UNIVERSITY OF SOUTHAMPTON

FACULTY OF SOCIAL, HUMAN AND MATHEMATICAL SCIENCES

Psychology

Volume 1 of 1

The utility of functioning in predicting health service use by people with common mental disorders

by

Conal Twomey

'Three Paper' thesis for the degree of Doctor of Philosophy

May 2017

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL, HUMAN AND MATHEMATICAL SCIENCES

Psychology

'Three Paper' thesis for the degree of Doctor of Philosophy

THE UTILITY OF FUNCTIONING IN PREDICTING HEALTH SERVICE USE BY PEOPLE WITH COMMON MENTAL DISORDERS

Conal Dominick Twomey

Background

Mental health services are under-resourced in countries around the world. It is essential therefore that the scarce resources are allocated fairly and efficiently. To facilitate this process, extensive knowledge of the variables associated with health service use (HSU) by people with mental disorders is required. As diagnoses and symptoms may not explain a large proportion of the variance in this HSU, there is a need to investigate the relevant associations of other variables. The term 'functioning' relates to physical and mental functions and the ability to undertake activities and participate in all of life. The studies described within this thesis were to determine the utility of functioning in predicting HSU by people with common mental disorders. It was hypothesised that functioning would have good utility in this prediction.

Methods

A systematic review was initially undertaken, followed by three empirical studies: one set in low- and middle-income countries (LMICs) and two within the English National Health Service (NHS).

Results

(1) Systematic Review: There is a lack of existing research into the association of functioning with HSU. The available evidence was mixed: three functioning variables (i.e. cognitive deficits, social withdrawal and activities of daily living) were significantly associated with HSU, but only in 55% of the assessments conducted on them.

(2) Cross-sectional study: For older people with depression living in LMICs (*n*= 4,590), functioning was significantly associated with one HSU outcome (i.e. hospital admission) but not another (i.e. "any community HSU").

(3) Historical cohort study: the seven functioning items from the HoNOS instrument were not associated with mental health service costs for patients with common mental disorders, over one year (*n*=1,343).

(4) Prospective cohort study: functioning predicted total NHS costs for patients with mood and anxiety disorders over a six month period (*n*=102).

Discussion

The mixed findings provide partial support for the hypothesised predictive ability of functioning. Narrow operationalisations of functioning evidently have limited utility in predicting HSU. However, the positive findings of the final study support the predictive ability of 'multi-domain' functioning. Multi-domain functioning information could therefore be useful for the formation of patient clusters within mental health payment systems.

Table of Contents

List of Tables

List of Figures

List of Accompanying Materials

The below papers arising from this thesis have been published in academic journals. Please email c.twomey@soton.ac.uk if you would like a copy of them.

Twomey, C. D., Baldwin, D. S., Hopfe, M., & Cieza, A. (2015). A systematic review of the predictors of health service utilisation by adults with mental disorders in the UK. *BMJ Open, 5*(7), e007575. doi:10.1136/bmjopen-2015-007575

Twomey, C. D., Prince, M., Cieza, A., Baldwin, D. S., & Prina, A. M. (2015). Cross-sectional associations of depressive symptom severity and functioning with health service use by older people in low-and-middle income countries. *International Journal of Environmental Research & Public Health, 12*(4), 3774-3792. doi:10.3390/ijerph120403774

Twomey, C., Prina, A. M., Baldwin, D. S., Das-Munshi, J., Kingdon, D., Koeser, L., . . . Cieza, A. (2016). Utility of the Health of the Nation Outcome Scales (HoNOS) in Predicting Mental Health Service Costs for Patients with Common Mental Health Problems: Historical Cohort Study. *PLoS One, 11*(11), e0167103. doi:10.1371/journal.pone.0167103

DECLARATION OF AUTHORSHIP

I, **Conal Twomey** declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Title: The utility of functioning in predicting health service use by people with common mental disorders

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published [for references, see ‘list of Accompanying Materials’ on page vi)

Signed:
Date: 01/05/2017

Acknowledgements

The research leading to all results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013/ under REA grant agreement n°316795

On a personal level, I greatly appreciate the help I received from so many people in completing this PhD thesis.

Thank you to everyone who gave up their time to work with me on the four research papers and for making all of the collaborations hugely enjoyable.

I would like to particularly acknowledge the people who worked most closely with me in Southampton over the last three years. Magda Nowak was incredibly helpful on a daily basis with administration, project planning and general support that went well beyond her job description. She was a pleasure to work alongside. My secondary supervisor, David Baldwin was extremely supportive, considerate, and conscientious. I consider myself very lucky to have worked with someone with such a deep understanding of clinical research, and working with him regularly in the clinic always put a spring in my step. Thank you to my primary supervisor Alarcos Cieza. I learned so much from working with Alarcos, on both a personal and professional level. Learning from her has increased my confidence and career ambitions hugely. I enjoyed our many conversations about research, and her humbleness and kindness was always evident throughout our time working together. She also did a great job running the MARATONE research consortium which provided me with exceptional research training and many happy memories. Thank you very much Magda, David and Alarcos!

Thank you to all my close friends who I am lucky to have. In particular, Liz, Richard and Polly were an amazing support and listened to any problems I had with the PhD with great patience! Thank you to everyone in the MARATONE consortium for the amazing memories! To my family - parents Dom and Anne and brother Gavin - it was difficult at times to be living in a different country but you were always fully behind me. Some of my best memories of Southampton were your visits, and it is great to be back with you in Ireland.

I'll only look back on this part of my life with great fondness. Thank you again everyone.

BACKGROUND AND AIMS

Overview

This thesis investigates the utility of functioning in the prediction of health service use (HSU) by people with common mental disorders. I begin by explaining why predicting HSU is important in general, while highlighting the challenges involved. I then provide a rationale for investigating the predictive utility of functioning.

The importance of predicting HSU

Many stakeholders with often conflicting interests are involved in the delivery of public health services. Patients seek the best obtainable care, doctors strive for autonomy regarding how services should be delivered, and policy-makers need to balance meeting high public demand with controlling health service expenditure (Essen, 2009a). To meet the demands of patients, doctors, and policy-makers, it is essential that health service resources are distributed to where they are needed most, in a systematic manner that ensures both fairness and efficiency.

To enable the fair and efficient distribution of health service resources, it is important to have good knowledge of the patient-related factors (e.g. diagnosis, symptoms, age, comorbidity) that can be used to predict HSU patterns. For example, regarding age, research has shown that HSU by people aged 65 or older is 3-5 times higher than that of younger people (Casey et al., 2003), and the ongoing demographic shift towards an older global population will increasingly place greater demands on health services (Prina, Deeg, Brayne, Beekman, & Huisman, 2012). The knowledge that age predicts HSU enables policy-makers to account for this in resource distribution, and may justify the allocation of additional resources towards services for older people, over the long term.

The importance of predicting HSU is reflected in ‘activity-based’ health service payment systems which, since the 1990s, have become a principal means of distributing resources in both high-income and low-and-middle income countries (Mathauer & Wittenbecher, 2013). These payment systems aim to achieve fairness and efficiency by financially incentivising health service providers to treat more patients, cut costs, and reduce waiting list times (Street & Maynard, 2007). In these systems, resources are allocated towards distinct patient ‘clusters’ (often termed diagnostic-related groups). Each patient treated by a health service provider is assigned to a specific cluster based on collected information about a range of patient-related variables which are associated with HSU (Mathauer & Wittenbecher, 2013). It is essential that patient clusters represent distinct clinical entities with distinct HSU patterns (Busse, Schreyögg, & Smith, 2006). This is because providers typically receive a fixed payment based on the cluster each patient is allocated to, with clusters with higher expected HSU generating higher payments than those with lower expected HSU (Marini & Street, 2007).

Therefore, in payment systems, the prediction of HSU by patients and patient clusters is important because it directly impacts the allocation of resources.

Despite the widespread use of 'activity-based' health service payment systems as a means to distribute health service resources, these systems often do not cover mental health services. Indeed, only a few high income countries (e.g., Australia, New Zealand, Canada, the Netherlands, Norway, USA, UK) have made any meaningful progress in the implementation of mental health payment systems, with widely varying methodologies (Mason, Goddard, Myers, & Verzulli, 2011). This indicates that predicting HSU by people with mental disorders is a challenging process. Barriers to this prediction include wide variations in the prognosis of psychiatric disorders and in treatment and care models (Appleby, Harrison, Hawkins, & Dixon, 2012).

Intuitively, it makes sense for psychiatric diagnosis and related symptom severity to be strongly associated with HSU; but previous evidence is mixed. Diagnosis has been shown to be a fairly weak predictor of HSU by people with mental disorders in various studies involving national-level datasets, accounting for around 10% of the variance (Elphick & Antony, 1996; English, Sharfstein, Scherl, Astrachan, & Muszynski, 1986; A. Macdonald & Elphick, 2012; Schumacher, Namerow, Parker, Fox, & Kofie, 1986). On the other hand, more recent large-scale studies indicate that diagnosis, in particular personality disorder, could be relatively robust predictor of this HSU (Jeremy Coid et al., 2002; J. Coid et al., 2009; Jeremy Coid, Yang, Tyrer, Roberts, & Ullrich, 2006; A. D. Tulloch, David, & Thornicroft, 2015). Regarding the association of symptom severity with HSU by people with mental disorders, there also exists inconsistency in large-scale studies involving diverse general and clinical populations. Some studies provide evidence for this association (C. Cooper et al., 2010; Feng, Yap, Kua, & Ng, 2009; Fleury, Grenier, Bamvita, & Caron, 2014; Hayward, Jordan, & Croft, 2010; Mills, Van Hooff, Baur, & McFarlane, 2012; Prina et al., 2015; Prina et al., 2013) but some do not (Beekman et al., 2002; Fischer et al., 2002; B. Y. Huang et al., 2000; Kullowatz, Kanniess, Dahme, Magnussen, & Ritz, 2007; Patel et al., 2006). Moreover, in the above studies that did find an association, the magnitude of effect tends to be small. For example, a recent meta-analysis of seven studies found an association of depressive symptom severity with hospital admission, with a small effect size (RR = 1.36) (Prina et al., 2015).

Overall, findings from previous studies indicate that although mental disorder diagnosis and related symptom severity are associated with HSU by people with mental disorders, much of the remaining variance in this HSU is accounted for by other variables. Indeed, studies have shown associations of several sociodemographic variables (e.g. age, gender, ethnicity, marital status, education, private health insurance) with HSU (Albanese et al., 2011; C. Cooper et al., 2010; Fleury et al., 2014; Foster, Meltzer, Gill, & Hinds, 2003; Kent, Fogarty, & Yellowlees, 1995; Vasiliadis, Lesage, Adair, Wang, & Kessler, 2007). Moreover, comorbidity has been shown to be associated with this HSU in several studies (Chollet, Saragoussi, Clay, & Francois, 2013; Fleury et al., 2014; Kent et al., 1995; Mills et al., 2012; Vasiliadis et al., 2007). In addition, the medical cost offset of receiving mental health interventions such as psychotherapy is well-documented (Clark, 2011; Hunsley, 2003; Sava, Yates, Lupu, Szentagotai, & David, 2009).

Despite previous research demonstrating that there are a range of additional variables associated with HSU by people with mental disorders, there are few systematic reviews on the topic, and these differ widely in terms of methodological approach (Hansson & Sandlund, 1992; Kent et al., 1995; Prina et al., 2015). This makes it difficult to ascertain the extent to which each of these additional variables are predictive of HSU, alongside diagnosis and symptom severity. Moreover, there may be important predictive variables that have been under-researched and are thus worthy of investigation. In this thesis, it is argued that one variable worthy of investigation in this regard is 'functioning'. In the next section, I explain why the association of functioning with HSU by people with mental disorders represents an important research gap, with implications for resource distribution.

Using functioning to predict health service use by people with mental disorders

In the *World Health Organization (WHO) International Classification of Functioning, Disability and Health (ICF),* 'functioning' is an encompassing term relating to the physiological and psychological functions of the body, and the ability to undertake activities and participate in all life situations (Cieza, Oberhauser, Bickenbach, Chatterji, & Stucki, 2014). In this widely-accepted theoretical framework, one's level of functioning is determined by interactions between health conditions and contextual factors (Figure 1) (Neubert, Sabariego, Stier-Jarmer, & Cieza, 2011; WHO, 2001). Functioning can be seen as a universal experience because at some point in life - and most markedly in later life - our ability to complete activities and participate will be somewhat limited by the impact of an experienced health condition (Cieza et al., 2014). Functioning also takes into account the hypothesis of "horizontal epidemiology", namely that common decrements in functioning (e.g. in sleeping, concentration, relationships) are often experienced across neurological and mental disorders (Cieza et al., 2015).

Due to its universality and multi-dimensionality, functioning arguably reflects an intuitive notion of overall health, and can be considered an appropriate descriptor of health in a given population (Cieza et al., 2014). Thus, it makes sense for functioning to be related to care needs and HSU. The utility of functioning in predicting HSU in mental disorders is worthy of investigation for various reasons: people with mental disorders experience substantial decrements in functioning and describe functional recovery as essential for remission (R. W. Lam, Parikh, Michalak, Dewa, & Kennedy, 2015; Zimmerman et al., 2006) poorer functioning has been shown to predict the recurrence of depressive and anxiety disorders (Rodriguez, Bruce, Pagano, & Keller, 2005) and functioning is increasingly being recognised as a priority in the treatment and assessment of mental disorders, as reflected in the new dimensional approach of the DSM-5 whereby it is rated alongside diagnostic severity (Gold, 2014; R. W. Lam et al., 2015).

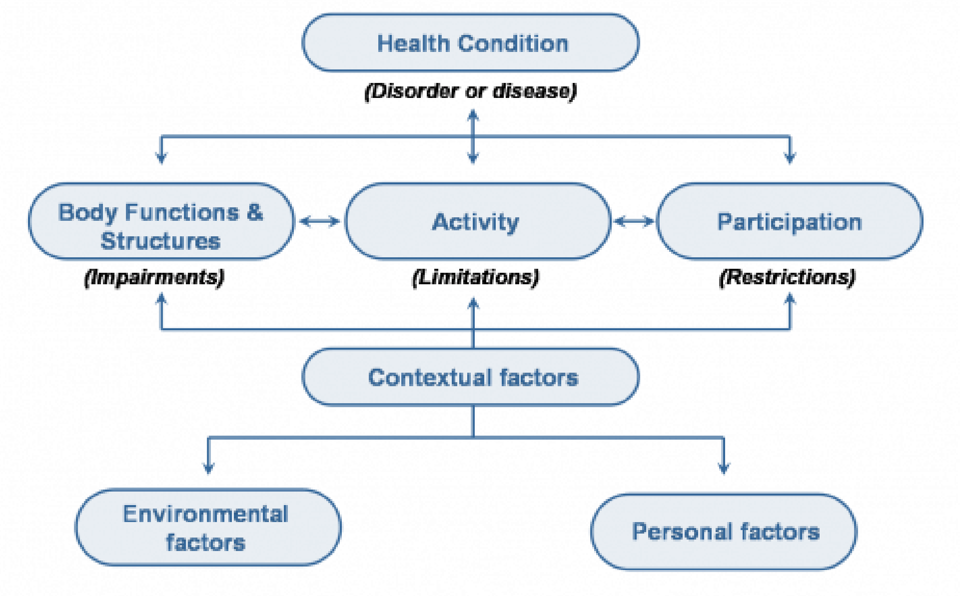


Figure 1. ICF framework (WHO, 2001).

The association of functioning with HSU has been found in numerous studies involving (predominantly older) adults with differing health problems, including neurological disability, physical disability, and stroke (Carpenter, Bobby, Kulinskaya, & Seymour, 2007; Chuang et al., 2003; Covinsky, Justice, Rosenthal, Palmer, & Landefeld, 1997; Disler, Roy, & Smith, 1993; Hopfe et al., 2016; Mayo et al., 2005; Perrin et al., 2011). Relatively few studies though, have investigated the association of of functioning with HSU by people with mental disorders. Some studies have yielded significant associations (C. Cooper et al., 2010; Fleury, Grenier, Bamvita, & Caron, 2011; A. D. Tulloch, Khondoker, Thornicroft, & David, 2014) but their comparibilty is limited due to differences in sample compositions (e.g. community dwellers, psychiatric inpatients), and their operationalisations of functioning (e.g. activities of daily living, global assessment of functioning) and HSU outcomes (e.g. primary care HSU, psychiatric admissions). Further research is therefore needed to ascertain the utility of functioning in prediciting HSU by people with mental disorders. This is an important research gap given that functioning - as a good indicator of overall health - potentially accounts for a substantial amount of variance in this HSU, making it useful for policy-makers and mental health service planning.

Limitations of the ICF

Appraisals of the utility of functioning in predicting HSU should be made in light of various limitations of the ICF classification system. One important limitation concerns the bio-psycho-social model underpinning it: it has been argued that a lack of coherence in the application of the model has led to an 'undisciplined eclecticism' whereby there exist no safeguards against either the dominance or the under-representation of any one of the three domains of bio, psycho, or social (Benning, 2015). Indeed, evidence supporting this argument can be drawn from the ICF, through the apparent under-representation and under-development of contextual factors (personal and environmental). Unlike other elements of the

ICF which are linked to other classification systems such as the WHO's International Classification of Diseases (ICD), there are no clear taxonomical classifications for contextual factors (Bornbaum, Doyle, Skarakis-Doyle, & Theurer, 2013). Moreover, various ICF-based self-report measures such as the WHODAS World Health Organization Disability Assessment Schedule II (WHODAS-II) (Rehm, Ustun, Saxena, Nelson, & Chatterji, 1999) and the PARADISE 24 (Cieza et al., 2015) do not directly address them. This under-representation of contextual factors is significant given the unequivocal evidence demonstrating their impact on both physical and mental health: factors such as age, ethnicity, gender, geographic location, service access, lifestyle factors, habits, coping mechanisms, behaviour patterns, personality, experienced stigma, and socio-economic status have the potential to profoundly influence an individual's morbidity, risk factors, functioning, quality of life, and HSU (Albanese et al., 2011; Bornbaum et al., 2013; Casey et al., 2003; Corrigan & Watson, 2002; Kessler et al., 2007; Kotov, Gamez, Schmidt, & Watson, 2010; Prince, Patel, et al., 2007).

Another important limitation of the ICF concerns an apparent lack of awareness of the system amongst many health professionals and patients: a survey conducted with six disability organisations in Sweden found that more than half of the (18) organisation representatives had very limited awareness of the ICF; moreover, representatives reported that very few of their patients were aware it (Lundalv, Tornbom, Larsson, & Sunnerhagen, 2015). Conversely, evidence suggests that the ICF has had considerable influence on research and clinical practice across the globe: between 2001 and 2009, over 670 papers from 34 countries were closely related to the ICF, in fields such as disability, rehabilitation, and employment (Cerniauskaite et al., 2011). Further research is needed to ascertain who is, and who is not, aware of the ICF, especially amongst the general population as opposed its proponents. Such investigations would be useful for two main reasons: (1) a key goal of the ICF is to provide a *global* language for the understanding of functioning, disability, and health (WHO, 2001) and widespread ignorance to the system prevents the realisation of this goal; (2) a potential lack of awareness of the ICF may make it more difficult for proponents of the ICF to argue a case for its shaping of public policy and health care.

The use of the term "health condition" represents another notable limitation of the ICF: it has been argued that "health state" is a more appropriate term due to its greater inclusivity and dynamism (Bornbaum et al., 2013). This argument is especially applicable to psychiatric disorders, given their considerable instability in terms of diagnosis and prognosis (Appleby et al., 2012).

Aims and specific objectives

The aim of the studies described within this thesis is to determine the utility of functioning in predicting HSU by people with common mental disorders. The findings in relation to this aim may help policy-makers and healthcare providers decide whether functioning warrants a more prominent role in health service planning, which could ultimately benefit patients. The focus on common mental disorders (e.g. depression, anxiety), is for two main reasons: (1) it adds precision to the findings, especially since people with more complex disorders are likely to have different backgrounds, service needs, and HSU patterns; (2) it was more feasible to collect and access data pertaining to people with common mental disorders than those with more complex problems.

There are two specific objectives in the thesis. First, to identify variables which might predict HSU by people with mental disorders. Work addressing this objective will serve as an introduction to the research area and will highlight where functioning can potentially fit into health service planning, alongside other important variables. The second specific objective is to determine the association of functioning with HSU by people with common mental disorders. Addressing this objective will provide evidence that directly relates to the overall aim and comprises the main part of this thesis. The thesis has four studies - a systematic review and three empirical studies - relating to the two specific objectives (Table 1) and these are described in the next sections.

Table 1. *Studies addressing specific objectives*

Objective	**Study Type**
Identify variables which might predict HSU by people with mental disorders	Systematic review
Determine association of functioning with HSU by people with common mental disorders	Cross-Sectional study Historical cohort study Prospective cohort study

Notes: HSU= health service use; NHS= National Health Service

It is hypothesised that functioning will have good utility in the prediction of HSU by people with common mental disorders. This predicted outcome is based on various reasons described earlier: functioning is arguably a good descriptor of overall health and care needs; people with mental disorders experience substantial decrements in functioning; poor functioning predicts relapse in mental disorders; and functioning evidently predicts HSU in the general population. Full support for this hypothesis would be shown through significant associations of functioning with HSU across the four included studies.

INTRODUCTION TO PAPER 1

Addressing the first specific objective of the thesis - to identify variables which might predict HSU by people with mental disorders - Paper 1 encompasses a systematic review of the predictors of HSU by adults with mental disorders in the UK. Conducting the review enabled me to gain knowledge of important predictors of HSU, and existing research gaps, in a structured and comprehensive manner; and this knowledge informed the design of further papers. The review was undertaken within the context of the ongoing development of a proposed payment system for NHS mental health services - *Mental Health Payment by Results (PbR)* (NHS, 2013). Learning about these policy developments was an additional benefit of conducting this review. The decision to limit the review to UK studies was made for two main reasons: (1) it facilitated closer application of the findings to the development of the UK's PbR system; (2) it was more feasible to conduct a review limited to UK only studies. A consequence of this decision is a reduction in the generalisability of the findings to countries other than the UK.

The study findings have been published in *BMJ Open* (C. D. Twomey, Baldwin, Hopfe, & Cieza, 2015) and the full manuscript appears in the next section. The PhD researcher was the lead author of the paper. Under supervision, the PhD researcher contribution included conceiving and planning the paper, undertaking statistical analysis, drafting all versions of the paper, and managing the submission and peer-review process.

PAPER 1: A systematic review of the predictors of health service utilisation by adults with mental disorders in the UK.

Abstract

Objectives
To identify variables which predict health service utilisation (HSU) by adults with mental disorders in the UK, and to determine the evidence level for these predictors.

Design
A narrative synthesis of peer-reviewed studies published after the year 2000. The search was conducted using four databases (i.e. *PsycINFO, CINAHL Plus with full text, MEDLINE,* and *EMBASE*) and completed on March 25th, 2014.

Setting
The majority of included studies were set in health services across primary, secondary, specialist, and inpatient care. Some studies used data from household and postal surveys.

Participants
Included were UK-based studies that predicted HSU by adults with mental disorders, Participants had a range of mental disorders including psychotic disorders, personality disorders, depression, anxiety disorders, eating disorders, and dementia.

Primary outcome
A wide range of HSU outcomes were examined, including GP contacts, medication usage, psychiatrist contacts, psychotherapy attendances, inpatient days, accident and emergency admissions, and 'total HSU'

Results
Taking into account study quality, 28 studies identified a range of variables with good preliminary evidence supporting their ability to predict HSU. Of these variables, comorbidity, personality disorder, age (heterogeneous age ranges), neurotic symptoms, female gender, a marital status of divorced, separated or widowed, non-white ethnicity, medication, high previous HSU, and activities of daily living were associated with increased HSU. Moreover, good preliminary evidence was found for associations of accessing a primary care psychological treatment service and medication use with decreased HSU.

Conclusions
The findings can inform decisions about which variables might be used to derive mental health clusters in 'payment by results' systems in the UK*.* The findings also support the need to investigate whether combining broad diagnoses with care pathways is an effective method for mental health clustering, and the need for research to further examine the association between mental health clusters and HSU.

Introduction

Many stakeholders with differing needs are involved in the delivery of public health services. Patients seek the best obtainable care, providers aim to deliver optimal care but also strive for self-regulation and autonomy, and policy-makers need to balance meeting high public demand with controlling health service expenditure.(Essen, 2009a) To meet these differing needs, fair and efficient health service payment systems are required. Contemporary, 'activity-based' payment systems aim to achieve this fairness and efficiency by financially incentivising competing providers to treat more patients, cut costs, and reduce waiting list times.(Street & Maynard, 2007)

In typical activity-based payment systems, resources are allocated towards distinct patient 'clusters' (or groups). These clusters are comprised of patients with similar clinical characteristics and expected health service utilisation (HSU) patterns.(Busse et al., 2006) Each patient treated by a health service provider is assigned to a specific cluster based on collected information about a range of 'case-mix' variables (e.g. diagnosis, comorbidity, age) which are associated with HSU.(Mathauer & Wittenbecher, 2013) Thereafter, health service providers receive a fixed payment based on the cluster each patient is allocated to, with clusters with higher expected HSU generating higher payments than those with lower expected HSU.(Marini & Street, 2007) Paying providers fixed payments based on 'clustering' of treated patients allows policy-makers to distribute resources in a systematic and equitable manner.(Busse et al., 2006)

In recent years, there have been ongoing efforts by the National Health Service (NHS) in England to develop (a potential UK-wide) activity-based payment system for its mental health services, in what is referred to as *Mental Health Payment by Results*. Initially, the system will cover secondary care services with various service types excluded (e.g. those relating to primary care psychotherapy, acquired brain injury, and autism).(NHS, 2013) A subject of much debate in this development surrounds how to define 'mental health clusters' for use in this system. In contrast to typical activity-based payment systems, diagnostic information has so far not been used to define these clusters. Instead, clusters have been defined using the newly-developed *Mental Health Clustering Tool* (MHCT). The MHCT assesses the domains of behaviour, symptoms, impairment, social functioning, and risk factors, and is used to assign patients to one of 21 clusters, falling under one of three broad 'super-classes' (non-psychotic, psychotic and organic).(A. J. Macdonald & Elphick, 2011)

One of the main reasons for not using diagnostic information for clustering in *Mental Health Payment by Results* was that mental disorder diagnosis was shown to be a poor predictor of HSU in studies involving national and multi-site trial datasets.(Elphick & Antony, 1996; English et al., 1986; A. Macdonald & Elphick, 2012; Schumacher et al., 1986) On the other hand, it has been argued that although mental disorder diagnosis alone is not sufficient for clustering purposes, information about broad diagnoses and care pathways can be combined, in a simple and practical manner, to form reliable clusters with homogenous resource patterns.(Kingdon et al., 2012) Moreover, the MHCT has also been criticised because its development did not take HSU and costs into account,(Bekas & Michev, 2013) and there currently exists very little evidence for the ability of the MHCT to predict HSU in patient populations.

In the context of the ongoing development of *Mental Health Payment by Results*, and the debate surrounding the use of diagnostic information and the MHCT, it is important to provide evidence that can inform decisions about which variables might be used to derive mental health clusters. To date, no UK-based systematic reviews informing this process have been undertaken. A review of relevant studies set in the UK would address UK-specific HSU patterns, increasing the applicability of findings to the *Mental Health Payment by Results* system. Therefore, the general objective of this systematic review is to identify variables with sufficient evidence supporting their ability to predict HSU. The review has two specific aims. First, to identify the variables examined in relation to the prediction of HSU by adults with mental disorders in the UK. Second, to determine the level of evidence that exists for identified predictors of this HSU.

Method

Inclusion and exclusion criteria

Only the following types of studies were included in the review: (1) observational and intervention studies that predicted HSU by adults with mental disorders. (For the purposes of this review, mental disorders included adults experiencing elevated symptoms of mental disorders, or adults formally diagnosed with a mental disorder. Studies with participants with intellectual disability were excluded due to the specific additional needs of this population which have to be met beyond the healthcare system (e.g. in the education or labour systems); (2) studies based in the UK, with UK participants; (3) peer-reviewed studies published in scientific journals, in the year 2000 or after. (This cut-off point was chosen so that included studies were approximately in line with the overall *Payment by Results* scheme introduced in 2003. Intervention costing studies which did not predict HSU were excluded).

Literature search

Based on these criteria, the first author searched four databases: *PsycINFO, CINAHL Plus with full text, MEDLINE,* and *EMBASE.* The final search was conducted on March 25th, 2014. Additional records were identified from hand-searching reference lists of included studies. Search terms and database subject headings related to HSU (i.e. health care utilisation [subject heading] OR health care utili* OR health service utili* OR health care use OR health service use) were combined with those terms for mental disorders (i.e. Mental disorders [subject heading] OR psychiatric) and the UK location (i.e. UK [subject heading] OR NHS). Due to the differing search procedures deployed by the four databases, slightly altered versions of this search strategy were used in each database. Independent screening of 20% of abstracts was undertaken by the third author. When the first author and third author disagreed regarding the screening outcome of an abstract, the abstract was included in screening at 'full-text' level (by the first author).

Data extraction

Data from included studies were extracted using an Excel spreadsheet. Extracted data pertained to basic study description, study design, records source, data collection times, participants, mental disorder investigated, operationalization of HSU outcomes, the prediction of HSU, and statistics. In addition, each study was assessed for quality using the STROBE statement(von Elm et al., 2008) (for observational studies) and the National Institute for Health and Clinical Excellence (NICE) checklist for Randomized Controlled Trials (RCTs).(NICE, 2009) The former is a checklist of 22 items related to the reporting of title (one item), introduction (two items), methods (nine items), results (five items), discussion (four items), and funding information (one item).(von Elm et al., 2008) The latter assesses bias in RCTs in four sections- selection bias, performance bias, attrition bias, and detection bias.(NICE, 2009)

Data analysis

Due to the heterogeneity in study designs, samples and mental disorders investigated, a meta-analysis was not possible. Narrative synthesis was deemed the most appropriate method of data analysis.

Results

Literature search flow

The literature search flow is displayed in Figure 1. In total, 1,364 records were identified. Database-searching yielded 1,347 records and hand-searching yielded 17 additional records. After duplicates were removed, 928 studies were screened at 'abstract' level. For screening of abstracts, there was a 94.1% agreement rate between the first author and the third author. After abstract screening, 133 studies were assessed for eligibility at 'full-text' level. 28 studies were included in the final review.

