al (2007).

Cut-off point between clinical and non-clinical populations for this study is CORE OM = 1.00 as recommended by Connell et al (2007). This value was used to delineate between two groups of participants: those above this threshold and those presenting to the service below this threshold. The Attender sample comprised of 82.1% (*n* = 1765) of participants presenting over the clinical threshold of a CORE OM score of 1.00.

4.3 Pre post therapy results - What are the outcomes in terms of symptom reduction

Pre post changes in CORE OM scores were calculated using two main approaches: firstly, each of the samples were subjected to paired-samples t tests in order to test for the statistical significance. Secondly, ESs were calculated for all of the samples. Raw ESs for each of the samples were calculated using the same calculation as that used in other primary care practiced based studies (e.g. Barkham et al., 2001; Barkham, Stiles, Connell, & Mellor-Clark, 2012; Connell et al., 2007; Stiles, Barkham, Mellor-Clark, & Connell, 2008). Pre post ESs were also calculated using the method as described in Dunlap et al. (1996) which uses the sample *n*, t statistic (t) and correlation (*r*) value obtained from the dependent t test results for each of the respective samples.

Methods for calculating both the RAW *ES* and $ES_{RepeatedMeasures}$ methods have been described in detail in the Method chapter.

Table 7

Pre post outcomes and effect sizes for all samples

The form returner sample (*n* = 1,395) contained only those participants with valid post therapy scores from the assessment stage. Hence, as none of these participants had pre therapy scores carried forward (i.e. all post therapy scores were obtained from participants at the end of therapy, not by application of an LOCF process). The RCSC rate was 46.4% (*n* = 648) and

reliable improvement rates was 64% (*n* = 894). The reliable deterioration was 1.9% (*n* = 27) and those with no reliable change 33.9% (*n* = 474). ESs increased significantly from those produced by the combined assessed and attender samples (appendix B). Both the raw and $ES_{Repeated\ Measure}$ effect sizes were 1.16.

Pre post analyses of all other samples as described by Barkham et al., (2012) are included in Appendices B and C (p. 218 and 219).

4.4 What are the outcomes in terms of treatment duration?

Certain criteria were required in the sample of participants used to investigate the relationship between the total number of sessions attended and its impact on rates of reliable and clinically significant change. Participants in the analyses had agreed a planned ending with their counsellor/therapist, attended more than two sessions and had completed pre therapy CORE OM forms and that post therapy data were actual results of questionnaire completions (not the result of a last observation carried forward process). Hence, the Completer sample (*n* = 1,335) was used for the following analyses and, the number of sessions attended value was as recorded by the counsellor/therapist on the End of Therapy (EOT) form. As the current study attempts to report the behaviour of factors associated therapy outcomes in a real world primary care psychotherapy service in a manner that is representative of a clinical reality, data for all session durations is included, i.e. from the minimum attended to the maximum recorded.

Table 8 shows the number of sessions attended by all participants in the current study along with the corresponding number of those participants were above the clinical cut off at the pre therapy stage. Also in table 8 are the numbers of participants who achieved RCSC as a percentage of all participants and only participants above pre therapy clinical cut off.

Table 8

Reliable and clinically significant improvement rates as a function of total number of session attended.

			Clients achieving RCSC		
Sessions Attended	All clients (*N*)	All clients above cut off (*n*)	*n*	As % of all clients	As % of clients above pre therapy cut off
1	0	0	*0*	0	0
2	9	7	6	66.7	85.7
3	21	8	7	33.3	87.5
4	29	18	12	41.4	66.7
5	51	42	30	58.8	71.4
6	80	67	40	50.0	59.7
7	137	101	58	42.3	57.4
8	275	227	136	49.5	59.9
9	510	416	235	46.1	56.5
10	70	60	33	47.1	55.0
11	35	32	17	48.6	53.1
12	34	31	14	41.2	45.2
13	31	29	13	41.9	44.8
14	13	12	8	61.5	66.7
15	5	5	1	20.0	20.0
16	5	5	2	40.0	40.0
17	11	9	5	45.5	55.6
18	5	5	4	80.0	80.0
19	4	4	1	25.0	25.0
20	2	2	1	50.0	50.0
21	4	3	1	25.0	33.3
22	3	3	1	33.3	33.3
23	1	1	0	0.0	0.0
Total	1335	1087	625		

For many participants from the Completers sample who began therapy above the clinical cut-off of a CORE OM pre therapy score of at or above 1.00 (*n* = 1,087) a trend emerged which is represented visually and discussed further later in relation to Figure 3. From the data in Table 8, it can be observed that two distinct groups emerge before and after session number 10, (represented by either side of the dotted line dividing the total number of sessions attended into those attending less and more than 10 sessions); 88.5% (n = 1,182) of all participants ended their therapy after the 10th session, while only 11.4% (*n* = 153) continued to attend therapy after the 10th session. Further, after session 9 there is a sudden, followed by a steep and steady increase in the numbers of participants ending therapy. With the exception of a reversal in this trend at session 17, the actual number decrease is very small. A large majority of participants decided to end their therapy after session number nine including those above the clinical cut-off, indicating that of all participants with planned endings who attended at least 2 therapy sessions (including the assessment session) and completed pre and post therapy CORE OM questionnaire and decided to end therapy after attending between 2 and 9 sessions (*n* = 1,112), an average of 64.3% ended therapy having achieved RCSC.

Figure 3 shows a graphic representation of the data pertaining to those participants above the pre therapy clinical cut off and how this group withdrew from the CIPC service in relation to the rate of RCSC achieved. The overall trend-line clearly shows a decrease in the percentage of participants achieving RCSC as the number of sessions increase and the subsequent number of overall participants decreases dramatically (as detailed earlier).

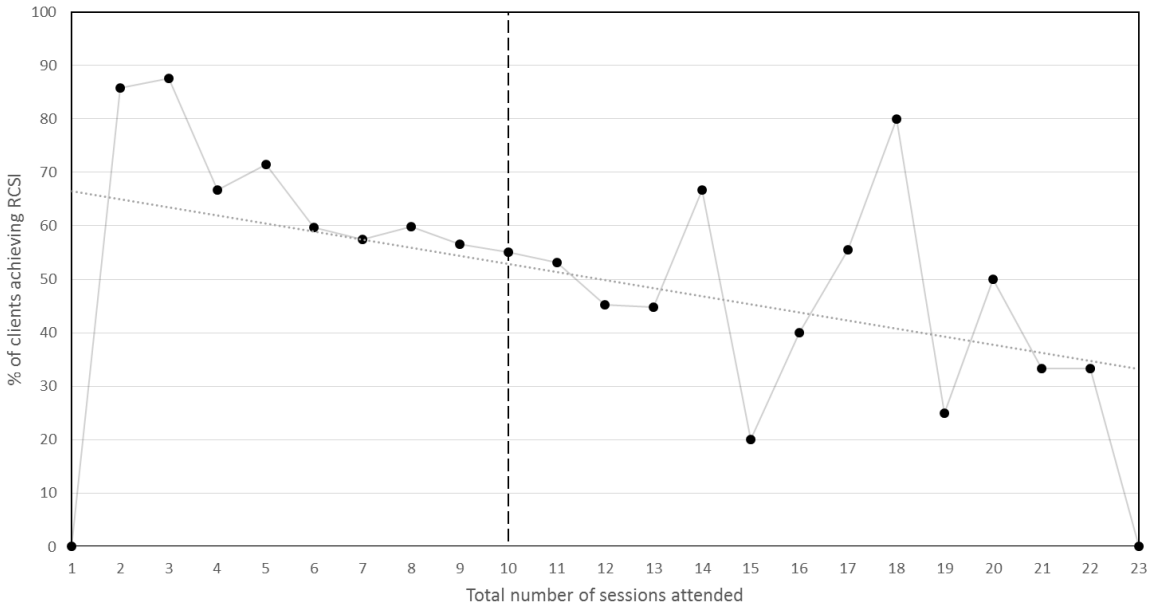


Figure 3: Graph showing reliable and significant improvement rates for all participants above clinical cut off level as a function of the total number of sessions attended (n = 1,087).

A preliminary t-test and series of cross-tabulations were conducted to assess whether there were any statistical differences between particular groups. The first independent t-test compared differences in RCSC based on the number of sessions attended. This result was significant (t = 3.417, df = 1085, p = .001) where the mean number of sessions attended by those who did achieve RCSC was 8.53 and those not achieving RCSC was 9.07. Results of the t-test comparing pre therapy CORE OM scores between those participants who did and those who did not achieve RCSC was also significant (t = 5.650, df = 1085, p <.0001) where those participants who did not achieve RCSC had a mean pre therapy CORE OM score of 1.83, and those who did had a mean a score of 1.67. Chi-square tests comparing the following variables for significant differences between RCSC groups (yes/no) by way of pre therapy risk levels for suicide, self-harm, harm to others or legal risk or duration of problem for problem types depression, anxiety, bereavement or self-esteem.

Based on these results a binary logistic regression analysis was performed on the Completers Sample, only participants above clinical cut off were included in the analysis as those below

the clinical cut off point of 1.00 could not achieve RCSC (*n* = 1,087). The regression was performed using the enter method and RCSC achieved (yes/no) as the dependent variable with total number of sessions attended and pre therapy CORE OM score included as a predictor variables. A total of 1,087 cases were included and analysed and the full model significantly predicted RCSC group membership (omnibus chi-square = 39.465, df = 2, $p <$ .0001).

The model accounted for between 3.6% and 4.8% of the variance in RCSC group membership, with 26.4% of participants not achieving RCSC correctly predicted. However, 84.6% of predictions for the participants who achieved RCSC were accurate. Overall 59.9% of predictions were accurate. Table 9 gives coefficients and the Wald statistic and probability values for the predictor variable. This shows that the number of sessions attended and pre therapy CORE OM scores reliably predicted RCSC achievement. The values of the coefficients reveal that an increase of one session in the total number of sessions attended was associated with a decrease in the odds of achieving RCSC by a factor of 0.93 (95% CIs [0.90, 0.98]), and that each SD decrease in CORE OM scores was associated with an increase in the odds of achieving RCSC by a factor of 0.49 (95% CIs [0.37, 0.64]).

Table 9

Logistic regression table showing session number as predictor of RCSC

Logistic regression table with number of session as predictor of participant RCSC achievement (n = 1087)							
Variable	*B*	S.E.	Wald	*p*	Odds ratio	95% CI fo odds ratio	
						LL	*UL*
The number of sessions attended	-0.07	0.025	8.051	.005	0.933	0.889	0.979
Pre therapy CORE OM score	-0.709	0.136	27.166	< .0001	0.492	0.377	0.642
Constant	2.154	0.316	46.436	< .0001	8.622		

4.5 Does the initial severity of participants' symptoms predict CORE OM outcomes?

One of the objectives of this study was to examine the relationship between symptom severity and the CORE OM. Several approaches to this were taken including; 1) a correlational analysis

of pre and post therapy CORE OM outcomes, 2) regression analysis of pre therapy severity levels as indicated by counsellor/therapist ratings on the Therapy Assessment Form (TAF) on CORE OM pre therapy scores in order to determine the overall impact of initial severity as measured by the CORE OM in relation to pre therapy scores. A multiple regression model was generated using depression and anxiety severity (i.e. scoring ordinal scale; 1 = *causing no difficulty,* 2 *= causing mild difficulty,* 3 *= causing moderate difficulty,* 4 *= causing severe difficulty*) as recorded by counsellor/therapists during assessment as predictor variables and pre therapy CORE OM scores as the dependent variable and 3), three different severity rating bands were formed by demarcating participants by three levels of severity; non-distressed (≤1.00), mildly, moderately, or moderate-to-severely distressed (scores between 1 and 2.5), and severely distressed (≥ 2.5), which was followed by an analysis of variance in CORE OM scores in relation to all three severity bands. Finally, t-tests were conducted to assess the level of difference between pre and post therapy in all three bands.

The first analysis showed a strong positive relationship between pre and post therapy CORE OM scores (r = .509, n = 1,397, p < .0005, one-tailed). This was a strong correlation accounting for 25.9% of the overall variation

Next, using the enter method, a significant model emerged: F (753) = 52.400, p < .0001. The model explained 12% of the variance (Adjusted R^2 = .120). Participants' predicted CORE OM score was equal to 0.810 + 0.172 (Depression severity) + 0.135 (Anxiety Stress severity) = 1.12, where all independent variables were coded as 1 = *causing no difficulty,* 2 *= causing mild difficulty,* 3 *= causing moderate difficulty,* 4 *= causing severe difficulty*. Participant's pre therapy CORE OM score increased by 0.172 for each unit on the depression severity scale and 0.135 for each unit on the anxiety severity scale. Both severity categories were significant predictors of pre therapy CORE OM levels. Table 10 gives further information for the predictor variables entered into the model.

Table 10

Regression table for pre therapy severity depression and anxiety levels predicting CORE OM score (n = 1,397)

			95% CI		
Variable	B	SE	*LL*	*UL*	*β*
Depression severity	0.172	0.026	0.120	0.243	.267*
Anxiety severity	0.135	0.029	0.079	0.175	.150*
Intercept	0.810	0.086	0.642	0.972	

Note. Adjusted R^2 = .120 *LL* = Lower Lmit *UL* = Upper Limit
*p < .0001

4.6 Are outcomes sustained six months after therapy has ended?

4.6.1 *Preliminary exploration of the data*

The group mean CORE OM scores were plotted in order to inspect the overall trend in outcomes at pre, post and six month follow up time points. All participants in the form returners data set who provided CORE OM data at three time points (i.e. pre and post therapy and at six month follow up) were included in the analysis (Pre and post therapy *n* = 1,397, six month follow up *n* = 240). The breakdown per severity category used later in the analyses comparing outcomes over all three time points was; Mild severity = *n* = 52, Mild to moderate and moderate to severe = 175, and the severe category *n* = 13. Pre therapy severity was rated by their counsellor/therapist after assessment and the total number of sessions attended was reported on each participant's end of therapy form.

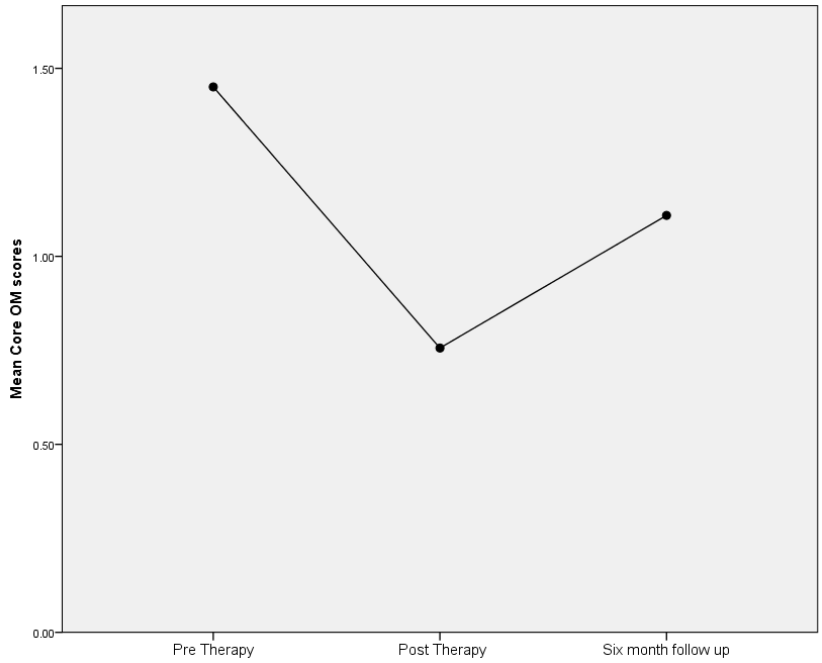


Figure 4: Mean CORE OM responses at pre therapy, post therapy, 6 month follow up and 12 month follow up (n = 240).

The resulting plot suggested an overall non-linear change trajectory for participants' CORE OM score patterns across the three time points. Bearing this in mind descriptive statistics for all three time points were examined to further determine the nature of differences in CORE OM change scores over all three time points and a sample inspection process as described by Heck et al. (2014) was followed in order to make the best decisions regarding model estimation and interpretation.

Table 11

Mean therapy scores at pre, post and six months after therapy (*n* = 240)

				95% CI	
Measurement occasion	*M*	*SD*	*SE*	*LL*	*UL*
Pre therapy	1.45	0.591	0.037	1.378	1.523
post therapy	0.75	0.572	0.035	0.686	0.826
Six month follow up	1.10	0.692	0.043	1.024	1.193

Note. M = mean; SD = Standard deviation; SE = Standard error; 95% CI = 95% Confidence interval; LL = Lower Limit; UL = Upper Limit

Examination of the mean difference between pre and post therapy shows an average decrease in participant CORE OM scores of 0.70, and an average increase between post therapy and six

months later of 0.35. This suggested that while overall benefit in the form of decreased levels of psychological distress were experienced by participants between pre and post therapy, this benefit appeared to have decreased six months after therapy had ended. Put another way, the trajectory of the decay curve observed between pre and post therapy appeared to have reversed and changed from a downward to an upward direction between post therapy and six months after therapy had ended. Bearing this in mind, it was decided that plotting linear growth trajectories could provide additional information regarding the range of participants' intercepts (i.e. pre therapy CORE OM scores) and also the steepness and overall trajectories of changes in CORE OM scores over time.

Heck et al. (2014) recommend that a visual inspection of a subset of data can provide additional information regarding important aspects of mathematical change over time. Hence, a random sub-sample of participants' pre, post and 6 month follow up scores were plotted and fit lines assuming a linear trend were applied. A visual inspection suggested a poor fit of fit lines to these participants CORE OM outcome data (*n* =17). As well as a wide variation in pre therapy CORE OM baseline scores (indicating varying pre therapy intercepts). There was also a

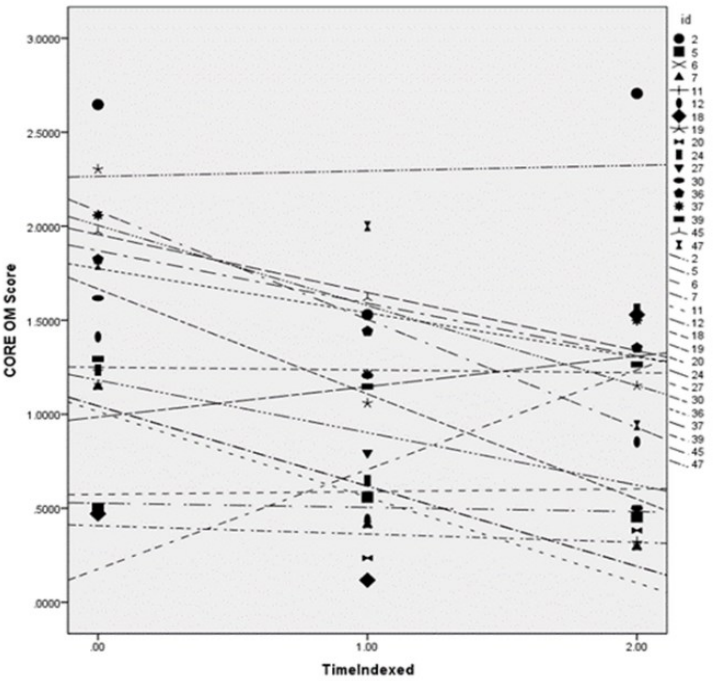


Figure 5: Plot diagram showing linear fit lines for 17 study participants at pre, post and six months after therapy.

wide variation in slopes between participants' pre therapy to six month follow up scores indicating that some improved at faster or slower rates than others (see figure 5).

Adding instead a quadratic term to the fit indices resulted in a much better fit for the trend of participants' data points. It was observed that while some clients' pre, post and six month measurements followed a roughly linear path (i.e. participants 45, 36 and 30), the majority did not, indicating that including a quadratic term long with the linear term to the subsequent hierarchical model could result in a better fitting model (see figure 6).

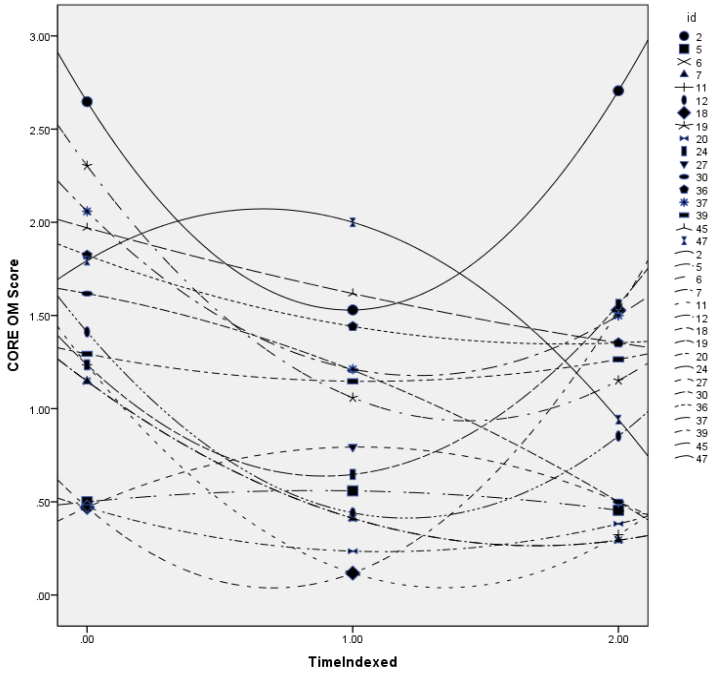


Figure 6: Plot diagram showing quadratic fit line for 17 study participants of pre post and six month

Once the for the same participants trajectories were graphed using a quadratic term the graph it became clear that its inclusion better described a majority of client pathways. Further, this visual representation suggested that a model with both random slopes (to account for varying rates of change between participants) and random intercepts (to account for the varying pre therapy CORE OM scores) might be the best representation of change for the sample of participants used in the subsequent MLM analysis (*n* = 240). The use of a quadratic term in order to model change due to psychotherapy was recommended by Barkham, Stiles and

Shapiro (1993). However, as highlighted by Heck et al. (2014) transforming polynomial functions (linear or quadratic) in the manner as described can introduce problems of collinearity between individual participants' CORE OM scores because of high levels of correlations between elements comprising the polynomial functions. As this model is heavily dependent on the correct estimation of the Time component and consequently the relationship between Time and change in CORE OM scores, a method of dealing with this issue was considered important. Hence, the polynomial component Time was further transformed to be orthogonal in order to help mitigate the collinearity effect.

The time-related within-subjects factor was first recoded in order to define the change in individuals over time. First, the linear time variable was coded 0 for Pre therapy, 1 for post therapy and 2 for six months after therapy. This method has the advantage of identifying clients' intercepts as pre therapy CORE OM scores which can be interpreted as the mean when all the predictors in the model are zero. For linear growth, the slopes were defined as change occurring between each interval point. The quadratic component was defined to capture the rate of change by squaring the linear time component resulting in values of 0 for pre therapy, 1 for post therapy and 4 for six months after therapy as per the method. These values were recoded from 0, 1, and 4 to 1, -2 and 1 respectively – resulting in the variables Orthtime and Orthquad were incorporated into the model as a higher order polynomials. Further, the transformation of the polynomials in this way makes the overall trend the focus of change, rather than change within the pre, post or six month follow up time intervals. Both the intercept and orthtime were entered into Model A as random effects, i.e. they were allowed to vary across participants (i.e. at level 2).

Table 12

Multi-level regression table estimating association between total number of sessions attended, pre therapy severity levels, CORE OM scores at pre therapy, post therapy and at six month follow up.

Level and Variable	Model A	Model B
Level 1-individual change over time		
Intercept (SE)	1.099 ** (0.322)	-0.342** (0.108)
Orthtime Linear time component Orthtime- slope (SE)	-0.173** (0.021)	0.149 (0.090)
Orthquad Quadratic term (SE)	0.174** (0.011)	0.173** (0.011)
Level 2 - between-subject variability		
Number of sessions attended (SE)		0.031* (0.010)
Pre therapy severity (SE)		0.641** (0.050)
Cross level interaction		
Orthtime * Pre therapy severity (SE)		-0.251** (0.041)
Orthtime * Number of sessions attended (SE)		0.017** (0.008)
Variance components		
Within TIME (L1) variance	0.173** (0.016)	
Pre therapy (SE)		0.076** (0.024)
Post therapy (SE)		0.200** (0.024)
Six month follow up (SE)		0.161** (0.045)
Intercept (L2) variance (SE)	0.196** (0.024)	0.090** (0.015)
Intercept-slope (L2) covariance (SE)	0.034** (0.011)	0.047** (0.014)
Slope (L2) variance (SE)	0.019 (0.012)	0.037** (0.016)
AIC	1200.739	934.776
Number of estimated parameters	7	13

SE = Standard Error, ** p <.001, * p <.05

3.8.1 *Multi levelling results*

Table 12 displays the results of the multilevel analysis. In the first model (Model A), a total of seven parameters were estimated including the intercept for the grand mean of the growth trend for CORE OM score across all participants included in the analysis (n = 243) (β_{0j} = 1.099; SD = 0.322). The linear time component (orthtime) (i.e. the mean linear rate of change across time) β_{1j} = -.0173 was significantly different from zero and negative, indicating a decrease in CORE OM scores for most participants at the outset of the current study. The quadratic term (orthquad) was also significantly different from zero and positive, indicating that, on average, participants' levels of distress (as indicated by scores on the CORE OM) first decreased and then increased.

4.6.2 *Main MLM results*

Both of the linear and quadratic estimates, which were significant and defined the shape of the decay within individuals, when combined with the intercept estimate produced an end

point CORE OM estimate of 1.1. Focussing on the variance in the residuals, the linear contrast varied randomly across individuals (Wald Z = 8.307, $p < .001$) and was significant, while the variance in the between level 2 slopes failed to reach significance but the covariance between the level 2 intercept and slope was significant. Overall Model A suggested that both the intercepts and the linear growth rates did vary significantly from zero across participants and the shape of that change was represented by an initial decrease in distress levels followed by a reversal in this trend to an increase in distress levels.

Model B sought to explain whether participants' pre therapy intercepts varied across individuals and whether the variation in the intercepts could be partly explained by the total number of sessions attended and the pre therapy level of the severity of participants' symptoms (as reported by their counsellor/therapist on the CORE om assessment form. In order to accomplish this, both the linear time and a quadratic growth parameters remained in the model as fixed effects, but the linear parameter was also allowed to vary across participants (i.e. added to the model as a random effect) in order to determine more about the shape of the decay trajectories. Level 2 predictors were added because the total number of sessions attended by participants and their initial level of severity were thought to explain differences in their CORE OM intercepts and linear growth rates. Additionally, cross level interaction terms were added in order to examine whether the variability in those participants at different levels of severity, i.e. were the linear decay rates different across severity levels, or the same. Further, a cross level interaction terms for the total number of sessions attended by participants was also added to examine the effect of this Level 2 between participant variable on the level 1 slope coefficient – participant decay rates.

Participants' CORE OM scores intercept estimate was -0.342. This was the participants' true grand mean CORE OM score adjusted for the total number of sessions they attended and their pre therapy level of severity. In this case the grand mean can be thought of as the CORE OM

score for participants who attended 2 sessions (i.e. the minimum number for any participants in the Form Returners sample) and were rated at the lowest severity level, i.e. 1 = causing minimal difficulty. Both predictors were related to differences in average CORE OM scores. The estimate for the number of sessions attended (i.e. 0.031) suggested that the grand mean CORE OM score for participants attending increased numbers of sessions was about -0.311 (SE = 0.010; $p < .05$), so those attending more sessions tended to have better outcomes up to some point (i.e. higher numbers of sessions were associated with lower CORE OM scores). The coefficient for levels of pre therapy severity was also a significant predictor related to differences in average CORE OM scores and the grand mean for participants CORE OM scores who had higher levels of pre therapy severity was 0.299 (SE = 0.050; $p < .001$).

In terms of whether differences in CORE OM scores were related to pre therapy severity levels or the total number of sessions attended, there was reason to suggest that rates of decay did vary significantly across the sample of participants. This was because the Wald *Z* test result for the variance of the level 2 (between participants) slope estimate which was significant (Wald $Z = 2.329$, $p = .02$). However, the linear time component (orthtime) was not significantly different from zero, but the quadratic component was still a significant contributor to the overall model. The linear interaction between CORE OM scores and pre therapy severity levels was also interpreted as a way to explain the variability between both variables. The interaction was significant and negative (-0.251, $p <.001$) and interpreted as meaning that participants with higher pre therapy severity levels demonstrated slightly more improvement than compared with participants with CORE OM scores around the grand mean. In terms of the cross level interaction between time and the total number of sessions attended; this was also significant (0.017, $p < .001$) indicating that participants who attended a higher number of total sessions had a higher rate of decay (i.e. greater reduction in CORE OM scores) over time compared with those participants who attended lower numbers of sessions. The orthtime quadratic polynomial was also significant implying that participants' rate of symptom decay as

measured by decreasing CORE OM scores slowed and reversed direction over time (0.173, $p <$ .001). Successively decreasing AIC values indicated that Model B (934.776) as a better fit to the data than Model A (1200.739).