Overview of included studies

To provide an overview of included studies, extracted data were summarised in two tables (Tables 1 and 2). Table 1 summarises observational studies of HSU, and Table 2 summarises studies of interventions (of both observational and experimental design) aiming to reduce HSU. As can be seen in both tables, the data source of included studies varied. Most frequently it included routine NHS service data or databases (n = 14), different versions of the *Adult National Psychiatric Morbidity Survey* (n = 6) and other household and postal surveys (n = 3). The sample composition also varied and included adults with a psychotic disorder (n = 7), personality disorder (n = 5), depression (n = 3), an anxiety disorder (n = 2), an eating disorder (n = 1), 'common mental health problems' (n = 2) and dementia (n = 1). It also included health service users (n = 6) and former adolescent psychiatric patients (n = 1). The quality of included studies was mixed. STROBE statement(von Elm et al., 2008) scores for observational studies (n = 25) ranged from 9-20 (mean *[M]* = 15.5; standard deviation *[SD]* = 3.05), out of a possible maximum score of 22. Of the three RCTs assessed using the NICE checklist,(NICE, 2009) two indicated the absence of bias, and one indicated the possible presence of bias. As can be seen in Tables 1 and 2, both the operationalisation of HSU outcomes and the identified predictors of HSU in individual studies varied widely.

Operationalisation of HSU outcomes

To determine the level of evidence for identified predictors of HSU, it was beneficial to first summarise the operationalisation of HSU outcomes across included studies. This summary is provided in Table 3. Across the 28 studies, 60 different HSU outcome variables were assessed 155 times in total: 24 of these related to primary care HSU, 79 to specialist HSU, 40 to inpatient HSU, and 12 to 'total and other' HSU. Across all categories apart from the 'total and other' HSU category, 65 outcomes related to mental health HSU and 78 related to general health HSU.

HSU outcomes used in three or more studies were: medication usage (*n* = 12); inpatient days (*n* = 9); accident and emergency (A & E) admissions (*n* = 8); inpatient admissions (*n* = 8); total HSU (*n* = 8); GP contacts (*n* = 7); GP contacts for psychological problems (*n* = 6); psychotherapy attendances (*n* = 6); community psychiatric nurse contacts (*n* = 5); psychiatrist contacts (*n* = 5); psychiatric inpatient admissions (*n* = 5); psychologist contacts (*n* = 5); nurse contacts (*n* = 4); outpatient attendances (*n* = 4); counsellor contacts (*n* = 3); and home carer visits (*n* = 3). Remaining HSU outcomes are shown in Table 3.

Summary of evidence for identified predictors of HSU

Table 4 provides a summary of the evidence for identified predictors of HSU. The table is structured as follows. First, identified predictors are categorised by 'demographics', 'diagnosis', 'interventions', 'symptoms', 'functioning', and 'behaviour'. Second, the table displays the number of times each identified predictor variable was assessed in relation to HSU, and the number of times each identified variable significantly predicted HSU (and vice versa). Third, using the broad categories of 'primary care HSU', 'specialist HSU', 'inpatient HSU', and 'total HSU', the table documents the operationalisation of HSU outcomes in relation to the prediction of HSU. Fourth, study quality information is provided to aid evaluation of the evidence. For simplicity, a study was arbitrarily deemed to be of 'satisfactory' quality if it scored $\geq$16 on the STROBE statement, (von Elm et al., 2008) or if bias was not present on three out of four domains on the NICE checklist for RCTs.(NICE, 2009)

As an overview, the review identified 31 predictor variables that were examined in relation to the prediction of HSU. By category, these were: twelve demographic variables, six intervention variables, five diagnostic variables, four symptom variables, three functioning variables and one behavioural variable.

The 12 demographic variables significantly predicted increased HSU 41 of 65 times assessed (63.1%). Six demographic variables predicted increased HSU in two or more assessments and in over 50% of assessments made. These variables, in order of frequency of increased HSU prediction, were: comorbidity (both mental and physical), age (heterogeneous age ranges), female gender, a marital status of divorced, separated or widowed, non-white ethnicity, and high previous HSU. Regarding the age variable, several heterogeneous age ranges (e.g. 35-54, 31-49, 35+, 50-64) were associated with increased HSU, thus it was not possible to draw conclusions relating to specific age ranges. Specific age ranges associated with increased HSU in individual studies are viewable in Table 1. As study quality was satisfactory in the vast majority of these assessments, it can be concluded that there exists good preliminary evidence for these six demographic variables in relation to the prediction of increased HSU.

The six intervention variables significantly predicted decreased HSU 10 of 17 times assessed (58.8%). Two intervention variables predicted decreased HSU in two or more assessments and in over 50% of assessments made. These variables, in order of frequency of decreased HSU prediction, were: accessing an *Improving Access to Psychological Therapies* (IAPT) service, and medication. As study quality was satisfactory in all but one these assessments (an assessment of IAPT), it can be concluded that there exists good preliminary evidence for both IAPT and medication in relation to the prediction of decreased HSU.

The five diagnostic variables significantly predicted increased HSU 13 of 15 times assessed (86.6%). Two diagnostic variables predicted increased HSU in two or more assessments and in over 50% of assessments made. These variables, in order of frequency of increased HSU prediction, were: personality disorder, and obsessive compulsive disorder. Whereas all (eight) assessments of personality disorder came from studies of satisfactory quality, none of the (four) assessments of obsessive compulsive disorder came from studies of satisfactory quality. Therefore, it can only be concluded that there exists good preliminary evidence for personality disorder in relation to the prediction of increased HSU.

The four symptom variables significantly predicted increased HSU 7 of 15 times assessed (46.6%). One symptom variable - neurotic symptoms- predicted increased HSU in six of six assessments made. Although two assessments came from studies of unsatisfactory quality, it can be concluded that there exists good preliminary evidence for neurotic symptoms in relation to the prediction of increased HSU.

The three functioning variables significantly predicted increased HSU 5 of 9 times assessed (55.6%). Two functioning variables predicted increased HSU in two or more assessments and in over 50% of assessments made. These variables, in order of frequency of increased HSU prediction, are: cognitive deficits and activities of daily living (ADLs). Whereas all (two) assessments of ADLs came from studies of satisfactory quality, none of the (three) assessments of cognitive deficits came from studies of satisfactory quality. Therefore, it can only be concluded that there exists good preliminary evidence for ADLs in relation to the prediction of increased HSU.

In the final variable category, a behavioural variable- self-harm- significantly predicted increased HSU one of one time assessed. This assessment came from a study of satisfactory quality. However, as just one assessment was undertaken, it cannot be concluded that there exists good preliminary evidence for self-harm in relation to the prediction of increased HSU.

In summary, taking into account frequency of prediction and study quality, several predictor variables have good preliminary evidence supporting their ability to predict HSU by adults with mental disorders in the UK. Of these variables (in order of frequency of prediction), comorbidity, personality disorder, age (heterogeneous age ranges), neurotic symptoms, female gender, a marital status of divorced, separated or widowed, non-white ethnicity, medication, high previous HSU, and activities of daily living were associated with increased HSU. Moreover, good preliminary evidence was found for associations of accessing a primary care psychological treatment service and medication use with decreased HSU. Figure 2 illustrates the relative frequencies of predictors of HSU, by category.

Discussion

Summary of main findings

Taking into account study quality, 28 studies identified a range of variables with good preliminary evidence supporting their ability to predict HSU. Of these variables, comorbidity, personality disorder, age (heterogeneous age ranges), neurotic symptoms, female gender, a marital status of divorced, separated or widowed, non-white ethnicity, medication, high previous HSU, and activities of daily living were associated with increased HSU. Moreover, good preliminary evidence was found for associations of accessing a primary care psychological treatment service and medication use with decreased HSU.

Comparison of main findings with other reviews

Few existing reviews of the predictors of HSU in mental health populations were available for comparison of results. Nevertheless, comorbidity- the most evidenced predictor of increased HSU in the present review- was also shown in a review of 72 studies to predict increased psychiatric service utilisation by 'heavy users' of psychiatric services.(Kent et al., 1995) This previous review found that several variables not examined by studies in our review (i.e. substance abuse, psychotic illness, isolation, homelessness, and social support) were predictive of increased psychiatric service utilisation. In line with the present review, another review of eight studies found that high previous utilisation predicted increased psychiatric service utilisation.(Hansson & Sandlund, 1992) On the other hand, this review found that the variables of living alone and psychosis diagnosis- not examined by studies in the present review- were predictive of increased psychiatric service utilisation.

Overall, the findings from previous reviews add robustness to our finding of good preliminary evidence for the variables of comorbidity and high previous HSU in relation to the prediction of increased HSU by adults with mental disorders in the UK. In addition, despite the sole focus of the previous reviews on psychiatric services which limits their comparability, it is possible that several additional variables- in particular, a psychosis diagnosis- may also predict increased HSU by adults with mental disorders in the UK.

Comparison of main findings with international studies of HSU

As the review was limited to UK studies only, it is informative to compare the findings with those from international studies of HSU by adults with mental disorders. Three recent international studies were chosen for comparative purposes because of their large samples comprising adults with a range of mental health problems.(Fleury et al., 2014; Mills et al., 2012; Vasiliadis et al., 2007) The first was set in Canada, and had a sample of 243 adults diagnosed with various mental disorders.(Fleury et al., 2014) In line with our review, it found that increased social withdrawal, female gender, and (mental disorder) comorbidity were associated with increased HSU. Additional predictors of increased HSU not identified by studies in our review were emotional problems, income, major depression diagnosis and alcohol dependence.

The second study was set in Australia and had a sample of 822 adults who had previously participated in a school-based epidemiological study in their youth.(Mills et al., 2012) In line with our review, it found that age (treated as continuous variable), comorbidity, and a marital status of divorced, were associated with increased HSU. Additional predictors of increased HSU not identified by studies in our review were psychological distress, affective disorder diagnosis, exposure to childhood trauma, while rural living predicted reduced HSU.

The third study(Vasiliadis et al., 2007) used data from a cross-national health survey and involved 8,688 adults from the USA and Canada. It found that comorbidity (various health comorbidities), female gender, and non-white ethnicity were associated with increased HSU. Additional predictors of increased HSU not identified by studies in our review were emotional problems, income, having a regular doctor, and having insurance.

The findings from these international studies add robustness to our finding of good preliminary evidence for the variables of comorbidity, female gender, and a marital status of divorced in relation to the prediction of HSU by adults with mental disorders in the UK. In addition, it is possible that several additional variables identified in international studies- in particular, emotional problems- may also predict HSU by adults with mental disorders in the UK.

Implications of findings for Mental Health Payment by Results

Our findings can inform the debate surrounding the use of diagnostic information and the MHCT for clustering purposes. The findings also highlight several additional variables that are worthy of consideration in the clustering process.

Regarding the use of diagnostic information, in contrast to previous large-scale studies which showed mental disorder diagnosis to be a poor predictor of increased HSU,(Elphick & Antony, 1996; English et al., 1986; Schumacher et al., 1986) the review yielded good preliminary evidence for personality disorder diagnosis in relation to the prediction of increased HSU. In addition, it is noted that diagnoses of psychosis, major depression and affective disorder were identified as predictors in previous reviews and international studies.(Fleury et al., 2014; Hansson & Sandlund, 1992; Kent et al., 1995; Mills et al., 2012)

Although methodological differences (e.g. in the operationalisation of HSU) in these reviews and studies mean that firm conclusions cannot be drawn, a possible explanation for the discrepancy in findings is that some but not other mental disorder diagnoses may be significantly associated with increased HSU. The uncertainty regarding the ability of mental disorder diagnoses to predict increased HSU means that this review neither refutes nor supports the argument that reliable mental health clusters can be formed by combining broad diagnoses with care pathways, in a simple and practical manner.(Kingdon et al., 2012)

Findings relating to the domains of the MHCT (i.e. behaviour, symptoms, impairment, social functioning, and risk factors) can aid assessments of its suitability for clustering purposes. Although some variables relating to these domains were examined, good preliminary evidence for the prediction of increased HSU was found for just two relevant variables- neurotic symptoms and ADLs. Therefore, this review does not provide sufficient evidence to settle the debate regarding the use of the MHCT. However, it highlights the need for further investigation of the link between the MHCT and increased HSU, especially since this link was not taken into account in the initial development of the MHCT.(Bekas & Michev, 2013)

Regarding additional variables worth considering in the clustering process, various demographic (i.e. comorbidity, age, female gender, marital status, non-white ethnicity, high previous HSU) and intervention (i.e. IAPT, medication) variables with good preliminary evidence relating to their ability to predict HSU were identified. Future research could investigate if adding these variables into the 'case mix' of the MHCT adds to the economic validity and reliability of mental health clusters. However, it is worth noting that variables that are predictive of HSU are not always suitable for clustering and resource allocation purposes. For example, concerning demographic variables, it could be argued that it would be unfair to distribute resources on the basis of increased HSU by females (relative to males). Similar arguments could be made regarding other population groupings with contrasting HSU levels (e.g. certain ethnic groups). Moreover, the benefit of using intervention variables for clustering purposes may be somewhat limited because it is relatively easy for providers to use these variables to 'game' the system (i.e. when patients are inappropriately and deliberately allocated to clusters that attract higher fixed payments) in order to generate additional revenue. (A. J. Macdonald & Elphick, 2011)

Methodological considerations

There is relevant research relating to HSU by people with mental disorders not included in this review. This was for various methodological reasons, for example, differing conceptualisations of HSU in investigations by Killapsy and Zi (Killaspy & Zis, 2013) and Trieman and Leff (Trieman & Leff, 2002). These studies focused on the stability of HSU over time and were excluded because they do not address our study question which concerns identifying predictive variables contributing to an increase or decrease in HSU. In addition, various methodological factors should be taken into account when interpreting our findings. First, the quality of included studies was mixed. Specifically, using arbitrarily cut-off points on the STROBE statement(von Elm et al., 2008) and the NICE checklist for RCTs,(NICE,

2009) 18 of the 28 studies (64.2%) were deemed to be of 'satisfactory' quality. This mixed quality limits the strength of conclusions that can be drawn. Second, there was wide heterogeneity in the operationalisation of HSU by included studies, which limits the validity of comparisons across studies. A possible reason for this heterogeneity is that 23 out of 28 (82%) of studies collected secondary data from NHS service databases or household surveys, and thus their operationalisation of HSU was constrained. Addressing this issue, the operationalisation of HSU in included studies was documented in considerable detail (Table 3). Third, the review was limited to UK studies only, meaning the list of identified variables is not exhaustive, and the findings may not be applicable to services in other countries. Indeed, this applicability is particularly limited given that only a few other countries (e.g. Australia, New Zealand, Canada, the Netherlands, Norway, USA) have made progress implementing mental health payment systems, using heterogeneous clustering and resource distribution methodologies. (Mason et al., 2011) Fourth, the majority of literature searching was undertaken by one study author. However, in order to minimise bias and error, 20% of abstracts were independently screened by another author. Fifth, the age variable was reported with heterogeneous age ranges across studies. Thus, conclusions in relation to specific age ranges could not be made. Finally, the study benefits from its thorough reporting process and use of structured checklists for assessments of study quality.

Additional future research directions

Five future research directions not already discussed in relation to *Mental Health Payment by Results* are provided. First, as the operationalisation of HSU in included studies was largely constrained by the use of secondary data from service databases, future HSU studies may benefit from the administration of measures such as the *Client Services Receipt Inventory*,(Beecham & Knapp, 2001) alongside secondary data. Second, an international systematic review of the predictors of HSU by mental health populations could provide a more comprehensive list of predictor variables. Third, the HSU of people with intellectual disabilities were not examined in this review due to the specific additional needs of this population which have to be met beyond the healthcare system. However, it is an important area of research since UK-based studies have highlighted the widespread failure of health services to make required additional accommodations (e.g. extended appointment hours) for this patient group, with no additional funding currently allocated for these purposes to NHS acute trusts.(Tuffrey-Wijne et al., 2014) Determining how the inadequate provision of additional accommodations impacts upon the HSU of people with intellectual disabilities could inform future decisions surrounding allocation of resources. Fourth, the review identified a number of variables (i.e. attending a community outreach service, attending a psychiatric liaison service, unspecified ICD-10 diagnosis, insomnia symptoms, self-harming behaviour) examined in relation to HSU in just one study yet predictive of HSU. Therefore, the associations of these variables with HSU could be explored in future research. Finally, further large-scale case register studies (including participants from shared service catchment areas) would address the study heterogeneity found in this review and provide more robust evidence on the predictors of HSU by people with mental disorders in the UK.

Conclusions

This review provides evidence that can inform decisions about which variables might be used to derive mental health clusters in the *Mental Health Payment by Results* system*.* Several variables- in particular comorbidity, female gender, age (heterogeneous age ranges) high previous HSU, and a marital status of divorced- have good preliminary evidence supporting their ability to predict HSU by adults with mental disorders in the UK, and thus are relevant for clustering purposes. The findings support the need to determine the association of the MHCT (and its domains of behaviour, symptoms, impairment, social functioning and risk factors) with HSU, the need to investigate whether combining broad diagnoses with care pathways is an effective alternative method for mental health clustering, and the need for research to further examine the association between existing mental health clusters and HSU. Overall, this review has highlighted important unresolved issues related to the *Mental Health Payment by Results* system. Addressing these issues could improve how health service resources are distributed, helping to ensure that people experiencing mental health problems can access the most appropriate services at their time of need.

Paper 1 Table 1. Observational studies of HSU (n=17)

Study	**Design**	**Data Source**	**Participants**				**Q**	**HSU outcomes**	**Predictors of increased HSU**	**Not predictive of increased HSU**
			Composition	***N***	**Age**	**%f**	**ST**			
(Button, Benson, Nollett, & Palmer, 2005)	Cohort	NHS eating disorders clinic	Eating disorder patients	147	26.3 (*SD* not stated)	96	9	• Total HSU	-	• Type of eating disorder diagnosis
(Byford, Barrett, Despiégel, & Wade, 2011)	Cohort	NHS primary care database	Depressed patients	88935	44.4 (*SD*= 16.75)	68	18	• A&E attendances • GP phone calls • GP visits • Inpatient days • Medication usage • Other specialist contacts • Psychiatrist contacts • Psychologist contacts	• Non-remission (after antidepressant treatment)	• Remission (after antidepressant treatment)
(Chollet et al., 2013)	Cohort	NHS primary care database	GAD patients	29131	48.5 (*SD*= 17.5)	67	18	• Total HSU	• Aged 31-49 • Aged 50-64 • High previous HSU • High previous medication use • Male • Two comorbidities	• Aged 18-30 • Aged >65 • Lower previous HSU • Lower previous medication use • FemaleNo, one, or three comorbidities
(J. Coid et al., 2009)	Cross-sectional	Adult Psychiatric Morbidity Survey	UK residents (some with BPD)	8397	16-74 (*M* not stated)	53	18	• Community psychiatric nurse contacts • Counsellor contacts • GP contacts for psychological problems • Psychiatric inpatient admission • Psychiatrist contacts • Total HSU	• Diagnosis of BPD	• No diagnosis of BPD

(Jeremy Coid et al., 2006)	Cross-sectional	Adult Psychiatric Morbidity Survey	UK residents with a PD	626	16-74 (*M* not stated)	56	17	• Community psychiatric nurse contacts • Counsellor contacts • GP contacts for psychological problems • Medication usage • Psychiatric inpatient admission • Psychiatrist contacts	• Cluster A, B, and C PD diagnoses • Comorbid mental disorder and substance abuse	• No comorbidity
(C. Cooper et al., 2010)	Cross-sectional	Adult Psychiatric Morbidity Survey	UK residents (some with CMPs)	7461	16+ (*M* not stated)	51	18	• PCT attendance • GP contact for psychological problems • Medication usage	• Aged >35 • ADLs • Widowed / divorced/ separated • Elevated neurotic symptoms • Female • Non-white ethnicity •	• Aged <35 • No ADLs • Marital status other than widowed / divorced/ separated • Non-elevated neurotic symptoms • Male • White ethnicity • Any home ownership status • Number of qualifications
(Claudia Cooper et al., 2013)	Cross-sectional	Adult Psychiatric Morbidity Survey	UK residents (some with CMPs)	22196	16+ (*M* not stated)	52	19	• PCT attendance • GP contact for psychological problems • Medication usage	• Aged 35-54 • Aged 75+ • Divorced/ separated/ widowed • Elevated neurotic symptoms • Female • Non-home owner • Non-white ethnicity •	• Aged 16-34 • Aged 55-74 • Marital status other than widowed / divorced/ separated • Male • Non-elevated neurotic symptoms • Home owner • White ethnicity

(Foster et al., 2003)	Cross-sectional	Government surveys: adults in private households; adults with psychosis in households and adults in mental institutions	Adults with psychosis	470	16-64 (*M* not stated)	NS	10	• GP contact for psychological problems • Psychiatric inpatient admission • Any service use for a psychological problem	• Aged 16-34 • Aged 45-54 • Elevated neurotic symptoms • White ethnicity	• Aged 35-44 • Aged 55-64 • Family circumstances • Gender • Household type • Living arrangements • Marital status • Non-elevated neurotic mental health symptoms • Non-white ethnicity • Occupation • Physical illness • Qualifications
(Hayward et al., 2010)	Cohort	Postal survey of a general practice population	GP attendees	2662	51.3 (*SD*= 17.18)	55	16	• GP contacts • Medication usage	• Insomnia symptoms • Comorbid anxiety or depression	• No insomnia symptoms • No comorbidity
(Keene & Rodriguez, 2007)	Cross-sectional	Databases: health authority, mental health population, and A&E population.	Health and mental health service users	625964	16+ (*M* not stated)	52	16	• A&E attendances	• Four typologies: (1) Young, male frequent attendees with self-harm and other injuries; (2) Young females with self-harm; (3) Older patients with multiple medical conditions; (4) Very old patients with cardiac conditions and fractures	-

(Knapp, McCrone, Fombine, Beecham, & Wostear, 2002)	Cross-sectional	Maudsley (NHS) psychiatric hospital data; patient interviews	Former Maudsley child and adolescent psychiatric patients	149	Not stated	61	11	• Client Services Receipt Inventory (Beecham & Knapp, 2001)	• Comorbidity of childhood depression and conduct disorder	• No childhood comorbidity
(Mohan et al., 2006)	Cohort	PRISM psychosis study set in Maudsley & Bethlem NHS trust area	White (group 1) and African Caribbean (group 2) patients with psychosis	140	40.55 (*SD*= 14.9)	49	18	• Client Services Receipt Inventory (Beecham & Knapp, 2001)	• Receiving intensive community treatment (for African Caribbean patients only)	• Ethnicity • Receiving intensive community treatment (for White patients only)
(Patel et al., 2006)	Cross-sectional data from a RCT	RCT data set in South London /Maudsley NHS trust area	Schizophrenia patients	85	26 (SD not stated)	26	13	• 'Other' • A&E attendances • CMHT contacts • Community psychiatric nurse contacts • Day care attendances • General medical ward attendances • GP contacts • Group PCT attendances • Home carer visits • Inpatient admissions • Inpatient days • Non-psychiatric outpatient attendances • Nurse contacts • Occupational therapist contacts • Psychiatric outpatient attendances • Psychiatrist contacts • Psychologist contacts • Sheltered workshop attendances	• Cognitive deficits	• Anti-social behaviour • Depression symptoms • No cognitive deficits • Positive symptoms • Social withdrawal

								• Specialist education attendances • Total inpatient service use		
(Torres et al., 2007)	Cross-sectional	Adult Psychiatric Morbidity Survey	UK residents with OCD	114	16-74 (*M* not stated)	65	13	• Any community service attendance • Counselling attendance • GP contact for psychological problems • Home carer visits • Medication usage • PCT attendance • Psychiatric inpatient admission • Community psychiatric nurse contacts • Psychiatric outpatient attendances • Psychiatrist contact • Psychologist contact • Support group attendances • Total HSU ('any kind of treatment')	• OCD diagnosis • OCD with co-morbid anxiety or depression	• No OCD diagnosis
(Ullrich & Coid, 2009)	Cross-sectional	Adult Psychiatric Morbidity Survey	UK residents with ASPD	245	16-74 (*M* not stated)	22	16	• Community psychiatric nurse contacts • GP contacts • Other nursing service contacts • Outreach worker contacts • Psychiatric inpatient admission • Psychiatrist contacts • Psychologist contacts • Support group attendances • Total HSU	• Comorbid Axis 1 mental disorders	• Comorbid personality disorders
(Walters, Buszewicz, Weich, & King, 2011)	Cohort	Seven NHS general practices.	Primary care patients with mild-to-moderate distress	250	46 (*SD* not stated)	71	20	• GP contacts	• ICD-10 disorders (apart from mixed anxiety and depression)	• Mixed anxiety and depression

(Wright, Gournay, Glorney, & Thornicroft, 2000)	Cross-sectional	NHS mental health services	Patients with functional psychosis and co-morbid substance abuse.	61	43.1 (*SD* not stated)	56	15	• Inpatient admissions • Inpatient days	-	• Dual diagnosis

Note: A&E= Accident and Emergency; ADLs= Activities of Daily Living restrictions; Age= mean age (if not stated, where possible, age range is stated); ASPD= Anti-social personality disorder; BPD= Borderline Personality Disorder; CMHT= Community Mental Health Team; CMP= common mental health problem; DBT= Dialectical Behaviour Therapy; f=female; GAD= Generalised Anxiety Disorder; GP= General Practitioner; HSU= health service utilisation; *M*= mean; NHS= National Health Service; NS= not stated; OCD= obsessive compulsive disorder; ST= STROBE statement (score range 0-22; 0 represents lowest quality and 22 represents highest quality) (von Elm et al., 2008); RCT= Randomised Controlled Trial; PCT= psychotherapy; PD= personality disorder. Q= Quality assessment

Paper 1 Table 2. Intervention studies of HSU (n=11)

Study	**Design**	**Data Source**	**Participants**				**Quality**		**HSU outcomes**	**Intervention**	**Control**	**Reduced HSU? (*p*<.05)**
			Composition	***N***	**Age**	**%f**	**ST**	**NC**				
(Amner, 2012)	Cohort	NHS service data	BPD patients availing of DBT	21	36.2 (*SD*= 10.87)	81	13	-	• Day care attendances • DBT attendances • Inpatient days • Nurse contacts • Outpatient attendances • PCT attendances • Total HSU	DBT	-	NO
(Ballard et al., 2002)	Quasi-experimental	Care facilities	Dementia patients	224	82.5 (*SD*= 7.1)	75	12	-	• GP contacts • Inpatient days	Psychiatric liaison	Usual care	YES
(Bateman & Fonagy, 2008)	RCT	NHS PD PCT unit	BPD patients	41	31.8 (*SD*= 6.23)	58	-	2	• A&E attendances • Outreach worker contacts • Inpatient days • Medication usage • PCT attendances • Psychiatric treatment days	Mentalisation-based treatment by partial hospitalisation	Usual care	YES
(Commander, Sashidharan, Rana, & Ratnayake, 2005)	Cohort	Assertive outreach service data	Outreach patients with schizophrenia, bipolar disorder or 'other' disorder	250	18-64 (*M* not stated)	26	12	-	• Compulsory admissions • Inpatient admissions • Inpatient days	Community outreach service use	-	YES
(de Lusignan, Chan, Parry, Dent-Brown, &	Cohort	NHS (IAPT and hospital service) data	IAPT attendees	1118	35.3 (SD=21.4)	50	15	-	• A&E attendances • Inpatient admissions • Inpatient days • Medication usage • Outpatient attendances • Sick notes issued	IAPT service	-	YES

Kendrick, 2012)												
(de Lusignan et al., 2013)	Case control	NHS (IAPT and hospital service) data	IAPT attendees with long term health conditions	1341	52.8 (SD= 11.15)	65	19	-	• A&E attendances • Inpatient admissions • Inpatient days • Medication usage • Outpatient attendances • Sick notes issued	IAPT service	-	YES
(Hayhurst, Brown, & Lewis, 2002)	Cohort	NHS University Hospitals service data	Patients receiving antipsychotic medication	126	42.55 (SD= 12.29)	35	16	-	• Inpatient admissions • Inpatient days	Clozapine	-	YES
(D. H. Lam, McCrone, Wright, & Kerr, 2005)	RCT	Maudsley & Bethlem NHS trust area	Patients with bipolar disorder	87	43.95 (*SD*= 11.45)	33	-	4	• Any community services attendance • Medication usage • Non-psychiatric inpatient days • Psychiatric inpatient days • Total HSU	Cognitive therapy (added to usual care)	Usual care	NO
(Shi et al., 2012)	Cohort	NHS primary care database	Depressed adults initiating duloxetine	909	49.6 (*SD*= 16.5)	67	17	-	• A&E attendances • Inpatient admissions • Medication usage • Referrals to specialists	Duloxetine	-	YES
(Wade et al., 2010)	Cohort	NHS primary care database	Depressed adults using escitalopram, venlafaxine, or generic SSRI	2485	43.1 (*SD*= 14.7)	60	18	-	• GP phone calls • GP visits • Inpatient admissions • Medication usage • Referrals to other specialist • Referrals to psychiatrist • Total HSU	Escitalopram	Generic SSRIs; venla-faxine	YES
(Woods et al., 2012)	RCT	Community-based RCT	Patients of NHS Memory Clinics and CMHTs	488	77.1 (*SD*= 7.3)	50	-	4	• A&E attendances • Care attendant contacts • Care manager contacts • Chiropodist contacts • CMHT contacts • Community psychiatrist contacts • Continuing care/respite contacts • Counsellor contacts • Day hospital contacts	Reminiscence group	Usual care	NO

- Dietician contacts
- Family support worker contacts
- GP contacts
- Health visitor contacts
- Home carer visits
- Informal/ voluntary care contacts
- Inpatient rehabilitation contacts
- General medical ward attendances
- NHS contacts
- Occupational therapist contacts
- Other inpatient ward contacts
- Outpatient attendances
- Physiotherapist contacts
- Nurse contacts
- Psychologist contacts
- Sitting scheme worker contacts

Note: A&E= Accident and Emergency; Age= mean age (if not stated, where possible, age range is stated); BPD= Borderline Personality Disorder; CMHT= community mental health team; DBT= Dialectical Behaviour Therapy; f=female; GP = General Practitioner; HSU= health service utilisation; IAPT= *Improving Access to Psychological Therapies* initiative; NC= NICE checklist for RCTs (score range = 0-4; 0 indicates bias and 4 indicates no bias) (NICE, 2009); NHS= National Health Service; ST= STROBE statement (score range 0-22; 0 represents lowest quality and 22 represents highest quality) (von Elm et al., 2008); PCT= psychotherapy; PD= personality disorder

Paper 1 Table 3. Frequency of HSU outcomes used across included studies (n=28).