4.6.3 *Mean CORE OM scores across pre, post and six month follow up points*

Mean CORE OM scores for each of the time points; pre therapy, post therapy and six month follow up are shown in Table 13 and figure 7. Overall (n = 240), participants reported a mild level of distress according to the CORE OM (M = 1.45; 95% CIs [1.37, 1.52]) at pre therapy assessment which dropped below the clinical cut off by post therapy (M = 0.75; 95% CIs [0.68, 0.82]). Six months after therapy however, this had risen to back above clinical cut off (M = 1.10; 95% CIs [0.1.01, 1.19]). CORE OM scores significantly improved between pre and post therapy (ΔM = -0.70; 95% CIs [0.62, 0.77]; $p <$ .001) and between the assessment session and six month follow up (ΔM = -0.75; 95% Cis [0.67, 0.84]; $p <$ .001). However, the increase in CORE OM scores between post therapy and six months later was significant (ΔM = .35; 95% CIs [-0.43, 0.27]; $p <$ .001). This analyses was conducted on all participant in the discrete pre therapy severity level categories, whereby some differences emerged.

Participants in the severe range at pre therapy (M = 2.63; 95% CIs [2.54, 2.72]) improved to levels just above the clinical cut off point (M = 1.08; 95% CIs: [0.70 - 1.50]). At the follow up period (i.e. six month later) this has risen back up to moderate levels (M = 1.54; 95% CIs: [0.93, 2.15]). Between the assessment session and post therapy, there was a significant reduction in CORE OM scores for those participants in the severe category (ΔM = -1.55; 95% CIs [-0.43, -0.27]; $p <$ .001). However, the change CORE OM score between post therapy and at six month follow up was also significant and increasing (ΔM = 0.446; 95% CIs [-0.88, .051]; p = .031). Between pre therapy and six month follow up a significant difference was also observed (ΔM = -1.46; 95% CIs [0.85, 2.07]; $p <$ .001)

Participants in the moderate range when they presented for their assessment session (M = 1.58; 95% CIs [1.52, 1.63]) decreased their CORE OM levels by post therapy to below the clinical cut off (M = 0.82; 95% CI [0.74, 0.91), which as observed in the changes in the severe category of participants, rose by the six month follow up time point (M = 1.21; 95% CIs [1.11, 1.31]). Change between pre and post therapy CORE OM scores was significant (Δ*M* = -0.76; 95% CIs [0.67, 0.84]; $p < .001$). Between post therapy and six month follow up and again, as in the group of severe participants at pre therapy assessment, there was a significant increase in CORE OM scores (Δ*M* = .386; 95% CIs [-0.29, -7.96]; $p < .001$). Between pre therapy and six months after therapy there was also a significant reduction in CORE OM scores (Δ*M* = -0.790; 95% CIs [0.69, 0.89]; $p < .001$).

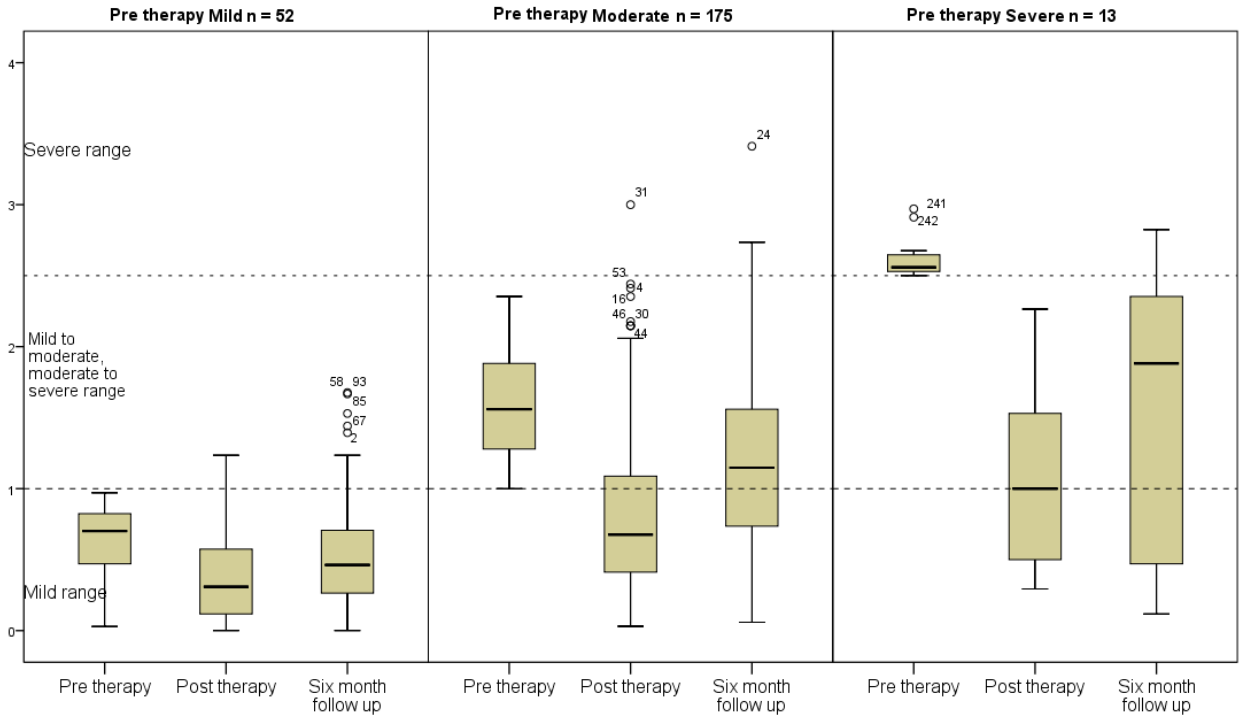


Figure 7: Box-plots showing CORE OM pre, post and six month follow up as a function of three severity levels; mild (< 1), mildly, moderately, or moderate-to-severely distressed (>1 to <2.5), and severe (> 2.5) (n = 240).

For those participants in the mild range (*n* = 52), who were sub clinical to begin with, pre therapy scores were mostly well below the cut-off point of 1.00 (M = 0.64; 95% CIs [0.58, 0.71]) and roughly halved by post therapy (M = 0.38; 95% CIs [0.30, 0.47]). Most of the

decrease between pre and post therapy was cancelled out by six months after therapy ended (M = 0.56; 95% CIs [0.44, 0.68]). In terms of the magnitude of the differences between the time points when CORE OM scores were measured, a significant difference was observed between pre and post therapy (Δ*M* = -0.26; 95% CIs [0.16, - 0.36]; *p* < .001). Between post therapy and six month follow up the same trend as observed in the other severity groups was observed, the decrease in CORE OM scores was reversed and scores rose significantly (Δ*M* = 0.17; 95% CIs [-0.28, 0.06]; *p* = .003). Between pre therapy and six month later there was a significant decrease in scores (Δ*M* = -0.443; 95% CIs: [0.33, 0.56]; *p* < .001).

Table 13

CORE OM scores at pre post and six months after therapy by baseline severity levels

Sample group, severity category and measurement stage					95% CI	
	n	M	SD	SE	*LL*	*UL*
Full sample	240					
Pre therapy		1.45	0.589	0.038	1.372	1.521
Post therapy		0.75	0.563	0.036	0.679	0.822
Six month follow up		1.10	0.697	0.045	1.013	1.189
Severe severity group	13					
Pre therapy		2.63	0.151	0.042	2.538	2.720
Post therapy		1.08	0.629	0.175	0.697	1.457
Six month follow up		1.54	1.012	0.281	0.932	2.154
Moderate severity group	175					
Pre therapy		1.58	0.364	0.027	1.526	1.634
Post therapy		0.82	0.561	0.042	0.740	0.907
Six month follow up		1.21	0.643	0.049	1.114	1.305
Mild severity group	52					
Pre therapy		0.64	0.230	0.032	0.579	0.707
Post therapy		0.39	0.313	0.043	0.297	0.472
Six month follow up		0.56	0.422	0.059	0.440	0.675

Note. Mean differences between measurement stages in all groups significantly differed from each other at the *p* < .05 level; M = mean; SD = Standard deviation; SE = Standard error; 95% CI = 95% Confidence interval; LL = Lower Limit; UL = Upper Limit

4.7 Are there differences in outcomes between therapeutic approaches?

Next followed an inspection of the proportions of therapy types indicated by counsellor/therapists as the main approach engaged in with their respective clients during the

study period. Table 14 shows the overall percentages of various types of therapeutic approaches undertaken during therapy. This showed therapists more often than not indicated that they employed more than one type of therapy approach with each client they saw. There were no indications of Structured Brief Therapy or Behavioural Therapy by any therapists.

Table 14

Breakdown of therapy approaches (n = 1,397)

	Responses	
TherapyType	*n*	%
Person Centred	884	26.4
Integrative	872	26
Supportive	586	17.5
Cognitive Behavioural	315	9.4
Psychodynamic	302	9.0
Systemic	236	7
Other	88	2.6
Psychoanalytic	57	1.7
Art	9	0.3
Total	3349	100.00%

Note 1 this analysis was conducted on the Form Returners sample.

Analysis showed the split between clients for whom therapists (*n* = 123) indicated only one versus more than one type of therapy approach was 32.9% and 63.5% respectively (*n* = 1,397), 3.6% of cases were missing or indicated as zero approaches applied. A breakdown of the total number of approaches employed by the frequency indicated is shown in table 15.

Table 15

Breakdown of therapy approaches by frequency indicated.

Number of therapies indicated	*n*	%
0	4	0.3
1	378	27.1
2	337	24.1
3	329	23.6
4	137	9.8
5	85	6.1
6	97	6.9
7	11	0.8
8	5	0.4
Total	1383	99.1

Note. n = 13 (0.9%) did not indicate any therapy type

Following the process outlined by Stiles et al. (2006; 2008) an attempt to form six individual groups was made. However, due to a lack of adequate numbers in the appropriate categories, exactly the same pairings as made by Stiles et al., (2008) could not be accomplished. This resulted in the CBT+1 and PSY+1 groups being omitted from further analyses as each pairing consisted of only five participants each. However two additional categories was added to the analyses: “Integrative” and “Integrative +1”. Hence the groups included in further analyses were: INT (integrative) and INT+1, CBT (cognitive and cognitive behavioural), PSY (Psychodynamic and psychoanalytic), PCT (Person Centred) and PCT+1.

These six separate groups were analysed by way of pre post paired sample t tests and effect sizes for each group were calculated in order to compare differences in the effectiveness of therapy between clients who received just one therapy approach and those who received more than one approach. As discussed in the methods section these categories were added to this analyses in order to make the results more relevant to the Irish psychotherapy sector. Also, Art Therapy could not be included in the analysis as the minimum number of other therapies it was combined with for use with particular participants was three, which excluded it by way of the restriction to one, or one plus one other therapy.

Table 16

Mean CORE OM scores for single and +1 therapy approaches and pre and post therapy. (n = 399)

		Pre therapy		post therapy		Pre/post difference		
Therapy approach	*n*	Mean	*SD*	Mean	*SD*	Mean difference	*Pooled SD*	Effect size
INT	194	1.53	0.592	0.77	0.516	0.76	0.555	1.37
CBT	39	1.40	0.575	0.67	0.441	0.73	0.512	1.42
PSY	55	1.68	0.580	0.91	0.551	0.77	0.566	1.36
PCT	51	1.57	0.641	0.92	0.573	0.65	0.608	1.07
INT+1	12	1.48	0.540	0.85	0.491	0.63	0.516	1.22
PCT+1	48	1.54	0.676	0.94	0.810	0.59	0.746	0.80
Total	399							

Note. Analysis carried out on the Form returners sample *(n*= 1397) INT, Integrative therapy; CBT, Cognitive, behavioural, or cognitive/behavioural therapy; PSY, psychodynamic/psychoanalytic therapy; PCT, Person Centred therapy; INT+1, INT combined with one other therapy; CBT+1, CBT combined with one other therapy ;PSY+1, PSY combined with one other therapy; PCT+1, PCT combined with systemic therapy (only cases where the pairing was unique)
Effect size- calculated as the mean difference divided by the pooled standard deviation .

A one-way analysis of variance (ANOVA) compared pre-therapy means across all six groups.

There was no significant difference in pre-treatment CORE OM scores between groups [*F*(5,

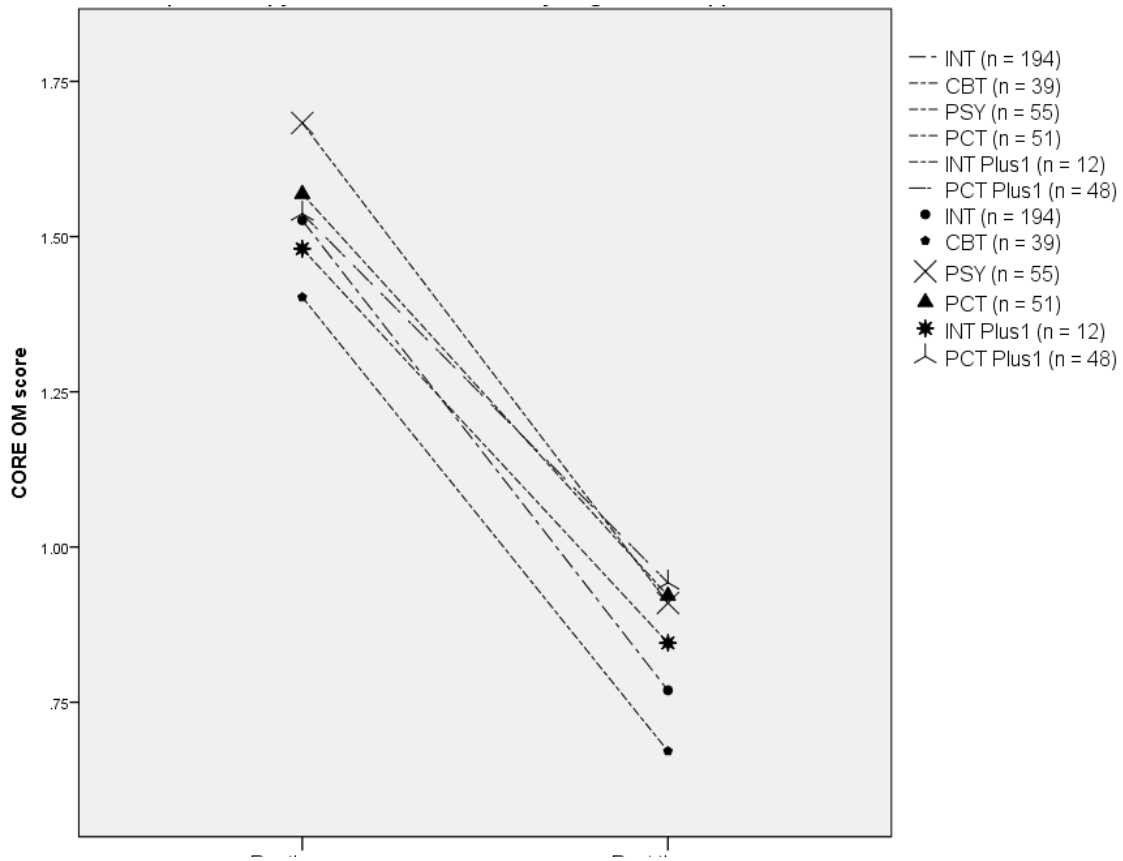


Figure 8: Pre post therapy mean outcomes for single versus more than one therapy approach

398) = 1.09, p = .368] indicating that all groups began treatment with roughly similar levels of distress as measured by the CORE OM.

This was followed by a repeated measures ANOVA showed a large overall within-participants main effect of all treatments [$F(1, 393) = 286.885$, $p < .0001$, partial $\eta^2 = 0.442$], indicating that improvement across treatments accounted for a large proportion of the variation in CORE OM scores between pre and post therapy.

A differential effect due to treatment type was assessed by an inspection of the pre post interaction effect between time and all approaches but was not significant [F (1, 5) = 0.894, p = .485 partial η^2 = .011] see Figure 8 for visual description of pre post outcomes across all six approaches.

The comparative effectiveness of single approaches versus single+ 1 forms of treatment was also not significant [$F(1, 397) = 2.821$, p = .094, partial η^2 0.007] see Figure 9 for a visual description of pre post outcomes for single versus +1 approaches.

Next, in keeping with the line of enquiry initiated by Stiles et al (2006; 2008) a 2x8 χ^2 comparing the

rates of RCSI (Reliable and Significant Improvement) versus non-RCSI across the six treatment groups was conducted. Up to this point the Form Returners sample used for this analysis comprised all participants in the sample. However, before conducting RCSI analysis all participants in all groups who did not score above the clinical cut-off point of 1.00 (as in

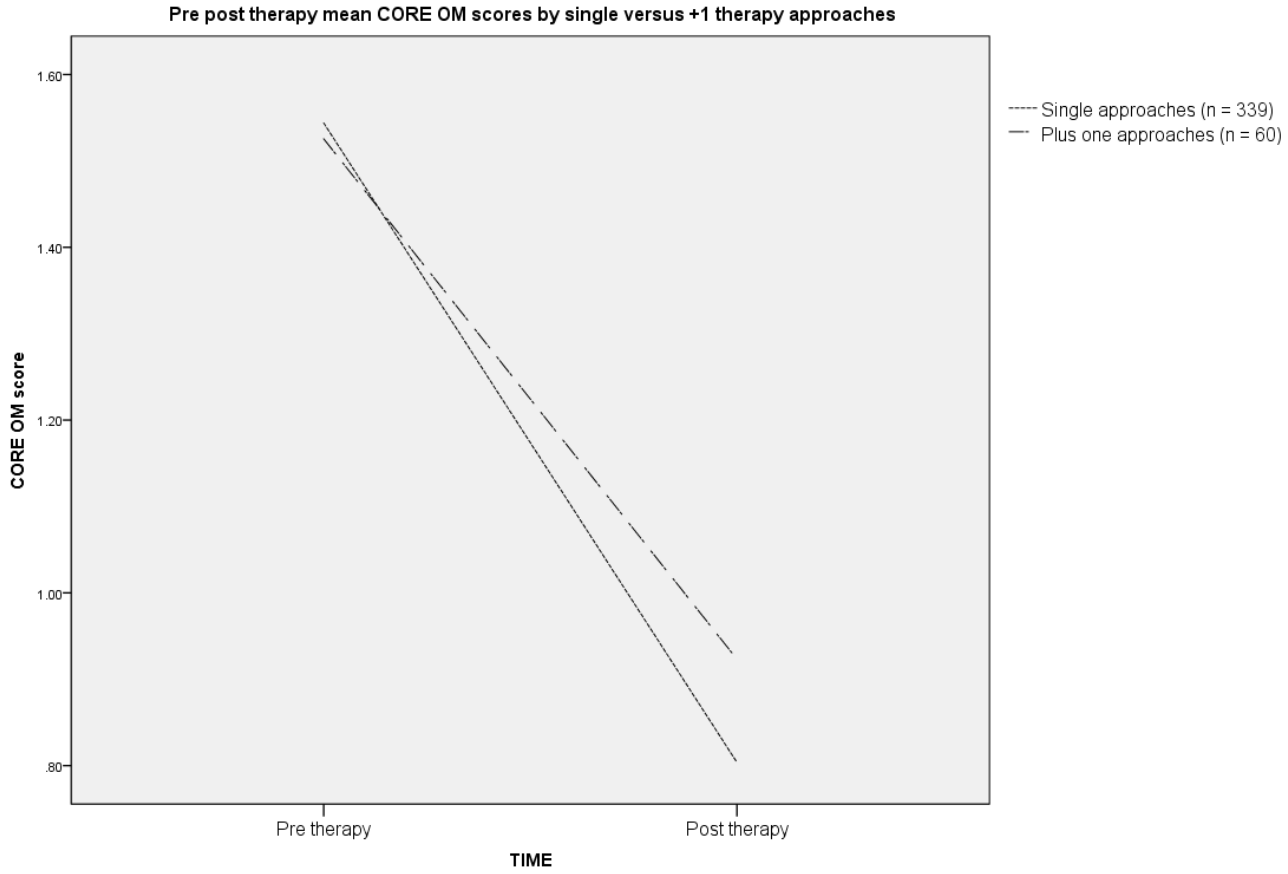


Figure 9: Figure showing interaction of single versus more than one therapy approach on CORE OM scores

Connell et al., 2007) were removed. This was because participants who had scored below the cut-off threshold of 1.00 could not achieve change which was clinically significant, i.e. drop from above to below the cut-off point. Removing these participants from all six groups (i.e. INT, CBT, PSY, PCT, INT+1 and PCT+1) from the Form Returners sample resulted in an N of 327 on which to conduct the analysis comparing rates of RCSI as well as reliable improvement only, no reliable change and reliable deterioration. Analysis showed that 81.3% of participants remaining were above the clinical cut off point of 1.00 on the CORE OM. The result of the Chi Square test was not significant [χ^2 (5) = 3.891, p = 0.565, n = 327]. This result was illustrative of the similar improvement rates across all six groups. Regardless of a differential effect of

therapy approach a significant proportion of participants achieved RCSC (56.3%). Participants who attended therapy with therapists indicating integrative as a single therapy approach achieved the highest rate of RCSI, followed by psychodynamic and psychoanalytic. Person centred plus one other approach was the combined therapy approach that achieved the highest rate of RCSI, followed by integrative in combination with one other approach.

Table 17

Breakdown of rates of RCSI for single versus more than one therapeutic approach (n = 327)

		RCSI		Reliable improvement only		No reliable change		Relaible deterioration	
Treatment group	*n*	Frequency	Row percent	Frequency	Row percent	Frequency	Row percent	Frequency	Row percent
INT	162	92	28.1	32	9.8	37	11.3	1	0.3
CBT	29	20	6.1	3	0.9	6	1.8	0	0.0
PSY	49	26	8.0	10	3.1	13	4.0	0	0.0
PCT	40	20	6.1	8	2.4	12	3.7	0	0.0
INT+1	10	7	2.1	0	0.0	3	0.9	0	0.0
PCT+1	37	19	5.8	3	0.9	14	4.3	1	0.3
Total	327	184	56.3	56	17.1	85	26.0	2	0.6

n = 327 patients whose pre-therapy CORE-OM scores were at or above the recommended cut-off of 1.00 (Connell et al. 2007).
RCSI, Reliable and clinically significant improvement; Reliable improvement only, decrease of 0.47 or more that did not fall below the cut-off of 1.00 on the CORE-OM;INT, Integrative therapy; CBT, Cognitive, behavioural, or cognitive/behavioural therapy; PSY, psychodynamic/psychoanalytic therapy; PCT, Person Centred therapy; INT+1, INT combined with one other therapy; PCT+1, PCT combined with systemic therapy (only cases where the pairing was unique)

4.8 What is correlation between CORE OM and Health Related Quality of Life?

As stated in the chapter describing the study methods, there are several aspects to the HRQOL measure important to determining whether or not participants had experienced statistical and clinical change. Firstly, of relevance when assessing changes in participants' self-reported general health are changes in the response patterns to question 1 - Would you say that in general your health is excellent, very good, good, fair, or poor? As laid out in the methods chapter detailing the process of evaluating HRQOL change (see page 106), a statistical approach to addressing this question was to conduct a Wilcoxon signed ranks test on the response data. However, before conducting the statistical test a visual aid in the form of a bar chart comparing pre and post levels of self-reported general health was generated in order to determine whether any changes in participants' perception of their general health between pre and post therapy had been reported at the aggregate level. A visual inspection of mean

responses for each of the Q1 response categories suggested that differences existed between pre and post therapy for each of the response categories (see figure 10).

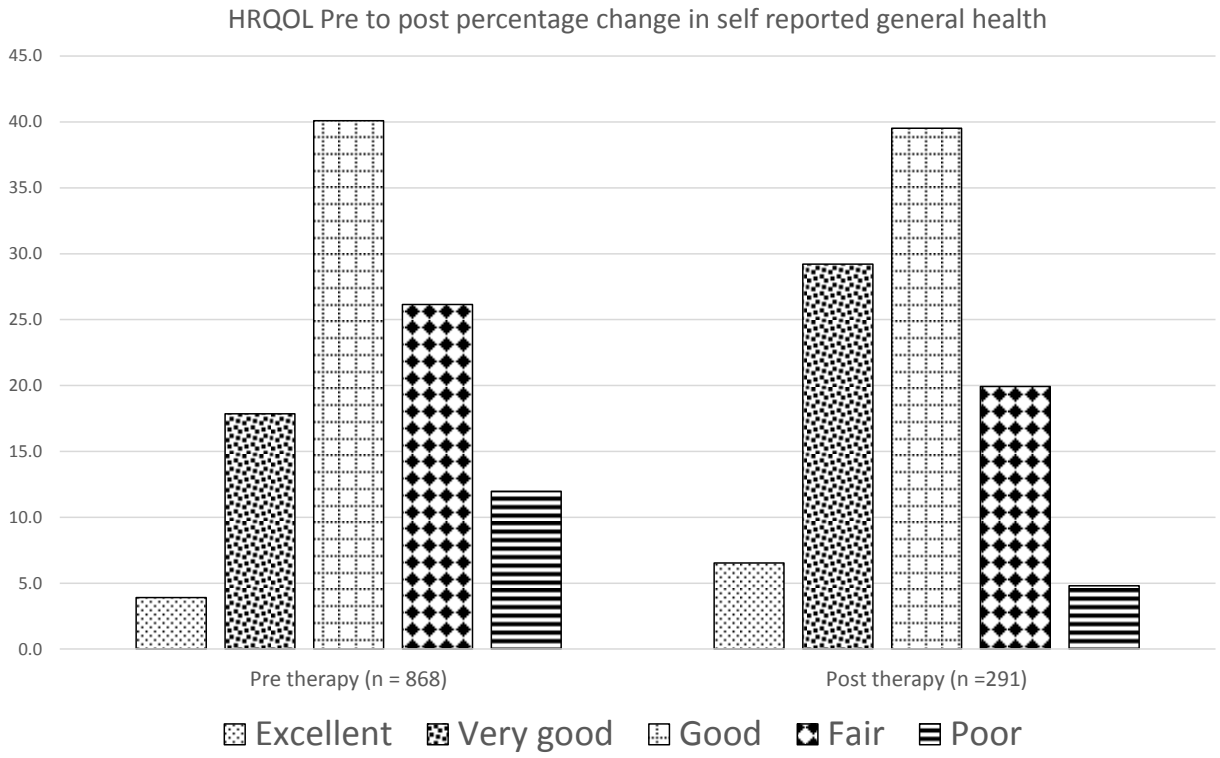


Figure 10. Pre post change in HRQOL Item 1 - General Health

The bar chart showed that all responses differences moved in a positive direction from the perspective of improvement for participants. The percentage of participants who reported their general health as Excellent increased between pre and post therapy from 3.9% to 6.5% and as Very Good from 17.9% to 29.2%. The percentage of responses for the Good category decreased slightly from 40.1% to 39.5%. Participants who reported their general health as Fair decreased as a percentage from 26.2% to 19.9% and the percentage reporting their general health as Poor at the assessment stage of the study decreased by over half from 12% to 4.8%. As these data were measured on an ordinal scale a Wilcoxon signed ranks test was conducted and confirmed there was a significant difference between pre and post therapy Q1 item responses (z = -5.915, N – Ties = 102, p <.0001, one-tailed).