Primary care HSU	***n***
Mental health	
GP contact(s) for psychological problems	6
Referrals to psychiatrist	1
Support group attendances	2
Total	9
*General health**	
GP contacts	7
GP phone calls	2
GP visits	2
Referrals to specialists	2
Sick notes issued	2
Total	15

Specialist HSU	***n***
Mental health	
Medication usage**	12
Psychotherapy attendance(s)	6
Community psychiatric nurse contacts	5
Psychiatrist contact(s)	5
Psychologist contact(s)	5
Outpatient attendances	4
Counsellor contacts	3
Day care attendances	2
CMHT contacts	2
DBT attendances	1
Psychiatric outpatient attendances	1
Psychologist / psychiatrist contacts	1
Total	47
*General health**	
Nurse contacts	4
Home carer visits	3
Any community service attendance	2
General medical ward attendances	2
Occupational therapist contacts	2
Outreach worker contacts	2
Care attendant contacts	1
Care manager contacts	1
Chiropodist contacts	1
Continuing care / respite contacts	1
Counselling attendance	1
Day hospital contacts	1
Dietician contacts	1
Family support worker contacts	1
Health visitor contacts	1
Informal/voluntary care contacts	1
Inpatient rehabilitation contacts	1
NHS contacts	1
Non-psychiatric outpatient attendances	1
Other inpatient ward contacts	1
Other nursing service contacts	1
Other specialist contacts	1
Physiotherapist contacts	1
Total	32

Inpatient HSU	***n***
Mental health	
Psychiatric inpatient admission(s)	5
Psychiatric inpatient days	1
Psychiatric outpatient attendances	1
Compulsory admissions	1
Psychiatric treatment days	1
Total	9
*General health**	
Inpatient days	9
A&E attendances	8
Inpatient admissions	8
Non-psychiatric inpatient days	2
Sheltered workshop attendances	1
Sitting scheme worker contacts	1
Specialist education attendances	1
Total inpatient service use	1
Total	31

Total and other HSU	***n***
Total HSU	8
‘Other’ HSU	1
Any service use for psychological problem	1
Client Services Receipt Inventory	2
Total	12

Summary totals	***n***
Primary Care HSU	24
Specialist HSU	79
Inpatient HSU	40
Total and other HSU	12
Mental health HSU	65
General health HSU	78
Types of outcome variables	60
Times outcomes assessed	155

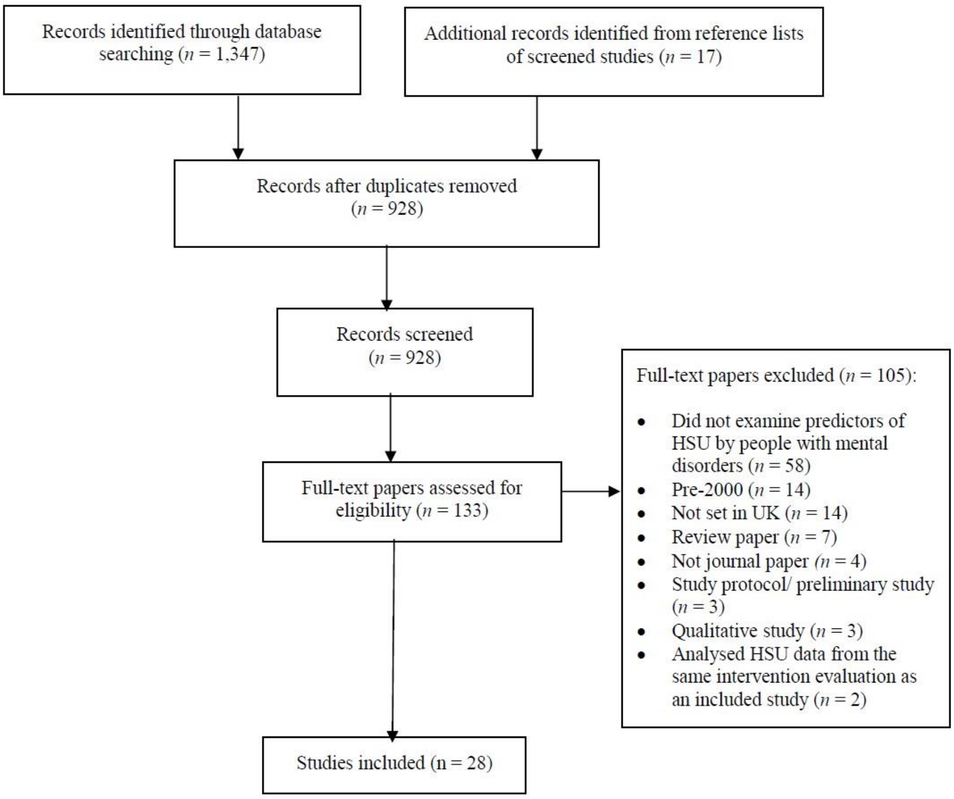
Note: * General health refers to HSU that was not specified as being directly linked to mental ill health. **Type of medication varied widely. A&E= Accident and Emergency; CMHT= community mental health team; Client Services Receipt Inventory (Beecham & Knapp, 2001); DBT= Dialectical Behaviour Therapy; HSU= health service utilisation. NHS= National Health Service.

Paper 1 Table 4. Summary of the evidence for examined predictors of HSU.

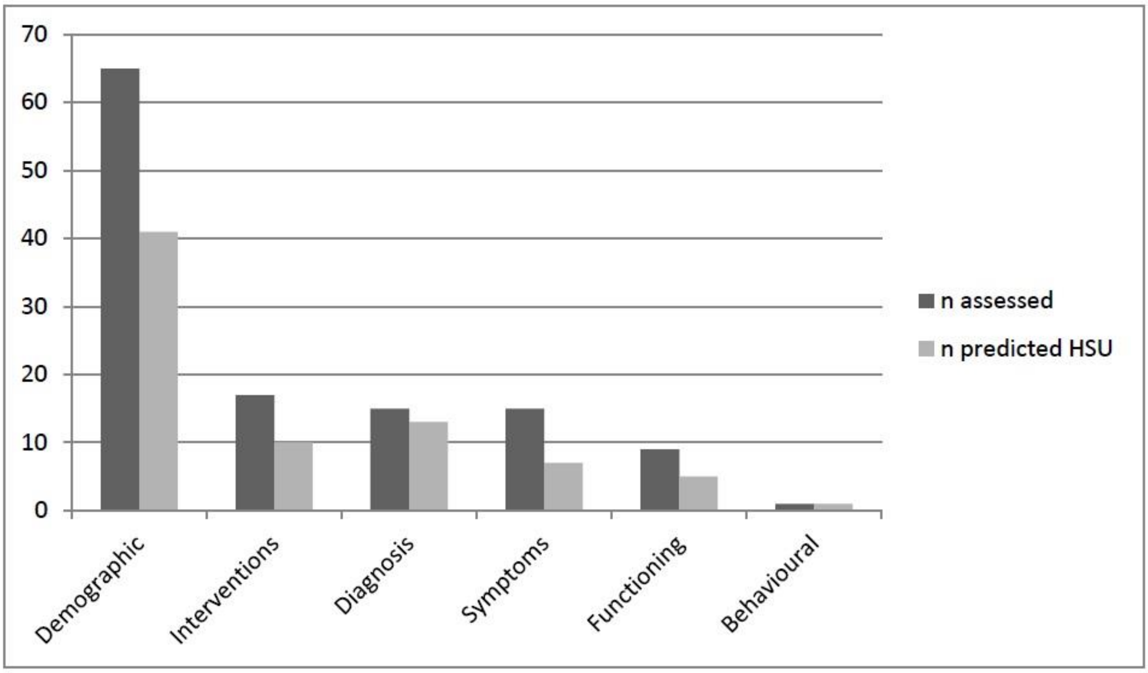
Predictor variables	***n* variable assessed**		**Assessments in which variable predicted HSU**													**Assessments in which variable did not predict HSU**												
		n	**HSU outcomes predicted (*n*)***												***n***	**HSU outcomes not predicted (*n*)***											
			Prim. Care			**Specialist**			**Inpatient**			**Total HSU**				**Prim. Care**			**Specialist**			**Inpatient**			**Total HSU**		
			n	**Quality**		***n***	**Quality**		***n***	**Quality**		***n***	**Quality**			***n***	**Quality**		***n***	**Quality**		***n***	**Quality**		***n***	**Quality**	
				+	**-**		**+**	**-**		**+**	**-**		**+**	**-**			**+**	**-**		**+**	**-**		**+**	**-**		**+**	**-**
Demographic (Chollet et al., 2013; J. Coid et al., 2009; Jeremy Coid et al., 2006; C. Cooper et al., 2010; Claudia Cooper et al., 2013; Foster et al., 2003; Hayward et al., 2010; Keene & Rodriguez, 2007; Knapp et al., 2002; Mohan et al., 2006; Torres et al., 2007; Ullrich & Coid, 2009; Wright et al., 2000)																											
Comorbidity**	15	14	*4*	**3**	1	*3*	**2**	1	*3*	**2**	1	*4*	**2**	2	1	*0*			*0*			*1*	**0**	1	*0*		
Age***	7	7	*2*	**2**	0	*2*	**2**	0	*2*	**1**	1	*1*	**1**	0	0												
Female gender	7	5	*2*	**2**	0	*2*	**2**	0	*1*	**1**	0	*0*			2	*0*			*0*			*1*	**0**	1	*1*	**1**	0
Male gender	7	2	*0*			*0*			*1*	**1**	0	*1*	**1**	0	5	*2*	**2**	0	*2*	**2**	0	*1*	**0**	1	*0*		
NW ethnicity	6	4	*2*	**2**	0	*2*	**2**	0	*0*			*0*			2	*0*			*0*			*0*			*2*	**1**	1
White ethnicity	6	1	*0*			*0*			*0*			*1*	**0**	1	5	*2*	**2**	0	*2*	**2**	0	*0*			*1*	**1**	0
Marital status****	5	4	*2*	**2**	0	*2*	**2**	0	*0*			*0*			1	*0*			*0*			*0*			*1*	**0**	1
Non-home owner	5	2	*1*	**1**	0	*1*	**1**	0	*0*			*0*			3	*1*	**1**	0	*1*	**1**	0	*0*			*1*	**0**	1
Qualifications	3	0	*0*			*0*			*0*			*0*			3	*1*	**1**	0	*1*	**1**	0	*0*			*1*	**0**	1
High prev. HSU	2	2	*0*			*0*			*0*			*2*	**2**	0	0	*0*			*0*			*0*			*0*		
Family situation	1	0	*0*			*0*			*0*			*0*			1	*0*			*0*			*1*	**0**	1	*0*		
Occupation	1	0	*0*			*0*			*0*			*0*			1	*0*			*0*			*1*	**0**	1	*0*		
Total	65	41	*13*	**12**	1	*12*	**11**	1	*7*	**5**	2	*9*	**6**	3	24	*6*	**6**	0	*6*	**6**	0	*5*	**0**	5	*7*	**3**	4
Intervention ***** (Amner, 2012; Ballard et al., 2002; Bateman & Fonagy, 2008; Commander et al., 2005; de Lusignan et al., 2012; de Lusignan et al., 2013; Hayhurst et al., 2002; D. H. Lam et al., 2005; Shi et al., 2012; Wade et al., 2010; Woods et al., 2012)																											
IAPT service	4	4	*1*	**1**	0	*1*	**1**	0	*2*	**1**	1	*0*			0	*0*			*0*			*0*			*0*		
Psychotherapy	4	1	*1*	**1**	0	*0*			*0*			*0*			3	*0*			*1*	**1**	0	*1*	**1**	0	*1*	**1**	0
Reminiscence grp.	4	0	*0*			*0*			*0*			*0*			4	*1*	**1**	0	*1*	**1**	0	*1*	**1**	0	*1*	**1**	0
Medication	3	3	*1*	**1**	0	*0*			*2*	**2**	0	*0*			0	*0*			*0*			*0*			*0*		
Comm. outreach	1	1	*0*			*0*			*1*	**1**	0	*0*			0	*0*			*0*			*0*			*0*		
Psychiatric liaison	1	1	*1*	**0**	1	*0*			*0*			*0*			0	*0*			*0*			*0*			*0*		

Total	17	10	*4*	**3**	1	*1*	**1**	0	*5*	**4**	1	*0*			7	*1*	**1**	0	*2*	**2**	0	*2*	**2**	0	*2*	**2**	0
Diagnostic (Button et al., 2005; J. Coid et al., 2009; Jeremy Coid et al., 2006; Torres et al., 2007; Walters et al., 2011)																											
PD	8	8	*2*	**2**	0	*2*	**2**	0	*2*	**2**	0	*2*	**2**	0	0	*0*			*0*			*0*			*0*		
OCD	4	4	*1*	**0**	1	*1*	**0**	1	*1*	**0**	1	*1*	**0**	1	0	*0*			*0*			*0*			*0*		
Unspec. ICD-10	1	1	*1*	**1**	0	*0*			*0*			*0*			0	*0*			*0*			*0*			*0*		
Eating disorder	1	0	*0*			*0*			*0*			*0*			1	*0*			*0*			*0*			*1*	**0**	1
MADD	1	0	*0*			*0*			*0*			*0*			1	*1*	**1**	0	*0*			*0*			*0*		
Total	15	13	*4*	**3**	1	*3*	**2**	1	*3*	**2**	1	*3*	**2**	1	2	*1*	**1**	0	*0*			*0*			*1*	**0**	1
Symptoms (C. Cooper et al., 2010; Claudia Cooper et al., 2013; Foster et al., 2003; Hayward et al., 2010; Patel et al., 2006)																											
Neurotic	6	6	*3*	**2**	1	*2*	**2**	0	*0*			*1*	**0**	1	0	*0*			*0*			*0*			*0*		
Depression	4	0	*0*			*0*			*0*			*0*			4	*1*	0	1	*1*	**0**	1	*1*	**0**	1	*1*	**0**	1
Positive******	4	0	*0*			*0*			*0*			*0*			4	*1*	0	1	*1*	**0**	1	*1*	**0**	1	*1*	**0**	1
Insomnia	1	1	*1*	**1**	0	*0*			*0*			*0*			0	*0*			*0*			*0*			*0*		
Total	15	7	*4*	**3**	1	*2*	**2**	0	*0*			*1*	**0**	1	8	*2*	0	2	*2*	**0**	2	*2*	**0**	2	*2*	**0**	2
Functioning (C. Cooper et al., 2010; Patel et al., 2006)																											
Social withdrawal	4	0	*0*			*0*			*0*			*0*			4	*1*	0	1	*1*	**0**	1	*1*	**0**	1	*1*	**0**	1
Cognitive deficits	3	3	*1*	**0**	1	*1*	**0**	1	*1*	**0**	1	*0*			0	*0*			*0*			*0*			*0*		
ADLs	2	2	*1*	**1**	0	*1*	**1**	0	*0*			*0*			0	*0*			*0*			*0*			*0*		
Total	9	5	*2*	**1**	1	*2*	**1**	1	*1*	**0**	1	*0*			4	*1*	0	1	*1*	**0**	1	*1*	**0**	1	*1*	**0**	1
Behavioural (Keene & Rodriguez, 2007)																											
Self-harm	1	1	*0*			*0*			*1*	**1**	0	*0*			0	*0*			*0*			*0*			*0*		

Note: *Most studies examined more than one health service utilisation outcome measure. **Both mental and physical comorbidity. ***Various heterogeneous age ranges predicted increased HSU in individual studies: 16-34; 31-49; 45-54; 50-64; >35; 35-54; 75+. **** Only divorced/ separated/ widowed marital statuses were predictive of increased HSU. *****If an intervention reduced HSU, it was counted as predicting HSU, and vice versa. ******Positive = positive symptoms associated with schizophrenia. ADLs= Activities of Daily Living ; Comm.= community; grp= group; HSU= health service utilisation; IAPT= *Improving Access to Psychological Therapies* initiative; ICD-10= International Classification of Diseases-10; MADD= mixed anxiety and depressive disorder; NW= non-white; PD=Personality Disorder; prev.= previous; Prim.= Primary; Unspec.= Unspecified; $^+$ = A score of $\geq$16 on STROBE statement,(von Elm et al., 2008) or $\geq$3 on NICE checklist for RCTs;(NICE, 2009) $^-$ = A score of $\leq$15 on STROBE checklist, or $\leq$2 on NICE RCT checklist).



Paper 1 Figure 1. Literature search flow



Paper 1 Figure 2. Frequency of HSU prediction by variable category

INTRODUCTION TO PAPER 2

Addressing the second specific objective of the thesis - to determine the association of functioning with HSU by people with common mental disorders - Paper 2 encompasses a cross-sectional study examining the associations of depressive symptom severity and functioning with HSU by older people (aged 65+) in low-and-middle income countries (LMICs). The decision to investigate this population was made for three main reasons: (1) older people consume substantially more health service resources than younger people, and their 'heavy', HSU has important implications for health service planning; (2) relatively little HSU research has been undertaken in LMICs so the findings may be impactful and influential in terms of policy development, especially since there is typically no universal healthcare coverage in these countries; (3) the large and comprehensive 10/66 Dementia Research Group population-based surveys dataset was available for analysis, ensuring the feasibility of the study. A consequence of investigating this population is a reduction in the generalisability of the findings to younger adults, and people residing in high income countries (such as the UK, the setting of the other studies in this thesis).

Functioning was measured using the 12-item WHODAS-II (Rehm et al., 1999) which was developed to be consistent with the ICF classification system, albeit without directly addressing contextual factors, and focusing mainly on the domains of activities and participation. The psychometric properties of the WHODAS-II are sound: internal consistency and test-retest reliability scores are high and its concurrent validity (with other measures of functioning and disability) is established in a wide range of clinical and general population samples (Andrews, Kemp, Sunderland, Von Korff, & Ustun, 2009). Of particular relevance to Paper 2 given its sample composition, the WHODAS-II has been previously evaluated in LMICs; here, findings from a Mokken analysis demonstrated that the WHODAS II is a unidimensional and hierarchical scale conforming with item response theory (IRT) principles (Sousa et al., 2010).

Paper 2 was undertaken in King's College London (KCL) where I spent three months on secondment from October 2014 to December 2014. The study findings have been published in the *International Journal of Environmental Research and Public Health* (C. D. Twomey, Prince, Cieza, Baldwin, & Prina, 2015). The PhD researcher was the lead author of the paper. Under supervision, the PhD researcher contribution included conceiving and planning the paper, undertaking statistical analysis (guided by a co-author), drafting all versions of the paper, and managing the submission and peer-review process.

PAPER 2: Cross-Sectional Associations of Depressive Symptom Severity and Functioning with Health Service Use by Older People in Low-and-Middle Income Countries

Abstract

Background
Comprehensive understanding of the determinants of health service use (HSU) by older people with depression is essential for health service planning for an ageing global population. This study aimed to determine the extent to which depressive symptom severity and functioning are associated with HSU by older people with depression in low and middle income countries.

Methods
A cross-sectional analysis of the 10/66 Dementia Research Group population-based surveys dataset. Participants (*n* = 4590) were those aged 65 or older, in the clinical range for depressive symptoms (defined as scoring four or more on the EURO-D), living in 13 urban and/or rural catchment areas in nine LMICs. Associations were calculated using Poisson regression and random-effects meta-analysis.

Results
After adjustment for confounding variables, (EURO-D) depressive symptom severity was significantly associated with "any community HSU" (Pooled Prevalence Ratios = 1.02; 95% CI = 1.01–1.03) but not hospital admission. Conversely, after adjustment, (WHODAS-II) functioning was significantly associated with hospital admission (Pooled PR = 1.14; 95% CI = 1.02–1.26) but not "any community HSU".

Conclusions
Depressive symptom severity does not explain a large proportion of the variance in HSU by older people with depression in LMICs. The association of functioning with this HSU is worthy of further investigation. In LMICs, variables related to accessibility may be more important correlates of HSU than variables directly related to health problems.

Introduction

Older people- here defined as people aged 65 years or older- consume substantially more health service resources than younger people. *Per capita*, health service use (HSU) by older people is 3–5 times higher than in younger people (Casey et al., 2003). In England, older people account for 45% of hospital admissions and nearly two thirds of hospital bed-days (Health, 2011; Wanless, 2006). Although better health patterns in old age should be taken into account (Lutz, Sanderson, & Scherbov, 2008), the ongoing demographic shift towards an older global population will increasingly place greater demands on health services (Prina et al., 2012). Therefore, a comprehensive understanding of the variables associated with HSU by older people is essential for effective health service planning over the long term.

One of the most common and burdensome health problems experienced by older people is depression. Up to 25% of older people suffer from major depression, dysthymic disorder, or subthreshold depression ((APA), 2010). Due to the ageing population, it has been projected that between the year 2000 and 2050, the number of older people with depression will increase by 117% (Heo, Murphy, Fontaine, Bruce, & Alexopoulos, 2008). Depression in older people is highly burdensome, and associated with poor quality of life, disability, and increased likelihood of physical health comorbidities (H. Huang et al., 2014). Moreover, depression has various biological effects (e.g., alteration of cardiac functioning, inflammation, impairments in cell-mediated immunity) and is associated with risk factors for chronic disease (e.g., smoking, poor diet, hypertension) (Prince, Patel, et al., 2007).

Given that the commonality and burden of depression in older people has important implications for health service planning, there has been a relative lack of research investigating the association of depressive symptom severity with HSU in older people. Moreover, findings from the few existing general population studies examining this association are somewhat inconsistent. Across HSU outcomes (e.g., community-based HSU, hospital admissions), most studies provide evidence for this association (Feng et al., 2009; Prina et al., 2015; Prina et al., 2013) but some do not (Beekman et al., 2002; Fischer et al., 2002; B. Y. Huang et al., 2000). Although a recent meta-analysis of seven studies found an association of depressive symptom severity with hospital admissions, the magnitude of this association was small (RR = 1.36) (Prina et al., 2015).

Previous research suggests that depressive symptom severity may not explain a large amount of the variance in HSU by older people with depression, and there is a need to determine the extent to which other variables may explain the remaining variance. One variable worthy of investigation in this regard is "functioning", an encompassing term relating to body functions, body structures and activities and participation (Escorpizo et al., 2013). Functioning is relevant because it reflects the burden of all health problems (which tends to increase with age), meaning that it is potentially associated with utilisation of a diverse range of health services. Various studies have demonstrated positive associations of functioning impairment (operationalised in differing ways) with HSU by clinical and general populations (C. Cooper et al., 2010; Covinsky et al., 1997; Patel et al., 2006) and depressive symptom severity in older people (Covinsky et al., 2010; Lenze et al., 2001; Schillerstrom, Royall, & Palmer, 2008). However, although one recent study included analysis of the association of physical impairments with HSU by older people (Albanese et al., 2011), few studies have so far directly examined the association of functioning with HSU by older people with depression (Prina et al., 2015).The majority of previous research has taken place in high income countries and there is a need to increase knowledge of the variables associated with HSU by older people with depression in low and middle income countries (LMICs), where momentum for health care reforms aimed at achieving universal health care coverage is growing (Mate et al., 2013). Available findings are mostly based on data from the general population (e.g., community-dwelling older people) rather than participants with probable depression. A sole focus on participants with probable depression would facilitate the examination of HSU patterns specifically related to depression, adding precision to subsequent findings. The present study encompasses a cross-sectional analysis of the 10/66 Dementia Research Group population-based surveys dataset (Prince, Ferri, et al., 2007) and addresses the three research gaps mentioned above. Our main aim was to determine the extent to which depressive symptom severity and functioning are associated with HSU by older people in LMICs. To explore further the (unclear) impacts of functioning on this HSU, we also aimed to determine if specific functioning difficulties are more related to this HSU than others.

Methods

Participants and Procedures

Prince *et al.* (Prince, Ferri, et al., 2007) have provided the full protocol of the 10/66 Dementia Research Group population-based surveys. Undertaken between 2003 and 2005, the surveys provide data from a sample of people aged 65 years and over living in 13 urban and/or rural catchment areas in nine LMICs (China, Cuba, Dominican Republic, India, Mexico, Nigeria, Peru, Puerto Rico, Venezuela) (Prince, Ferri, et al., 2007). Urban sites were of high density and low socioeconomic status, and rural sites were of low density, representing a traditional agrarian lifestyle. Selected sites had mapped boundaries, and systematic door-knocking identified eligible participants (*i.e.*, those aged 65 and over). Surveys consisted of questionnaires, structured clinical interviews, an informant interview, and a physical examination. Data collected pertained to demographics, physical health, mental health (e.g., depression, dementia), chronic disease risk factors, functioning impairment, HSU, care arrangements and caregiver strain. All studies were approved locally and also by the Institute of Psychiatry, King's College London. Data from over 17,000 participants are provided through the surveys. However, in the present study, only participants with elevated symptoms of depression (*n* = 4950) were included: participants had to score 4 or more on the EURO-D scale (Prince et al., 1999), described below (section 2.2.2).

Measures

Sociodemographics

Age, gender, education (stratified by education levels; e.g., "no education", "primary education", "secondary education"), marital status, household assets owned, pension coverage, private health insurance coverage, number of physical comorbidities (grouped as "none", "one to two" or "three and above"), "any diagnosis of dementia", and "previous ICD-10 depressive episode" were used to assess the sociodemographic background of participants.

Depression

Depressive symptoms were measured using the 12-item EURO-D scale (Prince et al., 1999), derived from the Geriatric Mental State examination (Copeland et al., 2002). The EURO-D addresses the symptoms of depressed mood, pessimism, suicidality, guilt, sleep, interest, irritability, appetite, fatigue, concentration, enjoyment, and tearfulness. Individual items are scored according to absence (0) or presence (1) of these symptoms and the total score ranges from 0–12. The EURO-D has been validated cross-culturally, in both high income countries and LMICs, with a score of four or more representing the presence of probable depression (Castro-Costa et al., 2007; Guerra et al., 2009; Prince et al., 1999).

Functioning

Functioning was measured using the 12-item WHODAS-II which addresses difficulties with: standing, taking care of household responsibilities, learning a new task, joining in community activities, emotionality, concentration, walking a long distance, washing, dressing and undressing, relating to unknown people, maintaining friendships and the performing of day-to-day work/school (Rehm et al., 1999). Individual items are scored on a five-point scale (1 = none; 2 = mild; 3 = moderate; 4 = severe, 5 = extreme/cannot do) and the total score is expressed as a percentage representing the degree of functional impairment experienced. This widely-used, reliable, and valid measure has been psychometrically evaluated in the 10/66 Dementia Research Group study population, generating a one-factor solution in most sites (Sousa et al., 2010).

HSU

HSU was measured using the Client Services Receipt Inventory (Beecham & Knapp, 2001), in a version adapted for usein LMICs (Chisholm, 2000). Participants were asked to recall whether or not they had a contact with any community health services (*i.e.*, primary care, hospital doctor/outpatient, private doctor, other community services, traditional healer) or if they had been admitted to hospital, in the previous three months.

Selection of Potential Confounding Variables

Based on previous research showing their associations with HSU by older people in LMICs (Albanese et al., 2011), potential confounding variables (*i.e.*, age, gender, education level, physical comorbidities, private insurance coverage, pension coverage and "any dementia diagnosis") were included in initial model-building for adjustment purposes. However, to ensure statistical power and precision, the decision was made to remove variables from the model. Thus, after sensitivity analysis showed that pension coverage and "any dementia diagnosis" did not explain meaningful levels of additional variance in outcomes, these variables were removed. It was not possible to include private insurance coverage because there were too few cases (or analysable participants) available to estimate parameters at various sites.

Statistical Analysis

All analyses were conducted using STATA 13 StataCorp LP, College Station, TX, USA),and release 3.3 of the 10/66 dataset. For each site, descriptive statistics summarised sociodemographics and previous HSU (which was also standardised for age, gender, and education, using the whole sample from all sites as the external standard population). Pooled, across-site, HSU descriptive data were also calculated through random-effects meta-analysis (Higgins & Green, 2011). For each site, the associations of depressive symptom severity, functioning and functioning items with HSU were determined using unadjusted and adjusted (for potential confounding variables) prevalence ratios (and 95% confidence intervals) modelled through Poisson regression analysis which controlled for household clustering. References in comparisons were a one-point increase on the EURO-D or a 10-point increase on the WHODAS-II, where appropriate. Poisson regression was chosen as it is appropriate for analysis of dichotomous count data and facilitates the calculation of prevalence ratios which are relatively easy to interpret. Pooled, across-site, unadjusted and adjusted prevalence ratios were then calculated through random-effects meta-analysis. Using random-effects meta-analysis allowed heterogeneity at site level to be accounted for in overall estimates. For meta-analyses, was used to determine the percentage of variance in study estimates attributable to heterogeneity.

Results

Sociodemographic and Clinical Characteristics of Participants

Table 1 summarises participant sociodemographic and clinical information for each study site. In total, 4590 participants scored 4 or more on the EURO-D scale (Prince et al., 1999) and were included in analyses. The smallest sample size was in "China (urban)" (*n* = 15) and the largest sample size was in the Dominican Republic (*n* = 757). In terms of age, 29.3% of participants were aged 65 to 69, 26.6% were aged 70–74, 20.5% were aged 75–79, and 23.3% were aged 80 or over. Most participants were female (71.5%). The majority of participants had received some education (73.5%) but relatively few had completed secondary education or above (17.6%). In terms of marital status, the majority were either married or widowed (81.2%). 65.9% of participants had five or more household assets (*i.e.*, car, television, refrigerator, telephone, mains water, mains electricity, plumbed toilet) and 45.7% had a pension, but 69.8% had no private insurance. Most participants had at least one physical comorbidity (70.6%). A minority had a dementia diagnosis (13.4%). Although all participants were in the clinical range for depressive symptoms (as per inclusion criteria), only a minority had a previous ICD-10 depressive episode (20.3%).

Previous HSU

Table 2 summarises previous HSU by participants. Across sites, the prevalence for contact with any community health service in the previous three months was 57.8% (95% CI = 46.6–69.0). After direct standardisation for age, gender and education (using the whole sample from all sites as the external standard population), the prevalence of HSU was 50.2% (95% CI = 36.6–63.7). Across sites, the crude prevalence for hospital admission in the previous three months was 3.6% (95% CI = 2.5–4.6) and the standardised prevalence was 3.4% (95% CI = 1.9–4.7).

Associations of Depressive Symptom Severity and Functioning with HSU

Table 3 reports the associations of depressive symptom severity (EURO-D) (Prince et al., 1999) and functioning (WHODAS-II) (Rehm et al., 1999) with HSU. The site "China (rural)" was excluded from these analyses because there were too few cases available to estimate parameters. For the "any community HSU" outcome, significant associations were found for both depressive symptom severity (Pooled PR = 1.02;95% CI = 1.01–1.03; I^2 = 14.0%) and functioning (Pooled PR = 1.02; 95% CI = 1.01–1.04; I^2 = 71.2%). After adjustment for age, gender, education, and physical comorbidities, the association remained significant for depressive symptom severity (Pooled PR = 1.02; 95% CI = 1.01–1.03; I^2 = 20.8%) but not functioning (Pooled PR = 1.01; 95% CI = 1.00–1.03; I^2 = 44.7%).