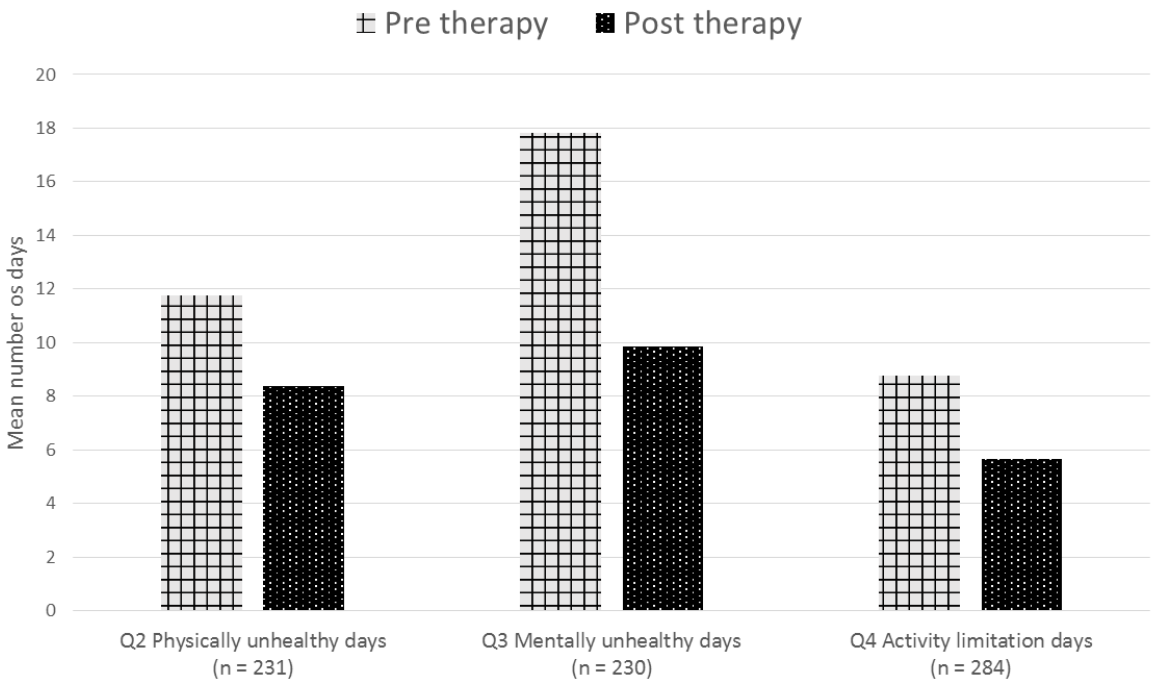


Figure 11. Pre post change in HRQOL Items 2, 3 and 4 - Physically unhealthy days, mentally unhealthy days and activity limitation days.

A visual inspection of the difference in the mean number of days between pre and post therapy measurement indicated decreases for all index items, Q2, Q3 and Q4. The pre to post index item means were calculated as part of between subjects t tests conducted for each of the items in order to determine if the changes were statistically significant. Next, the number of days from pre to post therapy reported for each of the HRQOL index items Q2, Q3 and Q4 were examined in order to provide an overview of any differences. The overall number of mentally and physically unhealthy days for each participant was summed and adjusted so the maximum number is limited to 30. This resulted in a mean number of unhealthy days for the Form Returners sample of 22.8 days (*n* = 6.5, SD = 9.40).

Once completed, statistical tests were conducted if necessary. For HRQOL items Q2 - perception of physical health, Q3 - perception of mental health and Q4 - perception of activity limitation, comparisons of pre and post therapy mean number of days for each question is reported for the participants included in the resulting t tests is reported.

A paired t test showed the average number of days participants reported feeling physically unhealthy during the previous 30 day period dropped by 3.4 days from 11.8 to 8.4, a statistically significant amount (t = 4.795, df = 230, p = <.0001, one-tailed). Mentally unhealthy days also reduced significantly from 17.9 to 9.9 days, a reduction of 8 days on average (t = 11.058, df = 229, p = <.0001, one-tailed) and the number of days participants reported their normal activity being limited due to their mental and/or physical problem was significantly less by the end of therapy, reducing by an average of 2.6 days (t = 4.325, df = 283, p = <.0001, one-tailed). The effect sizes for these changes ranged from 0.29 for the difference in pre to post physically unhealthy days, to 0.81 for mentally unhealthy days and 0.26 for activity limited days. See table 18

Table 18

Pre post therapy scores HR QOL for physically and mentally unhealthy days and activity limitation days

		Pre therapy		Post		
HR QOL Index item	*N*	Mean	*SD*	Mean	*SD*	*p*
Physically unhealthy days	231	11.77	11.62	8.37	10.20	<.0001
Mentally unhealthy days	230	17.80	9.87	9.84	8.97	<.0001
Activity limitation days	261	8.78	9.76	5.64	8.57	<.0001

Correlational analysis was conducted next to explore the relationships between the three pre therapy HRQOL index items reporting the number of days on a scale (0-30) and pre therapy CORE OM scores (Table 19)

Table 19

Correlations, mean score and standard deviations for CORE OM and HRQOL items 2, 3, and 4.

Variable	*N*	*M*	SD	1	2	3
1 Core OM	1397	1.54	0.601			
2 Q2 Physically unhealthy days	682	11.28	11.753	.323**		
3 Q3 Mentally unhelathy days	675	18.10	9.790	.498**	.297**	
4 Q4 Activity limitation days	774	9.56	10.541	.401**	.437**	.401**

Note. M and SD are used to represent mean and standard deviation respectively. ** indicates p <.0001.

All variables were significantly correlated at the pre therapy stage. The strongest correlation was observed between CORE OM scores and participants number of mentally unhealthy days (r = .498, N = 675, p <.0001) with CORE scores explained 25% of the variation in the number of

mentally unhealthy days. The next strongest correlation was between CORE OM scores and the number days when participants activity was limited due to the psychological problem for which they were seeking help (r = .401, N = 774, p <.0001), followed by the number of physically unhealthy days (r = .323, N = 682, p <.0001) with CORE OM scores explaining 16% and 10% of the variation respectively.

As there were significant and strong positive correlations between higher CORE OM pre therapy scores and HRQOL pre therapy Q2, Q3 and Q4 responses, i.e. higher numbers of days feeling unwell both physically and mentally and more days where activity was limited due to these problems, multiple regression analysis was conducted to further elucidate how these variables are related to one another and if the items had any predictive value in terms of pre therapy severity levels of anxiety and depression.

Next, as all HRQOL responses Q2, Q3 and Q4 were found to correlate quite strongly with pre therapy CORE OM scores and correlational analyses showed that the HR-QOL and CORE OM corresponded in terms of the direction of those correlations, a multiple linear regression analysis was calculated to predict pre therapy CORE OM scores as the dependent variable, based on the number of mentally unhealthy days, the number of physically unhealthy days and the number of days that participants felt their activity was limited.

Table 20

Logistic regression odds ratio results for HRQOL items predicting CORE OM score.

			B 95% CI			
Variable	*B*	SE	*LL*	*UL*	*β*	*p*
mentally unhealthy days	0.022	0.002	0.018	0.027	.360	<.0001
physically unhealthy days	0.006	0.002	0.002	0.010	.114	.003
activity limitation days	0.013	0.002	0.009	0.017	.228	<.0001
Intercept	0.971	0.044	0.884	1.058		<.0001

Note. Adjusted R^2 = .305 LL = Lower Lmit UL = Upper Limit

Using the enter method, a significant model emerged: F (3, 601) = 89.420, p <.0001. The model explained 30.5% of the variance (Adjusted R^2 = .305). Participants' predicted CORE OM score

was equal to 0.971 + 0.022 (mentally unhealthy days) + 0.006 (physically unhealthy days) + 0.013 (activity limitation days), where all independent variables were measured in number of days in the previous 30 day period. Participant's pre therapy CORE OM score increased by 0.022 for each mentally unhealthy 30 day period, 0.006 for each physically unhealthy 30 day period and 0.013 for each 30 day period their activity was limited due to their psychological problem. All three HRQOL items were significant predictors of pre therapy CORE OM levels. Table 21 gives further information for the predictor variables entered into the model.

In order to examine the level of correspondence between pre post therapy CORE OM score differences and the differences in the number of unhealthy (mental and physical) days and days when activity was limited, an additional correlational analysis was conducted which compared the pre post therapy differences between these variables.

Table 21

Correlations, means and standard deviations for pre post therapy score differences for CORE OM and HRQOL items 2, 3, and 4.

Correlations, means and standard deviations for pre post therapy score differences for CORE OM and HRQOL Questions 2,3 and 4

Variable	*N*	*M*	SD	1	2	3
1 Core OM	1397	-0.70	0.594			
2 Q2 Physically unhealthy days	207	-3.68	10.718	.108*		
3 Q3 Mentally unhelathy days	207	-7.69	11.170	.449**	.042¥	
4 Q4 Activity limitation days	250	-2.76	10.285	.299**	.315**	.313**

Note. M and SD are used to represent means and standard deviation of the difference respectively.
** indicates $p < .01$.
* $p = .122$
¥ $p = .585$

Pre post changes across all variables were positively correlated. The strongest statistically significant relationship was the pre post difference in CORE OM and the number of mentally unhealthy days reported ($r = .449$, $N = 207$, $p < .01$) followed by the number of days when activity was limited ($r = .299$, $N = 250$, $p < .01$). The pre post difference between the numbers of physically unhealthy days reported did not correlate with CORE OM pre post differences.

Finally, a binary logistic regression was performed in order to assess to what degree questions from the HRQOL measuring the number of days participants felt physically or mentally unwell (i.e. Q2 and Q3) and the number of days their activity was limited (i.e. Q3) predicted whether or not they achieved RCSC. The analysis used RCSC achieved (yes/no) as the binary dependent variable and the following predictor variables were included in the model HRQOL Qs 2 and 3, sex, currently on medication (yes/no) and unemployed (yes/no) in two blocks using the enter method.

Table 22

Logistic regression results for HRQOL items 2, 3 and 4, sex, medication status and unemployment status predicting RCSC achievement.

	Block 1 results							Block 2 results						
						95% CI for odds ratio							95% CI for odds ratio	
Variable	*B*	S.E.	Wald	*p*	OR	*LL*	*UL*	*B*	S.E.	Wald	*p*	OR	*LL*	*UL*
mentally unhealthy days	0.02	0.009	4.667	**.031**	1.02	1.002	1.039	0.021	0.009	5.077	**.024**	1.022	1.003	1.041
physically unhealthy days	-0.011	0.008	1.945	.163	0.989	0.974	1.004	-0.012	0.008	2.167	.141	0.988	0.973	1.004
activity limitation days	0.01	0.009	1.245	.264	1.01	0.992	1.028	0.012	0.009	1.572	.210	1.012	0.994	1.03
Sex (Base = 1 Male)								-0.379	0.187	4.104	**.043**	0.685	0.474	0.988
currentlyonmedication														
Yes								0.051	0.275	0.035	.853	1.052	0.614	1.804
No								0.016	0.276	0.003	.954	1.016	0.591	1.747
Unemployed (Base = 1 Yes)								0.337	0.263	1.638	.201	1.4	0.836	2.344
Constant	-0.49	0.179	7.518	**.006**	0.613			-0.737	0.373	3.901	**.048**	0.479		
Nagelkerke pseudo r-square	1.9%							3.40%						
-2 Log liklihood	825.630							818.717						

Note.OR = Odds ratio, significant predictors in bold

A total of 604 (43.2%) cases were analysed in the full model significantly RCSC achievement (omnibus chi-square = 8.48, df = 3, *p* = .037). The first Block of variables (Mentally and psychically unhealthy days and activity limitation days) accounted for between 1.4% and 1.9% of the variance in RCSC achievement, with 71.6% of participants not achieving RCSC correctly predicted and 34.3% of those achieving RCSC correctly predicted with an overall prediction accuracy rate of 54.3%. Of the three predictor variables included in Block 1 only mentally unhealthy days reliably predicted RCSC achievement. The value of the coefficient reveals that an increase of one mentally unhealthy day is associated with an increase in the odds of achieving RCSC by a factor of 1.02 (95% CIs [1.00, 1.04]).

Block 2 variables were introduced into the model and while the additional set of Block two variables (i.e. sex, current medication status and unemployment status) failed to reveal any significant additional associations and was not significant (omnibus chi-square for Block 2 = 6.913, df = 4, *p* = .141); the for overall model remained significant = 15.396, df = 7, *p* = .031). Of the predictor variables, the association between mentally unhealthy days and achieving RCSC remained significant, while being male also had a weak but significantly lesser chance of being associated with achieving RCSC. Table 22 gives coefficient, Wald statistics, associated degrees of freedom and probability values for each of the predictor variables.

4.9 Does gender or employment status have an effect on client outcomes?

A logistic regression was conducted using the enter method. The analysis used RCSC achieved (coded as yes/no) as the binary dependent variable and sex (coded as 1 = Male and 2 = Female) as the predictor variable. The result was marginal (omnibus chi-square = 3.83, df = 1, *p* = .050) showing a weak and non-significant association of a higher likely hood of achieving RCSC if participants were male (OR = .788, 95% CIs [0.62, 1.0]).

None of the employment categories had any association with the odds of achieving RCSC based on the results of a binary logistic regression analysis. This analysis used the same coding for RCSC achieved as that in the previous analysis. Employment statuses were coded as:

1 = Employed
2 = Unemployed and seeking work
3 = Student
4 = Long Term Sick, Disabled or Benefits
5 = Homemaker not working or actively seeking work
6 = Not receiving benefits and not working or actively seeking work
7 = Unpaid voluntary work, not working or actively seeking work
8 = Retired
9 = N/A
10 = Not Stated (Person asked but declined to provide a response).

The result was non-significant (omnibus chi-square = 4.65, df = 8, *p* = .794). For the sake of clarity, those participants self-identifying as unemployed were isolated and a comparison of

pre post CORE OM score differences between those unemployed participants and those in all other categories of employments statuses was completed by way of an independent t-test. This test was also non-significant (t = -0.238, df = 1393, p = .812, one-tailed).

Chapter 5: Discussion

5.1 Introduction

The Counselling in Primary Care (CIPC) service provides short term psychotherapy to adults in Ireland experiencing mild to moderate psychological problems. The vast majority of referrals originate in general practice, although referrals can come from other sources, e.g. physiotherapy services, occupational therapy among others.

This study arose from the opportunity to conduct the first empirical study of a national primary care counselling service in Ireland. The researcher was engaged by the CIPC service to conduct the study which recognised the value in practice based evidence gathering as a means of evaluating the effectiveness of the service. Overall effectiveness can be thought of as an aggregated representation of the CIPC therapy experience and includes client specific and service related factors. The broad objective of this study was to establish a profile of CIPC's overall effectiveness with a view to comparing outcomes with similar services internationally and also to examine certain therapeutic and service factors known to be important when evaluating therapy outcomes in large, publicly available nationally distributed primary care mental health services. While the evidence supporting the effectiveness of short term psychological therapy is in these settings is strong and plentiful, evidence for the role and behaviour of certain client and service related variables is less clear and scarcer. Hence, this study sough to provide some clarity to understanding clients' journeys through the CIPC service.

5.2 The CIPC Service – Client demographic and clinical profile

In terms of demographics, it is the case that historical and societal factors specific to the Irish context have been instrumental in determining the makeup of patients referred into CIPC services. For example, before the early 1990s Ireland had known very little inward migration.

However, in the 20 years between 1987 and 2007 the number of people from other countries migrating to Ireland increased from 17,200 per annum to over 150,000 (Central Statistics Office, 2019a). There are now over 508,000 people from the EU and other parts of the world living in Ireland – over 10% of the national population. However, the breakdown of participants recorded as being from a non-Irish background involved in the study was just 2.6%. This appears to be echoed somewhat by the IAPT service in the UK in which Black, Asian and other minority groups are also underrepresented (NHS, 2018), but not nearly as significantly as appears to be the case in Ireland according to these national proportions in CIPC, which are reported for the first time. Further, people from immigrant backgrounds show higher rates of psychosis, mood, post-traumatic stress and anxiety disorders than those from non-immigrant populations (Giacco et al., 2014). Hence, it is important that an active approach to ensuring equal access to the CIPC service be a priority in the future, as experience in other jurisdictions has shown that proportional representation of BME communities in counselling services is not normally achieved (du Plessis & Bor, 1999).

Also, 12.6% of participants in the current study were recorded as unemployed and seeking work. This represented a significantly smaller proportion of the total number who attended assessment and went on to consent to participate in the study. Hence the true impact of the difference that the therapy provided by the CIPC service could have made to this group, whom research has shown are disproportionately affected by mental health problems, may not have been completely elucidated. The Irish national unemployment rate fell from 4.8% to 4.0% over the course of the study (i.e. February 2018 – 2019) (Source Central Statistics Office, 2019b) and the relationship between being unemployed and the negative impact it can have mental health is well documented (Pharr et al., 2012). Hence, in terms of its own policy development the CIPC service should give regard to research in other jurisdictions associating increased utilisation of services with rising unemployment (Linn et al., 1985) for future service resource deployment management. Moreover, at the time of writing this thesis, the CIPC model of

service was set up only to accept referrals for members of the public who held valid medical cards. These are provided to those whose weekly income is below a certain threshold and allow for the free provision of certain health care services free of charge, including counselling as provided by CIPC. This accounts for proportion of study participants which is three times above the national unemployment rate.

In the current study, counsellor/therapists using the problem categories provided on the CORE Assessment Form as the method for indicating presenting problems, found anxiety related problems were by far the most common indicated at 80.3%, followed by depression 64.7% and interpersonal relationship problems at 44.7%. The order in which these top three presenting problems were indicated mirrors exactly that reported by Mullin et al. (2006) in their study involving more than 11,000 clients from 32 primary care NHS services in the UK and also Stiles et al. (2008). Other primary care services in Sweden (Werbart et al., 2013b), Norway (Knapstad et al., 2018) and the U.S.A. (Sawchuk et al., 2018) also report this hierarchy of problem presentation. However, a 2018 one year review of the IAPT service in the UK found the most common presenting issue was anxiety and stress disorders at 32% and depression at 26% (Moller et al., 2019), proportionally significantly less than those observed in the current study.

A possible explanation for this discrepancy could be the fact that IAPT reports focus on the information provided at the point of referral in order to collate presenting problems data. Whereas after a person is referred to CIPC, the determination of problem type occurs during the assessment session between the client and their counsellor/therapist. The disparity found in the current study between the reported percentages of presenting problem types found in the IAPT reports and non-IAPT services could be due to differences in how the GPs making the referrals and the counsellor/therapists conducting assessments define and make informal diagnoses of psychological problems. There is evidence that when it comes to making referral decisions involving patients presenting with psychological problems, GPs in Ireland have some,

or very little knowledge of specific psychological therapies and only 32% receive post graduate training in psychological therapies (Cullinan et al., 2016), a situation which appears to be mirrored in the UK, with only 46% of GP trainees taking up mental health training posts in 2016 (Millet, 2016). Hence, it could be that a lack of sufficient training of trainees GPs in psychological problems explains the disparity between the proportions of problem types reported by counsellor/therapists who are highly trained in assessing psychological problems and GPs, who are less trained in doing so. Interestingly in Ireland, those GPs who do receive post graduate training in psychological therapies also report a higher proportion of their patients as experiencing mental health problems, but are less likely to refer those patients onto psychiatric services (Copty & Whitford, 2005). The fact remains however, as this study has now verified within the Irish context, that anxiety and depression underlie the main reasons for people seeking help in primary care psychological service, as they do in many jurisdictions internationally (e.g. Clark, 2011; Knapstad et al., 2018; Sawchuk et al., 2018; Werbart et al., 2013).

In terms of assessing the severity of problems, using the coarsest method of doing so is the delineation of clients into clinical and non-clinical cases at the assessment stage. This study found that 82.1% of participants who attended at least an assessment and one therapy session (*n* = 2,151) were above the clinical threshold of 1.00 on the CORE OM. This is in line with other jurisdictions, e.g. in the UK when Mullin et al. (2006) who conducted a study of a large, primary care based psychological service found 89% of clients (n = 11,953) scored over this threshold. Coincidentally, in terms of comparison with IAPT or other jurisdictions using the PHQ-9 as a measure for depression Gilbody, Richards, & Barkham (2007) also recommended a cut-off of 10 (i.e. equivalent to 1.00 on the study by Mullin et al (2006)) between clinical and non-clinical populations. Using this metric, Gyani et al. (2013b) observed 81.3% of clients listed as having received some treatment were also above the clinical threshold and Knapstead et al.

(2018) found that 77.2% of their total sample (*n* = 1532) presented to Swedish primary mental health care clinics as clinical cases.

Two additional sources of the severity of clients' problems are recorded during assessment in a typical CIPC presentation; the first is a clinician estimation of the severity of the particular client's level of severity indicated on the Therapy Assessment Form (TAF) rated on a scale with four available options; *causing minimal difficulty, causing mild difficulty, causing moderate difficulty* and *causing severe difficulty*. The second source is derived from the CORE OM whereby the total score itself is constituted as a severity rating, i.e. higher scores signify higher levels of severity and as a way to standardise understanding and interpretation of the CORE OM scale in this way, it is common to find the application of banding the overall CORE OM score into a hierarchy ranging from non-distressed (< 10), mildly, moderately, or moderate-to-severely distressed (10 $\leq$ CORE-OM $\leq$25), and severely distressed ($\geq$ 25) and many studies have used a CORE OM score in excess of 2.5 to indicate severe levels of distress (Barkham, 2005; Bewick, McBride, & Barkham, 2006a; Mullin et al., 2006; Bewick et al., 2006a). While a detailed examination of differences between clinician and CORE OM derived ratings of severity is not an objective of this study, when discussing this aspect of the CIPC service profile it bears consideration that when it comes to assessment of risk using the same method as described above (i.e. converting the score of the risk items on the CORE OM versus clinician rated risk on a four item scale), the difference between clinician and client rated levels can be discordant by a significant degree – up to an average of 35%, i.e. the degree to which clinicians underrated symptom severity compared to ratings derived from the scoring of risk items on the CORE OM (Bewick, McBride, et al., 2006).

From the full sample of participants and based on their overall CORE OM score being $\geq$ 2.5 (*n* = 2,781) in this study 6.9% (*n* = 192) were in the severe range and from the sample of Form Returners (*n* = 1397) this fell only slightly to 6.2% (*n* = 87). Compared to ratings of the severity

of symptoms by clinicians for the top five presenting problems (in order of most to least indicated), was anxiety = 17.3%; interpersonal problems = 8.2%; bereavement = 7.5%; self-esteem = 6.6% and depression = 6.3%. Two noteworthy points are evident in light of the previous reference whereby clinicians appeared to underrate the level of risk compared to risk as derived from CORE OM risk scores as reported by Bewick et al. (2006). First, in contrast to the findings by Bewick et al. (2006) and with the exception of anxiety, in the current study the difference in percentages of participants rated as severe between clinician ratings and those derived from the CORE OM scores is minor (6.9% versus an average of 7.2% excluding anxiety). Second, the percentage of clients in the severe category for anxiety according to clinicians is over double that as derived from the next most prevalent problem type in the CORE OM severe band (17.3% versus 8.2% for interpersonal problems). This was not consistent with, and are in fact the opposite of findings by Berwick et al. (2006). Reasons for this may stem from an "apples and oranges" perspective, whereby the comparison of differences in ratings in the domain of risk between clinicians and clients is qualitatively a different question than that of differences between problem domains such as depression, anxiety, interpersonal problems etc. More important is the stark difference between the percentages of participants in this study indicated in the severe range of anxiety stress related symptoms compared to the other presenting problems. It appears that in the CIPC service, clients presenting with anxiety related symptoms appear to be over represented and more likely that their symptoms in the severe range.

5.3 Pre and post therapy outcomes – overall effectiveness and comparison with other services

The primary aims of all major publicly disseminated psychotherapy services is to help clients reduce the negative impact of the symptoms they resent to the service with and facilitate improvement in the overall quality of their lives. On the service side, improving service quality is an ongoing process (Clark, 2011; Health Service Executive & Irish College of General

Practitioners, 2006; Sawchuk et al., 2018; Werbart et al., 2013b). One of the objectives of the current study was to profile the CIPC service effectiveness and to do so in a number of different ways. The primary method was to compare differences between clients' aggregated pre and post therapy scores of psychological distress using the CORE OM. As shown in the results section, statistically significant reductions in CORE OM scores were observed in a large number of participants in this study – over 64% reliably improving with 46% also achieving clinically significant improvement.

Another objective of this study however, was also to address the specific question of how the effectiveness of the CIPC service (measured in the current study) compares with other large nationally distributed psychotherapy services. The answer to this question rests upon the ability to compare CIPC pre post differences in CORE OM scores garnered from this study with similar studies conducted in other jurisdictions. Table 23 summarises among other items, effect sizes, RCSC and combined improved and RCSC rates for this study along with the same results from six major primary care studies spread across the UK, Scandinavia and the U.S.A. and one meta-analysis conducted which have been cited throughout the course of this study thesis. These studies were discussed at length during the literature review and are used for comparative purposes here because like the current study they 1) largely focussed on evaluating clinical outcomes in services which were distributed across a large geographical area within their respective jurisdictions, 2) concentrated on evaluating outcomes following the delivery of psychotherapy under routine care conditions in primary care settings and 3) reported on a similar cohort of clients (i.e. heterogeneous demographically in the types and severity of presenting problems) and in some cases used the same primary outcome measure and service related forms i.e. Barkham et al. (2012), Stiles et al. (2008) and the majority of studies reported in Cahill et al. 2010).

Table 23

Comparison of effects sizes and RCSC rates between the current study and the main practice based studies.

Study name and year of publication	Country	Therapy approach	Measure(s)	N	Effect size*	Effect size calculation	RCSC rate* % (*n*)	Improved rate %
This study	Ireland	Mixed	CORE-OM	1397	1.16	(Pre-Post)/Pre SD	56.9 (*1138)*	64
Barkham et al 2012	UK	Mixed	CORE-OM	11651	1.47	(Pre-Post)/Pre SD	56.9 (*10397)*	63.3
Stiles 2008	UK	Mixed	CORE-OM	5613	1.39	(Pre-Post)/Pre SD	58.3 (*4954)*	77.7[d]
Cahill et al. 2010[a]	Multiple	Mixed	CORE-OM	10842	1.19[e]	N/A	56	82.4[d]
Gyani et al 2013	UK	CBT	PHQ-9 GAD-7	19395[c]	-	-	40.3	63.7
Knapstad 2018	Norway	CBT	PHQ-9	1295	1.09	(Pre-Post)/Pre SD	65 (*994*)	58
			GAD-7		1.03			
			PHQ-9		0.51			55
Sawchuk et al. 2018	USA	CBT	GAD-7	2577	0.57	Unknown	-	53.3
Werbart 2013[b]	Sweden	Mixed	GSI	175	0.83	(Post-Pre) ⁄	35 (*115*)	51
			SRH	177	1.20	SDpooled	41 (*116*)	45

Note. All effect size results are based on study samples meeting the following criteria: age ≥16, valid pre and post questionnaires returned, participants attended at least two sessions.

CORE-OM, Clinical Outcomes in Routine Evaluation; PHQ-9, Patient Health Questionnaire; GAD-7, Generalized Anxiety Disorder scale; SCL-90, Symptom Check List-90; GSI, The Global Severity Index; QOLI, The Quality of Life Inventory; SRH, Self-Rated Health.

*Results for combined clinical and non clinical samples, except for Stiles et al. (2008) and Sawchuck et al. (2018) who only reported outcomes for a clinical sample

**Results for clinical samples only as non-clinical participants are not valid for this calculation

a Meta-analysis of practice based studies conducted in UK primary care settings that included clinical and non-clinical groups in the analyses (n=10)

b Multiple approaches considered and reported separately

c This is an IAPT sample of one year's worth of data collection

d These studies reported RCSI and Improvement rates separately. They have been combined to be uniform with the rest of the results

e In order to simplify interpretation, this is the effct size calculated by Cahill et al (2013) after removal from the meta-analysis of Stiles (2008) and the earlier similar study by Stiles (2006)

- Not reported

In terms of effect sizes, the CIPC service is comparable with other services insofar as there was a very large effect size (i.e. far in excess of the 0.8 which is normally indicative of a large effect, see Howell, 2009). It is not unusual for practice based studies to produce large effect sizes as observed in meta analyses such as those observed in the current study (see Cahill et al., 2010). Overall effect sizes in all studies compared (including the current study) ranged from moderate, e.g. Cohen's *d* =0.51; Sawchuk et al. (2018) to very large, e.g. Cohen's *d* = 1.47; Barkham et al.(2012). It is reasonable to question such a wide variation in effect sizes for what appear to be generally similar populations experiencing the same types of problems and availing of psychotherapy in broadly similar settings. However, upon closer inspection of the table, a cluster of studies emerge were a mixed therapy approach and use of the CORE OM as the main outcome measure resulted in a tight range of effect sizes (Cohen's *d* = 1.16 – 1.47) and very similar RCSC rates to co-occur (56% – 58.3%), i.e. the current study, Cahill et al. (2010), Stiles et al. (2008) and Barkham et al. (2012). There are several reasons why this may have occurred.

First, these four studies meet all of the characteristics of what Shadish et al. (1997) describe as "Stage 3" studies, i.e. they meet all criteria used by Shadish et al. (1997) to identify studies carried out under "clinically representative conditions" up to that time. These studies included certain features in order to meet Stage 3 status; such as taking place in primary care settings or in mental health centres in the community, including participants which were heterogeneous in terms of their presenting problems, demographic and personal characteristics and the delivery of therapies without the use of treatment manuals or fidelity checks (for full list see Shadish et al., 1997). Hence, with such a considerable overlap of core characteristics between these four studies it is perhaps not entirely surprising that they produce a similar range of therapy outcomes.

Second, in terms of RCSC outcomes between these four studies, the cut-off points and degree of change required on the CORE OM in order to qualify study participants to meet RCSC criteria was also shared between all for studies; 1.00 on the CORE OM scale as a cut-off point (or 10 depending on the scale format used in each individual study) and a 0.5 point difference as a minimum for exhibiting reliable change (or 5.0 point difference, again depending on what scale format was used in each study). The only difference between studies in this respect was that the current study calculated an RCI of .47, as opposed to the .5 used in the other studies. Hence, with respect to how these four studies calculated RCSC rates we again observe very little (if any) significant variation in approaches. It is therefore not surprising that they should yield closely aligned rates of clients achieving RCSC. It could be argued that the advantages of use of a common measurement system (i.e. the CORE OM), coupled with the pursuit of further insight into a common question (i.e. the overall effectiveness of psychotherapy in primary care settings) is enabling a better understanding of psychotherapy effectiveness across jurisdictions. The very large overall N involved in these four studies combined (>29,500 participants), coupled with their closely aligned outcome metrics allows for a high degree of

generalisability of the outcome results to the population of people attending primary care psychotherapy services.

In terms of the remaining group of four studies (i.e. Gyani et al. (2013a), Knapstead et al. (2018), Sawchuk et al. (2018) and Werbart et al. (2013b)), one result in stark contrast to the other results are the effect sizes reported in Sawchuk et al (2018) for both the PHQ-9 and GAD-7 by which are roughly half those reported by the other three studies. This study to calculated effect sizes using "…Cohen's *d*, which is adjusted for paired samples".(2018, p. 7). However, the authors do not elaborate on which method of adjustment they used and many different methods of adjusting for repeated samples when calculating Cohen's *d* can be implemented (Durlak, 2009; Lakens, 2013; Seidel et al., 2014), one of which is reported in methods chapter of the current study and is discussed in more detail later in this chapter. It may have been the case that the alternative method of calculating effect size for the study by Sawchuk et al. (which was conditional on the repeated nature of pre post therapy scores) may have resulted in much lower effect sizes in relation to the other studies. Further, while the purpose of calculating effect sizes is to convey the magnitude of the impact of an intervention in a standardised way, i.e. without being measure specific (Goulet-Pelletier & Cousineau, 2018), the issue of how standardised the results actually are is one researchers, clinicians and other relevant stakeholders grapple with for good reason. As demonstrated in the current study by Werbart et al. (2013b) who used three different measures of subjective well-being (SWB) but even when restricting analyses to treatment completers only, still found effect size differences of up to .40 between the instruments measuring the same construct, i.e. subjective well-being. Moreover, significant differences in the methods and criteria used to specify caseness cut-off points and reliable change indices by Sawchuk et al. (2018) compared to Gyani et al. (2013b) and Knapstad et al., (see pages 45-46 for description), while utilising the same measures makes comparing and interpreting the outcomes relating to recovery, reliable change between these studies difficult.

This study also calculated effect sizes for CIPC service pre post therapy outcomes using a process of converting a *t* value to a Cohen's *d* as described by Dunlap et al. (1996) (see page 92 for full equations). These conversions were conducted in light of the finding by Seidel et al. (2014) who compared several different methods of calculating Cohen's *d* effect size, including those which took into account the non-independent nature of measurements taken from the same participants over more than one time point. They reported that the process of converting the t statistic derived from a paired sample t-test while incorporating the correlation estimate from the same test as proposed by Dunlap et al. (1996) produced "…consistently more moderate results than either the raw or severity adjusted methods." (2014, p. 7). This was indeed the case and can be observed in the table showing the range of effect sizes for all of the various samples constructed in accordance with the method used by Barkham et al., (2012) as shown in appendix C on page 219. It is clear that the range of raw, i.e. unconditional effect sizes is in excess of the range of effect sizes produced when implementing the conversion process by Dunlap et al (1996) - .35 and .30 respectively. The difference between the methods (0.5) is not very large.

Interestingly however, is the fact that the largest differences between the effect size calculation methods is evident between the samples where the Last Observation Carried Forward (LOCF) method was used to calculate post therapy scores (i.e. carrying forward the pre therapy scores for use as post therapy scores). Once the LOCF method was no longer in effect (i.e. all participants in the Form Returners and Completers samples registered both pre and post therapy scores actual pre and post therapy scores), effect sizes aligned between the calculation methods. Importantly, Barkham et al. (2012) acknowledge that the LOCF method utilised in their study produces post therapy scores which are conservative and assume no change between pre and post therapy for the clients concerned. Hence, this method should prevent against overestimating the effect of therapy since it has been clearly shown that more

clients attending primary care counselling services will improve due to therapy than remain unchanged or deteriorate after receiving therapy in a primary care setting (Cahill et al., 2010).

However, it should be noted, that concern regarding the analyses of data where pre therapy measurement scores are carried forward for use as post therapy scores has been expressed and the practice highlighted as questionable from a statistical validity point of view (see Lachin, 2016). Hence, while using the LOCF method in the current study to generate data for post therapy scores in order to calculate pre post differences and effect sizes was a way to remain consistent with the methodology used by Barkham et al. (2012) in terms of comparing outcomes, its use may be more appropriate when multiple measurements are provided in the context of recording sessional point data, as in Knapstad et al. (2018) who utilised multiple imputation as well as LOCF to account for missing data.

A final point to make regarding effect sizes is that the current study exceeded the 54% RCSC benchmark proposed in the meta-analysis by Cahill et al. (2010) along with four other studies in this comparison. When Shadish et al. (1997) conducted their comparison of effect sizes reported in meta analyses of studies to examine overall effectiveness, only one study from over 1000 included in the analyses met the Stage 3 criteria. As shown this number has increased dramatically, a testament to the growth in pragmatic and practice-based research approaches investigating the efficacy and effectiveness of psychotherapy in general. Bearing in mind the characteristical similarities between the samples in the current study and the comparison studies, the alignment of pre post effect sizes other outcome metrics generally supports the idea that the CIPC service in Ireland is not outside of normal effectiveness parameters in relation to other services.

5.4 Number of sessions - overall treatment duration

Particularly pertinent to the current study is that the duration of therapy sessions are brief and over-all duration of therapy provision to each client is time-limited. This a core feature of

similar primary care counselling services deployed at a national level in jurisdictions other than the CIPC service (Barkham, Stiles, Lambert, & Mellor-Clark, 2010; Connell, Grant, & Mullin, 2006). Further, guidelines formulated by the Health Service Executive (HSE), the Working Group on Mental Health in Primary Care and the Irish College of General Practitioners (ICGP) contained a stipulation that the service should provide counselling that was …"time limited and of short duration" (Health Service Executive & Irish College of General Practitioners, 2006, p. 9). Since concise care in both the ITT and per protocol samples in Meuldijk et al. (2016) led to improved outcomes which were essentially equivalent to that of standard care, it appears a correct determination by service providers providing counselling services on a large scale (i.e. CIPC, IAPT) to integrate time-limited and brief duration therapy episodes into the standard service model standard.

As sessional data were not available, an alternative approach to shedding some light on the question of how much therapy is enough**.** These data did lend themselves to examining the total number of sessions attended after which clients did not return for further therapy, and its relationship to the proportion of clients reaching RCSC at the end of their treatment. This approach will result in a different type of information returned than that provided in the studies by Barkham et al. (1996) and Stulz et al. (2013), who sought to an explanatory basis to their suggestion that as clients reached a level of improvement deemed as adequate, either by themselves or in conjunction with their therapists, they decided to end therapy, thus resulting in more or less consistent rates of RCSI being achieved for clients by the end of the treatment period. The current examination of the amount of therapy received by participants and how it relates to therapy outcomes will focus on the inter-relationship of these variables.

As mentioned during the literature review, some studies have examined the proportion of participants who have achieved RCSC as a function of the number of overall sessions attended (Elkin et al., 1995; Howard et al., 1986), and the current study approaches the question from a

similar perspective since sessional data were not available for analysis. The question of diminishing returns is more difficult to answer using pre post data only. It was observed that a relatively consistent percentage of CIPC users achieved RCSC regardless of session number – an average of 62% between session 2 and session number 12. However, the rate at which participants end their therapy is not consistent with 75% of clients ending therapy between the 6th and 9th session, but during this time the rate of RCSC achievement remains steady. Hence, questions remain such as; do clients leave after this number of sessions because they know therapy is time limited, or do they leave because they have improved enough to feel they can leave? At this time, the current study can only address the question "are rates of RCSC steady across treatment durations" and according to these data, yes they are. So the current study can speak to the question regarding consistency in rates of RCSC and when participants end therapy (because all Completers had a planned ending) what it cannot speak to is if there was a relationship between the progress participants made and the number of the session at which they decided to end therapy.

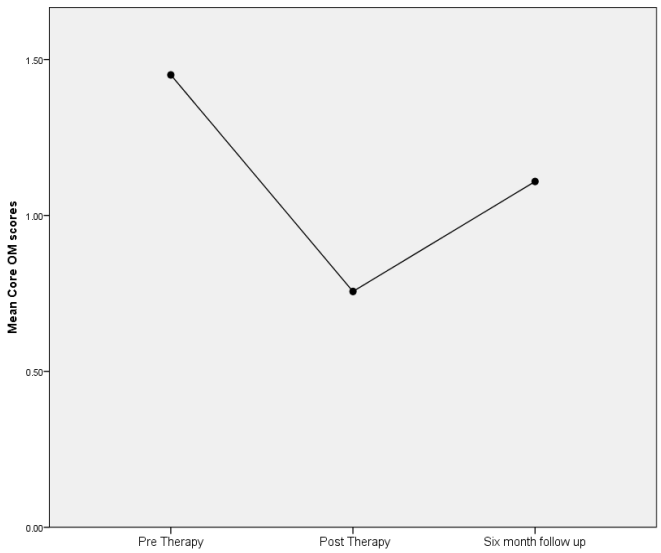
5.5 Initial severity of clients' problems and longer term effectiveness.

One of the aims of this current study was to investigate the relationship between; 1) pre therapy levels of symptom severity as indicated by counsellor/therapist rated severity and CORE OM scores and 2) the effect of pre therapy levels of severity as indicated by CORE OM scores and post therapy outcomes. Results of the linear regression analyses did show a relationship between the initial severity of participants' presenting depressive and anxiety related symptoms as indicated by counsellor/therapist rated severity and CORE OM scores. This result indicates a certain level of congruence between the severity indicators as completed by counsellor/therapists on the Therapy Assessment Form (TAF) and the expected behaviour of CORE OM scores with the sample of Form Returners (*n* = 1,397), i.e. as the severity of reported symptoms increased, so did CORE OM scores. The CORE OM grand mean average calculated in this analysis (0.81) was predicted to increase by a total of 0.31 for every

combined category increase on the depression and anxiety severity scales. The second relationship under scrutiny was the effect of between pre therapy levels of severity as indicated by CORE OM scores and post therapy outcomes. A number of previous practice based studies have used a process of banding participants into three severity categories which are commonly used by clinicians as aids during therapy (Barkham et al., 2012; Gyani et al., 2013b; Mullin et al., 2006; Ray-Sannerud et al., 2012). The current study banded participants as non-distressed (≤1.00), mildly, moderately, or moderate-to-severely distressed (scores between 1.00 and 2.5), and severely distressed (≥ 2.5) on the CORE OM scale in order to assess the effect of initial severity of symptoms on therapy outcomes (e.g. Barkham et al., 2012; Mullin et al., 2006). These severity category bands were shown by Barkham et al. (2005) to have good discriminative validity between psychological services delivered in primary care and secondary care settings. Further, previous studies have used severity bands as a way to examine whether pre to post therapy change in groups of participants at pre therapy changed as a function of the initial severity of their symptoms, or if it was an indicator of likely outcome (CORE Partnership, 2007; Elkin et al., 1995; Mullin et al., 2006), or whether they had an effect on eventual outcomes after low intensity interventions in the primary care setting (Bower et al., 2013).

Elkin et al. (1995) observed an interaction effect between the type of treatment, time and severity, but only at the high severity end of the scales on the measures they assessed. Bower et al. (2013) also found a small effect of initial severity. Results of the repeated measures ANOVA analysis conducted in the current study also appeared to support the suggestion that participants who presented to the service with higher levels of problem severity for depressive and anxiety made greater gains than those reporting with symptoms which were reported as mild or moderate. While this finding is to be expected as those clients experiencing higher levels of pre therapy symptomatology had more room to improve (Gyani et al., 2013a), there is value in the replication of that finding in line with other studies. Differential gains along lines of symptom severity were also examined by Saxon et al. (2008) who examined factors associated with poorer outcomes after counselling in an attempt to build a logistic regression model of CORE OM outcomes involving a range of predictors including pre therapy severity levels. However, while an association between having a higher CORE OM score and not achieving RCSC was observed in the study by Saxon et al., (2008), none of the models in their study constructed featured initial severity as a covariate that was a significant predictor of RCSC.

The current study has added to the general discussion regarding initial severity in terms of the finding that the levels of initial severity of clients presenting problems in the CIPC service and



their impact on outcomes of therapy appear to be similar to those observed in other jurisdictions.

5.6 Does the number of sessions and initial severity affect outcomes 6 months after therapy has ended?

A multi-level model (MLM) was constructed in order to investigate associations between CORE OM scores over time, the total number of sessions attended and pre therapy levels of severity reported by therapists for participants. This type of approach to representing change over time and during and after psychotherapy has become popular due to its flexibility in dealing with many types of hierarchical structure, i.e. clients clustered within therapists or other groups or as in this study, observations clustered within clients across time (Baldwin et al., 2014). Additional advantages lie in the ability of MLMs to accommodate the effects of correlation among different observations i.e. the between cluster variability as error terms called random effects and be more flexible regarding missing data (Heck et al., 2014).

The decay curve for participants was visually observed as curvilinear, not linear. There was a large decrease in negative symptoms as measured by the CORE OM between pre and post therapy followed by an increase of a lesser magnitude in reported symptoms six months after post therapy. The interpretation of the plots of a sub sample of participants' pre, post and six month time points suggested a quadratic function as an appropriate mathematical approach to describe the trajectories of change observed visually. Subsequently, linear and quadric terms were added where the linear term indicated the rate of change per unit of time and the quadratic term to indicate the rate of acceleration or deceleration of CORE OM scores. These terms were both recoded as orthogonal terms in order to mitigate issues of collinearity between the repeated measures (as recommended by Heck et al. (2014). An initial model was constructed which included the two predictor variables, followed by a second model that included both the level two predictors; pre therapy severity and the total number of sessions attended by participants. A linear cross level interaction between the transformed time

function (orthtime) and the total number of sessions attended and between orthtime and the variable pertaining to the pre therapy severity level was added to the next model to assess their association with CORE OM scores over time.

The initial Model A showed that the intercepts varied significantly across participants. Put another way, participants in the current study were heterogeneous in their levels of severity when they presented for assessment in CIPC, as evident in the significant level of variance in the level 2 (between-participant) intercepts. The statistically significant quadratic estimate in Model A (0.174, SE = 0.011) indicated that the shape of change was non-linear, decreasing initially and then increasing again toward the end of the study. This supports the idea that there was some degree of change over the course of the study period.

Change trajectories have been reported in the literature which are comparable to the one found in this study. For example, Baker et al (2002) reported that after the six month follow-up

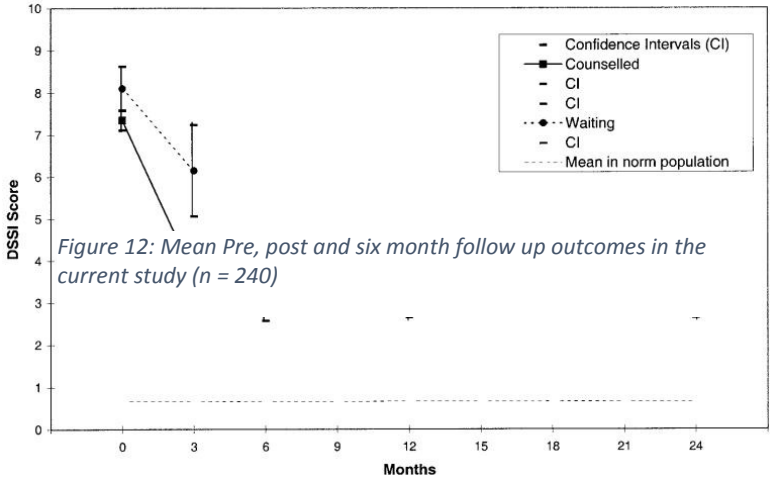


Figure 12: Mean Pre, post and six month follow up outcomes in the current study (n = 240)

Figure 13: Baker et al. (2002) - Severity of depression - Outcomes at pre post and 24 month follow up

period, the severity of symptoms for both depression and anxiety (as measured by the Delusion-Symptoms-States Inventory (DSSI; Bedford & Foulds, 1977) increased very slightly, but statistically remained significantly below the post therapy scores up to 24 months after therapy had ended (see figure 13). Baker et al. (2002) do comment that the gains (i.e. improvements) made are significantly over and above those made by the wait-list control

group, indicating a larger effect than a regression to the mean or spontaneous recovery. Ray-Sannerud et al. (2012) also reported long term follow up outcomes which showed gains being maintained two years after therapy had ended. In this study increasing scores on the Global Mental Health measure indicated improvement. This result repeated the findings by Baker et al. (2002) whereby the largest improvements were made during the course

of treatment then levelled off, but were maintained up to two years later (see figure 14).

There appear to be a distinct difference however between this main finding and that observed in this part of the current study. Mean differences in CORE OM scores between post and six

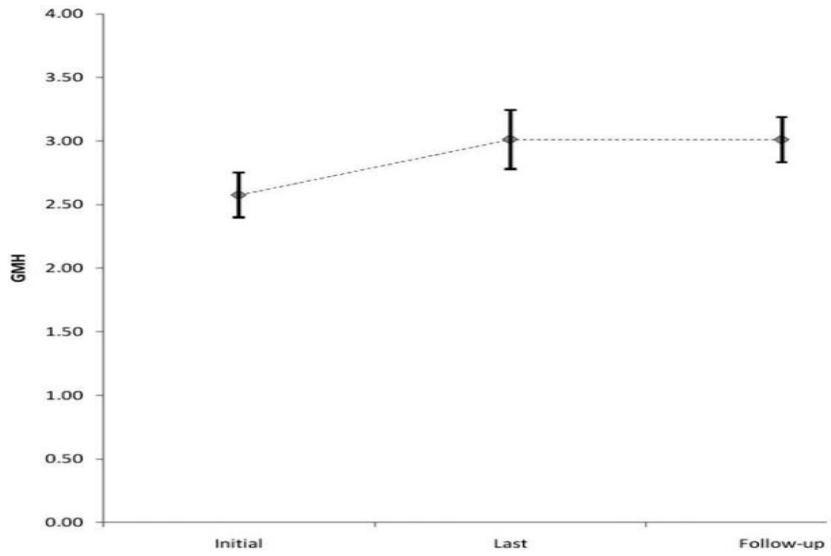


Figure 14: Ray-Sannerud et al (2012) Global Mental Health (GMH) scores at assessment, post therapy and two years after therapy

month follow up in the current study are large and significant (Δ*M* = .35; 95% CIs [0.43, 0.27]; *p* < .001). Based on the means and standard deviations provided in the respective studies, the calculated effect size for DSSI severity for depression as reported by Baker et al. (2002) between measurements at post therapy and at six months is Cohen's *d* = 0.178 and for the GMH measure in Ray-Sannerud et al (2012 Cohen's *d* = 0.002[1]. By way of contrast, the effect

[1] Effect sizes calculated for Baker et al. (2002) using means and standard deviation and for Ray-Sannerud et al. (2012) using standard deviation values derived from reported standard errors. Calculations made with RevMan Calculator by Drahota A. & Bellor E. at https://training.cochrane.org/resource/revman-calculator

size for the difference between post therapy and the six month follow up stage for the current study is Cohen's *d* = -0.552. This highlights a large and significant statistical difference in post therapy effects as reported by Ray-Sannerud (2012) and Baker et al. (2002) and in the current study which highlights the drift towards increasing levels of symptomatology for participants measured at the six month follow up stage (*n* = 240).

Also noteworthy, is the difference in change patterns during the follow up period between the current study and that reported by Ray-Sannerud et al. (2012) which were not only statistically significant, but indicative of a fundamental contrast between these results. Ray-Sannerud et al (2012) reported no difference in the number of sessions attended across the range of initial severity levels in their study ($F(3, 67) = 0.358$, $p = .783$, partial η^2 0.016). This was not the case in the current study as there was a positive relationship between the number of sessions attended and the severity of pre therapy symptoms ($r = .233$, $N = 258$, $p < .0001$, one tailed) and a significant difference between all three pre therapy severity categories ($F(2, 257) = 7.615$, $p = .001$) with those in the severe category attending a mean of 9.9 sessions, as opposed to a mean of 8.3 sessions for the mild to moderately and moderately to severe categories and 7.33 sessions for those in the mild category. From a service perspective, providing an increased number of sessions for those in higher pre therapy severity categories being beneficial in terms of outcomes has been previously flagged (see Shapiro et al., 1994) and is in keeping with the CIPC service policy whereby therapy is provided to fit client needs, i.e. it is client centred.

Results from meta-analyses of studies examining change over follow up periods of at least six months after short term psychotherapy usually do not find effects lasting longer than this. For example, Bower, Knowles, Coventry and Rowland (2011) found that over a medium to long term post therapy follow up period (i.e. up to six months) the benefits of counselling were no longer out performing those as attained from attendance at a GP. It could be argued that

comparisons between the current study and the results reported in these meta-analyses are not equal since the current study lacked a control condition. Further, according to the authors, many of the studies included were pragmatic RCTs, hence their ultimate objectives differed slightly from those of the current study for all of the reasons outlined during the review of the relevant literature in chapter two, e.g. RCTs primary focus being on extending understanding of change in psychotherapy through identification of causal mechanisms in tightly controlled settings in homogenous groups, as opposed to practice based studies were service and clients related factors are examined with the view to understanding their behaviour in the realistic setting under real world conditions (Barkham et al., 2008).

Further, while meta-analyses have generally found that four to six months after therapy the benefits of brief psychotherapy have been reduced to a level whereby participants would have found themselves with similar levels of distress had they remained in the care of their GP, participants are more satisfied with their engagement through the process of psychotherapy (Baker et al., 2002). Finally, studies which do pursue follow up data are not common, a fact highlighted by those attempting to conduct meta-analyses in this area (Cuijpers et al., 2016).

5.7 Are there differences in outcomes between and therapist Self-Declared Therapeutic Approaches?

Wampold (2013) encourages researchers and clinicians alike, to be grateful to Hans Eysenck (1952) for the withering assessment in his famous paper “The effects of psychotherapy: An Evaluation” in which he declared psychotherapy as being not only ineffective, but harmful to many who receive it. It is valid to question whether there would have been such a vast proliferation of psychotherapy research in the years since Eysenck’s assertion had he not made it, and if the direction of research would have led down the quite dichotomised path that is the specific versus common factors argument. There is some agreement that common, rather than specific therapeutic factors related to specific forms of therapy are responsible for the majority of change observed in clients as a result of psychotherapy (Wampold & Imel,

2015). However, while the discussion is far from over, it has been suggested that a move past this argument and investing energies in researching evolving psychotherapies would be a worthier pursuit (Mulder et al., 2017); a tacit recognition that further investment in "horse race" type research is a much less worthy pursuit (Norcross et al., 1993). Such is the objective of the current study, insofar as its focus is not on elucidating a superior outcome of one treatment over another, but rather on assessing the effectiveness of treatment of mild to moderate psychological problems in a very large, nationally distributed psychotherapy service, in spite of the application of many different and varied approaches taken by counsellor/ to help users of the service.

The examination of the relative outcomes between the types of therapeutic approaches taken by counsellor/therapists in the CIPC service yielded a result very much in keeping with the Dodo bird analogy (i.e. that all have won the race and so must have prizes). Overall there were nine separate types of therapeutic approaches indicated as applied by therapists during the current study period. Person Centred, Integrative and Supportive therapies accounted for 70% of all therapy type indications for the sample involved in the analysis of outcomes and types of therapy undertaken (n = 694). Further, collapsing down these groups into the single approach format required to conduct the analyses formed by far the largest group for comparison of the single approach therapies under examination (n = 195), followed by Psychodynamic and Psychoanalytic approaches which were also collapsed into a single group (n = 55) as were Cognitive and Cognitive Behavioural approaches (n = 39). The remaining therapy types: Systemic, Art and Other types accounted for just 2.9% of all approaches taken with participants. Interestingly, Structured Brief therapy and Behavioural therapy were not indicated as used by any therapists during the course of the study.

The analyses carried out in the current study replicated as far as possible that conducted by Stiles et al. (2008). The current study, like that carried out by Stiles et al. (2008) compared

outcomes between participants in a large, real-world counselling service who received a single therapeutic approach (as indicated by their counsellor/therapist) and a combination of approaches (fully described in the Method chapter). The main objective was to compare the results of both studies in order to assess if the findings in Stiles et al. (2008) were replicated across services (and jurisdictions). The results of the current study supported the general aforementioned hypothesis discussed insofar as no significant differences in outcomes were be found between therapies (and in this case between single or multiple approaches). The results in both studies were also broadly similar and are fully compared with each other in Table 24.

Table 24

Comparison of the current study single versus multiple approaches with Stiles et al. (2008).

Comparison of current study with Stiles et al. (2008) effect sizes, RCSI, reliable improvement only, no reliabe improvement and deterioration rates.										
			RCSI		Reliable Improvement		No reliable change		Relaible deterioration	
	Effect size		Row percent		Row percent		Row percent		Row percent	
Therapy approach	Current study	Stiles et al. (2008)	Current study	Stiles et al. (2008)	Current study	Stiles et al. (2008)	Current study	Stiles et al. (2008)	Current study	Stiles et al. (2008)
CBT	1.42	1.38	56.8	58.6	19.6	17.4	22.7	25.5	-	0.4
INT	1.35	1.39	69.0	58.6	10.3	18.0	20.7	21.8	-	1.5
PSY	1.36	1.29	53.1	54.4	20.4	19.7	26.5	24.1	-	1.8
PCT	1.07	1.40	50.0	59.2	20.0	19.9	30.0	20.0	-	0.9
INT+1	1.22	1.43	70.0	56.3	13.1	23.3	30.0	19.8	2.70	0.5
PCT+1	0.80	1.42	51.4	60.3	21.4	18.6	37.8	19.6	0.60	1.4

Note. Analysis carried out on the Form returners sample (*n*= 1397) INT, Integrative therapy; CBT, Cognitive, behavioural, or cognitive/behavioural therapy; PSY, psychodynamic/psychoanalytic therapy; PCT, Person Centred therapy INT+1, INT combined with one other therapy; PCT+1, PCT combined with one other therapy
Effect size, calculated as the mean difference divided by the pooled standard deviation.

None of the tests assessing differences, either between groups receiving a single, or those receiving more than one therapeutic approach reached statistical significance. Further, there were little differences (none statistically significant) in outcomes for participants between any of the different approaches. Hence, in line with Stiles et al. (2008) and other studies which have examined the question of the effectiveness of treatments delivered in primary care settings using pragmatic designs (e.g. King, Marston, & Bower, 2014; Pybis et al., 2017). As observed in Table 24, effect sizes calculated for all approaches between both studies are generally similar, with the exception of the INT+1 and PSY +1 categories where the current study observed significantly smaller effect sizes than those reported by Stiles et al. (2008).

Another point to be made regarding the outcomes reported by Stiles et al., (2008) were errors in the final effect sizes reported for all therapy approaches (i.e. single and +1). As described by Stiles et al., (2008) all effect sizes were calculated by obtaining the mean difference in per to post measures, and dividing the result by the pooled standard deviation. Using this method some of the reported effect sizes reported in Stiles et al., (2008, p. 680) are less than their true values, while others are above their original values. All of the original data have been included in Table 25. The effect sizes have been re-calculated and the correct values, along with the original reported values are provided in table 25.

Table 25

Table showing corrected effect size calculations for Stiles et al., (2008).