For the hospital admission outcome, no significant association was found for depressive symptom severity (Pooled PR = 1.11; 95% CI = 0.96–1.25; I^2 = 70.2%) but a significant association was found for functioning (Pooled PR = 1.18; 95% CI = 1.07–1.29; I^2 = 64.9%). After adjustment for age, gender, education, and physical comorbidities, similar results were found (Table 3).

To assess the validity of using the (EURO-D) (Prince et al., 1999) for measuring the association of depressive symptom severity with HSU, sensitivity analyses of this association were performed. Here depressive symptom severity was measured through "previous ICD-10 depressive episode". Similar to the previous analyses, depressive symptom severity ("previous ICD-10 depressive episode") was significantly associated with "any community HSU" (Pooled PR = 1.05; 95% CI = 1.01–1.10; I^2 = 44.5%) but not hospital admissions (Pooled PR = 1.37; 95% CI = 0.73–2.02; I^2 = 0%).

3.4. Associations of Specific Functioning Difficulties with "Any Community HSU"

Table 4 reports mutually-adjusted associations of the 12 functioning items (WHODAS-II) (Rehm et al., 1999) with the "any community HSU" outcome. The site "China (rural)" and the hospital admission outcome were excluded from analyses because there were too few cases available to estimate parameters. Only one item, "emotionally affected by health problems", was significantly associated with increased "any community HSU" (Pooled PR = 1.06; 95% CI = 1.03–1.09; I^2 = 31.7%). After adjustment for age, gender, education, and physical comorbidities, this association remained significant (Pooled PR = 1.04; 95% CI = 1.01–1.07; I^2 = 22.4%). The item "difficulty with washing whole body" was significantly associated with decreased "any community HSU", to a similar level before and after adjustment (Pooled PR = 0.93; 95% CI = 0.88–0.98; I^2 = 31.7%).

To account for the potential impact of depressive symptoms on yielded mutually-adjusted associations of specific functioning difficulties with "any community HSU", sensitivity analyses were conducted. Here WHODAS-II items relating to depressive symptoms (*i.e.*, item 5: emotionality; item 6: concentration) were removed from the adjusted models. These analyses found no additional significant associations of specific functioning difficulties with increased "any community HSU". However, "difficulty with washing whole body" remained significantly associated with decreased "any community HSU" (Pooled PR = 0.93; 95% CI = 0.89–0.98; I^2 = 37.0%).

Discussion

Summary of Findings

After adjustment for confounding variables, depressive symptom severity was significantly associated with "any community HSU" but not hospital admission. After adjustment, functioning was significantly associated with hospital admission but not "any community HSU". When investigating individual functioning difficulties, only "emotionally affected by health problems" was significantly associated with increased "any community HSU", whereas "difficulty with washing whole body" was associated with decreased "any community HSU".

Limitations and Strengths

Some methodological limitations should be considered when interpreting the findings.

The cross-sectional nature of the study meant it was not possible to determine if depressive symptom severity and functioning predict HSU over time. Relatedly, depression symptom severity and functioning data pertained to a retrospective one-month period whereas HSU data pertained to a retrospective three-month period. Therefore, yielded associations did not take into account potential fluctuations in depression symptom severity and functioning over the entire three-month investigative period. The WHODAS-II was used to measure functioning but due to the wide-ranging operationalisation of functioning in the literature, it is possible that other functioning measures (e.g., "activities of daily living") would have yielded markedly differing associations with HSU if used in the present study. Because hospital-based services are analogous to community-based services in various study sites, the outcome of "any community HSU" included services not typically set in the community in high-income countries: the findings relating to the hospital admissions outcome are more applicable to high income countries than those relating to the "any community HSU" outcome. Due to an insufficient number of available cases to estimate parameters, it was not possible to determine the association of functioning items from the WHODAS-II with hospital admission.

Finally, heterogeneity at site level varied in analyses of study outcomes and for some outcomes it was at a high level (Tables 3 and 4), perhaps reflecting differences in access to care. This heterogeneity limits somewhat the validity of findings from conducted meta-analyses. However, regression models adjusted for variables associated with access to care in LMICs such as education and pension coverage (Albanese et al., 2011), and similar access-related variables such as "number of assets" and private insurance were reported for each site.

The study benefits from its large sample size, representative of populations of older people living in urban and rural catchment areas in nine LMICs, with minimal missing data. The inclusion of only those participants in the clinical range for depressive symptoms adds precision to yielded associations, as compared with previous studies using general population samples. Moreover, the study is one of few investigations of the association of mental health and HSU, in LMICs. Therefore, the findings could be influential in a region in which momentum for health care reforms aimed at achieving universal health coverage is growing (Mate et al., 2013). Finally, sensitivity analyses of the association of depressive symptom severity (using "previous ICD-10 depressive episode") with HSU yielded similar results to analyses of this association involving the EURO-D (Prince et al., 1999). This adds validity to results concerning (EURO-D) depressive symptom severity and HSU.

Comparison of Findings with Previous Research

The mixed evidence for an association of depressive symptom severity with HSU by older people with depression reflects the inconsistent evidence from previous research (Prina et al., 2015). The significant association of depressive symptom severity with "any community HSU" is in line with some community-based general population studies (Feng et al., 2009; Fischer et al., 2002) but not others (Beekman et al., 2002). The absence of a similar association with hospital admissions contradicts results from a study involving older men with elevated depressive symptoms living in Australia (Prina et al., 2013) but corroborates results from a study involving older people registered with a health maintenance organisation in USA (Fischer et al., 2002). The significant association of depressive symptom severity with "any community HSU" yielded a small effect size (PR = 1.02), similar to the findings of a previous meta-analysis of the association of these symptoms with hospital admissions which also yielded a small effect size (RR = 1.36) (Prina et al., 2015). Overall, the findings suggest that depressive symptom severity does not explain a large amount of the variance in HSU by older people with depression, which highlights the need to investigate the influence of other variables on this HSU.

In relation to functioning, no previous research on the association of functioning with HSU by older people with depression was available for direct comparison of results. The absence of an association of functioning with "any community HSU" contradicts general population studies which found associations of physical impairments with community-based HSU by older people living in LMICs (Albanese et al., 2011), and "activities of daily living" with community-based HSU by adults of varying age in the general population (C. Cooper et al., 2010). These contrasting findings may be partly explained by the differences in the operationalisation of functioning across these studies. The finding of a significant association of functioning with hospital admissions in the present study accords with previous results showing that functioning (operationalised as "activities of daily living") is associated with hospital-based service use by medical inpatients (C. Cooper et al., 2010). However, the validity of this comparison is limited due to differing sample compositions and differing operationalisations of functioning. Overall, results from the present study and those from previous research provide preliminary evidence for an association of functioning with HSU by older people with depression, suggesting this association is worthy of further investigation in future research.

The finding that only one functioning item from the WHODAS-II was significantly associated with increased "any community HSU" is partly explained by the absence of an association of the WHODAS-II total score with "any community HSU". More associations relating to individual WHODAS-II functioning items would likely have been found if it was possible to undertake analysis using the outcome of hospital admission, because the WHODAS-II total score was significantly associated with this outcome. Nevertheless, the functioning item significantly associated with increased HSU concerned emotionality, and this is in line with the earlier finding from this study and previous research showing an association of depressive symptom severity with increased HSU (Prina et al., 2015). The finding that "difficulty with washing whole body" was associated with decreased "any community HSU" was unexpected, but this difficulty implies a lack of independence which could represent a barrier to accessing community-based health services. No previous research directly investigating an association of individual functioning items/ difficulties with HSU by older people with depression was available for further comparison of results.

Implications of Findings for Practice

The findings, added to those from previous research, have implications for health service planning. As these findings suggest that depressive symptom severity is not strongly related to HSU by older people with depression, the formation of relevant patient clusters (*i.e.*, groupings of patients with similar clinical characteristics and HSU patterns used in contemporary health service payment systems to allocate resources) should take into account additional patient-related variables. Moreover, health services should consider functioning to be a possible determinant of HSU by older people with depression, and functioning information should be investigated further to assess its suitability for patient clustering. The overall modest associations of depressive symptom severity and functioning with HSU suggest that health service planning for patients with depression is potentially a complex process. Only a few high income countries (e.g., Australia, New Zealand, Canada, the Netherlands, Norway, USA, UK) have made progress implementing mental health payment systems, but with widely varying methodologies (Mason et al., 2011). Therefore, it is important that more studies investigate the potential broad range of patient-related variables that may be related to HSU by older people with depression. The relatively modest associations also point to the complexity of estimating HSU by specific cohorts in LMICs with no universal health care coverage. In these settings, several sociodemographic variables relating to health service accessibility (e.g., private health insurance coverage, pension coverage, education) are associated with HSU by older people (Albanese et al., 2011) and these variables may explain substantially more of the variance in HSU than either symptom severity or functioning.

Future Research

Attempts to replicate the study using longitudinal data should make it possible to determine if depressive symptom severity and functioning predict HSU by older people with depression in LMICs over time. Based on our results and (the little) available evidence from longitudinal studies in various populations (Prina et al., 2015), it is expected that positive associations of depressive symptom severity and functioning with HSU would be found, although the magnitude of these associations may be small. Similar studies involving participants from younger age groups and high income countries would increase the applicability of findings to diverse health services. Because differing operationalisations of functioning may account for equivocal findings across studies, the association of differing functioning measures (as well as individual items) with HSU by older people with depression is worth investigating. The functioning measure—PARADISE-24 (Cieza et al., submitted- 2014) developed to account for the hypothesis of "horizontal epidemiology", namely that common functioning difficulties are typically experienced across all neurological and mental disorders, may be suitable for this endeavour. Findings relating to a study investigating the association of the PARADISE-24 with HSU would be applicable to a wide range of mental disorders (Cieza et al., submitted- 2014). Taking into account the range of integrated services needed to effectively treat mental health problems, future research investigating the link between mental health and HSU could include a wider range of HSU outcomes than used in the present study.

Paper 2 Table 1. Sociodemographic and clinical characteristics of participants.

Variable	***n* (%)**													
	All	**China (Urban)**	**China (Rural)**	**Cuba**	**Dominic. Republic**	**India (Urban)**	**India (Rural)**	**Mexico (Urban)**	**Mexico (Rural)**	**Nigeria**	**Peru (Urban)**	**Peru (Rural)**	**Puerto Rico**	**Vene-zuela**
n	4590	44	15	682	757	390	422	312	259	270	390	144	331	574
Age (MV)	5	0	0	1	0	2	0	1	0	0	0	0	0	1
65–69	1347 (29.3)	16 (36.3)	3 (20.0)	182 (26.6)	189 (24.9)	146 (37.4)	140 (33.1)	64 (20.5)	81 (31.2)	82 (30.3)	91 (23.3)	42 (29.1)	89 (26.8)	222 (38.6)
70–74	1224 (26.6)	11 (25.0)	8 (53.3)	151 (22.1)	173 (22.8)	130 (33.3)	152 (36.0)	106 (33.9)	65 (25.1)	76 (28.1)	99 (25.3)	40 (27.7)	69 (20.8)	144 (25
75–79	941 (20.5)	7 (15.9)	3 (20.0)	164 (24.0)	161 (21.2)	61 (15.6)	75 (17.7)	69 (22.1)	57 (22)	35 (12.9)	99 (25.3)	29 (20.1)	81 (24.4)	100 (17.4)
80–84	576 (12.5)	3 (6.8)	1 (6.6)	105 (15.4)	115 (15.1)	28 (7.1)	38 (9.0)	43 (13.7)	32 (12.3)	39 (14.4)	55 (14.1)	21 (14.5)	47 (14.2)	49 (8.5)
85–89	345 (7.5)	7 (15.9)	0	55 (8.0)	76 (10.0)	15 (3.8)	14 (3.3)	24 (7.6)	18 (6.9)	22 (8.1)	34 (8.7)	8 (5.5)	32 (9.6)	40 (6.9)
90+	152 (3.3)	0	0	24 (3.5)	43 (5.6)	8 (2.0)	3 (0.7)	5 (1.6)	6 (2.3)	16 (5.9)	12 (3.0)	4 (2.7)	13 (3.9)	18 (3.1)
Gender (MV)	7	0	0	0	0	0	0	0	0	0	0	0	0	0
Female	3282 (71.5)	27 (61.3)	6 (40.0)	535 (78.4)	573 (75.6)	261 (66.9)	232 (54.9)	235 (75.3)	196 (75.6)	165 (61.1)	282 (72.3)	92 (63.8)	262 (79.1)	416 (72.4)
Education (MV)	20	0	0	0	0	0	0	0	0	18	0	0	2	0
No education	1217 (26.5)	10 (22.7)	6 (40.0)	28 (4.1)	186 (24.5)	220 (56.4)	294 (69.6)	91 (29.1)	93 (35.9)	178 (65.9)	12 (3.0)	33 (22.9)	17 (5.1)	49 (8.5)
Some	1351 (29.5)	10 (22.7)	5 (33.3)	182 (26.6)	398 (52.5)	83 (21.2)	75 (17.7)	130 (41.6)	133 (51.3)	32 (11.8)	42 (10.7)	40 (27.7)	67 (20.2)	154 (26.8)
Completed primary	1155 (25.2)	14 (31.8)	4 (66.6)	247 (36.2)	108 (14.2)	60 (15.3)	40 (9.4)	59 (18.9)	25 (9.6)	32 (11.8)	157 (40.2)	56 (38.8)	82 (24.7)	271 (47.2)
Completed secondary	540 (11.8)	6 (13.6)	0	136 (19.9)	43 (5.6)	21 (5.3)	12 (2.8)	20 (6.4)	5 (1.9)	6 (2.2)	122 (31.2)	7 (4.8)	110 (33.2)	52 (9.0)
Completed tertiary	270 (5.8)	4 (9.0)	0	88 (12.9)	14 (1.8)	6 (1.5)	1 (0.2)	12 (3.8)	3 (1.1)	4 (1.4)	55 (14.1)	5 (3.4)	53 (16)	25 (4.3)
Unknown	37 (0.8)	0	0	1 (0.1)	8 (1.0)	0	0	0	0	0	2 (0.5)	3 (2.0)	0	23 (4.0)
Marital Status (MV)	47	0	0	1	4	0	0	0	0	18	1	0	2	21
Never married	265 (5.7)	0	0	69 (10.1)	37 (4.8)	8 (2.0)	2 (0.4)	16 (5.1)	10 (3.86)	9 (3.30)	42 (10.7)	10 (6.9)	15 (4.5)	47 (8.1)
Married/Cohabiting	1882 (41.0)	30 (68.1)	4 (26.6)	232 (34.0)	182 (24.0)	153 (39.2)	199 (47.1)	140 (44.8)	121 (46.72)	162 (60.0)	196 (50.2)	85 (59.0)	158 (47.7)	220 (38.3)
Widowed	1847 (40.2)	14 (31.8)	11 (73.3)	257 (37.6)	348 (45.9)	206 (52.8)	212 (50.2)	127 (40.7)	110 (42.47)	81 (30.0)	128 (32.8)	47 (32.6)	116 (35.0)	190 (33.1)

Divorced/Separated	549 (11.9)	0	0	123 (18.0)	186 (24.5)	23 (5.9)	9 (2.1)	29 (9.2)	18 (6.9)	0	23 (5.9)	2 (1.3)	40 (12)	96 (16.7)
Assets [1] (MV)	219	0	0	1	2	2	0	0	0	214	0	0	0	0
0	72 (1.6)	0	0	2 (0.3)	3 (0.4)	3 (0.8)	43 (10.2)	0	1 (0.4)	4 (1.5)	0	2 (1.4)	0	14 (2.4)
1–4	1274 (27.8)	3 (6.8)	5 (33.3)	72 (10.5)	281 (37.1)	265 (67.9)	351 (83.2)	36 (11.5)	155 (59.8)	43 (15.9)	9 (2.3)	50 (34.7)	4 (1.2)	560 (97.6)
5+	3025 (65.9)	41 (93.2)	41 (93.2)	607 (89.0)	471 (62.2)	120 (30.8)	28 (6.6)	276 (88.5)	103 (39.8)	9 (3.3)	381 (97.7)	92 (63.9)	327 (98.8)	0
Pension (MV)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Yes	2098 (45.7)	38 (86.4)	3 (20)	541 (79.3)	208 (27.5)	23 (5.9)	145 (34.4)	228 (73.1)	78 (30.1)	1 (0.4)	234 (60)	100 (69.4)	190 (57.4)	309 (53.8)
Private Insurance (MV)	63	0	0	2	1	2	0	0	0	29	2	0	2	25
None	3204 (69.8)	44 (100)	9 (60.0)	680 (99.7)	642 (84.8)	387 (99.2)	422 (100)	134 (42.9)	184 (71.0)	241 (89.2)	88 (22.5)	33 (22.9)	19 (5.7)	321 (55.9)
Physical Comorbidities (MV)	27	0	0	1	0	0	0	0	0	2	1	0	4	19
1–2	1954 (42.6)	18 (40.9)	8 (53.3)	382 (56.0)	342 (45.2)	147 (37.7)	204 (48.3)	135 (43.3)	113 (43.6)	139 (42.0)	166 (42.6)	74 (51.4)	37 (13.7)	189 (32.9)
3+	1287 (28.0)	19 (43.2)	4 (26.7)	122 (17.9)	310 (40.9)	28 (7.2)	101 (23.9)	92 (29.5)	85 (32.8)	131 (39.6)	117 (30.0)	23 (15.9)	7 (2.6)	248 (43.2)
Dementia (MV)	1	0	0	0	0	0	0	0	0	0	0	0	1	0
Any diagnosis	615	10	10	98	121	43	52	45	32	42	32	19	42	69
	(13.4)	(22.7)	(66.7)	(14.4)	(16.0)	(11.0)	(12.3)	(14.4)	(12.4)	(15.6)	(8.2)	(13.2)	(12.7)	(12.0)
Depression (MV)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ICD-10 depressive episode	931	3	6	142	274	37	124	46	42	5	84	16	46	106
	(20.3)	(6.8)	(40.0)	(20.8)	(36.2)	(9.5)	(29.4)	(14.7)	(16.2)	(1.9)	(21.5)	(11.1)	(13.9)	(18.5)

Note: MV = Number of missing values. [1] Household assets (*i.e.*, car, television, refrigerator, telephone, mains water, mains electricity, plumbed toilet).

Paper 2 Table 2. *Prevalence of previous health service utilisation.*

		Any Community Health Service [1] (MV = 0)				**Hospital Admission (MV = 27)**			
Country	***n***	**Crude**		**Standardised [2]**		**Crude**		**Standardised [2]**	
		%	**(95% CI)**	**%**	**(95% CI)**	**%**	**(95% CI)**	**%**	**(95% CI)**
China (urban)	44	59.0	(44.9–73.3)	37.1	(34.2–39.9)	13.6	(3.4–23.8)	6.7	(2.2–11.2)
China (rural)	15	6.6	(0–19.3)	2.1	(0–5.6)	0	-	0	-
Cuba	682	53.6	(49.9–57.5)	51.4	(46.0–56.9)	2.8	(1.5–4.0)	2.4	(1.2–3.7)
Dominican Republic	757	52.3	(48.7–55.9)	52.8	(49.0–56.6)	4.3	(2.9–5.8)	4.4	(2.8–6.1)
India (urban)	390	55.9	(50.8–60.9)	50.5	(45.7–55.3)	1.2	(0.2–2.3)	0.9	(0–1.8)
India (rural)	422	91.0	(88.3–93.7)	60.3	(57.1–63.6)	1.9	(0.6–3.1)	0.7	(0.2–1.2)
Mexico (urban)	237	79.4	(74.9–83.9)	73.7	(69.9–77.6)	4.1	(1.9–6.3)	3.4	(1.5–5.3)
Mexico (rural)	259	68.3	(62.6–74.0)	61.7	(57.6–65.7)	2.7	(0.7–4.6)	1.4	(0.5–2.3)
Nigeria	270	24.4	(19.4–29.4)	23.6	(19.2–27.9)	4.1	(1.7–6.5)	2.7	(1.4–4.1)
Peru (urban)	223	60.0	(54.9–65.0)	49.1	(43.9–4.4)	3.8	(1.8–5.8)	7.4	(6.4–8.4)
Peru (rural)	144	40.2	(32.2–48.4)	37.2	(32.1–42.2)	2.1	(−2.5–4.4)	1.8	(0.01–3.5)
Puerto Rico	331	83.9	(80.0–87.9)	82.1	(77.9–86.2)	6.1	(3.5–8.6)	4.5	(2.3–6.8)
Venezuela	574	68.8	(65.0–72.6)	70.4	(65.6–75.0)	6.4	(4.6–8.7)	5.4	(3.6–7.2)
Pooled estimate	4590	57.8	(46.6–69.0)	50.2	(36.6–63.7)	3.6	(2.5–4.6)	3.4	(1.9–4.7)

Notes: [1] Includes contacts with primary care, private doctor, hospital doctor, traditional healer, other services, in the previous 3 months; [2] Standardised for age (groups), gender, and education; MV = Number of missing values.

Paper 2 Table 3. Associations of depressive symptom severity (EURO-D total score) and functioning (WHODAS-II total score) with health service utilisation.

Country	***n***	**Prevalence Ratios (95% CI)**							
		Any Community Health Service [1] (MV = 0)				**Hospital Admission (MV = 27)**			
		Depression Severity		**Functioning [2]**		**Depression Severity**		**Functioning [2]**	
		Crude	**Adjusted [3]**	**Crude**	**Adjusted [3]**	**Crude**	**Adjusted [3]**	**Crude**	**Adjusted [3]**
China (urban)	44	0.92 (0.77–1.09)	0.92 (0.78–1.08)	0.97 (0.88–1.07)	0.99 (0.88–1.11)	1.35 (0.94–1.95)	1.30 (0.81–2.11)	1.52 (0.92–1.67)	1.21 (0.82–1.78)
Cuba	682	1.03 (0.99–1.07)	1.04 (1.00–1.08)	0.96 (0.93–0.99)	0.96 (0.93–1.01)	0.93 (0.73–1.19)	0.93 (0.73–1.19)	0.89 (0.75–1.07)	0.83 (0.68–1.02)
Dominican Republic	757	1.02 (0.98–1.06)	1.02 (0.98–1.06)	1.03 (1.00–1.05)	1.02 (0.98–1.04)	1.21 (1.04–1.40)	1.18 (1.01–1.37)	1.27 (1.14–1.42)	1.28 (1.11–1.47)
India (urban)	390	1.03 (0.98–1.09)	1.05 (1.00–1.10)	1.08 (1.03–1.11)	1.03 (0.98–1.07)	1.54 (1.29–1.84)	1.87 (1.54–2.27)	1.39 (1.06–1.81)	1.12 (0.84–1.48)
India (rural)	422	1.01 (0.99–1.03)	1.02 (1.00–1.04)	0.99 (0.97–1.01)	0.99 (0.97–1.01)	1.08 (0.82–1.43)	1.01 (0.77–1.33)	1.19 (0.81–1.76)	1.12 (0.76–1.63)
Mexico (urban)	237	1.01 (0.98–1.04)	1.01 (0.98–1.04)	1.00 (0.97–1.02)	1.00 (0.97–1.03)	0.91 (0.71–1.16)	0.93 (0.71–1.22)	1.22 (0.99–1.50)	1.26 (1.04–1.53)
Mexico (rural)	259	0.99 (0.94–1.05)	1.00 (0.94–1.05)	1.00 (0.97–1.04)	1.02 (0.99–1.05)	1.28 (0.95–1.74)	1.36 (0.95–1.95)	1.29 (1.09–1.52)	1.32 (1.09–1.61)
Nigeria	270	0.96 (0.87–1.05)	0.95 (0.87–1.04)	1.07 (0.98–1.17)	1.11 (1.02–1.21)	0.73 (0.57–0.94)	0.76 (0.58–0.98)	0.96 (0.77–1.17)	0.92 (0.73–1.17)
Peru (urban)	223	1.06 (1.01–1.10)	1.06 (1.01–1.10)	1.05 (1.01–1.09)	1.05 (1.02–1.09)	1.42 (1.11–1.82)	1.44 (1.06–1.94)	1.27 (1.09–1.49)	1.39 (1.18–1.64)
Peru (rural)	144	1.19 (1.03–1.36)	1.19 (1.04–1.36)	1.12 (1.03–1.23)	1.05 (0.95–1.14)	0.89 (0.45–1.76)	0.82 (0.43–1.52)	0.75 (0.47–1.20)	0.73 (0.42–1.28)
Puerto Rico	331	1.03 (1.00–1.05)	1.02 (1.00–1.05)	1.02 (0.99–1.04)	1.01 (0.98–1.04)	1.01 (0.78–1.31)	0.98 (0.75–1.27)	1.30 (1.12–1.52)	1.24 (1.02–1.51)
Venezuela	574	1.01 (0.97–1.05)	1.01 (0.98–1.04)	1.03 (1.00–1.06)	1.03 (0.99–1.05)	1.11 (0.93–1.33)	1.08 (0.91–1.30)	1.17 (1.03–1.34)	1.16 (1.00–1.35)
Meta-effect	4590	1.02 (1.01–1.03)	1.02 (1.01–1.03)	1.02 (1.00–1.04)	1.01 (1.00–1.03)	1.11 (0.96–1.25)	1.11 (0.95–1.26)	1.18 (1.07–1.29)	1.14 (1.02–1.26)
I^2		14.0%	20.8%	71.2%	44.7%	70.2%	70.7%	64.9%	64.4%

Notes: "China (rural)" was excluded because there were too few cases to estimate parameters; References in comparisons were a one point increase on the EURO-D or a 10-point increase on the WHODAS-II, where appropriate; **[1]** Includes contacts with primary care, private doctor, hospital doctor, traditional healer, other services, in the previous 3 months; **[2]** Because the total score is expressed as a percentage, WHODAS-II scores were adjusted to 10-point prevalence; **[3]** Adjusted for age (groups), gender, education, and physical comorbidities; MV = Number of missing values.

Paper 2 Table 4. Association of individual functioning difficulties (WHODAS-II items) with any community health service use [1].

No.	**Item**	**Pooled (Meta-Analysed) Prevalence Ratios (95% CI)**					
		Crude		**I^2 (%)**	**Adjusted[2]**		**I^2 (%)**
		PR	**(95% CI)**		**PR**	**(95% CI)**	
1	Difficulty with standing	1.02	(0.99–1.04)	0	1.01	(0.99–1.03)	0
2	Difficulty with taking care of household responsibilities	0.99	(0.96–1.01)	0	0.98	(0.96–1.01)	0
3	Difficulty with learning a new task	0.99	(0.96–1.01)	0	0.99	(0.97–1.02)	0
4	Difficulty with joining in community activities	0.99	(0.95–1.04)	60.2	1.00	(0.96–1.04)	55.2
5	Emotionally affected by health problems	1.06	(1.03–1.09)	31.7	1.04	(1.01–1.07)	22.4
6	Difficulty with concentrating on doing something for ten minutes	1.00	(0.98–1.03)	7.3	1.01	(0.98–1.03)	0
7	Difficulty with walking a long distance	1.01	(0.97–1.04)	60.4	1.00	(0.97–1.04)	51.5
8	Difficulty with washing whole body	0.93	(0.89–0.98)	28.9	0.93	(0.88–0.98)	37.6
9	Difficulty with getting dressed	1.01	(0.97–1.05)	1.7	1.01	(0.97–1.05)	0
10	Difficulty with unknown people	0.93	(0.85–1.01)	73.3	0.93	(0.86–1.01)	71.0
11	Difficulty with maintaining a friendship	1.02	(0.96–1.07)	54.4	1.01	(0.95–1.06)	49.2
12	Difficulty with day-to-day work/school	1.03	(0.99–1.06)	25.7	1.03	(0.99–1.06)	34.7

Notes: References in comparisons were a one point increase on individual WHODAS-II items; [1] Includes contacts with primary care, private doctor, hospital doctor, traditional healer, other services, in the previous 3 months; [2] Adjusted for age (groups), gender, education, and physical comorbidities; PR = Prevalence ratios.

INTRODUCTION TO PAPER 3

Addressing the second specific objective of the thesis - to determine the association of functioning with HSU by people with common mental disorders - Paper 3 encompasses an historical cohort study with a 12-month follow-up period. Using a comprehensive electronic repository of patient records from a relatively deprived public health service catchment area in South London, the paper investigated the utility of functioning in predicting mental health service costs for patients with common mental disorders. The study population was chosen because it is representative of relatively deprived patients with common mental disorders (residing in the UK) and thus the findings are applicable to similar populations elsewhere. In addition, the available database of patient records was large and comprehensive, ensuring statistical power and precision in the findings. The findings are less generalisable to people of higher socio-economic status, and people residing outside the UK.

Functioning was measured using the Health of the Nation Outcome Scales (HoNOS) (Wing et al., 1998). There are 7 items from the 12-item HoNOS that are representative of functioning and the HoNOS items are examined separately in Paper 3: thus, the associations of various aspects of functioning with HSU could be investigated. Although not their original purpose, the HoNOS are often operationalised as a composite 12-item scale (which provides a total HoNOS score) and this composite scale is the most widely used routine outcome measure in the mental health services of the UK, New Zealand and Australia (Lovaglio & Monzani, 2012). Internal consistency of the composite HoNOS is moderately high and concurrent validity with other clinician-rated instruments of symptoms and functioning is generally good (Pirkis et al., 2005). On the other hand, the HoNOS has various psychometric limitations: several factor structures have been proposed but none of these have acceptable fit (Speak & Muncer, 2015); Rasch analyses demonstrate the absence of an underlying construct in the composite scale (Lovaglio & Monzani, 2011, 2012); and there is no universal agreement regarding operationalisation of the HoNOS.

Paper 3 was undertaken in collaboration with colleagues from KCL where I spent three months on secondment from October 2014 to December 2014. The paper was mostly completed after this secondment period though, remotely. The paper has been published in the PLoS ONE journal (C. Twomey et al., 2016). The PhD researcher was the lead author of the paper. Under supervision, the PhD researcher contribution included conceiving and planning the paper, undertaking statistical analysis (guided by a co-author), drafting all versions of the paper, and managing the submission and peer-review process.

PAPER 3: Utility of the Health of the Nation Outcome Scales (HoNOS) in predicting mental health service costs for patients with common mental health problems: historical cohort study

Abstract

Background

Few countries have made much progress in implementing transparent and efficient systems for the allocation of mental health care resources. In England there are ongoing efforts by the National Health Service (NHS) to develop mental health 'payment by results' (PbR). The system depends on the ability of patient 'clusters' derived from the Health of the Nation Outcome Scales (HoNOS) to predict costs. We therefore investigated the associations of individual HoNOS items and the Total HoNOS score at baseline with mental health service costs at one year follow-up.

Methods

An historical cohort study using secondary care patient records from the UK financial year 2012-2013. Included were 1,343 patients with 'common mental health problems', represented by ICD-10 disorders between F32-48. Costs were based on patient contacts with community-based and hospital-based mental health services. The costs outcome was transformed into 'high costs' *vs* 'regular costs' in main analyses.

Results

After adjustment for covariates, 11 HoNOS items were not associated with costs. The exception was 'self-injury' with an odds ratio of 1.41 (95% CI 1.10-2.99). Population attributable fractions (PAFs) for the contribution of HoNOS items to high costs ranged from 0.6% (physical illness) to 22.4% (self-injury). After adjustment, the Total HoNOS score was not associated with costs (OR 1.03, 95% CI 0.99-1.07). However, the PAF (33.3%) demonstrated that it might account for a modest proportion of the incidence of high costs.

Conclusions

Our findings provide limited support for the utility of the self-injury item and Total HoNOS score in predicting costs. However, the absence of associations for the remaining HoNOS items indicates that current PbR clusters have minimal ability to predict costs, so potentially contributing to a misallocation of NHS resources across England. The findings may inform the development of mental health payment systems internationally, especially since the vast majority of countries have not progressed past the early stages of this development. Discrepancies between our findings with those from Australia and New Zealand point to the need for further international investigations.

Introduction

Few countries have made much progress in implementing transparent and efficient systems for the allocation of mental health care resources (Mason et al., 2011). In England there are ongoing efforts by the National Health Service (NHS) to develop mental health 'payment by results' (PbR). Healthcare providers receive funding for every patient treated, with the level of payment determined by the category (i.e. 'mental health cluster') to which each patient is assigned based on clinical characteristics and assessed needs (Mannion, Marini, & Street, 2008). There are 21 mental health clusters, organised under 'non-psychotic', 'psychotic' and 'organic' domains. Although mental health PbR financially rewards providers for volumes of work and thus may increase efficiency(Mannion et al., 2008), widespread criticism has contributed to repeated delays in its rollout (Wang, Shaw, & Middleton, 2015). Monitor - the NHS regulator - has highlighted that quality of care is not incentivized because provider funding is not linked to patient recovery (Monitor, 2015). The overall approach of using mental health clusters to determine the level of payment for each patient has been questioned given that pilot studies demonstrated their low resource homogeneity, and inferiority to an alternative statistically-derived model in reducing the variance in resource usage (*Health and Social Care Information Centre Casemix Service. Mental Health Casemix Classification Development: End Stage Report*, 2006; Alex D. Tulloch, 2012). In particular, the delayed rollout can be attributed to concerns surrounding the process of clustering patients and the validity of the mental health clusters (Wang et al., 2015).

Perhaps taking into account the lack of evidence for the utility of diagnostically-defined clusters in predicting costs, and the subsequent exclusion of mental health services from the original Medicare system of the United States (Cotterill & Thomas, 2004; Elphick & Antony, 1996; English et al., 1986; A. J. Macdonald & Elphick, 2011; Schumacher et al., 1986), NHS policy-makers have pursued a 'multi-domain' approach to clustering. Patients are assigned to mental health clusters using the Health of the Nation Outcome Scales (HoNOS)(Wing et al., 1998) and additional risk-based items. The principal assumption of the multi-domain approach is that patients who score similarly on clusters derived from the HoNOS have similar clinical needs and incur similar health service costs (Alex D. Tulloch, 2012). The HoNOS were initially proposed for clustering purposes because they "provided a suitable platform for a shared assessment tool in that [they were] already accepted and regarded as useful across the service, and nationally accepted and validated" (Self, Rigby, Leggett, & Paxton, 2008)(p38). As the HoNOS are used to measure clinical outcomes in the English Mental Health Minimum Dataset, the practicality of using the scales for clustering seems justifiable. However, the HoNOS were not originally designed to predict costs of care and their utility in this prediction was not considered in the decision to use them for clustering purposes (Bekas & Michev, 2013; Self et al., 2008; Alex D. Tulloch, 2012). As mental health PbR depends on the ability of HoNOS-derived clusters to reliably predict treatment costs (Wang et al., 2015), this omission means that important information pertaining to the optimal allocation of resources was not taken into account.

There is little existing evidence for the utility of the HoNOS in predicting costs for patients with mental disorders. Reports on ultimately unsuccessful attempts to implement casemix classification systems in Australia (Burgess, Pirkis, Buckingham, Eagar, & Solomon, 1999) and New Zealand (Eagar et al., 2004) provide some support for the instrument's utility; however, the evidence is tentative because both studies only demonstrate cost associations of mental health clusters, formed of the HoNOS and other 'casemix' variables such as age, ethnicity, diagnosis, and focus of care. Some studies have examined the associations of HoNOS scores with health service use as a proxy for costs, but the results are mixed with both positive (Andreas et al., 2010; Kisely, Campbell, Cartwright, Cox, & Campbell, 2010; A. D. Tulloch et al., 2015; A. D. Tulloch, Khondoker, Fearon, & David, 2012) and absent (Byrne, Hooke, & Page, 2010; Callaly, Trauer, Hyland, Coombs, & Berk, 2011; Goldney, Fisher, & Walmsley, 1998) associations. These limited findings neither support nor refute the utility of the HoNOS in predicting health service costs. There is an urgent need to address this unresolved issue, since a misallocation of resources could affect patients and clinical services across England.

We carried out an historical cohort study with a one-year follow up, using a comprehensive repository of anonymised electronic patient records, to investigate the utility of the HoNOS in predicting mental health service costs. We initially considered including participants with any psychiatric disorder but there was a large discrepancy in missing HoNOS data between patients with less severe disorders attending community-based services (~25%) and patients with more severe disorders who were using hospital-based care (>65%) so we therefore limited our analysis to patients with 'common mental health problems', represented by the recording of an ICD-10 disorder related to (unipolar) depression, anxiety, stress, adjustment or somatic problems (F32-F48). Although we use the term 'common mental health problems', our sample of secondary mental health care patients differs from those patients with milder and simpler difficulties treated predominately in primary care in the UK. HoNOS items are used for clustering within mental health PbR but it is commonplace for studies to deploy the total HoNOS score in analyses. To provide a comprehensive assessment of predictive utility, we separately investigated associations of individual HoNOS items and the total HoNOS score with costs.

Methods

Study Population

Data source

The data source was electronic patient records from the South London & Maudsley NHS Foundation Trust (SLaM) Biomedical Research Centre (BRC) Case Register (Stewart et al., 2009). SLaM provides comprehensive secondary mental health care to around 1.2 million residents of four relatively deprived London boroughs (Stewart et al., 2009). Via the Clinical Record Interactive Search (CRIS) system, the SLAM BRC Case Register allows secondary analysis of data from de-identified SLAM records on approximately 250,000 cases. The Case Register has received approval from the Oxford Research Ethics Committee C (reference 08/H0606/71+5) and full details of this approval process are provided elsewhere (Stewart et al., 2009) Further internal approval for this project was granted by the ethics committee of the SLaM BRC.

Inclusion criteria

We only included patients experiencing 'common mental health problems', represented by the recording of an ICD-10 disorder related to (unipolar) depression, anxiety, stress, adjustment or somatic problems (F32-F48). To enhance representativeness of common mental health problems, patients also needed to have a first SLaM contact in two types of community-based services primarily geared towards the treatment of less severe problems: 'Assessment Brief Therapy' and 'Mood, Anxiety and Personality Disorders'. Older adults (aged 65+) were excluded because their mental health service costs are based on a range of services not attended by working age adults (e.g. 'Mental Health of Older Adults and Dementia' service). The sample was followed up for one year from the start of their first treatment episode occurring within the UK financial year 2012-2013 (1 April 2012 to 31 March 2013).

Measures

Sample characteristics

Sample characteristics were assessed using the following variables: age (at baseline), gender, ethnicity, marital status, area-level deprivation based on the Index of Multiple Deprivation (Noble, Wright, Smith, & Dibben, 2006), ICD-10 diagnosis, borough where a SLaM mental health service was first accessed (not limited to the observation period), type of care at end of treatment episode, and days in contact with a SLaM mental health service in the year before baseline. The number of days between the start of the baseline treatment episode and HoNOS completion was also recorded.

Exposures: scores on the HoNOS at baseline

The HoNOS are a set of 12 scales measuring mental health-related problems in the domains of behaviour, symptoms, impairment, and social functioning (Wing et al., 1998). Scores on each scale/item range from 0 (no problem) to 4 (severe problem). Although not their original purpose, the HoNOS are often operationalised as a composite 12-item scale (which provides a total HoNOS score ranging from 0-48). The composite scale is the most widely used routine outcome measure in the mental health services of the UK, New Zealand and Australia (Lovaglio & Monzani, 2012). Internal consistency of the composite HoNOS is moderately high and concurrent validity with other clinician-rated instruments of symptoms and multi-domain functioning is generally good (although it is poor for self-rated instruments)(Pirkis et al., 2005). Several factor structures have been proposed but none of these have acceptable fit (Speak & Muncer, 2015). Rasch analyses demonstrate the absence of an underlying construct in the composite scale (Lovaglio & Monzani, 2011, 2012). Analyses involving individual HoNOS items have been undertaken in many studies (Eagar et al., 2004; A. D. Tulloch et al., 2015; A. D. Tulloch et al., 2014). There is no universal agreement regarding operationalisation of the HoNOS.

Outcome: mental health service costs at one year follow-up

Costing data were provided through SLaM internal financial records. We chose a one year follow-up period for two reasons: (1) in PbR, the interval for review of allocation to clusters covering less severe and more common disorders is typically one year; (2) the one year period is likely to even out the seasonal effects seen in admissions and mental health problems. For each patient, the mental health service costs outcome was calculated by adding costs of treatment by any community mental health team, whether generic or specialised, and the costs of any hospital treatment, taking as the start point the start of their initial treatment episode within the observation period. Both community mental health team and inpatient costs were calculated based on unit costs calculated at the individual team/ward level using SLaM internal financial data and the total relevant activity over that financial year (total face-to-face contacts for community mental health teams and total inpatient days for wards). On account of a highly skewed distribution, the outcome was transformed into 'high costs' (the top 10% of costs) and 'regular costs' (remaining 90% of costs). This cut-off was based on research demonstrating that a minority of 'high cost' patients (i.e. the top 5% - 10% for costs incurred) account for at least a 50% share of costs (Berk & Monheit, 2001; Lavik, 1983; Taube, Goldman, Burns, & Kessler, 1988), and similar cut-off points have been used in previous studies examining mental health service costs (Calver et al., 2006; Harrison-Read et al., 2002; Prina et al., 2014). In the current data set, 'high cost' patients accounted for a 58% share of costs.

Statistical analysis

Main analyses

All statistical analyses were undertaken using Stata 12. Descriptive statistics were used for sample characteristics. The utility of both individual HoNOS items and the Total HoNOS score (at baseline) in predicting costs (at one year follow up) were determined in separate analyses. HoNOS items were simultaneously entered into all relevant predictive models. Based on previous research showing their associations with mental health service costs (Durbin, Moineddin, Lin, Steele, & Glazier, 2015; C. D. Twomey, Baldwin, et al., 2015), additional adjustments were made for age, gender, marital status, ethnicity, area-level deprivation (Index of Multiple Deprivation, in quintiles, for the sample), and previous health service use (previous days in contact with a SLaM mental health service in the year before baseline). To account for possible differences in service configurations across locations, the borough where a SLaM mental health service was first accessed was an additional adjustment. To account for possible changes to HoNOS scores over time, 'the number of days between baseline and HoNOS completion' was also included in the models. Diagnosis was not adjusted for because patients with common mental health problems are thought to have similar needs, and preliminary analysis showed diagnosis was not related to costs.

As a precursor to the main analyses, t-tests were used to examine differences in both individual item and total HoNOS scores between 'regular cost' and 'high cost' patients. Associations of baseline HoNOS scores with 'one-year' costs were determined using unadjusted and adjusted odds ratios with 95% confidence intervals (CIs), modelled through logistic regression analysis. Population attributable fractions (PAFs)- representing the percentage decrease in the number of 'high cost' patients that would theoretically arise if a problem within a given HoNOS domain could be removed from the study population- were calculated and applied to fully-adjusted models. As the 'aflogit' command for PAFs is not supported for use with imputed data by Stata 12, PAFs were applied to complete-case analyses.

Missing data

Complete data pertaining to the costs outcome were available but HoNOS scores had sizeable missing data (i.e. 24.8%). We deployed multiple imputation by chained equations (MICE) to impute this missing data. Multiple imputation uses patterns in observed data to impute missing values, repeating this process multiple times to account for uncertainty in the imputed values (Lee & Simpson, 2014). MICE facilitated the sequential imputation of missing data for each HoNOS item via predictive mean matching. Imputation models included all variables known to predict missingness (i.e. means of contact, type of care at end of treatment episode, the costs outcome) and all other reported variables, apart from 'number of days between baseline and HoNOS completion' which also had sizeable missing data. 100 imputed datasets were created The number of imputation cycles was constrained by limited computational power; however, it resulted in the introduction of minimal standard error (i.e. Monte Carlo Error), as per guidelines (White, Royston, & Wood, 2011b). Checks between imputed and original values produced no anomalies. For final analyses of imputed datasets, estimates were combined using Rubin's rules (White, Royston, & Wood, 2011a).

Sensitivity analyses

Two sets of sensitivity analyses were undertaken: (1) complete case analyses using the fully-adjusted logistic regression models (for comparison of results with those derived from multiple imputation); (2) fully-adjusted associations of baseline HoNOS scores with a continuous 'one year' costs outcome, analysed using generalised linear models (GLM) with log link functions and gamma distributions (for comparison of results with those involving the dichotomized costs outcome). The use of 'GLM-log-gammas' accounted for the skewed distribution of the continuous costs outcome in latter analyses (Sabariego, Brach, & Stucki, 2011).

Supplementary analysis

Initially, we did not separate costs by admitted and non-admitted settings in analysis because such separation of costs is not prioritised in mental health PbR and only 5% of the sample had an inpatient admission in the follow-up period, limiting the potential influence of setting on the findings. However, previous international studies of casemix classification systems for mental health services have separated costs by setting. For additional comparison with these studies, we present the fully-adjusted associations of HoNOS scores with inpatient admission and community health service costs in a supplementary file (S1 Table).

Results

Sample characteristics

Table 1 provides a full summary of sample characteristics (n = 1,343). The sample was ethnically diverse with a large proportion living in deprived areas. Most participants remained under the care of CMHTs at the end of their first treatment episode within the observation period.

Associations of baseline HoNOS scores with costs at follow-up

The distribution of costs by HONOS items is presented in Table 2. 'High cost' patients had significantly higher HoNOS scores than 'regular cost' patients on items relating to self-injury ($t = 3.72$, $p < 0.01$) and on the Total HoNOS ($t = 2.35$, $p < 0.05$).

Table 3 displays the associations of individual HoNOS items and the Total HoNOS score with ('regular' vs 'high') costs. Only 'self-injury' was significantly associated with costs, with positive associations found in analyses involving adjustment for other HONOS items (odds ratio = 1.31; 95% CI 1.09 – 1.59) and full adjustment for all covariates (odds ratio = 1.41; 95% CI 1.15 – 1.72). The Total HoNOS score was significantly associated with costs in unadjusted analysis (odds ratio = 1.04; 95% CI 1.00 – 1.08) but not in fully adjusted analysis (odds ratio = 1.03; 95% CI 0.99 – 1.07). PAFs for the contribution of HoNOS items to high costs ranged from 0.6% (physical illness) to 22.4% (self-injury), with four items removed from the PAF model due to their negative associations with costs. The PAF for the Total HoNOS score was 33.3%

Sensitivity analyses

Table 4 reports (fully-adjusted) sensitivity analyses pertaining to the cost associations of individual HoNOS items and the Total HoNOS score. Complete case analyses yielded similar results to the main analyses: only the HoNOS 'self-injury' item was significantly associated with costs (odds ratio = 1.50; 95% CI 1.22 – 1.85). The Total HoNOS score was not associated with costs. In examination of the continuous costs outcome, 'GLM-log-gamma' analyses yielded largely similar results to the main analyses. Regarding HoNOS items, only 'self-injury' was significantly associated with costs (e^{β} = 1.17; 95% CI 1.01 – 1.36). However, a significant cost association was yielded for the Total HoNOS score (e^{β} = 1.03, 95% CI 1.01 – 1.06).

Discussion

Summary of findings

Findings pertaining to the associations of baseline HoNOS items and the Total HoNOS score with costs at one year follow up are summarised in turn. After adjustment for covariates, 11 of the 12 HoNOS items were not significantly associated with ('regular vs high') mental health service costs for patients with common mental health problems. The exception was the 'self-injury' item with an odds ratio of 1.41 (95% CI 1.15 – 1.72). PAFs for the contribution of HoNOS items to high costs ranged from 0.6% (physical illness) to 22.4% (self-injury). After adjustment, the Total HoNOS score was not significantly associated with costs in the main analysis, although the association was significant for total costs as a continuous outcome and in supplementary analyses which split costs by setting (S1 Table). Assuming that the observed effect was not accounted for by chance, the PAF of 33.3% demonstrated that it might account for a modest proportion of the incidence of high costs.

Limitations and strengths

As high levels of missing HoNOS data led to their exclusion, the findings are not applicable to patients with more severe mental disorders and service needs. This reflects the challenge of collecting comparable mental health service data across diverse settings and clinical populations. But the findings are applicable to the first two mental health clusters from PbR, which cover the most common mental health problems. The findings are less applicable to patients with milder and simpler 'common mental health problems' treated predominantly in primary care in the UK. The proportion of missing data for other clinical measures was very high (e.g. 98.5% for the CORE-OM, which measures subjective well-being, functioning and risk)(Evans et al., 2002), and this prevented the examination of potentially important predictors of costs. It also meant that it was not possible to assess whether it was the constructs rated by the HoNOS, or the format of the HoNOS, that accounted for the mostly absent cost associations. Most data pertaining to comorbidity were missing and thus its effect could not be examined, albeit that comorbidity is closely associated with mental health service costs (C. D. Twomey, Baldwin, et al., 2015). The modest PAFs for the contribution of exposures and covariates to high costs indicates that there are other determinants that have not been considered. The costs outcome did not capture the full range of health services typically accessed by people with common mental health problems (e.g. primary care psychological services) and stronger associations may have been yielded if it had been possible to incorporate such data.

The study benefits from its use of an established case register, which provided a large clinical sample from a defined catchment area covering a population of 1.2 million people, which is demographically and socio-economically similar to other deprived areas in London (Stewart et al., 2009).It is the first peer-reviewed study to directly investigate the associations of scores on the HoNOS with health service costs. Therefore, the findings may have important policy implications for the English NHS, which has assigned a key role to the HoNOS in mental health PbR.

The findings may also inform the development of mental health payment systems internationally, especially since the vast majority of countries have not progressed past the early stages of this development (Mason et al., 2011). Although the costs outcome was limited in scope, it provided an approximation of the costs used for reimbursement purposes in mental health PbR. This augments the applicability of findings to this system. Findings pertaining to dichotomised costs outcomes enabled intuitive statistical interpretations via odds ratios and PAFs. These were compared with associations for continuous costs outcomes in sensitivity analyses, increasing the validity of conclusions.

Comparison with other studies

Two studies, reporting on the development of casemix classification systems in Australia (Burgess et al., 1999) and New Zealand (Eagar et al., 2004), found that mental health clusters formed of the HoNOS alongside a range of casemix variables (e.g. age, ethnicity, diagnosis) were significantly associated with costs. However, the investigation of mental health clusters (i.e. categories to which patients are assigned based on their clinical characteristics and needs) rather than the HoNOS in these studies limits the comparisons that can be made. National differences in mental health service configurations also limit comparability. These issues aside, both of these studies had clinical samples comprising over 10,000 patients with a variety of mental disorders. It could be that their additional statistical power yielded associations that were not detectable in the present study. Their investigation of a broad spectrum of mental disorders may also explain the contrasting findings: the HoNOS is more often used in 'moderate-to-severe' clinical populations and may have better predictive ability in such populations than our sample which covered less severe disorders. Another explanation for the contrasting findings concerns confounding. The mental health clusters in the previous studies were partly defined using diagnosis. Diagnosis has been consistently found to be associated with costs (C. D. Twomey, Baldwin, et al., 2015) and accounted for a far greater amount of the variance in length of stay than any of the HoNOS items in a case-register study involving psychiatric inpatients (A. D. Tulloch et al., 2012). Therefore, the contrasting absence of cost associations for the HoNOS in the current study may be attributable to our sole inclusion of patients with common mental health problems, which negated the potential confounding effect of diagnosis on examined associations.

Comparisons with the mixed findings from previous health service use studies (Andreas et al., 2010; Byrne et al., 2010; Callaly et al., 2011; Goldney et al., 1998; Kisely et al., 2010; A. D. Tulloch et al., 2015; A. D. Tulloch et al., 2012) are limited by wide variations in the operationalisations of health service use outcomes in these studies (e.g. number of admissions, length of stay, outpatient clinic contacts). Moreover, the relevance of these studies to our research question is limited by their inability to provide a weighted summary of resource consumption through the use of costs outcomes. Overall, the results of the present study (mostly no associations) and previous research involving the HoNOS (mixed evidence for associations) highlight the need for further investigations of the utility of the HoNOS in predicting health service costs for patients with mental disorders.

Comparing the findings with previous research not involving the HoNOS, the association of self-injury with high costs corresponds with a previous report of an association of self-injury frequency with long-term health service costs (Sinclair, Gray, Rivero-Arias, Saunders, & Hawton, 2011). In supplementary analyses (S1 Table) self-injury predicted both inpatient admission and community costs but yielded a larger association with the former outcome than the latter. The strong utility of self-injury in the prediction of hospital based costs has also been demonstrated in a UK population-level investigation of accident and emergency (A&E) visits by people with varying mental disorders (Keene & Rodriguez, 2007). The lack of a cost association for depressive symptoms is also noteworthy, given the sample composition. It could be explained by the limited ability of one-item scales to capture depressive symptom severity (Blozik, Scherer, Lacruz, Ladwig, & group, 2013) or the relatively weak association of depressive symptoms with costs suggested in previous research (B. Y. Huang et al., 2000; Prina et al., 2015).

Implications for policy and future research

The predominant absence of cost associations for HoNOS items raises concerns about the decision by policy-makers to assign a key role to these items within mental health PbR, largely based on their presumed utility in predicting costs. These concerns are compounded by the lack of robust evidence from other studies for associations of HoNOS scores with costs (Andreas et al., 2010; Byrne et al., 2010; Callaly et al., 2011; Goldney et al., 1998; Kisely et al., 2010; A. D. Tulloch et al., 2015), operationalisation and validity issues (Lovaglio & Monzani, 2011; Speak & Muncer, 2015), and the fact that the HoNOS were designed to measure clinical outcome rather than need for care (Kingdon et al., 2012).

Our findings also highlight the need for assessments of alternative approaches to developing payment systems for mental health services. Monitor – the NHS regulator – has suggested that payments should be closely linked to agreed patient outcome standards to incentivise quality of care.(Monitor, 2015) Although the lack of utility of diagnostic related groups in predicting mental health service costs is well-documented (Cotterill & Thomas, 2004; Elphick & Antony, 1996; English et al., 1986; A. J. Macdonald & Elphick, 2011; Schumacher et al., 1986), it would be feasible to investigate the predictive ability of the combination of broad diagnostic categories with clinical pathways (Kingdon et al., 2012). Multi-domain approaches to patient clustering that make use of a wide range of patient-related variables alongside clinical outcome measures could also be examined. This approach has produced promising results in (ultimately unsuccessful) attempts to implement casemix classification systems in Australia and New Zealand (Burgess et al., 1999; Eagar et al., 2004). The recently developed Australian Mental Health Care Classification (AMHCC) system also clusters patients using multi-domain information, incorporating the HoNOS into its casemix classes (https://www.ihpa.gov.au/what-we-do/mental-health-care). Another proposal is for clinicians to judge the most appropriate care pathway option based on the detailed assessment of problems in nine domains of mental health and 12 domains of everyday living, using the MRC Needs For Care Assessment Schedule (Brewin, Wing, Mangen, Brugha, & MacCarthy, 1987; Wang et al., 2015). As associations of self-injury with high costs in this study are in line with previous research (Keene & Rodriguez, 2007; Sinclair et al., 2011), further investigation of the utility of self-injury information in payment systems is warranted. This is especially important as PAF analyses demonstrated that self-injury contributed substantially more to the incidence of high costs than other HoNOS items. Further research avenues could be generated through examination of international attempts to implement payment systems for mental health services (e.g., in Australia, New Zealand, Canada and the Netherlands)(Mason et al., 2011).

Future investigations of the associations of HoNOS scores with costs should address the limitations of the present study. For example, a costs outcome derived from patient contacts with the full range of health services would add generalisability. Examining a broader costs outcome would also raise the issue of whether or not mental health services should be compensated for their involvement in patient care for heavy consumers of primary care services. Greater integration of service sectors, in financing as well as commissioning, planning and delivery, would be required to facilitate this reimbursement approach. Costs information could be collected directly from participants using measures such as the Client Services Receipt Inventory (Beecham & Knapp, 2001), although this is less feasible in large-scale research, and the lack of data linkages between case-registers represents a further challenge (Garcia Alvarez et al., 2011). A more diverse sample composition (e.g. including people with severe mental disorders, and older adults) would enable a more comprehensive assessment of the potential utility of the HoNOS in predicting costs. Discrepancies between our findings with those from Australia (Burgess et al., 1999) and New Zealand (Eagar et al., 2004), point to the need for further international investigations. Given that the composite HoNOS is the most widely used routine outcome measure within NHS mental health services, future investigations of their utility are both feasible and necessary.

Paper 3 Table 1. Sample characteristics (N = 1343)

Variable	***n* (%)**	***M* (*SD*)**	***Median (25th, 75th centile)***
Age[1]		39.4 (11.9)	40.0 (29.0 – 49.0)
Missing	0		
Gender			
Male	586 (43.5)		
Female	757 (56.4)		
Missing	0		
Ethnicity			
White	677 (50.4)		
Black	248 (18.5)		
Other	354 (26.3)		
Missing	64 (4.8)		
Marital status			
Single	913 (68.0)		
Married or cohabiting	227 (16.9)		
Divorced, separated or widowed	160 (11.9)		
Missing	43 (3.2)		
Index of Multiple Deprivation score		34.4 (8.89)[2]	35.0 (27.9 – 40.7)
Missing	4 (0.3)		
ICD-10 Diagnosis			
Mood disorder (F32- F39)	864 (64.3)		
NSS disorder (F40-F48)	479 (35.7)		
Missing	0		
Borough where service first accessed			
Croydon	7 (0.5)		
Lambeth	400 (29.8)		
Lewisham	395 (29.4)		
Southwark	541 (40.3)		
Missing	0		
Under care of CMHT at end of episode	1202 (89.5)		
Missing	0		
Days in contact with SLaM in year before baseline		26.3 (66.89)	0 (0 – 1.0)
Missing	0		
Total HoNOS score[3]		11.2 (5.39)	10.0 (7.0 – 15.0)
Missing	333 (24.8)		
Days between HoNOS completion and baseline[4]		4.5 (10.35)	0 (0 – 2.0)
Missing	326 (24.3)		

Notes: [1]Age range is 18-64. [2]This score is within the top quintile for deprivation in the UK population. [3]Total HoNOS scores ranged from 0-32. [4]'Days between' ranged from 0-61. CMHT = Community Mental Health Team; NSS= neurotic, stress-related or somatoform; HoNOS = Health of the Nation Outcome Scales; IMD= Index of Multiple Deprivation.

Paper 3 Table 2. Differences in HoNOS scores between 'regular cost' and 'high cost' patients.[1,2]

Variable	M(SD)			*t* (regular vs high cost)
	All patients	Regular cost	High cost	
Individual scales				
(1) Behaviour (MV = 326; 24.3%)	0.56 (0.87)	0.54 (0.85)	0.69 (0.99)	1.61
(2) Self-injury (MV = 326; 24.3%)	0.64 (0.97)	0.61 (0.93)	0.96 (1.18)	3.72
(3) Drinking/ drug use (MV = 330; 24.6%)	0.42 (0.89)	0.42 (0.90)	0.34 (0.81)	0.92
(4) Cognitive Problems (MV = 326; 24.3%)	0.35 (0.68)	0.34 (0.67)	0.43 (0.79)	1.34
(5) Physical illness (MV = 326; 24.3%)	0.83 (1.17)	0.82 (1.17)	0.92 (1.23)	0.84
(6) Hallucinations or delusions (MV = 326; 24.3%)	0.39 (0.84)	0.38 (0.83)	0.48 (0.89)	1.12
(7) Depressive symptoms (MV = 326; 24.3%)	2.15 (0.91)	2.14 (0.91)	2.26 (0.93)	1.33
(8) Other mental health problems (MV = 326; 24.3%)	2.17 (0.98)	2.16 (0.99)	2.25 (0.97)	0.84
(9) Social relationships (MV = 326; 24.3%)	1.19 (1.11)	1.18 (1.10)	1.29 (1.19)	0.98
(10) Activities of daily living (MV = 327; 24.4%)	0.82 (1.03)	0.79 (1.02)	0.97 (1.16)	1.67
(11) Living conditions (MV = 327; 24.4%)	0.58 (1.02)	0.58 (1.02)	0.62 (1.04)	0.35
(12) Occupation and activities (MV = 328; 24.4%)	1.02 (1.14)	1.01 (1.14)	1.05 (1.21)	0.37
Total HoNOS (MV= 333; 24.8%)	11.15 (5.39)	11.01 (5.34)	12.30 (5.67)	2.35*

Notes: [1]'High cost' patients were those scoring in the top 10% of the sample for SLaM mental health service costs incurred. 'Regular cost' patients were the remainder of the sample. [2]Based on complete cases analyses. HoNOS = Health of the Nation Outcome Scales. MV = Missing values.

Paper 3 Table 3. Associations of individual HoNOS items and the Total HoNOS score (at baseline) with 'regular' vs 'high' mental health service costs (at one-year follow-up).