		Pre therapy		Post therapy		Mean pre/post difference			
Treatment group	*n*	Mean	SD	Mean	SD	Mean	SD	Reported effect size	Correct effect size
CBT	1045	17.3	6.7	8.6	6.5	8.7	6.8	1.38	1.32
PCT	1709	17.7	6.4	8.9	6.9	8.8	7.0	1.39	1.32
PDT	261	17.7	6.7	9.5	6.9	8.2	7.1	1.29	1.21
CBT+1	1035	17.3	6.0	8.4	5.7	8.9	6.2	1.40	1.52
PCT+1	1033	17.9	6.4	8.9	6.4	9.0	6.4	1.43	1.41
PDT+1	530	17.7	5.6	8.8	6.2	9.0	6.2	1.42	1.51

Cohen's d = (M2 - M1) ⁄ Sd_{pooled}, where $SD_{pooled} = \sqrt{((SD1^2 + SD2^2) / 2)}$

Since there was a basic equivalence in outcomes across therapy types found in the current study and as these results replicated the findings of Stiles et al. (2008), the next question is does this equate with the experiences of clients in similar services found in other parts of the world? Within large publicly delivered psychotherapy services across different international jurisdictions, a mix of policies can be found, with some recommending single line approaches as preferable and others (e.g. the CIPC service) allowing for multiple approaches (as specified on an accompanying list) to be utilised by service counsellor/therapists. For example, while Sawchuck et al. (2018) reported the initial findings of a large scale assessment of outcomes in primary care (*N* = 1,372), their focus was exclusively on the effectiveness of CBT to treat depression and anxiety. This appears to be because the treatment of anxiety disorders in primary care in the U.S.A. is conducted primarily through a Collaborative Care Model (CCM) guided by the Coordinated Anxiety Learning and Management (CALM) approach treatment

and provides CBT as the first line therapy approach (in conjunction with pharmacotherapy if so requested by the service user)(see Craske et al., 2011; Curran et al., 2012). For depression, the American Psychological Association (APA) offer guidelines which recommend cognitive/behavioural based therapies ((i.e. Cognitive, cognitive behavioural (CBT), and mindfulness-based, cognitive-therapy (MBCT)), as well as Interpersonal psychotherapy (IPT), Psychodynamic therapies and Supportive therapy and comment on the similarity of effectiveness found in the literature across all of these approaches (American Psychological Association, 2019).

However, while the comparative research between some of the specific psychotherapies cited in these APA guidelines was not recent but did report non-significant differences in overall outcomes across treatment approaches (i.e. Sotsky et al., 1991). In Sweden, Werbart et al. (2013b) reported that they found no differences in outcomes between the three treatment approaches; CBT, PDT and INT and they had focussed on these therapeutic approaches because they emerged as the most common approaches indicated by therapists. While in Norway, Knapstad et al. (2018) again was focussed on the effectiveness of CBT as this study was mainly occupied with replicating the IAPT model of care. It appears that the current study falls in line in terms of the overall finding of equivalency of the effectiveness of different types of therapy. It should also be stated that overall numbers for some combined therapeutic categories compiled for analyses in the current study (e.g., INT+1, *n* = 12) are low and the effect size and pre post differences should be considered in this context.

During the course of the current study of the total number of therapists who took part in this part of the analyses (*n* = 123), there appeared to an eclectic approach taken by two thirds. In total 63.5% indicated that they employed more than one approach and of those 52.5% indicated using between two and four different approaches, with the remaining 11% indicating between five and seven different approaches (32.9% indicated using one approach only, 2.7%

indicated zero and 0.9% did not enter any type of approach). Thus CIPC appears to offer, as a matter of course, a wide and varied collection of therapeutic approaches. As outlined during the review of the literature, this is as a result of service policy which was designed to be less prescriptive and reliant of a singular flavour of therapy with respect to service provision. This is in contrast to the original incarnation of the IAPT service, whereby the type of therapeutic approach offered to clients was strongly influenced by NICE recommendations for the treatment of depression and anxiety related disorders which went on to form the basis of the IAPT Service Specification (Williams, 2015). Established to provide the "…single source of authoritative and reliable guidance for service users, professionals and the public" (National Collaborating Centre for Mental Health, 2011), the NICE Guidelines for the Treatment of Common Mental Health Disorders recommend a stepped care approach including short and longer term psychological treatments. All treatments recommended by NICE for inclusion in a stepped care service were are selected by way of a review of the current evidence conducted by the relevant NICE guidance development group which at the time accepted the dominance of CBT RCT evidence. However, as NICE guidelines have been updated on foot of RCT evidence for the effectiveness of additional bone fide therapies, this has led to an expansion of the forms of therapeutic approach in the IAPT service in the UK in recent years. Relatedly, while serveral therapy approaches were utilised by the majority of therapists to treat individual clients, of the substantial minority (i.e. 32.9%) whom reported using a single approach it is reasonable to ask what proportion employed a transdiagnostic approach given the high incidence of anxiety and depressive symptoms in the sample. Transdiagnostic treatments are "those that apply the same underlying treatment principles across mental disorders, without tailoring the protocol to specific diagnoses (McEvoy et al., 2009, p. 21). Meta-analysis has shown transdiagnostic treatments to outperform control conditions for the treatment of anxiety, depression and quality of life (Newby et al., 2015). Hence, it would seem prudent for the CIPC service to be aware of current trends in this area of

research with the aim of continuous quality improvement of the therapeutic approaches being delivered to its clients in mind.

Also, the current study did not involve any measurement of therapeutic alliance. It was concerned with the *outcome* half of what Ardito and Rabellino (2011) refer to as the *outcome* and *process* dichotomy of psychotherapy research endeavours. This refers to examining the post therapy outcome in order to ascertain whether recipients of psychotherapy have improved, deteriorated or remained unchanged after receiving therapy. The process side focusses on examination of the processes involved in the therapeutic exchange between client and practitioner which can also be examined for changes which occur within both parties. Relevant to the current study were the differences in outcomes due to the mode of therapy applied by therapists, which appeared to be largely insignificant.

5.8 HRQOL (Health Related Quality of Life) – was it useful?

At the heart of the current study's assessment of change using the HRQOL was an examination of changes in clients' perceptions of their overall health, and the number of days they felt negatively impacted by the problem for which they sought help in CIPC. The goal was to determine whether the HRQOL act as a potential proxy measure of clients' level of severity for use by referrers of clients into the CIPC service.

The HRQOL has demonstrated good levels of correlation with certain sociodemographic and clinical factors associated with major depressive disorder in a sample of primary care attendees of a psychological services. Trivedi et al. (2006) showed that greater depressive symptomatology was associated with poorer HRQOL in a sample of people attending services in the U.S. (*N* = 1,397).

There were also some encouraging signs that the HRQOL may be of some use as a pre-assessment screening tool for the CIPC service. Linear regression results for the Form Returners sample (n = 1,397) showed a significant association between pre therapy CORE OM

levels and the number of days participants' perceived their psychological and physical health to be negatively impacted and the days their normal activity levels were reduced, i.e. participants reporting higher numbers of "bad" days had higher pre therapy CORE OM scores. This was quantifiable using the predicted overall increase in average CORE OM scores due to the number of mentally and physically unhealthy days and number of days activity was limited in the previous 30 day period and was 0.041 (i.e. 0.022 + 0.006 + for 0.013 respectively). Once this was adjusted to account for a maximum 30 day period, the combined mean number of mentally and physically unhealthy days (22.7) plus the average number of activity limited days (5.5) was 0.04. Hence, for the Form Returners sample, this increase represents a 0.04 increase in all predicted CORE OM scores at pre therapy due to the effect of HRQOL.

Based on this result, the Form Returners sample was re-analysed to investigate the additional number of clients who presented for assessment which would be considered clinical cases based on the influence of HRQOL on their CORE OM scores. Results showed that an additional 18 participants who presented with levels of pre therapy CORE OM scores in the sub-clinical range without the effect of HRQOL, would now be predicted to present with CORE OM scores in the clinical range (i.e. *n* = 18 presenting with CORE OM scores between 0.96 and 0.99) based on the effect of worse HRQOL. This number rose to 29 participants for the All Participants sample. This piece of analysis identified a cohort of participants at the fringe of being clinical cases when considered by CORE OM scores in isolation, but once these participants were considered with the additional impact of HRQOL scores of their COREOM scores, they were elevated into the category of clinical caseness. Hence, there is potential for the HRQOL to be utilised by CIPC service as a way of identifying potential clinical cases at the point of referral, mainly by GPs, rather than at the assessment stage of the clients' journey through the service. Further, this approach has the advantage of using a much shorter questionnaire (i.e. the HRQOL) saving considerable time in its administration.

Of importance when assessing pre to post changes using the HRQOL is the relevance of both the clinical and statistical changes which were observed. Most encouraging regarding the responses to the question related to participants' perceptions of their general health (I.e. Question 1 on the HRQOL questionnaire), was that the direction of change between pre and post therapy moved in a direction which exhibited real benefit for participants (i.e. reductions in negative perceptions) in all categories – *Excellent, Very Good, Good, Fair and Poor*. Studies which have examined the quality of therapy have suggested that the reliance on a single quantitative measure of clients' experiences after therapy may cause vital information regarding that experience to be lost (McElvaney & Timulak, 2013). The HRQOL offers an opportunity to expand the understanding of the subjective perception of one's health and give an account of the period for which they have been suffering before therapy, and how those perceptions may have changed after therapy has ended. Since the current study found some common ground between the HRQOL and the CORE OM outcomes both statistically and clinically, perhaps more of this vital information can be retained in a format (i.e. number of days) that is more readily understood and evident to both the client and their counsellor/therapist.

Further, The Centre for Disease Control (CDC) in the U.S. have reported that the HRQOL has recorded very high response rates to questions related to the number of days that mental and physical health was negatively affected and activity was limited (98.5% - 99.3%) by very large numbers of respondents (over 900,000). The high response rates coupled with the large N are used cited by the CDC in order to justify the claim that the method of capturing and representing a numerical estimation of HRQOL in this way is more easily understood by health policy funders and other stakeholders, and that the monthly timeframe it ask respondents to consider has also been found to be easily understood by people in general (Centers for Disease Control and Prevention, 2000). If this is the case, the data gathered in the current study pertaining to participants' perceptions of their general health and the number of days they are

affected by their mental and physical problems may have the potential to assist the CIPC service in a number of other ways. As the HRQOL has been shown to be sensitive to seasonal and time trends (Centers for Disease Control and Prevention, 2000), it may also have the potential to assist in the assessment of the impact of service wide change if it were administered on a regular basis and the results analysed for change over a suitable and relevant time period using the appropriate statistical approaches.

HRQOL items which probed the number of days that participants were negatively affected by their mental and physical problems, and how much their activity was limited correlated strongly with CORE OM outcomes in both the directions and magnitudes expected. In terms of expectations, it is not surprising that if participants feel better in terms lower levels of symptoms according to their CORE OM scores, they will inevitably report less days when they are negatively impacted by their problems caused by depression, anxiety etc. However, as mentioned during the literature review, definitions of recovery differ between service user and clinically based literature (Ronk et al., 2016) and what underlies this assessment of the linkages between the CORE OM and the HRQOL is a possible insertion of the HRQOL into the CIPC service as a plain language, user-friendly and easily understood tool by which to 1) gather clinically relevant information before a client engages with the service and 2) provide the opportunity at that stage for potential clients to meaningfully contribute to the beginning of their journey towards feeling better.

From the results of the analyses involving the HRQOL described earlier, the HRQOL has potential for use as a very short form screening tool which could: 1) be used by GPs as part of the referral process to highlight patient status *before* they opt into the service. 2) As an admin tool to monitor incoming clients numbers, i.e. the status of those referred but not opted in or 3) As an additional source of clinical information for counsellor/therapists during assessment,

treatment or at the end of therapy used to add context, or open up conversations about the particular areas highlighted by the HRQOL.

5.9 Does employment status or gender have an effect on client outcomes?

There was no evidence during the course of this study of differential outcomes, across any measure due to sex or employments status. As discussed in the literature review, in terms of research conducted in the primary care sector, some moderate effects have been indicated related to sex differences whereby women were observed to attend a higher number of session according to the CORE Therapy Assessment Form (TAF), but did not translate into significant statistical differences (Shepherd et al., 2005). This is not to say that differences based on both sex and employment status have not been discussed, the discussion however appears to centre mainly on help seeking differences (Paul & Moser, 2009) rather than differences as related to outcomes measured by the CORE OM.

5.10 Limitations

As evidenced during the literature review, policy decisions made in accordance with the objectives of each particular service determine the make-up of client populations attending large public health services. There some differences however, between the cohort of referred patients into the CIPC service and the IAPT service in the UK and those included in the large multi-site study by Sawchuk et al. (2018) in the U.S.A. In both of these jurisdictions, clients may avail of guided self-help and computerised CBT as early steps into treatment. However, neither of these interventions are available from CIPC nor from the Swedish service evaluated by Werbart et al. (2013b). Also, the minimum age of entry to available services for clients in some studies was 15 (Werbart et al., 2013b) or 16 (Barkham, Stiles, Connell, & Mellor-Clark, 2012) years, while all other comparable services required a minimum age of 18 for entry. Hence, study samples from across jurisdictions differ from one another in certain respects.

Further, the impact of other differences may be difficult to ascertain the impact on service factors such as drop-out and overall therapy outcomes.

While attrition effects many studies, practice base studies such as the current study accept client drop out as a signal that a proportion of participants are choosing to leave the service once they feel they no longer need therapy. Ideally this occurs in conversation with their therapist to bring a naturally agreed ending to their therapeutic relationship. However, since this does not always occur practice base studies deal with the phenomenon as part of the mechanism of conducting research in the real world. To begin with there was a 9.9% drop out rate of participants who consented to participate and actual pre therapy CORE OM forms completed (46.7% average across all HRQOL items). In terms of post therapy completion rates, an overall rate of 46.3% was achieved for the CORE OM questionnaire. Attrition rates of this magnitude are of great consequence to RCT and pragmatic style studies because they attempt to infer causation through manipulation and isolating the effects of particular variables while controlling for others. However, the impact of dropout on practice-based studies has different consequences and can, in certain circumstances be changed into benefits as demonstrated by Grant (2006) who built a benchmarking resource to help counsellor/therapists consider client attrition in a different way so they could more fully understand attrition, with a view to ultimately helping reduce the phenomenon in the overall service.

Another opportunity presented during this study was to fully track sessional outcome data with a large number of clients in the service in order to analyse the data mainly for tracking drop-out and outcome data together. However, due to inconsistency in the administering of outcome questionnaires throughout the CIPC service at a national scale this was not possible during the course of this study. This did impact the level of detail that could be applied to this particular analyses. Further, the multi-level analyses of the effect of initial severity and number of sessions on outcomes over the pre, post and six month follow up period would

have also benefitted from data at the level of individual sessions. Also, it is a fact that the current study, like all studies addressing the efficacy and effectiveness of psychotherapy, can only speak to the effects for those participants who remain in the service. Any data for those who decide to leave the service, whether it is because therapy has, or has not been successful, is not available and therefore cannot be interrogated.

Further and in relation to the analytic approach in the current study, it is noted that a large number of tests were conducted on the same data set (or sub-sets thereof). This process increases the chances of identifying false positives during the course of generating the resulting statistical analytic output (Heck et al. 2014). Bonferroni correction can be used in order to reduce the likelihood of obtaining false-positives (i.e. Type I errors) when multiple pair wise tests are performed in order to test multiple hypotheses.

5.11 Recommendations and future research

An issue which emerged during the course of the current study was concern regarding the quality of certain data inputted into the CORE OM system by counsellor/therapists. Certain items, fields and sections of the centralised web-based client data input system appeared to be misunderstood, or their overall purpose misinterpreted by counsellor/therapists. For example, the selection of therapy approaches sometimes appeared to be over populated, with almost 10% indicating the simultaneous use of four or more therapeutic approaches on a single client during a single episode. Moreover, certain approaches were not indicated as utilised at all during the course of the whole study, i.e. Structured Brief and Behavioural (i.e. not cognitive or CBT). Hence, there may be cause for an inspection of the CORENET system currently in use in Ireland with a view to customising it to better suit the Irish psychotherapy context. As the sole method of recording client information for the CIPC service, input options available on the CORENET system should accurately reflect the Irish psychotherapy environment in order to best facilitate counsellor/therapists in order to provide their

professional services in that context. This endeavour would however necessitate being mindful of the advantages and ease with which the current dataset can be compared with those using the same system in other jurisdictions, which is the foundation of the rationale of building practice-based networks. Hence, a careful and collaborative approach to such an endeavour would be advised in order to gain the advantages of a fully bespoke client data tracking system for Ireland, while maintaining the ability to easily share outcome data in a collaborative ways with researchers in other jurisdictions.

Additional research questions were omitted from the current study due to time and resource limitations. As mentioned, the collection of sessional data could not be fully implemented. The roll out of the CORE-10 (a shortened 10 item version of the full CORE OM 34) has been ongoing in the CIPC service and once fully adopted will provide the opportunity to more fully engage with the question of drop-out, therapy progress over the course of time and individual client progress tracking as a practitioner tool. Additionally, once fully implemented across the CIPC service, comparisons of outcomes between other measures of psychological distress such as the PHQ-9, GAD-7 and the CORE-10 and will be more easily facilitated because all three questionnaires will be of similar length (i.e. between seven and 10 questions) which can be important in terms of the varying sensitivity to change between longer and shorter measures.

The acquisition and analysis of follow up data in order to address numerous questions related to the longer terms effects of therapy are frequently aired (see Bower et al., 2011; Norcross et al., 1993; Pybis et al., 2017). The current study did include a follow up element which can in some ways address the question of the after effects of therapy. However, the current study pursued additional data at six and 12 month follow up periods, much of these data were not available in time for inclusion in the current report. These data will be submitted to the CIPC service over the course of the months ahead and will form the basis for a large scale long term follow up study.

Another element due for inclusion in the current study, but omitted due to time and resource constraints was an examination of the relationship between psychotropic medication prescription rates and therapy outcomes. This element of the study involved input from multiple stakeholders and the navigation of complex ethical compliance issues in addition to the construction of a data submission process that withstood scrutiny from a series of gatekeepers. This process took a great deal longer than anticipated and hence, while these data have begun to be received by the CIPC service, the entire psychotropic medication dataset along with paired outcomes belonging to those participants who consented to being involved will arrive for analyses by the researcher at a later date. This is not least due to the submission of those data also being subject to the imposition of a follow up period of one year after therapy has ended for the last client involved.

5.12 Conclusion

As a relatively new service (launched in 2013) funders and directors of the CIPC service sought an evaluation as a means of providing an empirical basis to the reporting of clinical outcomes as a result of the psychological services provided to clients. The study outputs were also required to be easily comparable with those from services of a similar profiles in other jurisdictions and easily utilised by clinicians, researchers, funders and managers of similar services elsewhere. As a result of these requirements, the primary aims of this study were to provide a profile of the service in terms clinical outcomes, compare outcomes of the CIPC service with those from services in other jurisdictions and examine in detail the behaviour of some of the important variables involved in the provision of therapy services on a large scale. A practice based approach was considered the most appropriate method of conducting the study as CIPC is a real world service dealing with real clients who present to the service with a high level of heterogeneity in terms of presenting problems, levels of distress and demographic characteristics.

As observed in the earlier section of the discussion, Ireland is a country changing both demographically and societally (see pages 149-152). One of the concerns highlighted by this evaluation was the lack of attendance of non-Irish clients in the CIPC service. Access to the service by people belonging to Black and Minority Ethnicity groups (BME) is very low (just 2.6% see Table 2) in relation to the real number of people from BME community in Ireland. Reasons for this may be simple and straightforward, i.e. lack of confidence in the GP community to refer BME patients or people in this community may not know that the service exists (Mental Health Reform, 2014). On the other hand, reasons could be complex and require multi-agency involvement in order to improve the situation. Whatever the solutions to this problem are needed, the problem has been starkly identified during the course of this evaluation.

This study also examined the HRQOL and its congruence with the CORE OM in terms representing the presence and level of distress using both measures' very different scale options and target outputs; the CORE OM with its 0 – 136 scale over 34 questions across four domains, and the HRQOL measuring a) general health, and days (out of the last 30) where b) physical health was not good, c) mental health was not good and d) activity was limited due to either or both of the previous two questions. Use of the HRQOL offered the opportunity to add dimensionality to the experiences of participants in the study and round out pre post CORE OM scores with the easily understood and comparable "number of days" concept. Of great encouragement to counsellor/therapists in the CIPC service is the concrete reduction in the time participants felt psychologically better post therapy – reduced by an average of 8 days. This represents a 25% reduction in real time for those experiencing distress because of the problems they presented to the service with, which is a very good indictment of the effect of therapy provided.

All HRQOL items moved in the right direction between pre and post therapy measurements in terms of demonstrating benefits for participants, but importantly mirrored increases and

decreases in CORE OM scores as verified by the large and significant correlational analyses results. This opens up the possibility of a justification of further and more detailed examination of the HRQOL as a pre assessment screening tool for use by CIPC referral agents. Currently, the CIPC service does not formally screen potential clients for the severity of their problems using a measurement tool. All referrals are screened by service coordinators prior to being allocated for assessment. However, anecdotal evidence suggests that this process differs between CIPC areas and the actual proportion of additional information supplied with referrals is currently unknown, but its successful incorporation with client information is generally reliant on the legibility of the documentation received from referring agents who are usually GPs. GPs handwriting has been studied and generally found to be markedly worse than that of other professions (Lyons et al., 1998), which can cause real problems in a medical care environment (Callen et al., 2008). Hence, additional information which could be gathered from referrers (including GPs) and passed to the CIPC service regarding details about patients' problems, i.e. the duration and severity in a format that is responsive, valid, self-reported and importantly brief, could be a welcome development in the CIPC referral process and the use of the HRQOL at the referral stage may provide such an opportunity.

Clinical outcomes as a result of therapy received in CIPC are very much in line with those reported in Sweden, Norway, The U.S.A. and the U.K. Bearing in mind the population under study across jurisdictions, it is not too surprising that this was the case. Importantly, the goals of CIPC are very much aligned with the largest service of its kind in the world situated geographically next door to Ireland in the U.K. (Clark, 2018a). The CIPC model is set to deal with mild to moderate problems occurring in the general population, those that do not require specialist treatment of the forms delivered in most secondary care services, where more severe experiences are normally dealt with. However, it remains the case, as emerged during this evaluation, that a significant proportion of clients seen by the CIPC service (14.7%; All Participants sample) as experiencing symptoms considered by their counsellor/therapists as

severe. This proportion is higher than that reported in a large scale service in the U.K. (*N* >12,000) where the number was 12.5%. However when one considers that the CIPC service exclusion criteria specify conditions which are accepted for referral to similar services in the U.K., (e.g. PTSD, OCD and eating disorders) this percentage 14.7% must be considered as being much higher than one would expect. Hence, it appears that the CIPC service deals with a high proportion of clients who engage with the service and are provided counselling and psychotherapy, but for whom the service may not have been specifically designed from the outset. Coupled with the results of the pre post and six month follow up – whereby those in the severe category were observed to make substantial gains between pre and post therapy, dropping from being in the severe range to just above the mild cut off point, but deteriorated between post therapy and six months later into the moderate to severe range once more, a number of questions arise. What is the proportion of unsuitable referrals arriving at CIPC, and how many may be accepted due to the lack of the availability of suitable service elsewhere? Are counsellor/therapists in CIPC accepting clients for treatment even when they may be unsuitable for the same reason? Is the CORE Therapy Assessment Form accurately capturing severity ratings or do counsellor/therapists have a vastly different conception of severity when rating the depth of clients' problems? There are more questions which arise from this piece of information which should be addressed in future studies.

This study completes part of the first national evaluation of CIPC service in Ireland. The researcher could find no other study of this type that incorporated a complete national dataset, making this study unique in its claim to being the first, geographically representative, national evaluation of counselling and psychotherapy services in a primary care setting, anywhere.

References

Alfredsson, J. (2002). Dagar med full hälsa i Gävleborgs län: Ett försök att mäta hälsorelaterad livskvalitet i befolkningen med hjälp av CDC:s frågebatteri [Days with full health in Gävleborg County: an attempt to measure health-related quality of life in the population with the help of the CDC's Health Days questionnaire]. *Gävleborg County Council. Community Medicine Gävleborg*, (2) 1104-7763.

Alkin, M. C., & Ellett, F. (1979). An inquiry into the nature of evaluation theorizing. *Studies in Educational Evaluation*, *5*(2), 151–156. https://doi.org/10.1016/0191-491X(79)90022-1

American Psychological Association. (2013). Recognition of psychotherapy effectiveness. *Psychotherapy*, *50*(1), 102–109. https://doi.org/10.1037/a0030276

American Psychological Association. (2019, February 11). *Clinical Practice Guideline for the Treatment of Depression*. Https://www.Apa.Org. https://www.apa.org/depression-guideline/index

Ammerman, A., Smith, T. W., & Calancie, L. (2014). Practice-Based Evidence in Public Health: Improving Reach, Relevance, and Results. *Annual Review of Public Health*, *35*(1), 47–63. https://doi.org/10.1146/annurev-publhealth-032013-182458

Baker, R., Allen, H., Gibson, S., Newth, J., & Baker, E. (1998). Evaluation of a primary care counselling service in Dorset. *The British Journal of General Practice: The Journal of the Royal College of General Practitioners*, *48*(428), 1049–1053.

Baker, R., Baker, E., Allen, H., Golden, S., Thompson, P., Hollingbery, T., Newth, J., & Gibson, S. (2002). A naturalistic longitudinal evaluation of counselling in primary care. *Counselling Psychology Quarterly*, *15*(4), 359–373.

Bakker, A., van Balkom, A. J., Spinhoven, P., Blaauw, B. M., & van Dyck, R. (1998). Follow-up on the treatment of panic disorder with or without agoraphobia: A quantitative review. *The Journal of Nervous and Mental Disease*, *186*(7), 414–419.

Baldwin, S. A., Berkeljon, A., Atkins, D. C., Olsen, J. A., & Nielsen, S. L. (2009). Rates of change in naturalistic psychotherapy: Contrasting dose–effect and good-enough level models of change. *Journal of Consulting and Clinical Psychology*, *77*(2), 203–211. https://doi.org/10.1037/a0015235

Baldwin, S. A., Imel, Z. E., Braithwaite, S. R., & Atkins, D. C. (2014). Analyzing multiple outcomes in clinical research using multivariate multilevel models. *Journal of Consulting and Clinical Psychology*, *82*(5), 920–930. https://doi.org/10.1037/a0035628

Barkham, M. (2005). Suitability and utility of the CORE-OM and CORE-A for assessing severity of presenting problems in psychological therapy services based in primary and secondary care settings. *The British Journal of Psychiatry*, *186*(3), 239–246. https://doi.org/10.1192/bjp.186.3.239

Barkham, M, Connell, J., Stiles, W. B., Miles, N. V., Margison, F., Evans, C., & Mellor-Clark, J. (2006). Dose-Effect Relations and Responsive Regulation of Treatment Duration: The Good Enough Level. *Journal of Consulting and Clinical Psychology*, *74*(1), 160–167. https://doi.org/10.1037/0022-006X.74.1.160

Barkham, M., Hardy, G. E., & Mellor-Clark, J. (2010). Improving Practice and Enhancing Evidence. In M. Barkham, G. E. Hardy, & J. Mellor-Clark (Eds.), *Developing and Delivering Practice-Based Evidence* (pp. 327–353). John Wiley & Sons, Ltd. http://doi.wiley.com/10.1002/9780470687994.ch13

Barkham, M, & Margison, F. (2007a). Practice-based evidence as a complement to evidence-based practice: From dichotomy to chiasmus. *Handbook of Evidence-Based Psychotherapies: A Guide for Research and Practice*, 443–476.

Barkham, M, & Margison, F. (2007b). Practice-based Evidence as a Complement to Evidence-based Practice: From Dichotomy to Chiasmus. In C. Freeman & M. Power (Eds.), *Handbook of Evidence-based Psychotherapies* (pp. 443–476). John Wiley & Sons Ltd. http://doi.wiley.com/10.1002/9780470713242.ch25

Barkham, M, Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., Benson, L., Connell, J., Audin, K., & McGrath, G. (2001). Service profiling and outcomes benchmarking using the CORE-OM: Toward practice-based evidence in the psychological therapies. *Journal of Consulting and Clinical Psychology*, *69*(2), 184–196. https://doi.org/10.1037/0022-006X.69.2.184

Barkham, M., Rees, A., Stiles, W. B., Shapiro, D. A., Hardy, G. E., & Reynolds, S. (1996). Dose–effect relations in time-limited psychotherapy for depression. *Journal of Consulting and Clinical Psychology*, *64*(5), 927–935. https://doi.org/10.1037/0022-006X.64.5.927

Barkham, M, Stiles, W. B., Connell, J., Twigg, E., Leach, C., Lucock M, Mellor-Clark J, Bower P, King M, Shapiro DA, Hardy GE, Greenberg L, & Angus L. (2008). Effects of psychological therapies in randomized trials and practice-based studies. *British Journal of Clinical Psychology*, *47*(4), 397–415.

Barkham, M., Stiles, W. B., Lambert, M. J., & Mellor-Clark, J. (2010). Building a Rigorous and Relevant Knowledge Base for the Psychological Therapies. In M. Barkham, G. E. Hardy, & J. Mellor-Clark (Eds.), *Developing and Delivering Practice-Based Evidence* (pp. 21–61). John Wiley & Sons, Ltd. http://doi.wiley.com/10.1002/9780470687994.ch2

Barkham, M., Stiles, W. B., & Shapiro, D. A. (1993). The shape of change in psychotherapy: Longitudinal assessment of personal problems. *Journal of Consulting and Clinical Psychology*, *61*(4), 667–677.

Barkham, Michael, Mullin, T., Leach, C., Stiles, W. B., & Lucock, M. (2007). Stability of the CORE-OM and the BDI-I prior to therapy: Evidence from routine practice. *Psychology and Psychotherapy: Theory, Research and Practice*, *80*(2), 269–278. https://doi.org/10.1348/147608306X148048

Barkham, Michael, Stiles, W. B., Connell, J., & Mellor-Clark, J. (2012). Psychological treatment outcomes in routine NHS services: What do we mean by treatment effectiveness? *Psychology & Psychotherapy: Theory, Research & Practice*, *85*(1), 1-16 16p. rzh. https://doi.org/10.1111/j.2044-8341.2011.02019.x

Batt, V., Nic Gabhainn, S., & Falvey, F. (2002). *Perspectives on the Provision of Counselling for Women in Ireland*. Women's Studies Centre and Centre for Health Promotion Studies, National University of Ireland, Galway.

Beck, A. T. (Ed.). (1979). *Cognitive therapy of depression*. Guilford Press.

Bedford, A., & Foulds, G. A. (1977). Validation of the Delusions-Symptoms-States Inventory. *British Journal of Medical Psychology*, *50*(2), 163–171. https://doi.org/10.1111/j.2044-8341.1977.tb02409.x

Bergin, A. E. (1971). The evaluation of therapeutic outcomes. In A. E. Bergin & S. L. Garfield (Eds.), *The handbook of psychotherapy and behaviour change* (pp. 217–270). Wiley.

Bewick, B. M., McBride, J., & Barkham, M. (2006). When clients and practitioners have differing views of risk: Benchmarks for improving assessment and practice. *Counselling and Psychotherapy Research*, *6*(1), 50–59. https://doi.org/10.1080/14733140600581481

Bewick, B. M., Trusler, K., Mullin, T., Grant, S., & Mothersole, G. (2006). Routine outcome measurement completion rates of the CORE-OM in primary care psychological therapies and counselling. *Counselling and Psychotherapy Research*, *6*(1), 33–40. https://doi.org/10.1080/14733140600581432

Bjorner, J. B., Kristensen, T. S., Orth-Gomér, K., Tibblin, G., Sullivan, M., Westerholm, P., & Swedish Council for Planning and Coordination of Research, F. (1996). *Self-rated health: A useful concept in research, prevention and clinical medicine*.

Blount, A. (2003). Integrated primary care: Organizing the evidence. *Families, Systems & Health: The Journal of Collaborative Family HealthCare*, *21*(2), 121–121.

Bor, R., & McCann, D. (1999). *The practice of counselling in primary care.* (R. Bor & D. McCann, Eds.; 1999-04031-000). Sage Publications Ltd; psyh. http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1999-04031-000&site=ehost-live

Bortolotti, B., Menchetti, M., Bellini, F., Montaguti, M. B., & Berardi, D. (2008). Psychological interventions for major depression in primary care: A meta-analytic review of randomized controlled trials. *General Hospital Psychiatry*, *30*(4), 293–302. https://doi.org/10.1016/j.genhosppsych.2008.04.001

Bourke, M., & Byrne, M. (2012). Evaluation of a pilot primary care adult mental health practitioner – delivered service. *The Irish Psychologist*, *38*(11), 262–268.

Bower, P. J., & Rowland, N. (2006). Effectiveness and cost effectiveness of counselling in primary care. In *Cochrane Database of Systematic Reviews*. John Wiley & Sons, Ltd. http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD001025.pub2/abstract

Bower, P, Knowles, S., Coventry, P. A., & Rowland, N. (2011). Counselling for mental health and psychosocial problems in primary care. *Cochrane Database of Systematic Reviews*, *9*, CD001025. https://doi.org/10.1002/14651858.CD001025.pub3

Bower, P., Kontopantelis, E., Sutton, A., Kendrick, T., Richards, D. A., Gilbody, S., Knowles, S., Cuijpers, P., Andersson, G., Christensen, H., Meyer, B., Huibers, M., Smit, F., Straten, A. van, Warmerdam, L., Barkham, M., Bilich, L., Lovell, K., & Liu, E. T.-H. (2013). Influence of initial severity of depression on effectiveness of low intensity interventions: Meta-analysis of individual patient data. *BMJ*, *346*, f540. https://doi.org/10.1136/bmj.f540

Bower, P, Rowland, N., Mellor Clark, J., Heywood, P., Godfrey, C., & Hardy, R. (2002). Effectiveness and cost effectiveness of counselling in primary care. In The Cochrane Collaboration (Ed.), *Cochrane Database of Systematic Reviews*. John Wiley & Sons, Ltd. https://doi.org/10.1002/14651858.CD001025

Bower, Peter, Wallace, P., Ward, E., Graffy, J., Miller, J., Delaney, B., & Kinmonth, A. L. (2009). Improving recruitment to health research in primary care. *Family Practice*, *26*(5), 391–397. https://doi.org/10.1093/fampra/cmp037

Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods*, *23*(3), 389–411. https://doi.org/10.1037/met0000159

Brown, G. S. J., Lambert, M. J., Jones, E. R., & Minami, T. (2005). Identifying highly effective psychotherapists in a managed care environment. *The American Journal of Managed Care*, *11*(8), 513–520.

Bryan, C. J., Corso, M. L., Corso, K. A., Morrow, C. E., Kanzler, K. E., & Ray-Sannerud, B. (2012). Severity of mental health impairment and trajectories of improvement in an integrated primary care clinic. *Journal of Consulting and Clinical Psychology*, *80*(3), 396–403. https://doi.org/10.1037/a0027726

Butler, M. (2002). *Evaluation in the Irish health sector*. Institute of Public Administration.

Butterworth, P., Leach, L. S., Strazdins, L., Olesen, S. C., Rodgers, B., & Broom, D. H. (2011). The psychosocial quality of work determines whether employment has benefits for mental health: Results from a longitudinal national household panel survey. *Occupational and Environmental Medicine*, *68*(11), 806–812. JSTOR.

Cacioppo, J. T., Semin, G. R., & Berntson, G. G. (2004). Realism, Instrumentalism, and Scientific Symbiosis: Psychological Theory as a Search for Truth and the Discovery of Solutions. *American Psychologist*, *59*(4), 214–223. https://doi.org/10.1037/0003-066X.59.4.214

Cahill, J., Barkham, M., & Stiles, W. B. (2010). Systematic review of practice-based research on psychological therapies in routine clinic settings. *British Journal of Clinical Psychology*, *49*(4), 421–453. https://doi.org/10.1348/014466509X470789

Cairns, M. (2014). Patients who come back: Clinical characteristics and service outcome for patients re-referred to an IAPT service. *Counselling and Psychotherapy Research*, *14*(1), 48–55. https://doi.org/10.1080/14733145.2013.770895

Callan, S., & Fry, B. (2012). *Commissioning effective talking therapies.* The Centre For Social Justice. https://www.centreforsocialjustice.org.uk/core/wp-content/uploads/2016/08/TalkingTherapiesPaper.pdf

Callen, J. L., Alderton, M., & McIntosh, J. (2008). Evaluation of electronic discharge summaries: A comparison of documentation in electronic and handwritten discharge summaries. *International Journal of Medical Informatics*, *77*(9), 613–620. https://doi.org/10.1016/j.ijmedinf.2007.12.002

Campbell-Sills, L., Norman, S. B., Craske, M. G., Sullivan, G., Lang, A. J., Chavira, D. A., Bystritsky, A., Sherbourne, C., Roy-Byrne, P., & Stein, M. B. (2009). Validation of a Brief Measure of Anxiety-Related Severity and Impairment: The Overall Anxiety Severity and Impairment Scale (OASIS). *Journal of Affective Disorders*, *112*(1–3), 92–101. https://doi.org/10.1016/j.jad.2008.03.014

Cape, J., Whittington, C., Buszewicz, M., Wallace, P., & Underwood, L. (2010). Brief psychological therapies for anxiety and depression in primary care: Meta-analysis and meta-regression. *BMC Medicine*, *8*, 38–50. https://doi.org/10.1186/1741-7015-8-38

Cartwright, N. (2007). Are RCTs the Gold Standard? *BioSocieties*, *2*(1), 11–20. https://doi.org/10.1017/S1745855207005029

Castonguay, L. G., Barkham, M., Lutz, W., & McAleavey, A. A. (2013). Practice-oriented research: Approaches and application. In M. J. Lambert (Ed.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (6th ed., pp. 85–133). John Wiley & Sons, Inc.

Centers for Disease Control and Prevention. (2000). *Measuring healthy days: Population assessment of health-related quality of life.* http://www.cdc.gov/hrqol/pdfs/mhd.pdf

Central Statistics Office. (2019a). *Population and Migration Estimates—April 2019* (Annual Population and Migration Estimates No. 20190829093831; pp. 1–10). Central Statistics Office.

https://pdf.cso.ie/www/pdf/20190829093831_Population_and_Migration_Estimates_April_2019_full.pdf

Central Statistics Office. (2019b, September 25). *Central Statistics Office—Labour Market Report Q2 2019*. HOME / STATISTICS / LABOUR MARKET / ARCHIVE / RELEASE ARCHIVE 2019. http://www.cso.ie/px/pxeirestat/Database/eirestat/Labour%20Force%20Survey%20Quarterly%20Series/Labour%20Force%20Survey%20Quarterly%20Series_statbank.asp?SP=Labour%20Force%20Survey%20Quarterly%20Series&Planguage=0

Chilisa, B., & Kawulich, B. (2012). Selecting a research approach: Paradigm, methodology, and methods. *C Wagner, B Kawulich, & M Garner, Doing Social Research: A Global Context*, 51–61.

Chin, W. Y., Chan, K. T. Y., Lam, C. L. K., Lam, T. P., & Wan, E. Y. F. (2015). Help-seeking intentions and subsequent 12-month mental health service use in Chinese primary care patients with depressive symptoms. *BMJ Open*, *5*(1). https://doi.org/10.1136/bmjopen-2014-006730

Christensen, L., & Mendoza, J. L. (1986). A method of assessing change in a single subject: An alteration of the RC index. *Behavior Therapy*, *17*(3), 305–308. https://doi.org/10.1016/S0005-7894(86)80060-0

Clark, D. M. (2011). Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: The IAPT experience. *International Review of Psychiatry*, *23*(4), 318–327. https://doi.org/10.3109/09540261.2011.606803

Clark, D. M. (2018a). Realizing the Mass Public Benefit of Evidence-Based Psychological Therapies: The IAPT Program. *Annual Review of Clinical*

Psychology, *14*(1), 159–183. https://doi.org/10.1146/annurev-clinpsy-050817-084833

Clark, D. M. (2018b). Realising the Mass Public Benefit of Evidence-Based Psychological Therapies: The IAPT Program. *Annual Review of Clinical Psychology*, *14*, 159–183. https://doi.org/10.1146/annurev-clinpsy-050817-084833

Clark, D. M., Layard, R., Smithies, R., Richards, D. A., Suckling, R., & Wright, B. (2009). Improving access to psychological therapy: Initial evaluation of two UK demonstration sites. *Behaviour Research and Therapy*, *47*(11), 910–920. https://doi.org/10.1016/j.brat.2009.07.010

Clift, S. (2012). Creative arts as a public health resource: Moving from practice-based research to evidence-based practice. *Perspectives in Public Health*, *132*(3), 120–127. https://doi.org/10.1177/1757913912442269

Cohen, D., Keller, DeVoe, & Davis, M. (2012). Characteristics and lessons learned from practice-based research networks (PBRNs) in the United States. *Journal of Healthcare Leadership*, *4*, 107. https://doi.org/10.2147/JHL.S16441

Comer, S. J., & Kendall, P. C. (2013). Methodology, Design, and Evaluation in Psychotherapy Research. In M. J. Lambert (Ed.), *Handbook of Psychotherapy and Behaviour Change* (6th ed., pp. 21–48). John Wiley & Sons Ltd.

Connell, J., Barkham, M., Stiles, W. B., Twigg, E., Singleton, N., Evans, O., & Miles, J. N. V. (2007). Distribution of CORE-OM scores in a general population, clinical cut-off points and comparison with the CIS-R. *The British Journal of Psychiatry*, *190*(1), 69–74. https://doi.org/10.1192/bjp.bp.105.017657

Connell, Janice, Grant, S., & Mullin, T. (2006). Client initiated termination of therapy at NHS primary care counselling services. *Counselling & Psychotherapy Research*, *6*(1), 60–67. https://doi.org/10.1080/14733140600581507

Cook, E. L., & Harman, J. S. (2008). A comparison of health-related quality of life for individuals with mental health disorders and common chronic medical conditions. *Public Health Reports (Washington, D.C.: 1974)*, *123*(1), 45–51. https://doi.org/10.1177/003335490812300107

Copty, M., & Whitford, D. L. (2005). Mental health in general practice: Assessment of current state and future needs. *Irish Journal of Psychological Medicine*, *22*(3), 83–86. https://doi.org/10.1017/S079096670000906X

CORE Partnership. (2007). *Is initial overall CORE-OM score an indicator of likely outcome?* (No. 1; CORE Partnership Occasional Paper). CORE IMS. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&ved=2ahUKEwi29u7pkavlAhWXQxUIHaTyDpYQFjACegQIABAC&url=http%3A%2F%2Fwww.coreims.co.uk%2Fsite_downloads%2FOP1-initial_CORE-OM_score.pdf&usg=AOvVaw3oRodxvYwO5hKZpfeVNCaE

CORE System Group. (1998). *CORE System (Information Management) Handbook*. CORE System Group. http://www.coreims.co.uk/About_Measurement_CORE_Tools.html

Corney, R., & Simpson, S. (2005). Thirty-six month outcome data from a trial of counselling with chronically depressed patients in a general practice setting. *Psychology & Psychotherapy: Theory, Research & Practice*, *78*(1), 127–138. https://doi.org/10.1348/147608304X21365

Craske, M. G., Stein, M. B., Sullivan, G., Sherbourne, C., Bystritsky, A., Rose, R. D., Lang, A. J., Welch, S., Campbell-Sills, L., Golinelli, D., & Roy-Byrne, P.

(2011). Disorder specific impact of CALM treatment for anxiety disorders in primary care. *Archives of General Psychiatry*, *68*(4), 378–388. https://doi.org/10.1001/archgenpsychiatry.2011.25

Creswell, J. W. (2003). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/10.1007/BF02310555

Cuijpers, P., Cristea, I. A., Karyotaki, E., Reijnders, M., & Huibers, M. J. H. (2016). How effective are cognitive behavior therapies for major depression and anxiety disorders? A meta-analytic update of the evidence. *World Psychiatry*, *15*(3), 245–258. https://doi.org/10.1002/wps.20346

Cullinan, V., Veale, A., & Vitale, A. (2016). Irish General Practitioner referrals to psychological therapies. *Irish Journal of Psychological Medicine*, *33*(02), 73–80. https://doi.org/10.1017/ipm.2015.17

Curran, G. M., Sullivan, G., Mendel, P., Craske, M. G., Sherbourne, C. D., Stein, M. B., McDaniel, A., & Roy-Byrne, P. (2012). Implementation of the CALM intervention for anxiety disorders: A qualitative study. *Implementation Science: IS*, *7*, 1–11. https://doi.org/10.1186/1748-5908-7-14

Das-Munshi, J., Goldberg, D., Bebbington, P. E., Bhugra, D. K., Brugha, T. S., Dewey, M. E., Jenkins, R., Stewart, R., & Prince, M. (2008). Public health significance of mixed anxiety and depression: Beyond current classification. *British Journal of Psychiatry*, *192*(3), 171–177. https://doi.org/10.1192/bjp.bp.107.036707

Davis, D., Corrin-Pendry, S., & Savill, M. (2008). A follow-up study of the long-term effects of counselling in a primary care counselling psychology service.

Counselling and Psychotherapy Research, *8*(2), 80–84. https://doi.org/10.1080/14733140802007863

De Charms, R., Levy, J., & Wertheimer, M. (1954). A Note on Attempted Evaluations of Psychotherapy. *Journal of Clinical Psychology*, *10*(3), 233–235.

DeFife, J., Drill, R., Beinashowitz, J., Ballantyne, L., Plant, D., Smith-Hansen, L., Teran, V., Werner-Larsen, L., Westerling, T. I., Yang, Y., Davila, M., & Nakash, O. (2015). Practice-based psychotherapy research in a public health setting: Obstacles and opportunities. *Journal of Psychotherapy Integration*, *25*(4), 299–312. https://doi.org/10.1037/a0039564

Demarteau, M. (2002). A Theoretical Framework and Grid for Analysis of Programme-evaluation Practices. *Evaluation*, *8*(4), 454–473. https://doi.org/10.1177/13563890260620649

Department of Health. (2012). *IAPT three-year report: The first million patients*. Department of Health London.

Deptartment of Health. (2006). *Vision for Change: Report of the Expert Group on Mental Health Policy*. Government Publications Office. http://health.gov.ie/future-health/mental-health-a-vision-for-change/a-vision-for-change/

Derogatis, L. R. (1993). *Brief symptom inventory: BSI*. Pearson.

Derogatis, L. R. (1996). *SCL-90-R: Symptom Checklist-90-R : Administration, Scoring, and Procedures Manual* (3rd ed.). NCS Pearson.

Des Groseilliers, I. B., Marchand, A., Cordova, M. J., Ruzek, J. I., & Brunet, A. (2013). Two-year follow-up of a brief dyadic cognitive-behavioral intervention designed to prevent PTSD. *Psychological Trauma: Theory, Research, Practice, and Policy*, *5*(5), 462–469. pdh. https://doi.org/10.1037/a0031967

Dhima, M., Paulusova, V., Lohse, C., Salinas, T. J., & Carr, A. B. (2014). Practice-Based Evidence from 29-Year Outcome Analysis of Management of the Edentulous Jaw Using Osseointegrated Dental Implants: Edentulous Jaw and Implants, 29-Year Outcome Analysis. *Journal of Prosthodontics*, *23*(3), 173–181. https://doi.org/10.1111/jopr.12084

Dickinson, L. M., & Basu, A. (2005). Multilevel Modeling and Practice-Based Research. *The Annals of Family Medicine*, *3*(suppl 1), S52–S60. https://doi.org/10.1370/afm.340

Driessen, E., Cuijpers, P., Hollon, S. D., & Dekker, J. J. M. (2010). Does pretreatment severity moderate the efficacy of psychological treatment of adult outpatient depression? A meta-analysis. *Journal of Consulting and Clinical Psychology*, *78*(5), 668–680. https://doi.org/10.1037/a0020570

du Plessis, P., & Bor, R. (1999). Evaluating counselling in primary health care. In R. Bor, D. McCann, R. (Ed) Bor, & D. (Ed) McCann (Eds.), *The practice of counselling in primary care.* (1999-04031-014; pp. 229–244). Sage Publications Ltd; psyh. http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1999-04031-014&site=ehost-live

Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, *1*(2), 170–177. https://doi.org/10.1037/1082-989X.1.2.170

Durlak, J. A. (2009). How to Select, Calculate, and Interpret Effect Sizes. *Journal of Pediatric Psychology*, *34*(9), 917–928. https://doi.org/10.1093/jpepsy/jsp004

Eatock, J. (2000). Counselling in primary care: Past, present and future. *British Journal of Guidance & Counselling*, *28*(2), 161–173. https://doi.org/10.1080/03069880050013476

Elkin, I., Gibbons, R., Shea, M. T., Sotsky, S. M., Watkins, J. T., Pilkonis, P. A., & Hedeker, D. (1995). Initial severity and differential treatment outcome in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Journal of Consulting and Clinical Psychology*, *63*(5), 841–847. https://doi.org/10.1037/0022-006X.63.5.841

Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., Glass, D. R., Pilkonis, P. A., Leber, W. R., & Docherty, J. P. (1989). National Institute of Mental Health Treatment of Depression Collaborative Research Program. General effectiveness of treatments. *Archives of General Psychiatry*, *46*(11), 971–982; discussion 983.

Emmelkamp, P. M. G., & Powers, M. B. (2010). Agoraphobia. In J. C. Thomas, & M. Hersen, (Eds.), *Handbook of Clinical Psychology Competencies* (1st ed., pp. 723–758). Springer-Verlag. http://link.springer.com.proxy.lib.ul.ie/referenceworkentry/10.1007/978-0-387-09757-2_26

Endicott, J., & Spitzer, R. L. (1978). A diagnostic interview: The schedule for affective disorders and schizophrenia. *Archives of General Psychiatry*, *35*(7), 837–844.

Evans, C., Margison, F., & Barkham, M. (1998). The contribution of reliable and clinically significant change methods to evidence-based mental health. *Evidence-Based Mental Health*, *1*(3), 70–72. https://doi.org/10.1136/ebmh.1.3.70

Evans, Chris, Connell, J., Barkham, M., Margison, F., McGRATH, G., Mellor-Clark, J., & Audin, K. (2002). Towards a standardised brief outcome measure: Psychometric properties and utility of the CORE—OM. *The British Journal of Psychiatry*, *180*(1), 51–60. https://doi.org/10.1192/bjp.180.1.51

Eysenck, H. J. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting Psychology*, *16*(5), 319–324. https://doi.org/10.1037/h0063633

Falkenström, F., Josefsson, A., Berggren, T., & Holmqvist, R. (2016). How much therapy is enough? Comparing dose-effect and good-enough models in two different settings. *Psychotherapy*, *53*(1), 130–139. https://doi.org/10.1037/pst0000039

Fournier, J. C., DeRubeis, R. J., Shelton, R. C., Hollon, S. D., Amsterdam, J. D., & Gallop, R. (2009). Prediction of Response to Medication and Cognitive Therapy in the Treatment of Moderate to Severe Depression. *Journal of Consulting and Clinical Psychology*, *77*(4), 775–787. https://doi.org/10.1037/a0015401

Frisch, M. B., Cornell, J., Villanueva, M., & Retzlaff, P. J. (1992). Clinical validation of the Quality of Life Inventory. A measure of life satisfaction for use in treatment planning and outcome assessment. *Psychological Assessment*, *4*(1), 92–101. https://doi.org/10.1037/1040-3590.4.1.92

Gibbard, I., & Hanley, T. (2008). A five-year evaluation of the effectiveness of person-centred counselling in routine clinical practice in primary care. *Counselling and Psychotherapy Research*, *8*(4), 215–222. https://doi.org/10.1080/14733140802305440

Gilbody, S., Richards, D., & Barkham, M. (2007). Diagnosing depression in primary care using self-completed instruments: UK validation of PHQ–9 and CORE–OM. *The British Journal of General Practice*, *57*(541), 650–652.

Glass, Gene V. (1972). The wisdom of scientific inquiry on education. *Journal of Research in Science Teaching*, *9*(1), 1–18. https://doi.org/10.1002/tea.3660090103

Glass, G.V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. Sage Publications.

Glass, G.V., & Worthen, B. R. (1971). Educational Evaluation and Research: Similarities and Differences. *Curriculum Theory Network*, *8/9*, 149–165. https://doi.org/10.2307/1179200

Goodman, D. C., Mick, S. S., Bott, D., Stukel, T., Chang, C., Marth, N., Poage, J., & Carretta, H. J. (2003). Primary Care Service Areas: A New Tool for the Evaluation of Primary Care Services. *Health Services Research*, *38*(1p1), 287–309. https://doi.org/10.1111/1475-6773.00116

Goulet-Pelletier, J.-C., & Cousineau, D. (2018). A review of effect sizes and their confidence intervals, Part I: The Cohen's d family. *The Quantitative Methods for Psychology*, *14*(4), 242–265. https://doi.org/10.20982/tqmp.14.4.p242

Graham, K., & Graham, C. (1996). The impact of primary care counselling on psychiatric symptoms. *Journal of Mental Health*, *5*(5), 515–524. https://doi.org/10.1080/09638239619202

Grant, S. (2006). Making sense of CORE System data: Attrition, effectiveness, concordance and information capture. *European Journal of Psychotherapy & Counselling*, *8*(2), 193–208. https://doi.org/10.1080/13642530600712569

Green, D., & Latchford, G. (2012). *Maximising the Benefits of Psychotherapy: A Practice-Based Evidence Approach*. John Wiley & Sons, Ltd. http://doi.wiley.com/10.1002/9781119967590

Green, L. W. (2008). Making research relevant: If it is an evidence-based practice, where's the practice-based evidence? *Family Practice*, *25*(suppl 1), i20–i24.

Griffiths, S., & Steen, S. (2013). Improving Access to Psychological Therapies (IAPT) programme: Setting key performance indicators in a more robust context: A new perspective. *The Journal of Psychological Therapies in Primary Care*, *2*(2), 133–141.

Gupta, S. K. (2011). Intention-to-treat concept: A review. *Perspectives in Clinical Research*, *2*(3), 109–112. https://doi.org/10.4103/2229-3485.83221

Guyon, H., Kop, J.-L., Juhel, J., & Falissard, B. (2018). Measurement, ontology, and epistemology: Psychology needs pragmatism-realism. *Theory & Psychology*, *28*(2), 149–171. https://doi.org/10.1177/0959354318761606

Gyani, A., Shafran, R., Layard, R., & Clark, D. M. (2013a). Enhancing recovery rates: Lessons from year one of IAPT. *Behaviour Research and Therapy*, *51*(9), 597–606. https://doi.org/10.1016/j.brat.2013.06.004

Gyani, A., Shafran, R., Layard, R., & Clark, D. M. (2013b). Enhancing recovery rates: Lessons from year one of IAPT. *Behaviour Research and Therapy*, *51*(9), 597–606. https://doi.org/10.1016/j.brat.2013.06.004

Hagerty, M. R., Cummins, R. A., Ferriss, A. L., Land, K., Michalos, A. C., Peterson, M., Sharpe, A., Sirgy, J., Ferriss, A. L., Hagerty, M. R., Cummins, R. A., Vogel, J., Land, K., Michalos, A. C., Peterson, M., Sharpe, A., & Vogel, J. (2001). Quality of Life Indexes for National Policy: Review and Agenda for Research. *Social Indicators Research*, *55*(1), 1–96. https://doi.org/10.1023/A:1010811312332

Haig, B. D. (2018). *The Philosophy of Quantitative Methods: Understanding Statistics*. Oxford University Press.

Hansen, N. B., Lambert, M. J., & Forman, E. M. (2006). The Psychotherapy Dose-Response Effect and Its Implications for Treatment Delivery Services. *Clinical Psychology: Science and Practice*, *9*(3), 329–343. https://doi.org/10.1093/clipsy.9.3.329

Harder, S., Koester, A., Valbak, K., & Rosenbaum, B. (2014). Five-Year Follow-Up of Supportive Psychodynamic Psychotherapy in First-Episode Psychosis: Long-Term Outcome in Social Functioning. *Psychiatry: Interpersonal & Biological Processes*, *77*(2), 155-168 14p. rzh. https://doi.org/10.1521/psyc.2014.77.2.155

Harnett, P., O'Donovan, A., & Lambert, M. J. (2010). The dose response relationship in psychotherapy: Implications for social policy. *Clinical Psychologist*, *14*(2), 39–44. https://doi.org/10.1080/13284207.2010.500309

Harvey, I., Nelson, S. J., Lyons, R. A., Unwin, C., Monaghan, S., & Peters, T. J. (1998). A randomized controlled trial and economic evaluation of counselling in primary care. *Br J Gen Pract*, *48*(428), 1043–1048.

Haynes, L., Service, O., Goldacre, B., & Torgerson, D. (2012). Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2131581

Health Service Executive, & Irish College of General Practitioners. (2006). *Guidance document on the provision of counselling in a primary care setting* (p. 33). Irish College of General Practitioners (ICGP). http://hdl.handle.net/10147/246453

Heck, R. H., Thomas, S. L., & Tabata, L. N. (2014). *Multilevel and longitudinal modeling with IBM SPSS* (Second edition). Routledge, Taylor & Francis Group.

Hemmings, A. (1997). Counselling in primary care: A randomised controlled trial. *Patient Education and Counseling*, *32*(3), 219–230. https://doi.org/10.1016/S0738-3991(97)00035-9

Hemmings, A. (2000). Counselling in primary care: A review of the practice evidence. *British Journal of Guidance & Counselling*, *28*(2), 233–252. https://doi.org/10.1080/03069880050013520

Hergenrather, K. C., Zeglin, R. J., McGuire-Kuletz, M., & Rhodes, S. D. (2015). Employment as a Social Determinant of Health: A Review of Longitudinal Studies Exploring the Relationship between Employment Status and Mental Health. *Rehabilitation Research, Policy, and Education*, *29*(3), 261–290.