Variable	**Odds Ratios (95% CI)**			**PAFs[1], in % (95% CI)**
	Model 1	**Model 2**	**Model 3**	**Model 3**
	(Crude)	**(Demographic)[2]**	**(Model 2 + Service)[3]**	**(Model 2 + Service)[3]**
HoNOS items[4]				
(1) Behaviour	1.17 (0.89 – 1.39)	1.10 (0.87 – 1.38)	1.06 (0.84 – 1.35)	4.5 (0.0 – 16.5)
(2) Self-injury	1.31 (1.09 – 1.59)	1.34 (1.10 – 1.63)	1.41 (1.15 – 1.72)	22.4 (9.5 – 33.3)
(3) Drinking/ drug use	0.84 (0.66 – 1.07)	0.85 (0.65 – 1.10)	0.86 (0.66 – 1.12)	-
(4) Cognitive Problems	1.11 (0.84 – 1.48)	1.10 (0.82 – 1.46)	1.08 (0.88 – 1.45)	3.6 (0.0 – 13.4)
(5) Physical illness	1.01 (0.86 – 1.35)	1.01 (0.83 – 1.29)	1.01 (0.83 – 1.21)	0.6 (0.0 – 14.5)
(6) Hallucinations or delusions	1.08 (0.63 – 1.91)	1.08 (0.85 – 1.36)	1.08 (0.85 – 1.38)	3.6 (0.0 – 12.3)
(7) Depressive symptoms	1.01 (0.79 – 1.28)	0.99 (0.77 – 1.28)	0.99 (0.77 – 1.28)	-
(8) Other mental health problems	0.98 (0.78 – 1.22)	0.97 (0.77 – 1.22)	0.98 (0.78 – 1.24)	-
(9) Social relationships	1.03 (0.85 – 1.26)	1.02 (0.83 – 1.25)	1.01 (0.82 – 1.23)	1.0 (0.0 – 20.4)
(10) Activities of daily living	1.11 (0.89 – 1.39)	1.11 (0.89 – 1.40)	1.06 (0.83 – 1.34)	7.1 (0.0 – 23.3)
(11) Living conditions	0.95 (0.78 – 1.16)	0.98 (0.80 – 1.21)	0.99 (0.80 – 1.22)	-
(12) Occupation and activities	0.95 (0.78 – 1.17)	0.94 (0.77 – 1.16)	0.95 (0.77 – 1.18)	-
Total HoNOS	1.04 (1.00 – 1.08)	1.04 (1.00 – 1.08)	1.03 (0.99 – 1.07)	33.3 (0.2 – 55.5)

Notes: [1]Due to computational necessity, population attributable fractions (PAFs) are based on complete case analyses. [2]Adjusted for the following demographic variables: age, gender, marital status, ethnicity, and Index of Multiple Deprivation score (in quintiles). [3]Adjusted for the following health service variables in addition to Model 1: service borough, previous days in contact with a SLaM mental health service, and days between baseline and HoNOS completion. [4]Items were adjusted for each other in all analyses. HoNOS = Health of the Nation Outcome Scales.

Paper 3 Table 4. Sensitivity analyses for adjusted models[1] *predicting mental health service costs.*

	Odds Ratios (95% CI)	**e^{β} (95% CI)**
Variable	**Complete cases**[2]	**Total costs (continuous)**
Individual scales[3]		
(1) Behaviour	1.05 (0.82 – 1.35)	1.05 (0.89 – 1.22)
(2) Self-injury	1.50 (1.22– 1.85)	1.17 (1.01 – 1.36)
(3) Drinking/ drug use	0.85 (0.65 – 1.11)	0.91 (0.77 – 1.06)
(4) Cognitive Problems	1.07 (0.79 – 1.45)	1.04 (0.84 – 1.28)
(5) Physical illness	1.01 (0.83 – 1.23)	0.94 (0.83 – 1.07)
(6) Hallucinations or delusions	1.06 (0.82 – 1.35)	1.04 (0.87 – 1.25)
(7) Depressive symptoms	0.95 (0.72 – 1.25)	1.04 (0.88 – 1.23)
(8) Other mental health problems	0.99 (0.78 – 1.26)	0.97 (0.83 – 1.14)
(9) Social relationships	0.97 (0.79 – 1.20)	1.07 (0.94 – 1.22)
(10) Activities of daily living	1.07 (0.84 – 1.36)	1.14 (0.97 – 1.33)
(11) Living conditions	0.99 (0.79 – 1.24)	1.01 (0.87 – 1.17)
(12) Occupation and activities	0.96 (0.77 – 1.19)	0.94 (0.82 – 1.07)
Total HoNOS	1.03 (0.99 – 1.07)	1.03 (1.01 – 1.06)

Notes: [1]Adjusted for demographic and health service variables listed in Table 3. [2] Dichotomized ('regular' vs 'high') costs outcome. [3]Items were adjusted for each other in all analyses. HoNOS = Health of the Nation Outcome Scales. e^{β} = Ratio of means, percentage increase in mean cost per unit increase in the predictor variable.

Paper 3 S1 Table. Associations of individual HoNOS items and the Total HoNOS score with inpatient admission and 'regular' vs 'high' community mental health service costs.[1]

Variable	**Odds Ratios (95% CI)**	
	Inpatient admission	Community costs
Individual scales[2]		
(1) Behaviour	1.19 (0.98 – 1.64)	1.06 (0.85 – 1.33)
(2) Self-injury	1.46 (1.11 – 1.93)	1.27 (1.04 – 1.55)
(3) Drinking/ drug use	0.98 (0.71 – 1.34)	0.80 (0.61 – 1.06)
(4) Cognitive Problems	1.09 (0.72 – 1.66)	1.07 (0.81 – 1.43)
(5) Physical illness	0.95 (0.72 – 1.26)	1.02 (0.84 – 1.22)
(6) Hallucinations or delusions	1.07 (0.76 – 1.50)	1.14 (0.91 – 1.44)
(7) Depressive symptoms	0.89 (0.61 – 1.28)	1.10 (0.85 – 1.41)
(8) Other mental health problems	0.82 (0.59 – 1.15)	1.09 (0.86 – 1.37)
(9) Social relationships	1.03 (0.77 – 1.39)	1.18 (0.96 – 1.44)
(10) Activities of daily living	1.16 (0.82 – 1.64)	1.00 (0.80 – 1.25)
(11) Living conditions	1.14 (0.86 – 1.53)	1.04 (0.86 – 1.28)
(12) Occupation and activities	1.02 (0.75 – 1.40)	0.94 (0.77 – 1.15)
Total HoNOS	1.06 (1.01 – 1.11)	1.06 (1.02 – 1.10)

Notes: [1]Adjusted for demographic and health service variables listed in Table 3. [2]Items were adjusted for each other in all analyses. HoNOS = Health of the Nation Outcome Scales.

INTRODUCTION TO PAPER 4

Addressing the second specific objective of the thesis - to determine the association of functioning with HSU by people with common mental disorders - Paper 4 encompasses a prospective cohort study with a 6-month follow-up period. The study involved a convenience sample and was set within a mood and anxiety disorders outpatient clinic in Southampton. It examined the association of functioning with HSU by NHS patients with mood and anxiety disorders, over 6 months. What set this paper aside from the others included in this thesis is that it involved primary data collection: we were not restricted by the availability of existing data and had full control over the study design and methodology, albeit within the parameters of what was feasible. Therefore, we could deploy a new ICF-based measure of functioning - PARADISE-24 - and through the use of both patient records and self-reported data, we were able to capture the full range of health services typically accessed by NHS patients within the HSU outcome. On the other hand, the use of a convenience sample limited the generalisability of the findings in terms of gender, ethnicity and socio-economic status: 61% of participants were female; the vast majority of participants were White; and most participants did not reside in relatively deprived geographical areas.

The PARADISE 24 measure of functioning was developed using probabilistic test theory and tested on over 700 participants with nine different neuropsychiatric disorders residing in four European countries (Italy, Poland, Spain and Finland)(Cieza et al., 2015). The reported psychometric properties of the metric are sound: as per infit mean square statistics, all items score in the (0.7 to 1.3) range for good item fit and the internal reliability of the instrument, indicated by the person–separation index (which has a maximum score of 1.0 and is analogous to Cronbach's Alpha) is 0.92 (Cieza et al., 2015). This 24-item self-report instrument covers functioning decrements in the following domains which can be mapped ICF classification system: psychological (12 items); activities and participation (10 items); pain (1 item); and sexual activities (1 item) (Cieza et al., 2015). As with the ICF-based WHODAS-II instrument used in Paper 2, the PARADISE 24 metric does not directly address contextual factors as per the ICF.

A manuscript describing Paper 4 is under review in the *International Clinical Psychopharmacology* journal. The PhD researcher was the lead author of the paper. Under supervision, the PhD researcher contribution included conceiving and planning the paper, data collection and management, statistical analysis, drafting all versions of the paper, and managing the submission and peer-review process.

PAPER 4: Utility of functioning in predicting costs of care for patients with mood and anxiety disorders: a prospective cohort study

Abstract

Background
Development of payment systems for mental health services has been hindered by limited evidence for the utility of diagnosis or symptoms in predicting costs of care. We investigated the utility of functioning information in predicting costs for patients with mood and anxiety disorders.

Methods
A prospective cohort study involving 102 adult patients attending a tertiary referral specialist clinic for mood and anxiety disorders. The main outcome was total costs, calculated by applying unit costs to healthcare use data.

Results
After adjusting for covariates, a significant total costs association was yielded for functioning (e^{β} = 1.02; 95% CI 1.01 – 1.03) but not depressive symptom severity nor anxiety symptom severity. When we accounted for the correlations between the main independent variables by constructing an abridged functioning metric, a significant total costs association was again yielded for functioning (e^{β} = 1.04; 95% CI 1.01 – 1.09) but not symptom severity.

Conclusions
The utility of functioning in predicting costs for patients with mood and anxiety disorders was supported. Functioning information could be useful within mental health payment systems.

Introduction

Mental disorders account for a sizeable share of the global burden of disease (~7%) but services for their treatment remain underfunded in most countries (Murray et al., 2012; Saxena, Sharan, & Saraceno, 2003). To meet the needs of patients it is important that scarce mental health service resources are allocated systematically and efficiently (Essen, 2009b). This could be achieved by healthcare payment systems which financially incentivise hospitals, clinics and other providers to treat as many patients as possible (Street & Maynard, 2007). Providers receive a fixed payment for every patient treated, adjusted for the "cluster" that patients are assigned to, based on clinical characteristics and background. The primary purpose of patient clusters is to offer an accurate estimation of the costs of treating a given patient and so the variables used in their definition must have good utility in predicting costs of care.

Diagnosis has been important in defining clusters because it facilitates their understanding as distinct clinical entities, and the utility of diagnostically-defined clusters in predicting costs of care for the general population is well-established (Busse et al., 2006; Mathauer & Wittenbecher, 2013). However, predicting costs of care in mental disorders is complex, largely due to instability in diagnosis and prognosis and wide variations in treatment and care models (Appleby et al., 2012). Moreover, the utility of diagnostically-defined clusters in predicting costs of care in mental disorders has been refuted in various large-scale studies, and this has contributed to a lack of progress in developing payment systems for mental health services around the world (Cotterill & Thomas, 2004; Elphick & Antony, 1996; English et al., 1986; A. J. Macdonald & Elphick, 2011; Mason et al., 2011; Schumacher et al., 1986). It seems necessary therefore to investigate alternatives to clustering psychiatric patients according to diagnosis. The governments of Australia, New Zealand and England have pursued a 'multi-domain' approach for defining patient clusters, principally using the Health of the Nations Outcome Scales (HoNOS). However, concerns over the validity and predictive ability of the HoNOS-based clusters have delayed the implementation of arising payment systems in these countries (Burgess et al., 1999; Eagar et al., 2004; A. Macdonald & Elphick, 2012; Wang et al., 2015). For example, pilot studies in the English National Health Service (NHS) have demonstrated the low resource homogeneity of HoNOS-based clusters, and their inferiority to an alternative statistically-derived model in reducing the variance in resource usage (*Health and Social Care Information Centre Casemix Service. Mental Health Casemix Classification Development: End Stage Report*, 2006; Alex D. Tulloch, 2012).

In the framework of the International Classification of Functioning, Disability and Health (ICF), 'functioning' is an encompassing term relating to physiological and psychological health and the ability to undertake daily activities and participate in various life domains (Cieza et al., 2014). At some point in life, everybody will experience decrements in functioning, and common decrements occur across varying health conditions (Cieza et al., 2014; Cieza et al., 2015). The utility of functioning in predicting healthcare costs of mental disorders merits investigation for various reasons: people with mental disorders typically experience substantial decrements in functioning and describe functional recovery as essential for remission (R. W. Lam et al., 2015; Zimmerman et al., 2006); poorer functioning may predict recurrence of depressive and anxiety disorders (Rodriguez et al., 2005); and functioning is increasingly recognised as a priority in the treatment and assessment of mental disorders, as reflected in the new dimensional approach of the DSM-5 whereby it is rated alongside diagnostic severity (Gold, 2014; R. W. Lam et al., 2015). The utility of functioning in predicting costs in the general population was supported in a recent review (Hopfe et al., 2016) but its utility in predicting costs in mental disorders is unclear - there is mixed evidence from investigations which deployed various domain-specific operationalisations of functioning and uncosted healthcare use outcomes (C. Cooper et al., 2010; Patel et al., 2006; C. D. Twomey, Baldwin, et al., 2015; C. D. Twomey, Prince, et al., 2015).

We carried out a cohort study set within a NHS tertiary referral specialist clinic, to investigate the utility of functioning in predicting costs for patients with mood and anxiety disorders. Functioning is measured using a new ICF-based metric (the PARADISE 24)(Cieza et al., 2015) that captures its multi-domain nature in an overall summary score. The PARADISE 24 differs from a previously developed ICF-based metric of multi-domain functioning (i.e. the WHODAS-II) in that it includes items relating to symptoms of psychological problems (e.g. anxiety and depression). We investigated the association of baseline functioning with total NHS costs at six-month follow up, and how functioning performed in comparison to depressive and anxiety symptom severity in costs prediction. We also investigated if the potential predictive ability of functioning was driven by decrements related to the ICF domains of psychological health or 'activities and participation'.

Methods

Participants

The study involved a convenience sample of adult patients attending a NHS tertiary referral specialist service for mood and anxiety disorders. Patients with cognitive, memory, or literacy difficulties that prevented their provision of data were excluded. A sample size of 103 was required to detect a medium effect size (at 80% power) while entering seven predictive variables into General Linear Models (GLM) described below (Faul, Erdfelder, Lang, & Buchner, 2007).

Procedure

The NHS London Queen Square Research Ethics Committee (reference: 14/LO/1900) and the University of Southampton (reference: 12086) provided ethics approval. Patients were invited to participate by means of a letter and accompanying information sheet. Participation involved completing questionnaires in a baseline data collection meeting, providing information about recent healthcare use by telephone at follow up, and agreeing that author CT could access electronic patient records held by the clinic. Baseline data collection meetings principally took place in the clinic after routine consultations, and, occasionally, at more convenient times in the clinic and by telephone. Participants were compensated for their time with a £10 shopping voucher.

Measures

Sample characteristics

Sample characteristics were assessed at baseline using age, gender, ethnicity, marital status, educational, employment status, Index of Multiple Deprivation score(Noble et al., 2006), general health comorbidity (Self-Administered Comorbidity Questionnaire)(Sangha, Stucki, Liang, Fossel, & Katz, 2003), psychiatric comorbidity, ICD-10 diagnosis, depressive and anxiety symptom severity, functioning, and clinician-rated severity of illness (Clinical Global Impression Scale)(Guy, 1976).

Predictor variables: Depressive and anxiety symptom severity, and functioning

Depressive and anxiety symptom severity were measured using the two seven-item subscales from the Hospital Anxiety and Depression Scale (HADS), a psychometrically-sound instrument that is widely used in clinical populations (Bjelland, Dahl, Haug, & Neckelmann, 2002; Zigmond & Snaith, 1983). Functioning was measured using PARADISE 24, a metric developed using probabilistic test theory and tested on over 700 participants with nine different neuropsychiatric disorders residing in four European countries (Italy, Poland, Spain and Finland)(Cieza et al., 2015). The reported psychometric properties of the metric are sound: as per infit mean square statistics, all items score in the (0.7 to 1.3) range for good item fit and the internal reliability of the instrument, indicated by the person–separation index (which has a maximum score of 1.0 and is analogous to Cronbach's Alpha) is 0.92 (Cieza et al., 2015). This 24-item self-report instrument covers functioning decrements in the following domains: psychological (12 items: 'not feeling rested and refreshed'; 'loss of interest'; 'appetite'; 'sleeping'; 'irritability'; 'slowed down'; 'feeling sad, low or depressed'; 'worry or anxiety'; 'not being able to cope'; 'concentration'; 'remembering to do important things'; 'making decisions'); activities and participation (10 items: 'starting and maintaining a conversation'; 'walking a long distance'; 'grooming or dressing, toileting or eating'; 'staying by yourself for a few days'; 'looking after your health'; 'initiating and maintaining a friendship'; 'getting along with people who are close to you'; 'day-to-day work or school'; 'managing your money'; 'joining in community activities'); pain (1 item); and sexual activities (1 item). Each item is scored on a three-point scale representing the level of decrements in functioning: 0 (None); 1 (Some); 2 (A lot). The raw score ranges from 0-48 before transformation into a more intuitive scale ranging from 0-100 (Cieza et al., 2015).

Outcome: Total NHS costs at six-month follow up

The primary outcome was NHS total costs, with secondary analyses conducted on the sub-categories of mental health service costs and general health service costs. Costs were calculated in two stages. First, we counted the number of contacts patients had with different NHS providers, using the combination of electronic patient records and an adapted version of the Client Service Receipt Inventory (CSRI)(Beecham & Knapp, 2001) administered by author CT at three month intervals. The electronic patient records provided data on mental health service contacts. The CSRI also covered mental health service contacts –

including contacts not documented in electronic patient records – but it was mainly used for other types of service contacts (e.g. general practitioner). Second, we converted the counted NHS contacts into monetary values (Pounds Sterling; £) by applying unit costs, principally those from NHS reference costs for 2014-2015 (DoH, 2015). As all required unit costs were not available in this source, some were based on 2015 costs provided by Personal Social Service Research Unit (PSSRU)(Curtis & Burns, 2015), and internal financial records. Table 1 details the unit costs used in this study.

Statistical analysis

Main analyses

Analyses were undertaken with STATA 13 (StataCorp LP, College Station, TX, USA). Descriptive statistics were used for sample characteristics. The separate associations of baseline depressive symptom severity, anxiety symptom severity and functioning with costs at six-month follow up were determined using unadjusted and adjusted exponentiated coefficients (with 95% confidence intervals; CIs) modelled through 'GLM-log-gamma' analysis which accounted for the skewness in the costs outcome. The skewness can be simply illustrated by the presence of markedly higher mean costs (£3899) than median costs (£1595) GLM-log-gamma is widely regarded as the analysis of choice for predicting skewed costs outcomes (with few zero values), largely because it shares the benefits of log or Box-Cox transformation while facilitating ease of interpretation of coefficients and avoiding back-transformation issues (Gregori et al., 2011). The exponentiated coefficients indicated the percentage increase in mean costs per unit increase in the specified covariate. For illustrative purposes, an exponentiated coefficient of 1.00 means a 0% increase in mean costs per unit increase in a specified covariate where as a coefficient of 1.10 indicates a 10% increase.

We sought to investigate how functioning performed in comparison to depressive and anxiety symptom severity in predicting costs. However, exploratory analysis showed that the validity of this comparison was limited by strong correlations between scores on the PARADISE 24 metric and the depression ($r = 0.73$) and anxiety ($r = 0.69$) subscales of the HADS. Thus, these variables were not entered together in statistical models. Their correlation was unsurprising because depressive and anxiety symptoms are part of functioning according to the ICF and are therefore included in the PARADISE 24 metric. To enable a more refined comparison of predictive ability, we removed items from the PARADISE 24 metric that were analogous to depressive and anxiety symptoms (i.e. items 1-9 and item 11) and analysed the association of this 'PARADISE 14' metric with costs. This procedure was not based on correlations between individual PARADISE 24 items with the HADS, rather we focused on accounting for the overlap in the content of the measures (e.g. depressive symptoms). This procedure also allowed us to investigate the whether the potential predictive ability of functioning was driven by decrements related to the ICF domains of psychological health or 'activities and participation'. Raw scores were used for analyses involving the truncated PARADISE 14 instrument because its scores could not be converted onto the same 100 point metric scale as the longer PARADISE 24.

Selection of covariates

Based on previous research showing their associations with mental health service costs (Durbin et al., 2015; C. D. Twomey, Baldwin, et al., 2015), initial adjustments were made for age, gender, marital status, ethnicity, employment status, area-level deprivation, general health comorbidity, psychiatric comorbidity, clinician-rated severity of illness, NHS costs incurred in the three months prior to baseline, functioning, depressive symptom severity and anxiety symptom severity. To safeguard statistical power, we subsequently removed several covariates that (1) were not associated with costs in exploratory analysis and (2) yielded *p* values >.20 in this association. These variables were i.e. gender, marital status, ethnicity, employment status, area-level deprivation and clinician-rated severity of illness.

Missing data

Missing cost data (three participants, 2.9%) arose due to a participant death, and two dropouts from follow up. To preserve statistical power, we imputed these missing data using multiple imputation by chained equations with a predictive mean-matching model. Multiple imputation uses patterns in observed data to impute missing values, repeating this process multiple times to account for uncertainty in the imputed values (Lee & Simpson, 2014), Imputation models included all predictive variables entered in the GLM models. A total of 100 imputed datasets were created, resulting in the introduction of minimal standard error, as per guidelines (White et al., 2011b). Checks between imputed and original values produced no anomalies. Estimates were combined using Rubin's rules (White et al., 2011a).

Results

Participation

Out of 115 clinic patients approached, 103 (90%) initially agreed to take part. One patient dropped out before providing data, leaving 102 as the final sample size. During follow up one participant died and two could not be contacted.

Sample characteristics

Table 2 provides a full summary of sample characteristics (*n* = 102). The mean age was 50.6 years. Sixty-one percent of the population were female and the vast majority were White. Marital and employment status varied. Most participants did not reside in relatively deprived geographical areas, but a sizeable minority did. Most had comorbidities, and most were diagnosed with an ICD-10 depressive disorder. Functioning scores were normally distributed. About half of participants were deemed to be "moderately ill" according to Clinical Global Impression Scale scores. The mean of total costs accrued during follow-up was £3899 (*SD* = 7997), with a median of £1595.

Associations of baseline predictor variable scores with costs at follow-up

Internal consistency for the three main predictor variables was high: PARADISE 24 (α = .93); HADS-depression (α = .86); HADS-anxiety (α = .84). Table 3 summarises the unadjusted and adjusted associations. In unadjusted models, significant total costs associations were yielded for functioning (e^{β} = 1.05; 95% CI 1.03 – 1.07) and depressive symptom severity (e^{β} = 1.10; 95% CI 1.02 – 1.18), but not for anxiety symptom severity (e^{β} = 1.08; 95% CI 0.98 – 1.18). After adjusting for age, baseline costs, and comorbidity in separate models, a significant total costs association was yielded for functioning (e^{β} = 1.02; 95% CI 1.01 – 1.03) but not depressive symptom severity (e^{β} = 1.03; 95% CI 0.98 – 1.07) nor anxiety symptom severity (e^{β} = 1.03; 95% CI 0.98 – 1.07). In the final model which accounted for the strong correlations between functioning and HADS scores by removing items from the PARADISE 24 metric that were analogous to depressive and anxiety symptoms, a similar pattern emerged: a significant total costs association was yielded for functioning (e^{β} = 1.04; 95% CI 1.01 – 1.09) but not depressive symptom severity (e^{β} = 0.99; 95% CI 0.94 – 1.05) nor anxiety symptom severity (e^{β} = 0.98; 95% CI 0.93 – 1.04). The latter finding also supported the predictive ability of functioning decrements related to 'activities and participation', over and above decrements in psychological health. In additional analyses that split the costs outcome into categories, all predictor variables had greater utility in predicting mental health service costs than general health service costs.

Discussion

Summary of main findings

After adjusting for covariates in separate models, a significant total costs association was yielded for functioning but not depressive symptom severity nor anxiety symptom severity. Interpreting the magnitude of the significant association, for every one point increase in mean PARADISE 24 score there was a 2% increase in costs. Since the PARADISE 24 has 100 points, this can be considered as a relatively strong association. In the final model which accounted for the strong correlations between functioning and HADS scores by removing items from the PARADISE 24 metric that were analogous to depressive and anxiety symptoms, a similar pattern emerged: a significant total costs association was yielded for functioning but not symptom severity. The latter finding also supported the predictive ability of functioning decrements related to 'activities and participation', over and above decrements in psychological health. All predictor variables had greater utility in predicting mental health service costs than general health service costs.

Limitations and strengths

Our study is the first to predict healthcare costs for people with mental disorders using an ICF-based measure of functioning. Sample representativeness was strengthened by a high participation rate (90%), low dropout rate (3%), and use of multiple imputation. However, statistical power was limited by the sample size of 102, which reduced the number of covariates which could be included in statistical models. The study was set within a tertiary referral specialist clinic, so the sample is not typical of all patients with mood or anxiety disorders: compared with the characteristics of patients attending secondary care mental health services in the Lambeth region of London (*n*=266,169)(Stewart et al., 2009), the sample had a higher mean age and more participants of female gender and White ethnicity. Furthermore, the service is led by author DSB who makes treatment based on clinical judgment: different treatment decisions may be made in other services which could limit the generalisability of our findings. Unavoidable practical issues meant that electronic patient records only covered contacts with mental health services and the remaining health service use data was collected using the CSRI which may have been subject to recall errors. The precision and applicability of unit costs data is limited due to various data access issues: unit costs data were not available for all types of NHS contacts and thus it was necessary to approximate the costs of some contacts using available unit costs from similar services (Table 1); the data were based on national averages and may not be applicable to certain NHS services; and the data had to be extracted from two different sources. The six-month timeframe for the analysis of costs does not take into account possible seasonal effects on resource use to the same degree as studies lasting over one year. The duration of our timeframe was constrained by the amount of costs data that needed to be collected using the CSRI (at 3 month intervals) and other practical considerations.

Comparison with other studies

Comparisons of our findings with those from relevant previous studies are tentative because these studies had differing clinical populations and deployed various domain-specific operationalisations of functioning and uncosted healthcare use outcomes (C. Cooper et al., 2010; Patel et al., 2006; C. D. Twomey, Baldwin, et al., 2015; C. D. Twomey, Prince, et al., 2015). Domain-specific operationalisations of functioning may be less representative of care needs than the multi-domain PARADISE 24 while uncosted healthcare use outcomes do not provide a weighted summary of resource consumption and are therefore less precise than costed outcomes.

Nevertheless, our findings which support the predictive utility of functioning correspond with those of a cross-sectional study (*n* = 7461) whereby 'activities of daily living' was associated with the number of psychotherapy and GP attendances by people with 'common mental disorders' (C. Cooper et al., 2010) but not with those of a cohort study (*n* = 85) whereby social functioning impairment was not associated with healthcare costs for people with schizophrenia (Patel et al., 2006). The utility of functioning in predicting costs in the general population has also been supported in a recent review (Hopfe et al., 2016). The lack of strong support for the utility of depressive and anxiety symptom severity in predicting costs accords with evidence from numerous studies involving large-scale and national datasets (Cotterill & Thomas, 2004; Elphick & Antony, 1996; English et al., 1986; A. J. Macdonald & Elphick, 2011; Prina et al., 2015; Schumacher et al., 1986; C. D. Twomey, Baldwin, et al., 2015).

Potential implications

Our findings support the utility of functioning in predicting costs for patients with mood and anxiety disorders, and this may have implications for health policy-makers. The PARADISE 24 metric benefits from its theoretical underpinnings in the ICF and is short, easy-to-use, and applicable across mental disorders. It should be noted though that the overlap of items from the PARADISE 24 with measures of diagnostic and symptom severity and quality of life needs to be taken into account to prevent the confounding of estimates in analysis. Our findings supporting the utility of functioning decrements relating to ‘activities and participation’ in costs prediction show that adding functioning information to existing diagnostically-defined clusters may improve their predictive ability, as has been demonstrated in the general population (Hopfe et al., 2016).

Future research

A more diverse clinical sample would enable a more complete assessment of the predictive utility of functioning. The use of a more comprehensive case-register could increase the validity of the costs outcome - although the widespread absence of data linkage between primary care, secondary care, and hospital case-registers represents a drawback (Garcia Alvarez et al., 2011). The HoNOS instrument has been proposed for use in the mental health PbR system of the English NHS (Lovaglio & Monzani, 2011, 2012; Self et al., 2008; Speak & Muncer, 2015; Wang et al., 2015) and a future study directly comparing its utility in with that of the PARADISE 24 metric in costs prediction might inform policy debates in England and other countries. The PARADISE 24 was developed as a self-report measure that can be used across all mental disorders, but for more complex and severe disorders, ‘self-report’ is often not possible: future investigations of the psychometric properties and practicality of a clinician-rated version of the PARADISE 24 metric would be welcome. Future research could explore alternative approaches to developing payment systems for mental health services: for example, Monitor – the NHS regulator – has suggested payments should be closely linked to agreed patient outcome standards rather than costs, to incentivise quality of care (Monitor, 2015).

Paper 4 Table 1. Unit costs used in study

Contact	**Source of unit cost**	**Terminology used in source**	**Cost**
Mental health			
Clinical psychologist	NHS Reference Costs	Clinical psychology	£210
CMHT other	PSSRU	Community-based nurse (mental health)	£67
Crisis resolution team	PSSRU	Crisis resolution team member (mental health)	£189
Day care attendances	PSSRU	Local authority social services (mental health)	£32
Drug and alcohol misuse	NHS Reference Costs	Alcohol outpatient attendances	£81
ECT	Southern Health NHS FT	Southampton ECT Team	£151
IAPT (primary care)	NHS Reference Costs	IAPT, Adult and Elderly	£94
Psychiatric hospital days	NHS Reference Costs	Care clusters unit cost per occupied bed day[1]	£493
Psychiatric liaison	NHS Reference Costs	Liaison psychiatry	£142
Psychiatrist	NHS Reference Costs	Adult Mental Illness	£225
Social worker	PSSRU	Social worker (adult services)	£55
General health			
A&E attendance	NHS Reference Costs	A&E attendance[2]	£132
Dentist	NHS Reference Costs	General dental service	£77
Dermatologist	NHS Reference Costs	Dermatology	£97
District nurse	NHS Reference Costs	District nurse	£37
Dietician / nutritionist	NHS Reference Costs	Dietician	£83
GP	PSSRU	GP patient contact (11.7 minutes)	£44
Hospital bed days	NHS Reference Costs	Excess bed days	£303
Hospital Doctor (general)	NHS Reference Costs	General medicine (hospital-based consultant)	£149
Hospital Nurse	PSSRU	Hospital-based nurse: band 5	£88
Nurse (GP practice)	PSSRU	Nurse (GP practice)	£47
Occupational therapist	NHS Reference Costs	Occupational therapist	£73
Orthopaedic surgeon	NHS Reference Costs	Trauma & orthopaedics	£104
Optometry	NHS Reference Costs	Optometry	£93
Physiotherapist	NHS Reference Costs	Physiotherapist	£52
Podiatrist	NHS Reference Costs	Podiatrist, Tier 1, General Podiatry	£40

Note: [1]The mean of the care clusters unit cost per occupied bed day was calculated and used. A&E = Accident & Emergency. [2]The mean cost of different types of A&E attendances was calculated and used. CMHT other = Community Mental Health Team: Nurse / Support Worker / Assistant Psychologist / Care co-ordinator / Occupational therapist. ECT= Electroconvulsive therapy. FT = Foundation Trust. GP = General Practitioner IAPT = Improving Access to Psychological Therapies initiative. NHS = National Health Service. PSSRU= Personal Social Service Research Unit.