Hjermstad, M. J., Fayers, P. M., Bjordal, K., & Kaasa, S. (1998). Health-related quality of life in the general Norwegian population assessed by the European Organization for Research and Treatment of Cancer Core Quality-of-Life Questionnaire: The QLQ=C30 (+ 3). *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, *16*(3), 1188–1196.

Horn, S. D., & Gassaway, J. (2007). Practice-based evidence study design for comparative effectiveness research. *Medical Care*, *45*(10), S50–S57.

Horvath, A. O. (2013). You can't step into the same river twice, but you can stub your toes on the same rock: Psychotherapy outcome from a 50-year perspective. *Psychotherapy*, *50*(1), 25–32. https://doi.org/10.1037/a0030899

Horvath, P. (1988). Placebos and common factors in two decades of psychotherapy research. *Psychological Bulletin*, *104*(2), 214–225. https://doi.org/10.1037/0033-2909.104.2.214

Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose–effect relationship in psychotherapy. *American Psychologist*, *41*(2), 159–164. https://doi.org/10.1037/0003-066X.41.2.159

Howell, D. C. (2009). *Statistical Methods for Psychology*. Cengage Learning.

HSRC, N. C. S. (2003). *SENCS : Survivors' Experiences of the National Counselling Service. For adults who experienced childhood abuse.* Department of Psychology, Royal College of Surgeons in Ireland.

Jackson, P. R., & Warr, P. (1987). Mental Health Of Unemployed Men In Different Parts Of England And Wales. *British Medical Journal (Clinical Research Edition)*, *295*(6597), 525–525. JSTOR.

Jacobson, N S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*(1), 12–19. https://doi.org/10.1037/0022-006X.59.1.12

Jacobson, Neil S., Follette, W. C., & Revenstorf, D. (1986). Toward a standard definition of clinically significant change. *Behavior Therapy*, *17*(3), 308–311. https://doi.org/10.1016/S0005-7894(86)80061-2

Jacobson, N.S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy*, *15*(4), 336–352. https://doi.org/10.1016/S0005-7894(84)80002-7

Kabisch, M., Ruckes, C., Seibert-Grafe, M., & Blettner, M. (2011). Randomized Controlled Trials. *Deutsches Ärzteblatt International*, *108*(39), 663–668. https://doi.org/10.3238/arztebl.2011.0663

Karyotaki, E., Smit, Y., de Beurs, D. P., Henningsen, K. H., Robays, J., Huibers, M. J. H., Weitz, E., & Cuijpers, P. (2016). The Long-Term Efficacy of Acute-Phase Psychotherapy for Depression: A Meta-Analysis of Randomized Trials. *Depression and Anxiety*, *33*(5), 370–383. https://doi.org/10.1002/da.22491

King, M., Marston, L., & Bower, P. (2014). Comparison of non-directive counselling and cognitive behaviour therapy for patients presenting in general practice with an ICD-10 depressive episode: A randomized control trial. *Psychological Medicine*, *44*(09), 1835–1844. https://doi.org/10.1017/S0033291713002377

Klerman, G. L., Weissman, M. M., Rounsaville, B. J., & Chevron, E. S. (1984). *Interpersonal Psychotherapy Of Depression* (1 edition). Basic Books.

Knapstad, M., Nordgreen, T., & Smith, O. R. F. (2018). Prompt mental health care, the Norwegian version of IAPT: Clinical outcomes and predictors of change in a multicenter cohort study. *BMC Psychiatry*, *18*. https://doi.org/10.1186/s12888-018-1838-0

Kopta, S. M., Howard, K. I., Lowry, J. L., & Beutler, L. E. (1994). Patterns of symptomatic recovery in psychotherapy. *Journal of Consulting and Clinical Psychology*, *62*(5), 1009–1016. https://doi.org/10.1037/0022-006X.62.5.1009

Koufogiannakis, D. (2012). Considering the place of practice-based evidence within Evidence Based Library and Information Practice (EBLIP). *Library and Information Research*, *35*(111), 41–58. https://doi.org/10.29173/lirg486

Kraus, D. R., Bentley, J. H., Alexander, P. C., Boswell, J. F., Constantino, M. J., Baxter, E. E., & Castonguay, L. G. (2016). Predicting therapist effectiveness from their own practice-based evidence. *Journal of Consulting and Clinical Psychology*, *84*(6), 473–483. https://doi.org/10.1037/ccp0000083

Kusek, J. Z., & Rist, R. C. (2004). *Ten steps to a results-based monitoring and evaluation system: A handbook for development practitioners*. World Bank.

La Bash, H., Galovski, T., & Stirman, S. W. (2019). Adapting Evidence-Based Psychotherapies While Maintaining Fidelity. *Current Treatment Options in Psychiatry*, *6*(3), 198–209. https://doi.org/10.1007/s40501-019-00177-9

Lachin, J. M. (2016). Fallacies of last observation carried forward analyses. *Clinical Trials (London, England)*, *13*(2), 161–168. https://doi.org/10.1177/1740774515602688

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00863

Lambert, M. J. (2013). The Efficacy and Effectiveness of Psychotherapy. In M. J. Lambert (Ed.), *Handbook of Psychotherapy and Behavior Change* (6th ed., pp. 169–218). Wiley.

Lambert, M. J., Hansen, N. B., & Finch, A. E. (2001). Patient-focused research: Using patient outcome data to enhance treatment effects. *Journal of Consulting and Clinical Psychology*, *69*(2), 159–172.

Leichsenring, F. (2004). Randomized controlled versus naturalistic studies: A new research agenda. *Bulletin of the Menninger Clinic*, *68*(2), 137–151. https://doi.org/10.1521/bumc.68.2.137.35952

Lenhard, W., & Lenhard, A. (2016). Calculation of Effect Sizes. *Psychometrica*. https://doi.org/10.13140/RG.2.1.3478.4245

Linn, M. W., Sandifer, R., & Stein, S. (1985). Effects of unemployment on mental and physical health. *American Journal of Public Health*, *75*(5), 502–506. https://doi.org/10.2105/AJPH.75.5.502

Lipsey, M. W., & Wilson, D. (2000). *Practical Meta-Analysis* (1 edition). SAGE Publications, Inc.

Locher, C., Gaab, J., & Blease, C. (2018). When a Placebo Is Not a Placebo: Problems and Solutions to the Gold Standard in Psychotherapy Research. *Frontiers in Psychology*, *9*, 2317. https://doi.org/10.3389/fpsyg.2018.02317

Lyons, R., Payne, C., McCabe, M., & Fielder, C. (1998). Legibility of Doctors' Handwriting: Quantitative Comparative Study. *BMJ: British Medical Journal*, *317*(7162), 863–864.

Margison, F. (2010). Handbook of evidence-based psychodynamic psychotherapy: Bridging the gap between science and practice. *Psychodynamic Practice*, *16*(1), 105–107. https://doi.org/10.1080/14753630903458390

Margison, F. R., McGrath, G., Barkham, M., Clark, J. M., Audin, K., Connell, J., & Evans, C. (2000). Measurement and psychotherapy Evidence-based practice and practice-based evidence. *The British Journal of Psychiatry*, *177*(2), 123–130. https://doi.org/10.1192/bjp.177.2.123

Martin, E., Hawkins, S., Hicks, T., & O'Flynn, M. (2008). Psychological services for adults in primary care. *The Irish Psychologist*. http://www.lenus.ie/hse/handle/10147/271913

Matt, G. E. (1989). Decision rules for selecting effect sizes in meta-analysis: A review and reanalysis of psychotherapy outcome studies. *Psychological Bulletin*, *105*(1), 106–115. https://doi.org/10.1037/0033-2909.105.1.106

McElvaney, J., & Timulak, L. (2013). Clients' experience of therapy and its outcomes in 'good' and 'poor' outcome psychological therapy in a primary care setting: An exploratory study. *Counselling and Psychotherapy Research*, *13*(4), 246–253. https://doi.org/10.1080/14733145.2012.761258

McLeod, J. (2001). An administratively created reality: Some problems with the use of self-report questionnaire measures of adjustment in counselling/psychotherapy outcome research. *Counselling and Psychotherapy Research*, *1*(3), 215–226. https://doi.org/10.1080/14733140112331385100

McMillan, D., & Morley, S. (2010). Single Case Quantitative Methods for Practice-Based Evidence. In Michael Barkham, G. E. Hardy, & J. Mellor-Clark (Eds.), *Developing and Delivering Practice-Based Evidence* (pp. 109–138). John Wiley & Sons, Ltd. http://doi.wiley.com/10.1002/9780470687994.ch5

Melchert, T. P. (2016). Leaving behind our preparadigmatic past: Professional psychology as a unified clinical science. *American Psychologist*, *71*(6), 486–496. https://doi.org/10.1037/a0040227

Mellor-Clark, J. (2000). Developing practice and evidence for counselling in primary care: The agenda. *British Journal of Guidance & Counselling*, *28*(2), 253–266. https://doi.org/10.1080/03069880050013539

Mellor-Clark, J. (2006). Developing CORE performance indicators for benchmarking in NHS primary care psychological therapy and counselling services: An editorial introduction. *Counselling & Psychotherapy Research*, *6*(1), 1–2. psyh. https://doi.org/10.1080/14733140600581176

Mellor-Clark, J., Barkham, M., Connell, J., & Evans, C. (1999). Practice-based evidence and standardized evaluation: Informing the design of the CORE system. *European Journal of Psychotherapy & Counselling*, *2*(3), 357–374. https://doi.org/10.1080/13642539908400818

Mellor-Clark, J., Connell, J., Barkham, M., & Cummins, P. (2001). Counselling outcomes in primary health care: A CORE system data profile. *European Journal of Psychotherapy, Counselling and Health*, *4*(1), 65–86. psyh. https://doi.org/10.1080/13642530110040118

Mental Health Reform. (2014). *Ethnic Minorities and Mental Health: A position paper*. Mental Health Reform. https://www.mentalhealthreform.ie/wp-

content/uploads/2014/10/Ethnic-Minorities-and-Mental-Health-A-position-paper_WEB.pdf

Meuldijk, D., Carlier, I. V. E., van Vliet, I. M., van Veen, T., Wolterbeek, R., van Hemert, A. M., & Zitman, F. G. (2016). The clinical effectiveness of concise cognitive behavioral therapy with or without pharmacotherapy for depressive and anxiety disorders; a pragmatic randomized controlled equivalence trial in clinical practice. *Contemporary Clinical Trials*, *47*, 131–138. https://doi.org/10.1016/j.cct.2015.12.021

Millet, D. (2016, November 1). *Most GP trainees miss out on mental health training placements | GPonline* [General practitioner training and information]. Most GP Trainees Miss out on Mental Health Training Placements. https://www.gponline.com/gp-trainees-miss-mental-health-training-placements/mental-health/article/1413996

Moller, N. P., Ryans, G., Rollings, J., & Barkham, M. (2019). The 2018 UK NHS Digital annual report on the Improving Access to Psychological Therapies programme: A brief commentary. *BMC Psychiatry*, *19*(1), 1–5. https://doi.org/10.1186/s12888-019-2235-z

Moorhouse, A., & Caltabiano, M. L. (2007). Resilience and unemployment: Exploring risk and protective influences for the outcome variables of depression and assertive job searching. *Journal of Employment Counseling*, *44*(3), 115–125.

Moras, K. (2002). Research in Psychotherapy. In M. Hersen & W. Sledge (Eds.), *Encyclopedia of Psychotherapy* (pp. 525–545). Academic Press. https://doi.org/10.1016/B0-12-343010-0/00183-5

Moriarty, D. G., Zack, M. M., & Kobau, R. (2003). The Centers for Disease Control and Prevention's Healthy Days Measures – Population tracking of perceived

physical and mental health over time. *Health and Quality of Life Outcomes*, *1*(1), 37. https://doi.org/10.1186/1477-7525-1-37

Moum, T. A. (1998, June 18). *Mode of administration and response style effects in self-reported health and health behavior: An experimental broad population study.* National Conference on Sociology, Longyearbyen, Spitsbergen, Norway.

Mulder, R., Murray, G., & Rucklidge, J. (2017). Common versus specific factors in psychotherapy: Opening the black box. *The Lancet Psychiatry*, *4*(12), 953–962. https://doi.org/10.1016/S2215-0366(17)30100-1

Mullin, T., Barkham, M., Mothersole, G., Bewick, B. M., & Kinder, A. (2006). Recovery and improvement benchmarks for counselling and the psychological therapies in routine primary care. *Counselling and Psychotherapy Research*, *6*(1), 68–80. https://doi.org/10.1080/14733140600581515

Mulrow, C., & Oxman, A. (1997). Cochrane collaboration handbook. *The Cochrane Library*, *4*.

Mundt, J. C. (2002). The Work and Social Adjustment Scale: A simple measure of impairment in functioning. *The British Journal of Psychiatry*, *180*(5), 461–464. https://doi.org/10.1192/bjp.180.5.461

National Center for Chronic Disease Prevention and Health Promotion. (2002). *Health-Related Quality of Life—Puerto Rico, 1996—2000*. https://www.cdc.gov/mmwr/preview/mmwrhtml/mm5108a3.htm

National Collaborating Centre for Mental Health. (2011). *Common mental health disorders: The NICE guideline on identification and pathways to care*. The British Psychological Society and The Royal College of Psychiatrists. https://www.nice.org.uk/guidance/cg123/evidence/full-guideline-pdf-181771741

Newby, J. M., McKinnon, A., Kuyken, W., Gilbody, S., & Dalgleish, T. (2015). Systematic review and meta-analysis of transdiagnostic psychological treatments for anxiety and depressive disorders in adulthood. *Clinical Psychology Review*, *40*, 91–110. https://doi.org/10.1016/j.cpr.2015.06.002

NHS. (2018). *The Improving Access to Psychological Therapies Manual* (No. 2; pp. 1–78). The National Collaborating Centre for Mental Health. https://www.rcpsych.ac.uk/docs/default-source/improving-care/nccmh/the-iapt-manual--final--republished-7-3-18.pdf?sfvrsn=a192d1af_0

NHS Digital. (2018). *Psychological Therapies: Annual Report on the use of IAPT services, England, further analyses on 2016-17* (p. 34). The Health and Social Care Information Centre. https://files.digital.nhs.uk/publication/s/n/psyc-ther-ann-rep-2016-17_add.pdf

NICE. (2011). *Generalised anxiety disorder in adults: Management in primary, secondary and community care*. British Psychological Society : Royal College of Psychiatrists.

Norcross, J. C., Glass, C. R., Arnkoff, D. B., Lambert, M. J., Shoham, V., Stiles, W. B., Shapiro, D. A., Barkham, M., & Strupp, H. H. (1993). Research directions for psychotherapy integration: A roundtable. *Journal of Psychotherapy Integration*, *3*(2), 91–131. https://doi.org/10.1037/h0101184

OECD, & European Union. (2018). *Health at a Glance: Europe 2018: State of Health in the EU Cycle*. OECD. https://doi.org/10.1787/health_glance_eur-2018-en

Ogrodniczuk, J., & Staats, H. (2002). [Psychotherapy and gender: Do men and women require different treatments?]. *Zeitschrift Für Psychosomatische Medizin Und Psychotherapie*, *48*(3), 270–285.

Otto, M. W. (2002). The Dose and Delivery of Psychotherapy: A Commentary on Hansen et al. *Clinical Psychology: Science and Practice*, *9*(3), 348–349. https://doi.org/10.1093/clipsy.9.3.348

Ounpuu, S., Krueger, P., Vermeulen, M., & Chambers, L. (2000). Using the U.S. Behavior Risk Factor Surveillance System's Health-related Quality of Life survey tool in a Canadian city. *Can J Public Health*, *91*(1), 67–72.

Paul, K. I., & Moser, K. (2009). Unemployment impairs mental health: Meta-analyses. *Journal of Vocational Behavior*, *74*(3), 264–282. https://doi.org/10.1016/j.jvb.2009.01.001

Pearce, W., & Raman, S. (2014). The New Randomised Controlled Trials (RCT) Movement in Public Policy: Challenges of Epistemic Governance. *Policy Sciences*, *47*(4), 387–402.

Percevic, R., Lambert, M. J., & Kordy, H. (2006). What is the predictive value of responses to psychotherapy for its future course? Empirical explorations and consequences for outcome monitoring. *Psychotherapy Research*, *16*(3), 364–373. https://doi.org/10.1080/10503300500485524

Petticrew, M., Chalabi, Z., & Jones, D. R. (2012). To RCT or not to RCT: Deciding when 'more evidence is needed' for public health policy and practice. *J Epidemiol Community Health*, *66*(5), 391–396. https://doi.org/10.1136/jech.2010.116483

Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology*, *48*(1), 85–112. https://doi.org/10.1016/j.jsp.2009.09.002

Pharr, J. R., Moonie, S., & Bungum, T. J. (2012). *The Impact of Unemployment on Mental and Physical Health, Access to Health Care and Health Risk Behaviors*

[Research article]. International Scholarly Research Notices. https://doi.org/10.5402/2012/483432

Phillips, D. C., & Burbules, N. C. (2000). *Postpositivism and educational research*. Rowman & Littlefield Publishers.

Pirkis, Jane, Burgess, P., Coombs, T., Clarke, A., Jones-Ellis, D., & Dickson, R. (2005). Routine measurement of outcomes in Australia's public sector mental health services. *Australia and New Zealand Health Policy*, *2*, 8. https://doi.org/10.1186/1743-8462-2-8

Pirkis, JE, Blashki, G., Murphy, A., Hickie, I., & Ciechomski, L. (2006). The contribution of general practice based research to the development of national policy: Case studies from Ireland and Australia. *Australia and New Zealand Health Policy*, *3*, 4. https://doi.org/10.1186/1743-8462-3-4

Poghosyan, H., Darwish, S. A., Kim, S. S., & Cooley, M. E. (2016). The association between social support and smoking status in cancer survivors with frequent and infrequent mental distress: Results from 10 US states, 2010. *Journal of Cancer Survivorship*, *10*(6), 1078–1088. https://doi.org/10.1007/s11764-016-0551-6

Potter, M. A., Quill, B. E., Aglipay, G. S., Anderson, E., Rowitz, L., Smith, L. U., Telfair, J., & Whittaker, C. (2006). Demonstrating Excellence in Practice-Based Research for Public Health. *Public Health Reports*, *121*(1), 1–16. https://doi.org/10.1177/003335490612100102

Pybis, J, Saxon, D., Hill, A., & Barkham, M. (2017). The comparative effectiveness and efficiency of cognitive behaviour therapy and generic counselling in the treatment of depression: Evidence from the 2(nd) UK National Audit of psychological therapies. *BMC Psychiatry*, *17*(1), 215. https://doi.org/10.1186/s12888-017-1370-7

Pybis, Jo, Saxon, D., Hill, A., & Barkham, M. (2017). The comparative effectiveness and efficiency of cognitive behaviour therapy and generic counselling in the treatment of depression: Evidence from the 2nd UK National Audit of psychological therapies. *BMC Psychiatry*, *17*, 215. https://doi.org/10.1186/s12888-017-1370-7

Ray-Sannerud, B. N., Dolan, D. C., Morrow, C. E., Corso, K. A., Kanzler, K. E., Corso, M. L., & Bryan, C. J. (2012). Longitudinal outcomes after brief behavioral health intervention in an integrated primary care clinic. *Families, Systems, & Health*, *30*(1), 60–71. https://doi.org/10.1037/a0027029

Richards, D. A., & Borglin, G. (2011). Implementation of psychological therapies for anxiety and depression in routine practice: Two year prospective cohort study. *Journal of Affective Disorders*, *133*(1), 51–60. https://doi.org/10.1016/j.jad.2011.03.024

Richards, DA, & Suckling, R. (2009). Improving access to psychological therapies: Phase IV prospective cohort study. *The British Journal of Clinical Psychology*, *48*(Pt 4), 377–396. https://doi.org/10.1348/014466509X405178

Richards, Derek, Richardson, T., Timulak, L., Viganò, N., Mooney, J., Doherty, G., Hayes, C., & Sharry, J. (2016). Predictors of depression severity in a treatment-seeking sample. *International Journal of Clinical and Health Psychology*, *16*(3), 221–229. https://doi.org/10.1016/j.ijchp.2016.02.001

Ronk, F. R., Hooke, G. R., & Page, A. C. (2016). Validity of clinically significant change classifications yielded by Jacobson-Truax and Hageman-Arrindell methods. *BMC Psychiatry*, *16*. https://doi.org/10.1186/s12888-016-0895-5

Rowland, N., Bower, P., Mellor-Clark, J., Heywood, P., & Hardy, R. (2000). Counselling in primary care: A systematic review of the research evidence.

British Journal of Guidance & Counselling, *28*(2), 215–231. https://doi.org/10.1080/03069880050013511

Roy-Byrne, P., Craske, M. G., Sullivan, G., Rose, R. D., Edlund, M. J., Lang, A. J., Bystritsky, A., Welch, S. S., Chavira, D. A., Golinelli, D., Campbell-Sills, L., Sherbourne, C. D., & Stein, M. B. (2010). Delivery of Evidence-Based Treatment for Multiple Anxiety Disorders in Primary Care: A Randomized Controlled Trial. *JAMA*, *303*(19), 1921–1928. https://doi.org/10.1001/jama.2010.608

Sawchuk, C. N., Craner, J. R., Berg, S. L., Smyth, K., Mack, J., Glader, M., Burke, L., Haggerty, S., Johnson, M., Miller, S., Sedivy, S., Morcomb, D., Heredia, D., Williams, M. W., & Katzelnick, D. J. (2018). Initial outcomes of a real-world multi-site primary care psychotherapy program. *General Hospital Psychiatry*, *54*, 5–11. https://doi.org/10.1016/j.genhosppsych.2018.06.005

Saxon, D, Firth, N., & Barkham, M. (2017). The Relationship Between Therapist Effects and Therapy Delivery Factors: Therapy Modality, Dosage, and Non-completion. *Administration and Policy in Mental Health and Mental Health Services Research*, *44*(5), 705–715. https://doi.org/10.1007/s10488-016-0750-5

Saxon, D, Ivey, C., & Young, T. (2008). Can CORE assessment data identify those clients less likely to benefit from brief counselling in primary care? *Counselling & Psychotherapy Research*, *8*(4), 223–230. rzh.

Saxon, David, & Barkham, M. (2012). Patterns of therapist variability: Therapist effects and the contribution of patient severity and risk. *Journal of Consulting and Clinical Psychology*, *80*(4), 535–546. https://doi.org/10.1037/a0028898

Schennach, R., Möller, H.-J., Obermeier, M., Seemüller, F., Jäger, M., Schmauss, M., Laux, G., Pfeiffer, H., Naber, D., Schmidt, L. G., Gaebel, W., Klosterkötter, J.,

Heuser, I., Maier, W., Lemke, M. R., Rüther, E., Klingberg, S., Gastpar, M., Musil, R., … Riedel, M. (2016). Challenging the understanding of significant improvement and outcome in schizophrenia - the concept of reliable and clinically significant change methods: Outcome in Schizophrenia. *International Journal of Methods in Psychiatric Research*, *25*(1), 3–11. https://doi.org/10.1002/mpr.1476

Schroer, S., & Macpherson, H. (2009). Acupuncture, or non-directive counselling versus usual care for the treatment of depression: A pilot study. *Trials*, *10*, 3. https://doi.org/10.1186/1745-6215-10-3

Scott, J., & Sensky, T. (2003). Methodological aspects of randomized controlled trials of psychotherapy in primary care. *Psychological Medicine*, *33*(2), 191–196. https://doi.org/10.1017/S0033291702007262

Seekles, W., Cuijpers, P., Kok, R., Beekman, A., van Marwijk, H., & van Straten, A. (2013). Psychological treatment of anxiety in primary care: A meta-analysis. *Psychological Medicine*, *43*(2), 351–361. psyh. https://doi.org/10.1017/S0033291712000670

Segev, Z., Arif, A. A., & Rohrer, J. E. (2012). Activity Limitations and Healthcare Access as Correlates of Frequent Mental Distress in Adults 65 Years and Older: A Behavioral Risk Factor Surveillance Study—2008. *Journal of Primary Care & Community Health*, *3*(1), 17–22. https://doi.org/10.1177/2150131911412380

Seidel, J. A., Miller, S. D., & Chow, D. L. (2014). Effect size calculations for the clinician: Methods and comparability. *Psychotherapy Research*, *24*(4), 470–484. https://doi.org/10.1080/10503307.2013.840812

Seob Kim, Y., Yoon, D., Byun, J., Park, H., Lee, A., Hyun Kim, I., Lee, S., Lim, H.-S., & Park, R. W. (2017). Extracting information from free-text electronic patient

records to identify practice-based evidence of the performance of coronary stents. *PLOS ONE*, *12*, e0182889. https://doi.org/10.1371/journal.pone.0182889

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton, Mifflin and Company.

Shadish, W. R., Matt, G. E., Navarro, A. M., Siegle, G., Crits-Christoph, P., Hazelrigg, M. D., Jorm, A. F., Lyons, L. C., Nietzel, M. T., Robinson, L., Prout, H. T., Smith, M. L., Svartberg, M., & Weiss, B. (1997). Evidence that therapy works in clincally representative conditions. *Journal of Consulting and Clinical Psychology*, *65*(3), 355–365. https://doi.org/10.1037/0022-006X.65.3.355

Shapiro, D. A., Barkham, M., Rees, A., Hardy, G. E., Reynolds, S., & Startup, M. (1994). Effects of treatment duration and severity of depression on the effectiveness of cognitive-behavioral and psychodynamic-interpersonal psychotherapy. *Journal of Consulting and Clinical Psychology*, *62*(3), 522–534. https://doi.org/10.1037/0022-006X.62.3.522

Shean, G. (2014). Limitations of Randomized Control Designs in Psychotherapy Research. *Advances in Psychiatry*, *2014*, 1–5. https://doi.org/10.1155/2014/561452

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., & Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *The Journal of Clinical Psychiatry*, *59 Suppl 20*, 22-33;quiz 34-57.

Shepherd, M., Ashworth, M., Evans, C., Robinson, S., Rendall, M., & Ward, S. (2005). What factors are associated with improvement after brief psychological

interventions in primary care? Issues arising from using routine outcome measurement to inform clinical practice. *Counselling & Psychotherapy Research*, *5*(4), 273–280. rzh.

Shih, M., & Simon, P. A. (2008). Health-related quality of life among adults with serious psychological distress and chronic medical conditions. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *17*(4), 521–528. https://doi.org/10.1007/s11136-008-9330-9

Simpson, S., Corney, R., Fitzgerald, P., & Beecham, J. (2000). A randomised controlled trial to evaluate the effectiveness and cost-effectiveness of counselling patients with chronic depression. *Health Technology Assessment*, *4*(36), 1–83.

Slabaugh, S. L., Shah, M., Zack, M., Happe, L., Cordier, T., Havens, E., Davidson, E., Miao, M., Prewitt, T., & Jia, H. (2017). Leveraging Health-Related Quality of Life in Population Health Management: The Case for Healthy Days. *Population Health Management*, *20*(1), 13–22. https://doi.org/10.1089/pop.2015.0162

Smith, M., & Glass, G. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, *32*(9), 752–760. https://doi.org/10.1037/0003-066X.32.9.752

Smith, M., Glass, G., & Miller, T. (1980). *The benefits of psychotherapy*. Johns Hopkins University Press.

Sotsky, S. M., Glass, D. R., Shea, M. T., Pilkonis, P. A., Collins, J. F., Elkin, I., Watkins, J. T., Imber, S. D., Leber, W. R., & Moyer, J. (1991). Patient predictors of response to psychotherapy and pharmacotherapy: Findings in the NIMH Treatment of Depression Collaborative Research Program. *The American Journal of Psychiatry*, *148*(8), 997–1008. https://doi.org/10.1176/ajp.148.8.997

Speer, D. C. (1998). *Mental health outcome evaluation*. Academic Press.

Health Act, Pub. L. No. 1, 58 (1971). http://www.irishstatutebook.ie/eli/1970/act/1/enacted/en/print.html

Stepped Care. (2017). Wellbeing Info. http://wellbeinginfo.org/self-help/mental-health/stepped-care/

Stiles, W.B., Barkham, M., Twigg, E., Mellor-Clark, J., & Cooper, M. (2006). Effectiveness of cognitive-behavioural, person-centred and psychodynamic therapies as practised in UK National Health Service settings. *Psychological Medicine*, *36*(04), 555. https://doi.org/10.1017/S0033291706007136

Stiles, William B., Barkham, M., Mellor-Clark, J., & Connell, J. (2008). Effectiveness of cognitive-behavioural, person-centred, and psychodynamic therapies in UK primary-care routine practice: Replication in a larger sample. *Psychological Medicine*, *38*(05), 677–688. https://doi.org/10.1017/S0033291707001511

Stulz, N., Lutz, W., Kopta, S. M., Minami, T., & Saunders, S. M. (2013). Dose–effect relationship in routine outpatient psychotherapy: Does treatment duration matter? *Journal of Counseling Psychology*, *60*(4), 593–600. https://doi.org/10.1037/a0033589

Sullivan, G. M., & Feinn, R. (2012). Using Effect Size—Or Why the P Value Is Not Enough. *Journal of Graduate Medical Education*, *4*(3), 279–282. https://doi.org/10.4300/JGME-D-12-00156.1

Trivedi, M. H., Rush, A. J., Wisniewski, S. R., Warden, D., McKinney, W., Downing, M., Berman, S. R., Farabaugh, A., Luther, J. F., Nierenberg, A. A., Callan, J. A., & Sackeim, H. A. (2006). Factors associated with health-related quality of life among outpatients with major depressive disorder: A STAR*D report. *The Journal of Clinical Psychiatry*, *67*(2), 185–195.

Unsworth, G., Cowie, H., & Green, A. (2012). Therapists' and clients' perceptions of routine outcome measurement in the NHS: A qualitative study. *Counselling & Psychotherapy Research*, *12*(1), 71–80. a9h.

Wampold. (2013). The good, the bad, and the ugly: A 50-year perspective on the outcome problem. *Psychotherapy*, *50*(1), 16–24. https://doi.org/10.1037/a0030570

Wampold, B. E., Ahn, H., & Kim, D. M. (2000). Meta-analysis in the social sciences: *Asia Pacific Education Review*, *1*(1), 67–74. https://doi.org/10.1007/BF03026147

Wampold, B. E., Flückiger, C., Del Re, A. C., Yulish, N. E., Frost, N. D., Pace, B. T., Goldberg, S. B., Miller, S. D., Baardseth, T. P., Laska, K. M., & Hilsenroth, M. J. (2017). In pursuit of truth: A critical examination of meta-analyses of cognitive behavior therapy. *Psychotherapy Research*, *27*(1), 14–32. https://doi.org/10.1080/10503307.2016.1249433

Wampold, B. E., & Imel, Z. E. (2015). *The Great Psychotherapy Debate: The Evidence for What Makes Psychotherapy Work*. Routledge.

Ward, E., King, M., Lloyd, M., Bower, P., Sibbald, B., Farrelly, S., Gabbay, M., Tarrier, N., & Addington-Hall, J. (2000). Randomised controlled trial of non-directive counselling, cognitive-behaviour therapy, and usual general practitioner care for patients with depression. I: Clinical effectiveness. *BMJ*, *321*(7273), 1383–1388. https://doi.org/10.1136/bmj.321.7273.1383

Ward, F. (2012). Bridging the gap in our mental health services. *Forum*, 20–22.

Watson, B., Abidi, S., & Cochrane, W. (2004). Using randomised controlled trials (RCTs) to test service interventions: Issues of standardisation, selection and generalisability. *Nurse Researcher*, *11*(3), 28–42.

Watzke, B., Rüddel, H., Jürgensen, R., Koch, U., Kriston, L., Grothgar, B., & Schulz, H. (2010). Effectiveness of systematic treatment selection for psychodynamic and cognitive–behavioural therapy: Randomised controlled trial in routine mental healthcare. *The British Journal of Psychiatry*, *197*(2), 96–105. https://doi.org/10.1192/bjp.bp.109.072835

Weiss, C. H. (1972). *Evaluation research: Methods for assessing program effectiveness*. Prentice-Hall.

Weitz, E. S., Hollon, S. D., Twisk, J., van Straten, A., Huibers, M. J. H., David, D., DeRubeis, R. J., Dimidjian, S., Dunlop, B. W., Cristea, I. A., Faramarzi, M., Hegerl, U., Jarrett, R. B., Kheirkhah, F., Kennedy, S. H., Mergl, R., Miranda, J., Mohr, D. C., Rush, A. J., … Cuijpers, P. (2015). Baseline Depression Severity as Moderator of Depression Outcomes Between Cognitive Behavioral Therapy vs Pharmacotherapy: An Individual Patient Data Meta-analysis. *JAMA Psychiatry*, *72*(11), 1102. https://doi.org/10.1001/jamapsychiatry.2015.1516

Werbart, A., Levin, L., Andersson, H., & Sandell, R. (2013a). Everyday evidence: Outcomes of psychotherapies in Swedish public health services. *Psychotherapy*, *50*(1), 119–130. pdh. https://doi.org/10.1037/a0031386

Werbart, A., Levin, L., Andersson, H., & Sandell, R. (2013b). Everyday evidence: Outcomes of psychotherapies in Swedish public health services. *Psychotherapy*, *50*(1), 119–130. https://doi.org/10.1037/a0031386

Westfall, J. M., Mold, J., & Fagnan, L. (2007). Practice-Based Research—"Blue Highways" on the NIH Roadmap. *JAMA*, *297*(4), 403. https://doi.org/10.1001/jama.297.4.403

WHO. (2004). *International Statistical Classification of Diseases and Related Health Problems*. World Health Organization.

WHO. (2014). *Basic documents.* (48th ed.). World Health Organization. http://apps.who.int/gb/bd/PDF/bd48/basic-documents-48th-edition-en.pdf#page=7

Williams, C. H. J. (2015). Improving access to psychological therapies (IAPT) and treatment outcomes: Epistemological assumptions and controversies. *Journal of Psychiatric and Mental Health Nursing*, *22*(5), 344–351. https://doi.org/10.1111/jpm.12181

Winter, B. (2013). Linear models and linear mixed effects models in R with linguistic applications. *ArXiv:1308.5499 [Cs]*. http://arxiv.org/abs/1308.5499

Wootton, B. M., Bragdon, L. B., Steinman, S. A., & Tolin, D. F. (2015). Three-year outcomes of adults with anxiety and related disorders following cognitive-behavioral therapy in a non-research clinical setting. *Journal of Anxiety Disorders*, *31*, 28–31. https://doi.org/10.1016/j.janxdis.2015.01.007

Zachar, P. (2010). The abandonment of latent variables: Philosophical considerations. *Behavioral and Brain Sciences*, *33*(2–3), 177–178. https://doi.org/10.1017/S0140525X10000841

Appendices

Client demographics and presenting problems for the all clients

	All clients referred (*N* = 5,025)	All participants sample (*n* = 2,781)	Assessed sample (*n* = 2,507)	Form returner sample (*n* = 1,397)	Completor sample (*n* = 1,335)
Patient demographics					
Age *M (SD)*	43.6 *(15.1)*	42.6 *(14.6)*	42.4 (*14.5*)	43.6 *(14.3)*	43.9 *(14.4)*
Sex					
Female % (*N*)	74 *(3721)*	74.9 *(2082)*	74.7 (*1873*)	73.4 *(1026)*	73.3 *(979)*
Male % (*N)*	26 *(1304)*	25.1 *(699)*	25.3 (634)	26.6 *(371)*	26.7 *(356)*
Ethnicity % (*N*)					
White Irish	68.6 *(3448)*	73.2 *(2036)*	73.9 *(1853*)	74.7 *(1044)*	74.7 *(1044)*
White Irish Traveller	1.2 *(60)*	0.6 *(18)*	0.6 (*15*)	0.4 *(6)*	0.4 *(6)*
Any other White Background	10.6 *(531)*	11.4 *(317)*	11.6 (*292*)	0.5 *(7)*	0.5 *(7)*
Black, or Black Irish - African	0.6 *(30)*	0.5 *(14)*	0.5 (*13*)	0.4 *(5)*	0.4 *(5)*
Black, or Black Irish - any other Black background	0.7 *(34)*	0.5 *(14)*	0.5 (*13*)	0.5 *(7)*	0.5 *(7)*
Asian or Asian Irish - Chinese	0.3 *(15)*	0.3 *(8)*	0.3 (*7*)	0.3 *(4)*	0.3 *(4)*
Asian or Asian Irish - any other Aisan background	0.5 *(24)*	0.3 *(9)*	0.3 (*8*)	0.4 *(5)*	0.4 *(5)*
Other including mixed background	1.1 *(57)*	1 *(27)*	1 (*26*)	0.9 *13)*	0.9 *13)*
N/A	16.3 *(821)*	12.1 *(337)*	11.1 (*279*)	10.9 *(152)*	10.9 *(152)*
No Data	0.1 *(5)*	0.1 *(1)*	0.1 *(1)*	0.1 *(1)*	0.1 *(1)*
Employment status % (*N*)					
Employed	28.6 (*1439*)	32.7 (*910*)	31.2 (*781*)	30.2 (*422*)	29.9 (*3*)
Unemployed and seeking work	12.1 (*608*)	12.6 (*351*)	13.3 (*333*)	13.7 (*192*)	13.7 (*183*)
Student	5.4 (*272*)	5.6 (*155*)	6 (*151*)	6.2 (*151*)	6.2 (*83*)
Long Term Sick, Disabled or Benefits	15.2 (*766*)	15.5 (*430*)	15.4 (*387*)	17 (*237*)	17.4 (*232*)
Homemaker not working or actively seeking work	15.5 (*781*)	16.9 (*471*)	17 (*425*)	15.7 (*220*)	15.4 (*235*)
Not receiving benefits and not working or actively seeking work	0.4 (*21*)	0.4 (*11*)	0.4 (*10*)	0.5 (*7*)	0.5 (*7*)
Unpaid voluntary work, not working or actively seeking work	0.7 (*33*)	0.8 (*23*)	0.8 (*21*)	1.1 (*16*)	1.2 (*16*)
Retired	8.2 (*414*)	7.3 (*202*)	7.1 (*178*)	7.8 (*109*)	8.2 (*109*)
N/A	13.5 (*680*)	8 (*223*)	8.7 (*217*)	7.6 (*108*)	7.4 (*99*)
No Data	0.2 (5)	0.2 (5)	0.2 (*4*)	0.1 (2)	0.1 (2)
Presenting problems %					
Anxiety	74.2	80.3	80.3	80.5	80.4
Depression	54.9	59.3	60	59.7	60.1
Interpersonal/relationship	40.1	44.7	44.7	46.8	46.9
Self-esteem	35.4	38.7	39	38.6	38.4
Breavement/loss	30	31.1	30.3	30.4	30.5
Trauma/abuse	22.1	23.5	23.7	22.1	21.9
Work/academic	13.6	15.1	15.1	16.1	16.1
Living/welfare	19.8	20.2	20.8	20.3	20.3

Appendix B – Table showing pre post outcomes for all samples

Pre post therapy raw and adjusted effect sizes calculated for Assessed, Attenders, Form Returner and Completers samples

		Pre therapy		Post therapy								95% CI		
Sample	*N*	Mean	*SD*	Mean	*SD*	Pre-post difference	*t*	df	*r*	*p*	RAW *ES*	LL	UL	*ES* Dunlap et al. (1996)
Attender sample	2151	1.56	0.609	1.00	0.673	0.57	41.608	2150	.520	<.0001	0.92	0.539	0.592	0.88
Assessed sample	2507	1.56	0.606	1.05	0.689	0.51	40.950	2506	.538	<.0001	0.84	0.488	0.537	0.89
Form returner sample	1397	1.54	0.601	0.84	0.599	0.69	43.651	1396	.509	<.0001	1.16	0.663	0.725	1.16
Completer sample	1335	1.54	0.598	0.84	0.592	0.70	42.864	1334	.497	<.0001	1.17	0.668	0.732	1.17

Note. Full Participant sample = All clients age > 18 and who consented to participate in the study

Assessed sample = Full participant sample and all participants with valid pre therapy CORE OM scores

Attenders sample = Assessed sample and attended at least 2 therapy sessions (assessment session plus one therapy session)

Form returner sample = Attenders sample and therapist completed pre therapy CORE form and valid post therapy CORE OM,

Completer sample =Form returner sample and planned ending recorded by therapist.

RAW ES = ((Pre therapy mean CORE OM score - Post therapy Mean CORE OM score) / pre therapy SD))

ES Dunlap et al. (1996) see page XX for full equations

Recovery and improvement rates in sub-samples of study participants

Sample	*N*	Recovery rate: RCSI criterion		Reliable improvement rate		Reliable deterioration rate		No reliable Change	
		N	%	*n*	%	*n*	%	*n*	%
Attender sample*	2,151	823	38.3	1161	54.0	74	3.4	912	42.4
Form returner sample	1,397	648	46.4	894	64.0	27	1.9	474	33.9
Completer sample	1,335	625	46.8	860	64.4	26	1.9	447	33.5
Clinical samples (pre therapy CORE OM ≥ 1.00)									
Attender sample*	1765	823	46.6	1083	61.4	49	2.8	630	35.7
Form returner sample	1138	648	56.9	830	72.9	14	1.2	293	25.7
Completer sample	1087	625	57.5	799	73.5	13	1.2	274	25.2

Note. Attender Sample = therapist completed pre therapy CORE form, age ≥ 18, valid pre therapy CORE OM results.

Form returner sample = attender sample and valid post therapy CORE OM.

Completer sample = form returner sample and planned ending recorded by therapist.

*Last Observation Carried Forward applied to impute missing post therapy CORE OM scores

Appendix D – CORE Therapy Assessment Form

THERAPY ASSESSMENT FORM V2

Site ID (letters / numbers) **Age**

Client ID **Male** **Female**

Sub Codes (TH ID number / SC2 numbers / SC3 numbers) **Employment**

Referrer(s) **Ethnic Origin**

Referral date (D D / M M / Y Y Y Y)

First assessment date attended (D D / M M / Y Y Y Y)

Last assessment date (D D / M M / Y Y Y Y)

Total number of assessments

Previously seen for therapy in this service? Yes / No — Episode

Months since last episode

Is this a follow-up/review appointment? Yes / No

Relationships/support *Please tick as many boxes as appropriate*

- Living alone (not including dependents)
- Living with partner
- Caring for children under 5 years
- Caring for children over 5 years
- Living with parents/guardian
- Living with other relatives/friends
- Full time carer (of disabled/elderly etc)
- Living in shared accommodation (eg lodgings)
- Living in temporary accommodation (eg hostel)
- Living in institution/hospital
- Other

Current/previous use of services for psychological problems?
Please tick as many boxes as appropriate

		Concurrent	< 12 mths	> 12 mths
Primary	GP or other member of primary care team (eg practice nurse, counsellor)			
Secondary	In primary care setting			
	In community setting			
	In hospital setting on sessional basis			
	Day care services (eg day hospital)			
	Hospital admission < = 10 days			
	Hospital admission > = 11 days			
Specialist	Psychotherapy/psychological treatments from specialist team (sessional)			
	Attendance at day therapeutic programme			
	Inpatient treatment			
Other	Counsellor in eg voluntary, religious, work, educational setting			

Is the client currently prescribed medication to help with their psychological problem(s)? Yes ☐ No ☐

If yes, please indicate type of medication:

Anti-psychotics (neuroleptics/major tranquillizers) ☐ Anti-depressants ☐ Anxiolytics/Hypnotics (minor tranquillizers) ☐ Other ☐

Survey : 76

© CORE System Trust: http://www.coreims.co.uk/copyright.pdf
Supported by www.coreims.co.uk

Page : 1

Brief description of reason for referral

Identified Problems/Concerns

Severity		< 6 months	6-12 months	> 12 months	Recurring/ continuous
☐	Depression	☐	☐	☐	☐
☐	Anxiety/Stress	☐	☐	☐	☐
☐	Psychosis	☐	☐	☐	☐
☐	Personality Problems	☐	☐	☐	☐
☐	Cognitive/Learning	☐	☐	☐	☐
☐	Eating Disorder	☐	☐	☐	☐
☐	Physical Problems	☐	☐	☐	☐
☐	Addictions	☐	☐	☐	☐

Severity		< 6 months	6-12 months	> 12 months	Recurring/ contin.
☐	Trauma/abuse	☐	☐	☐	☐
☐	Bereavement/loss	☐	☐	☐	☐
☐	Self esteem	☐	☐	☐	☐
☐	Interpersonal/relationship	☐	☐	☐	☐
☐	Living/Welfare	☐	☐	☐	☐
☐	Work/Academic	☐	☐	☐	☐
☐	Other *(specify below)*	☐	☐	☐	☐

Risk

	None	Mild	Mod	Sev
Suicide	☐	☐	☐	☐
Self Harm	☐	☐	☐	☐
Harm to others	☐	☐	☐	☐
Legal/Forensic	☐	☐	☐	☐

ICD-10 CODES

	F/Z	Main code	Sub-code		F/Z	Main code	Sub-code
1	☐	☐☐	• ☐	3	☐	☐☐	• ☐
2	☐	☐☐	• ☐	4	☐	☐☐	• ☐

What has the client done to cope with/avoid their problems? *Please tick, and then specify actions*

Positive actions ☐

Negative actions ☐

Assessment outcome ***(tick one box only)***

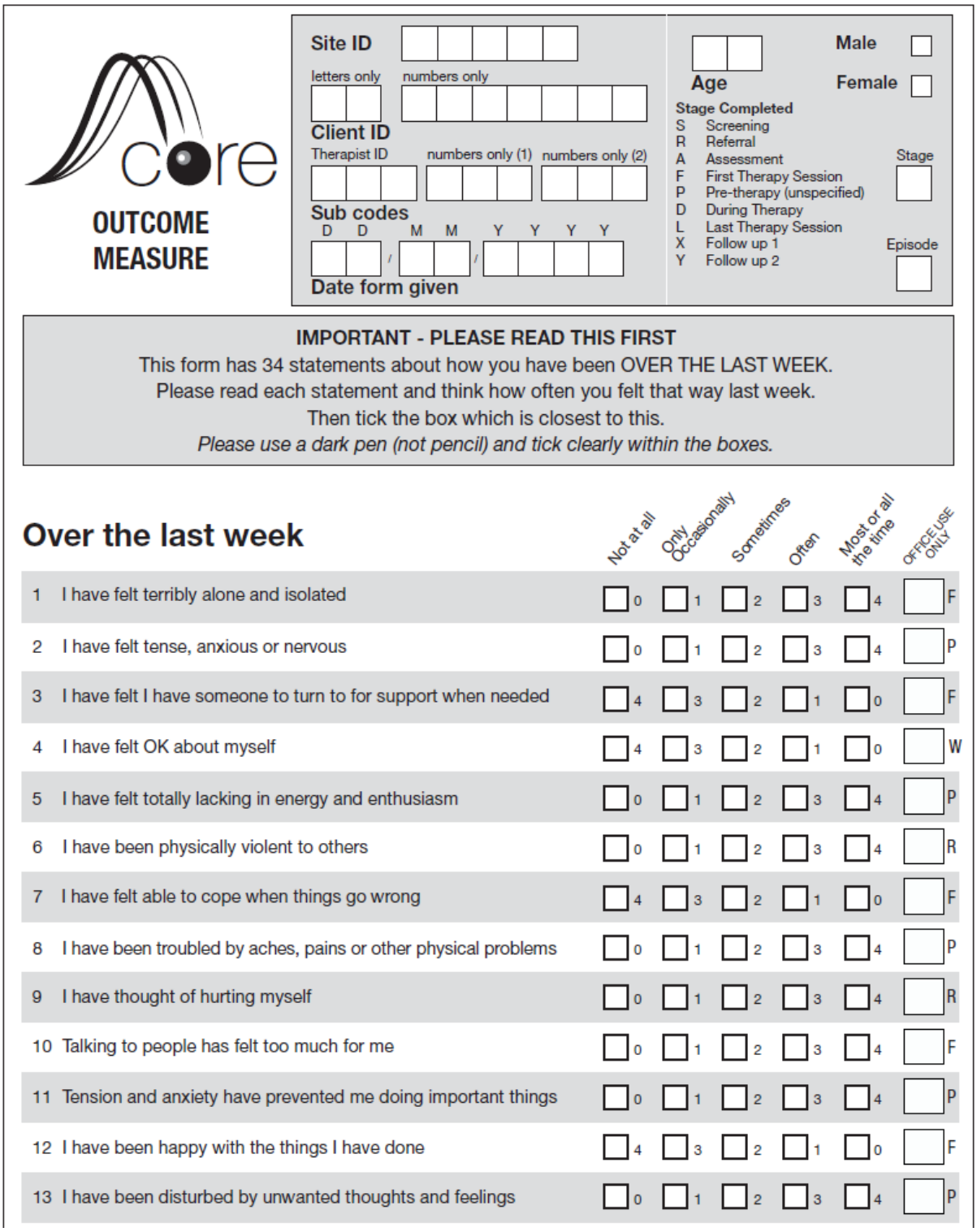
- Assessment/one session only ☐
- Accepted for therapy ☐
- Accepted for trial period of therapy ☐
- Long consultation ☐
- * Referred to other service ☐
- * Unsuitable for therapy at this time ☐

*** If the client is not entering therapy give brief reason**

39192

© CORE System Trust: http://www.coreims.co.uk/copyright.pdf
Supported by www.coreims.co.uk

Appendix E – CORE Outcome Measure 34



core

OUTCOME MEASURE

Site ID

letters only

numbers only

Client ID

Therapist ID

numbers only (1)

numbers only (2)

Sub codes

Date form given

D D / M M / Y Y Y Y

Age

Male

Female

Stage Completed

S Screening
R Referral
A Assessment
F First Therapy Session
P Pre-therapy (unspecified)
D During Therapy
L Last Therapy Session
X Follow up 1
Y Follow up 2

Stage

Episode

IMPORTANT - PLEASE READ THIS FIRST

This form has 34 statements about how you have been OVER THE LAST WEEK.
Please read each statement and think how often you felt that way last week.
Then tick the box which is closest to this.
Please use a dark pen (not pencil) and tick clearly within the boxes.

Over the last week

		Not at all	Only Occasionally	Sometimes	Often	Most or all the time	OFFICE USE ONLY
1	I have felt terribly alone and isolated	0	1	2	3	4	F
2	I have felt tense, anxious or nervous	0	1	2	3	4	P
3	I have felt I have someone to turn to for support when needed	4	3	2	1	0	F
4	I have felt OK about myself	4	3	2	1	0	W
5	I have felt totally lacking in energy and enthusiasm	0	1	2	3	4	P
6	I have been physically violent to others	0	1	2	3	4	R
7	I have felt able to cope when things go wrong	4	3	2	1	0	F
8	I have been troubled by aches, pains or other physical problems	0	1	2	3	4	P
9	I have thought of hurting myself	0	1	2	3	4	R
10	Talking to people has felt too much for me	0	1	2	3	4	F
11	Tension and anxiety have prevented me doing important things	0	1	2	3	4	P
12	I have been happy with the things I have done	4	3	2	1	0	F
13	I have been disturbed by unwanted thoughts and feelings	0	1	2	3	4	P

14 I have felt like crying	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4		W

Please turn over

© CORE System Trust: http://www.coreims.co.uk/copyright.pdf
Supported by www.coreims.co.uk

Over the last week

	Not at all	Only Occasionally	Sometimes	Often	Most or all the time	OFFICE USE ONLY	
15 I have felt panic or terror	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4		P
16 I made plans to end my life	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4		R
17 I have felt overwhelmed by my problems	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4		W
18 I have had difficulty getting to sleep or staying asleep	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4		P
19 I have felt warmth or affection for someone	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0		F
20 My problems have been impossible to put to one side	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4		P
21 I have been able to do most things I needed to	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0		F
22 I have threatened or intimidated another person	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4		R
23 I have felt despairing or hopeless	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4		P
24 I have thought it would be better if I were dead	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4		R
25 I have felt criticised by other people	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4		F
26 I have thought I have no friends	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4		F
27 I have felt unhappy	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4		P
28 Unwanted images or memories have been distressing me	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4		P
29 I have been irritable when with other people	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4		F
30 I have thought I am to blame for my problems and difficulties	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4		P

31 I have felt optimistic about my future	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0	W
32 I have achieved the things I wanted to	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0	F
33 I have felt humiliated or shamed by other people	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	F
34 I have hurt myself physically or taken dangerous risks with my health	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	R

THANK YOU FOR YOUR TIME IN COMPLETING THIS QUESTIONNAIRE

	(W)	(P)	(F)	(R)	All items	All minus R
Total Scores	☐	☐	☐	☐	☐	☐
Mean Scores (Total score for each dimension divided by number of items completed in that dimension)	☐	☐	☐	☐	☐	☐

Survey: 151 © CORE System Trust: http://www.coreims.co.uk/copyright.pdf Page: 1
Supported by www.coreims.co.uk

Appendix F – CORE End of Therapy Form

END OF THERAPY FORM v.2

Site ID

Client ID letters numbers

Sub Codes Therapist ID SC4 numbers SC5 numbers

Date therapy commenced D D / M M / Y Y Y Y

Date therapy completed D D / M M / Y Y Y Y

Number of sessions planned

Number of sessions attended

Number of sessions unattended

What type of therapy was undertaken with the client? *Please tick as many boxes as appropriate*

Psychodynamic	☐	Person-centred	☐
Psychoanalytic	☐	Integrative	☐
Cognitive	☐	Systemic	☐
Behavioural	☐	Supportive	☐
Cognitive/Behavioural	☐	Art	☐
Structured/Brief	☐	Other *(specify below)*	☐

What modality of therapy was undertaken with the client? *Please tick as many boxes as appropriate*

Individual	☐	Family	☐
Group	☐	Marital/Couple	☐

What was the frequency of therapy with the client?

More than once weekly	☐	Less than once weekly	☐
Weekly	☐	Not at a fixed frequency	☐

Which of the following best describes the ending of therapy?

Unplanned ☐		**Planned** ☐	
Due to crisis	☐	Planned from outset	☐
Due to loss of contact	☐	Agreed during therapy	☐
Client did not wish to continue	☐	Agreed at end of therapy	☐
Other unplanned ending *(specify below)*	☐	Other planned ending *(specify below)*	☐

Survey : 78

© CORE System Trust: http://www.coreims.co.uk/copyright.pdf
Supported by www.coreims.co.uk

Page : 1

Review of Identified Problems/Concerns

Severity		Therapy issue	Severity		Therapy issue
☐	Depression	☐	☐	Trauma/Abuse	☐
☐	Anxiety/Stress	☐	☐	Bereavement/Loss	☐
☐	Psychosis	☐	☐	Self esteem	☐
☐	Personality Problems	☐	☐	Interpersonal/relationship	☐
☐	Cognitive/Learning	☐	☐	Living/Welfare	☐
☐	Physical Problems	☐	☐	Work/Academic	☐
☐	Eating Disorder	☐	☐	Other *(specify below)*	☐
☐	Addictions	☐			

Risk

	None	Mild	Mod	Sev
Suicide	☐	☐	☐	☐
Self Harm	☐	☐	☐	☐
Harm to others	☐	☐	☐	☐
Legal/Forensic	☐	☐	☐	☐

Contextual Factors

	Poor	Moderate	Good
Motivation	☐	☐	☐
Working Alliance	☐	☐	☐
Psychological Mindedness	☐	☐	☐

Benefits of Therapy

	Improved Yes	Improved No	Not addressed
Personal insight/understanding	☐	☐	☐
Expression of feelings/problems	☐	☐	☐
Exploration of feelings/problems	☐	☐	☐
Coping strategies/techniques	☐	☐	☐
Access to practical help	☐	☐	☐
Other benefits	☐		

Tick box and then specify below

	Improved Yes	Improved No	Not addressed
Control/planning/decision making	☐	☐	☐
Subjective well-being	☐	☐	☐
Symptoms	☐	☐	☐
Day to day functioning	☐	☐	☐
Personal relationships	☐	☐	☐

Has contact with this service resulted in a change of medication? Yes ☐ No ☐ Not applicable ☐

If yes, is this change likely to be of benefit to the client? Yes ☐ No ☐

Details of change: Started ☐ Discontinued ☐ Increased ☐ Decreased ☐ Modified ☐

Has the client been given a follow-up appointment? Yes ☐ No ☐

Number of months until appointment ☐☐

© CORE System Trust: http://www.coreims.co.uk/copyright.pdf
Supported by www.coreims.co.uk

Appendix G – Health Related Quality of Life Questionnaire

Please rate the following:

1. Would you say that in general your health is:
   a. Excellent ☐
   b. Very Good ☐
   c. Good ☐
   d. Fair ☐
   e. Poor ☐

   or

   Don't Know/Not Sure ☐

2. Now thinking about your physical health, which includes physical illness and injury, for how many days during the past 30 days was your physical health not good?

   Insert number of days here ☐

   or

   a. None ☐
   b. Don't Know/Not Sure ☐

3. Now thinking about your mental health, which includes stress, depression, and problems with emotions, for how many days during the past 30 days was your mental health not good?

   Insert number of days here ☐

   or

   a. None ☐
   b. Don't Know/Not Sure ☐

4. During the past 30 days, for about how many days did poor physical or mental health keep you from doing your usual activities, such as self-care, work, or recreation?

   Insert number of days here ☐

   or

   c. None ☐
   d. Don't Know/Not Sure ☐