Paper 4 Table 2. Sample characteristics (N = 102)

Variable	***n* (%)**	***M* (*SD*)**	***Median (25th, 75th centile)***
Age[1]		50.6 (13.5)	52.0 (42.8 – 60.3)
Gender			
Male	40 (39.2)		
Female	62 (60.8)		
Ethnicity			
White	96 (94.1)		
Non-white	6 (5.9)		
Marital status			
Single	28 (27.5)		
Married or in civil union	54 (52.9)		
Divorced, separated or widowed	20 (19.6)		
Employment status			
In paid employment	35 (34.3)		
Unemployed or unable to work	41 (40.2)		
Retired	23 (22.6)		
Student	3 (2.9)		
Index of Multiple Deprivation decile[2]		6.3 (2.6)	6.0 (4.0 – 8.0)
General health comorbidity (SCQ)		5.3 (4.6)	5.0 (2.0 – 8.0)
ICD-10 Diagnosis			
Depressive disorder (F31-F34; F38.10)	55 (53.9)		
Anxiety disorder (F40-F42)	16 (15.7)		
Bipolar disorder (F31)	28 (27.5)		
Other	3 (2.9)		
Psychiatric comorbidity			
Yes	50 (49.0)		
No	52 (51.0)		
Clinical Global Impression score		3.7 (0.8)	4.0 (3.0 – 4.0)

Notes: [1]Age range is 18-79. [2]The lower the decile score, the higher the relative deprivation in the area. SCQ = Self-Administered Comorbidity Questionnaire. NHS = National Health Service.

Paper 4 Table 3. Associations of baseline HADS-depression, HADS-anxiety, HADS-Total and functioning (PARADISE 24 and 'PARADISE 14') with '6 month' costs (n = 102)

***Predictor variable* and covariates**	e^{β} **(95% CI)** **MH service costs**	**GH service costs**	**Total costs**
HADS- depression			
Unadjusted	1.11 (1.02 – 1.21)	1.02 (0.95 – 1.09)	1.10 (1.02 – 1.18)
Age, baseline total costs, comorbidity[1]	1.05 (0.99 – 1.11)	0.98 (0.94 – 1.03)	1.03 (0.98 – 1.07)
Further adjustment for anxiety, 'PARADISE 14' scores	1.02 (0.95 – 1.08)	0.99 (0.93 – 1.05)	0.99 (0.94 – 1.05)
HADS anxiety			
Unadjusted	1.09 (0.97 – 1.23)	0.99 (0.92 – 1.07)	1.08 (0.98 – 1.18)
Age, baseline total costs, comorbidity[1]	1.04 (0.99 – 1.10)	0.96 (0.92 – 1.01)	1.03 (0.98 – 1.07)
Further adjustment for depression, 'PARADISE 14' scores	0.99 (0.93 – 1.06)	0.95 (0.89 – 1.01)	0.98 (0.93 – 1.04)
Functioning			
Unadjusted	1.05 (1.03 – 1.08)	1.06 (0.99 – 1.03)	1.05 (1.03 – 1.07)
Age, baseline total costs, comorbidity[1]	1.03 (1.01 – 1.05)	1.00 (0.98 – 1.01)	1.02 (1.01 – 1.03)
***Functioning ('PARADISE 14')**[2,3]*			
Unadjusted	1.14 (1.07 – 1.20)	1.03 (0.98 – 1.08)	1.11 (1.06 – 1.16)
Age, baseline total costs, comorbidity[1]	1.06 (1.02 – 1.11)	0.99 (0.96 – 1.02)	1.04 (1.01 – 1.07)
Further adjustment for depression, anxiety	1.06 (1.00 – 1.11)	1.02 (0.97 – 1.06)	1.04 (1.01 – 1.09)

Notes: e^{β} = Exponentiated coefficients - modelled using 'GLM-log-gammas'. CGI= Clinical Global Impressions scale. GH = General Health; MH= Mental Health; HADS = Hospital Anxiety and Depression Scale. [1]Adjustments were made for both general health and psychiatric comorbidity. [2]The PARADISE 14 was formed after removing items from the PARADISE 24 considered to be symptoms of mood disorders (i.e. items 1-9 and item 11). [3]Raw scores were used for analysis involving this truncated instrument.

DISCUSSION

Summary of main findings

The papers described within this thesis together investigated the utility of functioning in predicting HSU by people with common mental disorders. It was hypothesised that functioning would have good utility in this prediction. Paper 1 - a systematic review of the predictors of HSU by adults with mental disorders in the UK - identified a lack of existing research surrounding the association of functioning with HSU. The available evidence was mixed: three functioning variables (i.e. cognitive deficits, social withdrawal and activities of daily living) were significantly associated with HSU, but only in 55% of assessments conducted. Paper 2 - a cross-sectional analysis of the associations of functioning with HSU by older people in LMICs (n = 4,950) - also yielded mixed evidence in relation to the predictive utility of functioning. Here, functioning was significantly associated with one HSU outcome (i.e. hospital admission) but not another (i.e. "any community HSU"). Paper 3 - a historical cohort study investigating the utility of the HoNOS in predicting mental health services costs for patients with common mental health problems (n = 1,343) – did not support the predictive utility of functioning: all seven functioning items from the HoNOS were not associated with costs. Paper 4 - a prospective cohort study investigating the utility of functioning in predicting costs for patients with mood and anxiety disorders (n = 102) - provided support for functioning's predictive utility: unlike depressive and anxiety symptom severity, functioning predicted total NHS costs over a six month period. Overall, the mixed findings across the four papers provide partial support for the hypothesised predictive ability of functioning. The ambiguity reflects the challenge of predicting HSU, and warrants considered explanation.

Generalisability of the findings

Before further discussion of the overall findings, I outline the limits to their generalisability. Regarding Paper 1 (i.e. the systematic review), restricting its scope to UK studies facilitated closer application of its findings to the development of the UK's PbR system but it lessened generalisability to other countries. Regarding Paper 2 (i.e. the cross-sectional study), the investigation of an under-researched population residing in countries without universal healthcare coverage (i.e. older adults with elevated depressive symptoms in LMICs) increased the likelihood that the findings will influence policy development, but it reduced generalisability to younger adults and people residing in high income countries such as the UK. Regarding Paper 3 (i.e. the historical cohort study), the study population was chosen because it is representative of relatively deprived patients with residing in the UK and thus the findings are applicable to similar populations elsewhere; less so to people of higher socio-economic status, and people residing outside the UK. Regarding Paper 4 (i.e. the prospective cohort study), primary data collection meant that we were not restricted by the availability of existing data and had control over the study design and methodology; but the the use of a convenience sample limited the generalisability of the findings: for example, in terms of gender, ethnicity and socio-economic status the sample was not representative of the general UK population. In light of the marked generalisability issues across the four papers, interpretations of the overall findings should be made with caution: the programme of research is best thought of as exploratory in nature; it may encompass a pre-requisite to the development of a more systematic and clearly-focused programme.

Explaining the ambiguity in the findings

Why was the support for the predictive ability of functioning in the final study not seen in the earlier studies? There are various plausible explanations. The systematic review was limited to UK studies and only identified a few investigations of the association of functioning with HSU. If the review had broadened its focus beyond the UK to include international investigations, it undoubtedly would have identified more relevant studies, and it is possible that stronger support for functioning's predictive ability may have been identified. The decision to limit the review to UK studies allowed the findings to be closely linked to the development of the PbR system in the UK, but this was largely a pragmatic decision. In retrospect, the pragmatic hurdle to the inclusion of international studies could have been overcome by narrowing the scope of the review to the sole investigation of the predictive ability of functioning (and not other variables), or by only including studies with one targeted HSU outcome (e.g. hospital admission). As is stands, the findings of the review only provide tentative evidence in relation to the utility of functioning in predicting HSU.

A potential reason for the findings of the cross-sectional study - that contrasted with the positive findings of the final cohort study - concerns its setting. Unlike the final cohort study which was set in a high-income country (i.e. the UK), the cross-sectional study was set across nine LMICs, which typically had no universal healthcare coverage and relatively poor service infrastructure (e.g. a lack of adequate public transport). In LMICs, several sociodemographic variables relating to individual access to health services (e.g., private health insurance coverage, pension coverage, education) are very important determinants of HSU (Albanese et al., 2011). Moreover, their importance may partly explain the lack of strong support yielded for utility of functioning in this prediction. Another reason for the discrepancy in findings concerns the operationalisation of HSU outcomes. In the final cohort study a weighted summary of resource consumption was provided through the use of healthcare costs but in the cross-sectional study a crude categorical HSU outcome variable was deployed: participants were asked to recall whether or not they had a contact with any community health services or if they had a hospital admission, in the three months prior to baseline. It is plausible that the crudeness of the latter HSU outcomes decreased the chances of strong functioning-HSU associations being detected, although the study did have a very large sample size which may have offset this limitation somewhat (*n* = 4,950).

The absence of yielded costs associations for all seven functioning times from the HoNOS in first cohort study directly contrasted with the positive findings of the final cohort study. The HoNOS is being proposed for use in the PbR system of the English NHS, so investigation of its predictive ability allowed us to undertake a paper with direct policy implications. The validity of the HoNOS, though, has been much criticised: several factor structures have been proposed but none of these have acceptable fit (Speak & Muncer, 2015); Rasch analyses demonstrate the absence of an underlying construct in the composite scale (Lovaglio & Monzani, 2011, 2012); there is no universal agreement regarding its operationalisation; and previous research investigating its utility in predicting costs has yielded mixed evidence (Andreas et al., 2010; Callaly et al., 2011; Goldney et al., 1998; Kisely et al., 2010; A. D. Tulloch et al., 2015; A. D. Tulloch et al., 2012). It is possible therefore that the questioned validity and utility of the HoNOS instrument was an important contributory factor in the absence of yielded costs associations. Moreover, individual HoNOS items may not be representative of overall functioning impairment and thus may have limited utility in predicting costs. Indeed, one-item impairment and severity scales have generally been found to have limited validity and predictive utility (Blozik et al., 2013). In comparison, the final study utilised the ICF-based PARADISE 24 metric that has better psychometric qualities than the HoNOS and captures the multi-domain nature of functioning in an overall summary score (Cieza et al., 2015). Another explanation for the contrasting results of this study concerns rating systems: the HoNOS metric is rated by clinicians, yet it has been argued that to understand the true impact of health conditions on individual functioning, it is essential to collect information directly from people experiencing such conditions (Cieza et al., 2015). Therefore, it is possible that in comparison to self-report measures such as the PARADISE 24, the clinician-rated HoNOS may be less representative of functioning and HSU. The representativeness of the HoNOS was limited further by the manner of data collection in the study: HoNOS data were extracted from routine patient records which were completed by busy clinicians, many of whom may not have been highly motivated to score the HoNOS with high levels of precision. Potential inaccuracies in HoNOS scores may have contributed to the failure of functioning items to significantly predict costs.

The PARADISE 24 metric - deployed in the final cohort study - is based on the ICF and covers functioning in a variety of domains: psychological (12 items), activities and participation (10 items), pain (1 item) and sexual activities (1 item) (Cieza et al., 2015). The multi-dimensionality of the PARADISE 24 metric contrasts with the domain-specific measures of functioning reported by the systematic review and the HoNOS items deployed in the first cohort study. Its composition also contrasts, albeit less so, with the WHODAS-II (Rehm et al., 1999) deployed in the cross-sectional study. Like the PARADISE 24 metric, the WHODAS-II is grounded in the ICF and covers functioning in a variety of domains (cognition, mobility, self-care, getting along with others, life activities, participation); but there are no items relating to psychological problems such as anxiety and depressive symptoms, and the WHODAS-II has half as many items (12) as the PARADISE 24 metric (24). Overall, due to its broadness and incorporation of psychological problems, the PARADISE 24 metric deployed in the final study may be more representative of overall functioning than the ‘narrower’ instruments deployed in earlier studies, and this may explain its greater utility in predicting HSU.

Alternatively, the mixed findings may indicate that while functioning has some utility in predicting HSU for people with common mental disorders, the level of this utility may be limited. Moreover, the systematic review and cross-sectional studies highlighted the importance of a wide range socio-demographic predictors (e.g. age, gender, comorbidity, access to healthcare). Functioning may have a similar level of predictive utility as psychiatric diagnosis or symptom severity, which have yielded small HSU associations in numerous large-scale studies (Cotterill & Thomas, 2004; Elphick & Antony, 1996; English et al., 1986; A. J. Macdonald & Elphick, 2011; Prina et al., 2015; Schumacher et al., 1986; C. D. Twomey, Baldwin, et al., 2015). Undoubtedly, the prediction of HSU by people with common mental disorders represents a challenge for both researchers and policy-makers.

Implications for policy-makers

The development of payment systems for mental health services has been hindered by a lack of strong evidence for the utility of diagnosis, symptoms or alternative approaches (e.g. the HoNOS, as deployed by the English NHS) in predicting HSU (Cotterill & Thomas, 2004; Elphick & Antony, 1996; English et al., 1986; A. J. Macdonald & Elphick, 2011; Prina et al., 2015; Schumacher et al., 1986; C. D. Twomey, Baldwin, et al., 2015; Wang et al., 2015). Therefore, the findings partially supporting the utility of functioning in predicting HSU for patients with common mental disorders may have important implications for policy-makers.

If functioning is to be used in payment systems, the way in which it is measured is an important consideration. Domain-specific and narrow operationalisations of functioning evidently have limited utility in predicting HSU. On the other hand, the PARADISE 24 metric captures the multi-dimensionality of functioning and thus is likely to be more representative of overall health, care needs, and HSU. The PARADISE 24 metric also benefits from its theoretical underpinnings in the ICF and is short, easy-to-use, and applicable across mental disorders. The metric could therefore be useful for the formation of patient clusters within mental health payment systems.

As a wide range of variables together predict the HSU of people with common mental disorders, it seems prudent for policy-makers to explore multi-domain approaches to the definition of patient clusters within payment systems. Functioning in combination with diagnosis, symptom severity and socio-demographic variables could be used in this definition. There exists some encouraging evidence in relation to this approach: investigations of attempts to implement casemix classification systems in Australia (Burgess et al., 1999) and New Zealand (Eagar et al., 2004) support the validity and predictive utility of patient clusters formed of the HoNOS and other ‘casemix’ variables (e.g. age, ethnicity, diagnosis, and focus of care). The recently developed Australian Mental Health Care Classification (AMHCC) system also clusters patients using multi-domain information, incorporating functioning - via the HoNOS - into its casemix classes (https://www.ihpa.gov.au/what-we-do/mental-health-care). Another ‘multi-domain’ proposal is for clinicians to judge the most appropriate care pathway option based on the detailed assessment of problems in nine domains of mental health and 12 domains of everyday living, using the MRC Needs For Care Assessment Schedule (Brewin et al., 1987; Wang et al., 2015).

Finally, the findings supporting the predictive ability of functioning decrements relating to ‘activities and participation’ - arising from analyses involving the WHODAS II (Paper 2) and PARADISE 24 (Paper 4) measures - point to the possibility that adding functioning information to existing diagnostically-defined clusters may improve their predictive ability. Findings from a recent review support the credibility of this suggestion: across 10 included studies with large general population samples, adding functioning information to existing patient clusters increased their precision in the prediction of costs (Hopfe et al., 2016)

Implications for clinicians

Previous research indicated that people with mental disorders experience substantial decrements in functioning (R. W. Lam et al., 2015; Zimmerman et al., 2006) and that poorer functioning predicts the recurrence of depressive and anxiety disorders (Rodriguez et al., 2005). Alongside this previous research, the (partial) support for the utility of functioning in predicting HSU yielded in the current thesis highlights the need for clinicians to recognise functioning as a priority in the treatment and assessment of mental disorders, as reflected in the new dimensional approach of the DSM-5 whereby it is rated alongside diagnostic severity (Gold, 2014; R. W. Lam et al., 2015). Considerable work is needed to achieve greater recognition here though: the authors of a recent review of the effectiveness of psychotherapy and pharmacotherapy for functioning concluded - based on modest treatment effects - that future tailoring of interventions is warranted in order to address functional impairment in people wuth mental disorders (Kamenov, Twomey, Cabello, Prina, & Ayuso-Mateos, 2017). One promising example of such an intervention is behavioural activation for depression: a simple and emprically-supported intervention which broadly aims to improve functioning in the areas of ‘activities and participation’ in everyday life, in order to lift mood (Ekers et al., 2014).

Future research

Regarding Paper 1, a systematic review not restricted by international region investigating the utility of functioning in predicting HSU by people with mental disorders would help to clarify the uncertainty in relation to functioning’s predictive utility. A meta-analysis would also be beneficial, although validity and statistical feasibility would likely be limited by wide variations in the HSU outcomes deployed by identified studies. A solution would be to restrict the focus of the meta-analysis to one HSU outcome, as in a recent meta-analysis of the association of depressive symptom severity with hospital admissions (Prina et al., 2015). The HSU of people with intellectual disabilities was not examined in Paper 1, largely for reasons related to precision in analysis. Future research surrounding of the prediction of HSU should be directed towards this population though, especially given their considerable care needs.

Attempts to replicate the cross-sectional study (Paper 2) using longitudinal data would make it possible to determine if functioning predicts HSU by older people with depression in LMICs over time. Such replications would have greater validity if they provided a weighted summary of resource consumption through the use of costs outcomes rather than the specific uncosted outcomes we deployed. This is especially the case given that – as highlighted in Paper 1 - there exist at least 60 different HSU outcomes which have limited comparability with each other (C. D. Twomey, Baldwin, et al., 2015). More generally, the importance of socio-demographic variables related to personal access to care in the prediction of HSU in LMICs - resulting from the absence of universal healthcare and poor infrastructure - should be investigated further. Moreover, Paper 2 could be replicated within high income countries where personal access variables are less likely to confound the relationship between functioning and HSU than in LMICs (Albanese et al., 2011; C. D. Twomey, Prince, et al., 2015).

Regarding Paper 3, the HoNOS instrument is being proposed for use in the mental health PbR system of the English NHS (Lovaglio & Monzani, 2011, 2012; Self et al., 2008; Speak & Muncer, 2015; Wang et al., 2015) and a study directly comparing its utility with that of the PARADISE 24 metric in costs prediction might inform policy debates in England and other countries. Discrepancies between the largely null results of Paper 3 with those from Australia (Burgess et al., 1999) and New Zealand (Eagar et al., 2004), point to the need for further international investigations of the utility of the HoNOS (and its functioning items) in the prediction of HSU. This paper's scope was limited by the sole use of a mental health costs outcome that was necessitated by data availability; future investigations in the area could examine other costs typically incurred by people with mental health disorders (e.g. via primary care), although the lack of data linkages between case-registers represents an impediment to this data collection (Garcia Alvarez et al., 2011). Given that the composite HoNOS is the most widely used routine outcome measure within NHS mental health services, future investigations of their utility are both feasible and necessary.

Owing to its convenience sample, Paper 4 can be considered a pilot study of the utility of the PARADISE 24 metric in predicting HSU: future studies involving more representative clinical samples would enable a more complete assessment of the predictive ability. One location where such a study could be undertaken is Finland: ten Finnish community organisations for substance abuse and mental ill health are currently collaborating on a project involving the administration of the PARADISE 24 metric and measures relating HSU to over 1,000 patients (Pitkänen, Levola, Tourunen, Kaskela, & Holopainen, 2016). The PARADISE 24 was developed as a self-report measure that can be used across all mental disorders, but for more complex and severe disorders, 'self-report' is often not possible: future investigations of the predictive utility of a clinician-rated version of the PARADISE 24 metric would be welcome.

The decision to focus on common mental disorders in this thesis added precision to the findings, but it was taken for mainly pragmatic reasons - it was more feasible to collect and access data pertaining to people with common mental disorders than those with more complex problems. More diverse clinical samples - involving people with both 'serious mental illness' (e.g. psychosis) and common mental disorders - would enable a more comprehensive assessment of the utility of functioning in predicting HSU. Future research should also investigate if multi-domain measures of functioning have greater predictive utility then domain-specific measures of functioning, as indicated by the findings across the four studies. Such research may yield greater support for the hypothesised predictive ability of functioning than the overall findings of this thesis.

The functioning measures deployed in this programme of research - the PARADISE 24 (Cieza et al., 2015), WHODAS-II (Rehm, Ustun, Saxena, Nelson, & Chatterji, 1999) and the HoNOS - (Wing et al., 1998) - did not address the contextual factors of functioning, as per the ICF. These factors (e.g. age, gender, ethnicity, geographic location, and perceived stigma) have been evidenced to have a profound impact on health and HSU, in both the general population and for people with mental disorders (Albanese et al., 2011; Bornbaum et al., 2013; Casey et al., 2003; Corrigan & Watson, 2002; Kessler et al., 2007; Kotov et al., 2010; Prince, Patel, et al., 2007). Therefore: (1) it is likely that the incorporation of contextual factors within deployed measures of functioning would have led to a greater proportion of HSU being accounted for, across the four papers; (2) future research should investigate the feasibility of developing an ICF-based measure of functioning that incorporates contextual factors. The challenge here is that the bio-psycho-social model underpinning the ICF has arguably been subject to 'undisciplined eclecticism', with no safeguards in place against the under- and over- representation of its different elements (Benning, 2015). Indeed, unlike other elements of the ICF which are better developed, there are no clear taxonomical classifications for contextual factors (Bornbaum et al., 2013). Developing functioning measures that incorporate the diverse array of contextual factors - inconsistently measured in existing research – is easier said than done.

CONCLUSIONS

The mixed findings across the four papers provide partial support for the hypothesised utility of functioning in predicting HSU for people with common mental disorders. The ambiguity in the findings is reflective of various methodological issues: the systematic review was limited to UK studies and identified relatively few investigations of the association of functioning with HSU; the cross-sectional study was set in LMICs where sociodemographic variables relating to access to health services may have confounded the association of functioning with HSU; in the historical cohort study, potential inaccuracies related to HoNOS scores may have contributed to the failure of functioning items to predict costs; due to its broadness and incorporation of psychological problems, the PARADISE 24 metric deployed in the final cohort study may be more representative of overall functioning than the 'narrower' instruments deployed in earlier studies, and this may explain its greater yielded utility in predicting HSU.

Due to these important methodological issues, the programme of research is probably best thought of as exploratory in nature. Nevertheless, the findings may have important implications for policy-makers in relation to the prediction of HSU: for example, narrow operationalisations of functioning evidently have limited predictive ability; but the ICF-based PARADISE 24 metric which captures the multi-dimensionality of functioning in an easy-to-use 24-item scale appears to have strong predictive ability. Moreover, the findings (partially) supporting functioning's predictive ability - alongside previous research demonstrating that people with mental disorders experience substantial decrements in functioning and that poorer functioning predicts the recurrence of depressive and anxiety disorders (Rodriguez et al., 2005) - emphasise the need to consider functioning as a priority in in the treatment and assessment of mental disorders.

The programme of research has also identified several avenues for future research. For example, a systematic review not restricted by international region would help to clarify the uncertainty in relation to functioning's predictive utility. Moreover, the cross-sectional study could be replicated within high income countries where personal access variables are less likely to confound the relationship between functioning and HSU than in LMICs. The HoNOS instrument is being proposed for use in the mental health PbR system of the English NHS and a study directly comparing its utility with that of the PARADISE 24 metric in costs prediction might inform policy debates in England and other countries. Finally, the functioning measures deployed in this programme of research did not address the contextual factors of functioning, as per the ICF. These factors have been evidenced to have a profound impact on health and HSU and future research should investigate the feasibility of developing an ICF-based measure of functioning that incorporates them.

GLOSSARY OF TERMS

HoNOS	Health of the Nations Outcome Scales
HSU	health service use
ICF	International Classification of Functioning, Disability & Health
LMICs	low and middle income countries
NHS	National Health Service
PbR	Payment by Results
WHO	World Health Organisation
WHODAS-II	World Health Organisation Disability Assessment Schedule

REFERENCES

American Psychiatric Assoication (APA) (2010). Practice guideline for the treatment of patients with major depressive disorder, third edition. Retrieved from http://www.guideline.gov/content.aspx?id=24158

Albanese, E., Liu, Z., Acosta, D., Guerra, M., Huang, Y., Jacob, K. S., . . . Prince, M. J. (2011). Equity in the delivery of community healthcare to older people: findings from 10/66 Dementia Research Group cross-sectional surveys in Latin America, China, India and Nigeria. *BMC Health Services Research, 11*, 153-153. doi:10.1186/1472-6963-11-153

Amner, K. (2012). The effect of DBT provision in reducing the cost of adults displaying the symptoms of BPD. *British Journal of Psychotherapy, 28*(3), 336-352. doi:10.1111/j.1752-0118.2012.01286.x

Andreas, S., Harries-Hedder, K., Schwenk, W., Hausberg, M., Koch, U., & Schulz, H. (2010). Is the Health of the Nation Outcome Scales appropriate for the assessment of symptom severity in patients with substance-related disorders? *J Subst Abuse Treat, 39*(1), 32-40. doi:10.1016/j.jsat.2010.03.011

Andrews, G., Kemp, A., Sunderland, M., Von Korff, M., & Ustun, T. B. (2009). Normative Data for the 12 Item WHO Disability Assessment Schedule 2.0. *PLoS One, 4*(12), e8343. doi:10.1371/journal.pone.0008343

Appleby, J., Harrison, T., Hawkins, L., & Dixon, A. (2012). *Payment by Results- how can payment systems help deliver better care?* Retrieved from London:

Ballard, C., Powell, I., James, I., Reichelt, K., Myint, P., Potkins, D., . . . Barber, R. (2002). Can psychiatric liaison reduce neuroleptic use and reduce health service utilization for dementia patients residing in care facilities? *International Journal of Geriatric Psychiatry, 17*(2), 140-145. doi:10.1002/gps.543

Bateman, A., & Fonagy, P. (2008). 8-Year Follow-Up of Patients Treated for Borderline Personality Disorder: Mentalization-Based Treatment Versus Treatment as Usual. *American Journal of Psychiatry, 165*, 631-638.

Beecham, J., & Knapp, M. (2001). Costing psychiatric interventions. In G. Thornicroft, C. Brewin, & J. K. Wing (Eds.), *Measuring Mental Health Needs* (2 ed., pp. 163-183). London: Gaskell.

Beekman, A. T. F., Penninx, B. W. J. H., Deeg, D. J. H., Beurs, E. d., Geerlings, S. W., & Tilburg, W. v. (2002). The impact of depression on the well-being, disability and use of services in older adults: a longitudinal perspective. *Acta Psychiatrica Scandinavica, 105*(1), 20-27. doi:10.1034/j.1600-0447.2002.10078.x

Bekas, S., & Michev, O. (2013). Payment by results: validating care cluster allocation in the real world. *The Psychiatrist, 37*(11), 349-355. doi:10.1192/pb.bp.112.041780

Benning, T. B. (2015). Limitations of the biopsychosocial model in psychiatry. *Advances in Medical Education and Practice, 6*, 347-352. doi:10.2147/AMEP.S82937

Berk, M. L., & Monheit, A. C. (2001). The Concentration Of Health Care Expenditures, Revisited. *Health Affairs, 20*(2), 9-18. doi:10.1377/hlthaff.20.2.9

Bjelland, I., Dahl, A. A., Haug, T. T., & Neckelmann, D. (2002). The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *Journal of Psychosomatic Research, 52*(2), 69-77.

Blozik, E., Scherer, M., Lacruz, M., Ladwig, K.-H., & group, t. K. s. (2013). Diagnostic utility of a one-item question to screen for depressive disorders: results from the KORA F3 study. *BMC Family Practice, 14*(1), 198.

Bornbaum, C. C., Doyle, P. C., Skarakis-Doyle, E., & Theurer, J. A. (2013). A critical exploration of the International Classification of Functioning, Disability, and Health (ICF) framework from the perspective of oncology: recommendations for revision. *Journal of Multidisciplinary Healthcare, 6*, 75-86. doi:10.2147/JMDH.S40020

Brewin, C. R., Wing, J. K., Mangen, S. P., Brugha, T. S., & MacCarthy, B. (1987). Principles and practice of measuring needs in the long-term mentally ill: the MRC Needs for Care Assessment. *Psychol Med, 17*(04), 971-981. doi:doi:10.1017/S0033291700000787

Burgess, P., Pirkis, J., Buckingham, W., Eagar, K., & Solomon, S. (1999). Developing a casemix classification for specialist mental health services. *Casemix Quarterly, 1*(4), 4-20.

Busse, R., Schreyögg, J., & Smith, P. C. (2006). Editorial: Hospital case payment systems in Europe. *Health Care Management Science, 9*(3), 211-213. doi:10.1007/s10729-006-9039-7

Button, E. J., Benson, E., Nollett, C., & Palmer, R. L. (2005). Don't forget EDNOS (eating disorder not otherwise specified): Patterns of service use in an eating disorders service. *Psychiatric Bulletin, 29*(4), 134-136.

Byford, S., Barrett, B., Despiégel, N., & Wade, A. (2011). Impact of treatment success on health service use and cost in depression: longitudinal database analysis. *Pharmacoeconomics, 29*(2), 157-170. doi:10.2165/11537360-000000000-00000

Byrne, S. L., Hooke, G. R., & Page, A. C. (2010). Readmission: a useful indicator of the quality of inpatient psychiatric care. *Journal of Affective Disorders, 126*(1-2), 206-213. doi:10.1016/j.jad.2010.02.135

Callaly, T., Trauer, T., Hyland, M., Coombs, T., & Berk, M. (2011). An examination of risk factors for readmission to acute adult mental health services within 28 days of discharge in the Australian setting. *Australasian Psychiatry, 19*(3), 221-225. doi:10.3109/10398562.2011.561845

Calver, J., Brameld, K. J., Preen, D. B., Alexia, S. J., Boldy, D. P., & McCaul, K. A. (2006). High-cost users of hospital beds in Western Australia: a population-based record linkage study. *Medical Journal of Australia, 184*(8), 393-397.

Carpenter, I., Bobby, J., Kulinskaya, E., & Seymour, G. (2007). People admitted to hospital with physical disability have increased length of stay: implications for diagnosis related group re-imbursement in England. *Age Ageing, 36*(1), 73-78. doi:10.1093/ageing/afl148

Casey, B., Oxley, H., Whitehouse, E., Antolin, P., Duval, R., & Leibfritz, W. (2003). *Policies for an ageing society: recent measures and areas for further reform*. (369). OECD, France: OECD Publishing.

Castro-Costa, E., Dewey, M., Stewart, R., Banerjee, S., Huppert, F., Mendonca-Lima, C., . . . Prince, M. (2007). Prevalence of depressive symptoms and syndromes in later life in ten European countries: the SHARE study. *British Journal of Psychiatry, 191*, 393-401. doi:10.1192/bjp.bp.107.036772

Chisholm, D. (2000). Integration of mental health care into primary care: Demonstration cost-outcome study in India and Pakistan. *British Journal of Psychiatry, 176*(6), 581-588. doi:10.1192/bjp.176.6.581

Chollet, J., Saragoussi, D., Clay, E., & Francois, C. (2013). A clinical research practice datalink analysis of antidepressant treatment patterns and health care costs in generalized anxiety disorder. *Value in Health, 16*(8), 1133-1139.

Chuang, K. H., Covinsky, K. E., Sands, L. P., Fortinsky, R. H., Palmer, R. M., & Landefeld, C. S. (2003). Diagnosis-related group-adjusted hospital costs are higher in older medical patients with lower functional status. *Journal of the American Geriatrics Society, 51*(12), 1729-1734.

Cieza, A., Oberhauser, C., Bickenbach, J., Chatterji, S., & Stucki, G. (2014). Towards a minimal generic set of domains of functioning and health. *BMC Public Health, 14*, 218. doi:10.1186/1471-2458-14-218

Cieza, A., Sabariego, C., Anczewska, M., Ballert, C., Bickenbach, J., Cabello, M., . . . Consortium, P. (2015). PARADISE 24: A Measure to Assess the Impact of Brain Disorders on People's Lives. *PLoS One, 10*(7), e0132410. doi:10.1371/journal.pone.0132410

Cieza, A., Sabariego, C., Anczewska, M., Ballert, C., Bickenbach, J. E., Cabello, M., . . . Chatterji, S. (submitted- 2014). PARADISE-24: a measure to assess the impact of brain disorders on people's lives. *PLoS One*.

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. doi:10.3109/09540261.2011.606803

Coid, J., Petruckevitch, A., Bebbington, P., Brugha, T., Bhugra, D., Jenkins, R., . . . Singleton, N. (2002). Ethnic differences in prisoners. II: Risk factors and psychiatric service use? *The British Journal of Psychiatry, 181*(6), 481-487. doi:10.1192/bjp.181.6.481

Coid, J., Yang, M., Bebbington, P., Moran, P., Brugha, T., Jenkins, R., . . . Ullrich, S. (2009). Borderline personality disorder: health service use and social functioning among a national household population. *Psychological Medicine, 39*(10), 1721-1731. doi:10.1017/S0033291708004911

Coid, J., Yang, M., Tyrer, P., Roberts, A., & Ullrich, S. (2006). Prevalence and correlates of personality disorder in Great Britain. *British Journal of Psychiatry, 188*(5), 423-431. doi:10.1192/bjp.188.5.423

Commander, M., Sashidharan, S., Rana, T., & Ratnayake, T. (2005). North Birmingham assertive outreach evaluation. Patient characteristics and clinical outcomes. *Social Psychiatry and Psychiatric Epidemiology, 40*(12), 988-993.

Cooper, C., Bebbington, P., McManus, S., Meltzer, H., Stewart, R., Farrell, M., . . . Livingston, G. (2010). The treatment of Common Mental Disorders across age groups: results from the 2007 Adult Psychiatric Morbidity Survey. *Journal of Affective Disorders, 127*(1-3), 96-101. doi:10.1016/j.jad.2010.04.020

Cooper, C., Spiers, N., Livingston, G., Jenkins, R., Meltzer, H., Brugha, T., . . . Bebbington, P. (2013). Ethnic inequalities in the use of health services for common mental disorders in England. *Social Psychiatry and Psychiatric Epidemiology, 48*(5), 685-692. doi:10.1007/s00127-012-0565-y

Copeland, J. R., Prince, M., Wilson, K. C., Dewey, M. E., Payne, J., & Gurland, B. (2002). The Geriatric Mental State Examination in the 21st century. *Int J Geriatr Psychiatry, 17*(8), 729-732. doi:10.1002/gps.667

Corrigan, P. W., & Watson, A. C. (2002). Understanding the impact of stigma on people with mental illness. *World Psychiatry, 1*(1), 16-20.

Cotterill, P. G., & Thomas, F. G. (2004). Prospective payment for Medicare inpatient psychiatric care: assessing the alternatives. *Health Care Financing Review, 26*(1), 85-101.

Covinsky, K. E., Justice, A. C., Rosenthal, G. E., Palmer, R. M., & Landefeld, S. (1997). Measuring Prognosis and Case Mix in Hospitalized Elders. The Importance of Functional Status. . *Journal of General Internal Medicine, 12*, 203–208.

Covinsky, K. E., Yaffe, K., Lindquist, K., Cherkasova, E., Yelin, E., & Blazer, D. G. (2010). Depressive symptoms in middle age and the development of later-life functional limitations: the long-term effect of depressive symptoms. *Journal of the American Geriatrics Society, 58*(3), 551-556. doi:10.1111/j.1532-5415.2010.02723.x

Curtis, L., & Burns, A. (2015). Unit Costs of Health & Social Care. *Personal Social Service Research Unit*.

de Lusignan, S., Chan, T., Parry, G., Dent-Brown, K., & Kendrick, T. (2012). Referral to a new psychological therapy service is associated with reduced utilisation of healthcare and sickness absence by people with common mental health problems: a before and after comparison. *Journal of Epidemiology & Community Health, 66*(6), 1-6. doi:10.1136/jech.2011.139873

de Lusignan, S., Chan, T., Tejerina Arreal, M. C., Parry, G., Dent-Brown, K., & Kendrick, T. (2013). Referral for psychological therapy of people with long term conditions improves adherence to antidepressants and reduces emergency department attendance: Controlled before and after study. *Behaviour Research and Therapy, 51*(7), 377-385. doi:10.1016/j.brat.2013.03.004

Disler, P. B., Roy, C. W., & Smith, B. P. (1993). Predicting hours of care needed. *Archives of Physical and Medical Rehabilitation, 74*(2), 139-143.

DoH. (2015). NHS reference costs 2014–2015. *Department of Health*.

Durbin, A., Moineddin, R., Lin, E., Steele, L. S., & Glazier, R. H. (2015). Examining the relationship between neighbourhood deprivation and mental health service use of immigrants in Ontario, Canada: a cross-sectional study. *BMJ Open, 5*(3), e006690. doi:10.1136/bmjopen-2014-006690

Eagar, K., Gaines, P., Burgess, P., Green, J., Bower, A., Buckingham, B., & Mellsop, G. (2004). Developing a New Zealand casemix classification for mental health services. *World Psychiatry, 3*(3), 172-177.

Ekers, D., Webster, L., Van Straten, A., Cuijpers, P., Richards, D., & Gilbody, S. (2014). Behavioural Activation for Depression; An Update of Meta-Analysis of Effectiveness and Sub Group Analysis. *PLoS One, 9*(6), e100100. doi:10.1371/journal.pone.0100100

Elphick, M., & Antony, P. (1996). Casemix groupings for psychiatry: Strengths and weaknesses of `Version 2.0 Healthcare Resource Groups' (HRGs). *Journal Of Mental Health (Abingdon, England), 5*(5), 443-450.

English, J. T., Sharfstein, S. S., Scherl, D. J., Astrachan, B., & Muszynski, I. L. (1986). Diagnosis-related groups and general hospital psychiatry: the APA Study. *American Journal of Psychiatry, 143*(2), 131-139.

Escorpizo, R., Kostanjsek, N., Kennedy, C., Robinson Nicol, M. M., Stucki, G., & Ustan, T. B. (2013). Harmonizing WHO's International Classification of Diseases (ICD) and International Classification of Functioning, Disability and Health (ICF): importance and methods to link disease and functioning. *BMC Public Health, 13*, 742-746.

Essen, A. M. v. (2009a). New hospital payment systems: Comparing medical strategies in The Netherlands, Germany and England. *Journal of Health Organisation and Management, 23*(3), 304-318. doi:10.1108/14777260910966735

Essen, A. M. v. (2009b). New hospital payment systems: Comparing medical strategies in The Netherlands, Germany and England. *Journal of Health Organization and Management, 23*(3), 304-318. doi:doi:10.1108/14777260910966735

Evans, C., 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. *Br J Psychiatry, 180*, 51-60.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191.

Feng, L., Yap, K. B., Kua, E. H., & Ng, T. P. (2009). Depressive symptoms, physician visits and hospitalization among community-dwelling older adults. *Int Psychogeriatr, 21*(3), 568-575. doi:10.1017/S1041610209008965

Fischer, L. R., Wei, F., Rolnick, S. J., Jackson, J. M., Rush, W. A., Garrard, J. M., . . . Nitz, N. M. (2002). Geriatric Depression, Antidepressant Treatment, and Healthcare Utilization in a Health Maintenance Organization. *Journal of the American Geriatrics Society, 50*, 307–312.

Fleury, M. J., Grenier, G., Bamvita, J. M., & Caron, J. (2011). Mental health service utilization among patients with severe mental disorders. *Community Ment Health J, 47*(4), 365-377. doi:10.1007/s10597-010-9320-6

Fleury, M. J., Grenier, G., Bamvita, J. M., & Caron, J. (2014). Determinants and patterns of service utilization and recourse to professionals for mental health reasons. *BMC Health Services Research, 14*(161).

Foster, K., Meltzer, H., Gill, B., & Hinds, K. (2003). The circumstances of adults with a psychotic disorder. *International Review of Psychiatry, 15*(1-2), 84-90. doi:10.1080/0954026021000045985

Garcia Alvarez, L., Aylin, P., Tian, J., King, C., Catchpole, M., Hassall, S., . . . Holmes, A. (2011). Data linkage between existing healthcare databases to support hospital epidemiology. *J Hosp Infect, 79*(3), 231-235. doi:10.1016/j.jhin.2011.06.016

Gold, L. H. (2014). DSM-5 and the assessment of functioning: the World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0). *Journal of the American Academy of Psychiatry and the Law, 42*(2), 173-181.

Goldney, R. D., Fisher, L. J., & Walmsley, S. H. (1998). The Health of the Nation Outcome Scales in psychiatric hospitalisation: a multicentre study examining outcome and prediction of length of stay. *Australian and New Zealand Journal of Psychiatry, 32*(2), 199-205.

Gregori, D., Petrinco, M., Bo, S., Desideri, A., Merletti, F., & Pagano, E. (2011). Regression models for analyzing costs and their determinants in health care: an introductory review. *International Journal for Quality in Health Care, 23*(3), 331-341. doi:10.1093/intqhc/mzr010

Guerra, M., Ferri, C. P., Sosa, A. L., Salas, A., Gaona, C., Gonzales, V., . . . Prince, M. (2009). Late-life depression in Peru, Mexico and Venezuela: the 10/66 population-based study. *British Journal of Psychiatry, 195*(6), 510-515. doi:10.1192/bjp.bp.109.064055

Guy, W. (1976). The Clinical Global Impression Scale *ECDEU Assessment Manual for Psychopharmacology-Revised* (pp. 218-222). d. Rockville, MD: US Department of Health, Education and Welfare, ADAMHA, MIMH Psychopharmacology Research Branch.

Hansson, L., & Sandlund, M. (1992). Utilization and patterns of care in comprehensive psychiatric care organizations. A review of studies and some methodological considerations. *Acta Psychiatrica Scandinavica, 86*(4), 255-261.

Harrison-Read, P., Lucas, B., Tyrer, P., Ray, J., Shipley, K., Simmonds, S., . . . Hickman, M. (2002). Heavy users of acute psychiatric beds: randomized controlled trial of enhanced community management in an outer London borough. *Psychological Medicine, 32*(03). doi:10.1017/s0033291702005305

Hayhurst, K. P., Brown, P., & Lewis, S. W. (2002). The cost-effectiveness of clozapine: A controlled, population-based, mirror-image study. *Journal of Psychopharmacology, 16*(2), 169-175. doi:10.1177/026988110201600208

Hayward, R., Jordan, K. P., & Croft, P. (2010). Healthcare use in adults with insomnia: A longitudinal study. *British Journal of General Practice, 60*(574), 334-340.

Health and Social Care Information Centre Casemix Service. Mental Health Casemix Classification Development: End Stage Report. (2006). Health and Social Care Information Centre.

Health, D. o. (2011). Hospital Episode Statistics (HES). *National Health Service, NHS: London*.

Heo, M., Murphy, C. F., Fontaine, K. R., Bruce, M. L., & Alexopoulos, G. S. (2008). Population projection of US adults with lifetime experience of depressive disorder by age and sex from year 2005 to 2050. *International Journal of Geriatric Psychiatry, 23*(12), 1266-1270. doi:10.1002/gps.2061

Higgins, J. G. T., & Green, S. (2011). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011].

Hopfe, M., Stucki, G., Marshall, R., Twomey, C. D., Üstün, B. T., & Prodinger, B. (2016). Capturing patients' needs in casemix: a systematic literature review on the value of adding functioning information in reimbursement systems. *BMC Health Services Research, 16*(1), 1-17. doi:10.1186/s12913-016-1277-x

Huang, B. Y., Cornoni-Huntley, J., Hays, J. C., Huntley, R. R., Galanos, A. N., & Blazer, D. G. (2000). Impact of depressive symptoms on hospitalization risk in community-dwelling older persons. *Journal of the American Geriatrics Society, 48*(10), 1279-1284.

Huang, H., Menezes, P. R., da Silva, S. A., Tabb, K., Barkil-Oteo, A., & Scazufca, M. (2014). The association between depressive disorders and health care utilization: results from the Sao Paulo ageing and health study (SPAH). *General Hospital Psychiatry, 36*(2), 199-202. doi:10.1016/j.genhosppsych.2013.11.003

Hunsley, J. (2003). Cost-effectiveness and medical cost-offset considerations in psychological service provision. *Canadian Psychology, 44*, 61–73.

Kamenov, K., Twomey, C., Cabello, M., Prina, A. M., & Ayuso-Mateos, J. L. (2017). The efficacy of psychotherapy, pharmacotherapy and their combination on functioning and quality of life in depression: a meta-analysis. *Psychological Medicine, 47*(3), 414-425. doi:10.1017/s0033291716002774

Keene, J., & Rodriguez, J. (2007). Are mental health problems associated with use of Accident and Emergency and health-related harm? *European Journal of Public Health, 17*(4), 387-393. doi:10.1093/eurpub/ckl248

Kent, S., Fogarty, M., & Yellowlees, P. (1995). A review of studies of heavy users of pyschiatric services. *Psychiatric Services, 46*(12), 1247-1253.

Kessler, R. C., Angermeyer, M., Anthony, J. C., R, D. E. G., Demyttenaere, K., Gasquet, I., . . . Ustun, T. B. (2007). Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry, 6*(3), 168-176.

Killaspy, H., & Zis, P. (2013). Predictors of outcomes for users of mental health rehabilitation services: a 5-year retrospective cohort study in inner London, UK. *Social Psychiatry & Psychiatric Epidemiology, 48*(6), 1005-1012. doi:10.1007/s00127-012-0576-8

Kingdon, D., Solomka, B., McAllister-Williams, H., Turkington, D., Gregoire, A., Elnazer, H., . . . Naeem, F. (2012). Care clusters and mental health Payment by Results. *British Journal of Psychiatry, 200*(2), 162. doi:10.1192/bjp.200.2.162

Kisely, S., Campbell, L. A., Cartwright, J., Cox, M., & Campbell, J. (2010). Do the Health of the Nation Outcome Scales measure outcome? *Canadian Journal of Psychiatry - Revue Canadienne de Psychiatrie, 55*(7), 431-439.

Knapp, M., McCrone, P., Fombine, E., Beecham, J., & Wostear, G. (2002). The Maudsley long-term follow-up of child and adolescent depression: 3. Impact of comorbid conduct disorder on service use and costs in adulthood. *British Journal of Psychiatry, 180*(1), 19-23. doi:10.1192/bjp.180.1.19

Kotov, R., Gamez, W., Schmidt, F., & Watson, D. (2010). Linking "big" personality traits to anxiety, depressive, and substance use disorders: a meta-analysis. *Psychological Bulletin, 136*(5), 768-821. doi:10.1037/a0020327

Kullowatz, A., Kanniess, F., Dahme, B., Magnussen, H., & Ritz, T. (2007). Association of depression and anxiety with health care use and quality of life in asthma patients.

Respiratory Medicine, 101(3), 638-644. doi:http://dx.doi.org/10.1016/j.rmed.2006.06.002

Lam, D. H., McCrone, P., Wright, K., & Kerr, N. (2005). Cost-effectiveness of relapse-prevention cognitive therapy for bipolar disorder: 30-month study. *British Journal of Psychiatry, 186*, 500-506. doi:10.1192/bjp.186.6.500

Lam, R. W., Parikh, S. V., Michalak, E. E., Dewa, C. S., & Kennedy, S. H. (2015). Canadian Network for Mood and Anxiety Treatments (CANMAT) consensus recommendations for functional outcomes in major depressive disorder. *Annals of Clinical Psychiatry, 27*(2), 142-149.

Lavik, N. J. (1983). Utilization of mental health services over a given period. *Acta Psychiatrica Scandinavica, 67*(6), 404-413. doi:10.1111/j.1600-0447.1983.tb09721.x

Lee, K. J., & Simpson, J. A. (2014). Introduction to multiple imputation for dealing with missing data. *Respirology, 19*(2), 162-167. doi:10.1111/resp.12226

Lenze, E. J., Rogers, J. C., Martire, L. M., Mulsant, B. H., Rollman, B. L., Dew, M. A., . . . Reynolds, C. F., III. (2001). The Association of Late-Life Depression and Anxiety With Physical Disability: A Review of the Literature and Prospectus for Future Research. *The American Journal of Geriatric Psychiatry, 9*(2), 113-135. doi:10.1097/00019442-200105000-00004

Lovaglio, P. G., & Monzani, E. (2011). Validation aspects of the health of the nation outcome scales. *International Journal of Mental Health Systems, 5*, 20. doi:10.1186/1752-4458-5-20

Lovaglio, P. G., & Monzani, E. (2012). Health of the nation outcome scales evaluation in a community setting population. *Quality of Life Research, 21*(9), 1643-1653. doi:10.1007/s11136-011-0071-9

Lundalv, J., Tornbom, M., Larsson, P. O., & Sunnerhagen, K. S. (2015). Awareness and the arguments for and against the International Classification of Functioning, Disability and Health among representatives of disability organisations. *International Journal Of Environmental Research And Public Health, 12*(3), 3293-3300. doi:10.3390/ijerph120303293

Lutz, W., Sanderson, W., & Scherbov, S. (2008). The coming acceleration of global population ageing. *Nature, 451*(7179), 716-719. doi:10.1038/nature06516

Macdonald, A., & Elphick, M. (2012). Care clusters and mental health Payment by Results (Author's reply). *British Journal of Psychiatry, 200*(2), 163. doi:10.1192/bjp.200.2.163

Macdonald, A. J., & Elphick, M. (2011). Combining routine outcomes measurement and 'Payment by Results': will it work and is it worth it? *British Journal of Psychiatry, 199*(3), 178-179. doi:10.1192/bjp.bp.110.090993

Mannion, R., Marini, G., & Street, A. (2008). Implementing payment by results in the English NHS: Changing incentives and the role of information. *Journal of Health Organisation and Management, 22*(1), 79-88. doi:10.1108/14777260810862425

Marini, G., & Street, A. (2007). *The administrative costs of payment by results*. Retrieved from

Mason, A., Goddard, M., Myers, L., & Verzulli, R. (2011). Navigating uncharted waters? How international experience can inform the funding of mental health care in England. *Journal of Mental Health, 20*(3), 234-248. doi:10.3109/09638237.2011.562261

Mate, K. S., Sifrim, Z. K., Chalkiou, K., Cluzeau, F., Cutler, D., Kimball, M., . . . Baker, P. (2013). Improving health system quality in low- and middle-income countries that are expanding health coverage: a framework for insurance. *International Journal for Quality in Health Care, 25*(5), 497–504. doi:10.1093/intqhc/mzt053

Mathauer, I., & Wittenbecher, F. (2013). Hospital payment systems based on diagnosis-related groups: experiences in low- and middle-income countries. *Bulletin of the World Health Organisation, 91*(10), 746-756A. doi:10.2471/BLT.12.115931

Mayo, N. E., Nadeau, L., Levesque, L., Miller, S., Poissant, L., & Tamblyn, R. (2005). Does the addition of functional status indicators to case-mix adjustment indices improve prediction of hospitalization, institutionalization, and death in the elderly? *Medical Care, 43*(12), 1192-1202.

Mills, V., Van Hooff, M., Baur, J., & McFarlane, A. C. (2012). Predictors of mental health service utilisation in a non-treatment seeking epidemiological sample of Australian adults. *Community Ment Health J, 48*(4), 511-521. doi:10.1007/s10597-011-9439-0

Mohan, R., McCrone, P., Szmukler, G., Micali, N., Afuwape, S., & Thornicroft, G. (2006). Ethnic differences in mental health service use among patients with psychotic disorders. *Social Psychiatry and Psychiatric Epidemiology, 41*(10), 771-776.

Monitor. (2015). Open consultation: proposed changes to local payment rules covering mental health services in the NHS. *2016/17 national tariff proposals*. Retrieved from https://www.gov.uk/government/consultations/proposed-changes-to-local-payment-rules-covering-mental-health-services-in-the-nhs

Murray, C. J., Vos, T., Lozano, R., Naghavi, M., Flaxman, A. D., Michaud, C., . . . Memish, Z. A. (2012). Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. The *Lancet, 380*(9859), 2197-2223. doi:10.1016/s0140-6736(12)61689-4

Neubert, S., Sabariego, C., Stier-Jarmer, M., & Cieza, A. (2011). Development of an ICF-based patient education program. *Patient Education and Counselling, 84*(2), e13-17. doi:10.1016/j.pec.2010.07.021

NHS. (2013). *Mental Health Payment by Results Guidance for 2013-14*. Leeds: NHS.

NICE. (2009). *The guidelines manual*. Retrieved from London:

Noble, M., Wright, G., Smith, G., & Dibben, C. (2006). Measuring Multiple Deprivation at the Small-Area Level. *Environment and Planning A, 38*(1), 169-185. doi:10.1068/a37168

Patel, A., Everitt, B., Knapp, M., Reeder, C., Grant, D., Ecker, C., & Wykes, T. (2006). Schizophrenia patients with cognitive deficits: factors associated with costs. *Schizophrenia Bulletin, 32*(4), 776-785.

Perrin, N., Stiefel, M., Mosen, D. M., Bauck, A., Shuster, E., & Dirks, E. M. (2011). Self-reported health and functional status information improves prediction of inpatient admissions and costs. *American Journal of Managed Care, 17*(12), e472-e478.

Pirkis, J. E., Burgess, P. M., Kirk, P. K., Dodson, S., Coombs, T. J., & Williamson, M. K. (2005). A review of the psychometric properties of the Health of the Nation Outcome Scales (HoNOS) family of measures. *Health Qual Life Outcomes, 3*, 76. doi:10.1186/1477-7525-3-76

Pitkänen, T., Levola, J., Tourunen, J., Kaskela, T., & Holopainen, A. (2016). *Aivotoiminnan häiriöiden yhteydessä yleisesti koetut psykososiaaliset vaikeudet. PARADISE24-kyselyn tutkimusperusta.* Helsinki: Kela, Sosiaali- ja ter veystur van tutkimuksia.

Prina, A. M., Cosco, T. D., Dening, T., Beekman, A., Brayne, C., & Huisman, M. (2015). The association between depressive symptoms in the community, non-psychiatric hospital admission and hospital outcomes: A systematic review. *Journal of Psychosomatic Research, 78*(1), 25-33. doi:10.1016/j.jpsychores.2014.11.002

Prina, A. M., Deeg, D., Brayne, C., Beekman, A., & Huisman, M. (2012). The association between depressive symptoms and non-psychiatric hospitalisation in older adults. *PLoS One, 7*(4), e34821. doi:10.1371/journal.pone.0034821

Prina, A. M., Huisman, M., Yeap, B. B., Hankey, G. J., Flicker, L., Brayne, C., & Almeida, O. P. (2013). Association between depression and hospital outcomes among older men. *CMAJ, 185*(2), 117-123. doi:10.1503/cmaj.121171

Prina, A. M., Huisman, M., Yeap, B. B., Hankey, G. J., Flicker, L., Brayne, C., & Almeida, O. P. (2014). Hospital costs associated with depression in a cohort of older men living in Western Australia. *General Hospital Psychiatry, 36*(1), 33-37. doi:10.1016/j.genhosppsych.2013.08.009

Prince, M., Ferri, C. P., Acosta, D., Albanese, E., Arizaga, R., Dewey, M., . . . Uwakwe, R. (2007). The protocols for the 10/66 dementia research group population-based research programme. *BMC Public Health, 7*, 165. doi:10.1186/1471-2458-7-165

Prince, M., Patel, V., Saxena, S., Maj, M., Maselko, J., Phillips, M. R., & Rahman, A. (2007). No health without mental health. *The Lancet, 370*(9590), 859-877. doi:10.1016/s0140-6736(07)61238-0

Prince, M., Reischies, F., Beekman, A. T., Fuhrer, R., Jonker, C., Kivela, S. L., . . . Copeland, J. R. (1999). Development of the EURO-D scale--a European, Union initiative to compare symptoms of depression in 14 European centres. *British Journal of Psychiatry, 174*(4), 330-338. doi:10.1192/bjp.174.4.330

Rehm, J., Ustun, T. D., Saxena, S., Nelson, C. B., & Chatterji, S. (1999). On the development and psychometric testing of the WHO screening instrument to assess disablement in the general population. *International Journal of Methods in Psychiatr Research, 8*(2), 110-122.

Rodriguez, B. F., Bruce, S. E., Pagano, M. E., & Keller, M. B. (2005). Relationships among psychosocial functioning, diagnostic comorbidity, and the recurrence of generalized anxiety disorder, panic disorder, and major depression. *Journal of Anxiety Disorders, 19*(7), 752-766. doi:10.1016/j.janxdis.2004.10.002

Sabariego, C., Brach, M., & Stucki, G. (2011). Determinants of major direct medical cost categories among patients with osteoporosis, osteoarthritis, back pain or fibromyalgia undergoing outpatient rehabilitation. *Journal of Rehabilitation & Medicine, 43*(8), 703-708. doi:10.2340/16501977-0845

Sangha, O., Stucki, G., Liang, M. H., Fossel, A. H., & Katz, J. N. (2003). The Self-Administered Comorbidity Questionnaire: a new method to assess comorbidity for clinical and health services research. *Arthritis and Rheumatism, 49*(2), 156-163. doi:10.1002/art.10993

Sava, F. A., Yates, B. T., Lupu, V., Szentagotai, A., & David, D. (2009). Cost-effectiveness and cost-utility of cognitive therapy, rational emotive behavioral therapy, and fluoxetine (Prozac) in treating depression: a randomized clinical trial. *Journal of Clinical Psychology, 65*(1), 36-52. doi:10.1002/jclp.20550

Saxena, S., Sharan, P., & Saraceno, B. (2003). Budget and financing of mental health services: baseline information on 89 countries from WHO's project atlas. *Journal of Mental Health Policy and Economics, 6*(3), 135-143.

Schillerstrom, J. E., Royall, D. R., & Palmer, R. F. (2008). Depression, Disability and Intermediate Pathways: A Review of Longitudinal Studies in Elders. *Journal of Geriatric Psychiatry and Neurology, 21*(3), 183-197. doi:10.1177/0891988708320971

Schumacher, D. N., Namerow, M. J., Parker, B., Fox, P., & Kofie, V. (1986). Prospective payment for psychiatry-feasibility and impact. *New England Journal of Medicine, 315*(21), 1331-1336.

Self, R., Rigby, A., Leggett, C., & Paxton, R. (2008). Clinical Decision Support Tool: A rational needs-based approach to making clinical decisions. *Journal of Mental Health, 17*(1), 33-48. doi:10.1080/09638230701505806

Shi, N., Cao, Z., Durden, E., Schacht, A., Torres, A., & Happich, M. (2012). Healthcare utilization among patients with depression before and after initiating duloxetine in the United Kingdom. *Journal of Medical Economics, 15*(4), 672-680.

Sinclair, J. M., Gray, A., Rivero-Arias, O., Saunders, K. E., & Hawton, K. (2011). Healthcare and social services resource use and costs of self-harm patients. *Social Psychiatry and Psychiatric Epidemiol, 46*(4), 263-271. doi:10.1007/s00127-010-0183-5

Sousa, R. M., Dewey, M. E., Acosta, D., Jotheeswaran, A. T., Castro-Costa, E., Ferri, C. P., . . . Prince, M. J. (2010). Measuring disability across cultures--the psychometric properties of the WHODAS II in older people from seven low- and middle-income countries. The 10/66 Dementia Research Group population-based survey. *International Journal of Methods in Psychiatric Research, 19*(1), 1-17. doi:10.1002/mpr.299

Speak, B., & Muncer, S. (2015). The structure and reliability of the Health of the Nation Outcome Scales. *Australasian Psychiatry, 23*(1), 66-68. doi:10.1177/1039856214563851

Stewart, R., Soremekun, M., Perera, G., Broadbent, M., Callard, F., Denis, M., . . . Lovestone, S. (2009). The South London and Maudsley NHS Foundation Trust Biomedical Research Centre (SLAM BRC) case register: development and descriptive data. *BMC Psychiatry, 9*(1), 51.

Street, A., & Maynard, A. (2007). Activity based financing in England: the need for continual refinement of payment by results. *Health Economics Policy & Law, 2*(Pt 4), 419-427. doi:10.1017/S174413310700429X

Taube, C. A., Goldman, H. H., Burns, B. J., & Kessler, L. G. (1988). High users of outpatient mental health services, I: Definition and characteristics. *Amercian Journal of Psychiatry, 145*(1), 19-24.

Torres, A. R., Prince, M. J., Bebbington, P. E., Bhugra, D. K., Brugha, T. S., Farrell, M., . . . Singleton, N. (2007). Treatment seeking by individuals with obsessive-compulsive disorder from the British Psychiatric Morbidity Survey of 2000. *Psychiatric Services, 58*(7), 977-982.

Trieman, N., & Leff, J. (2002). Long-term outcome of long-stay psychiatric in-patients considered unsuitable to live in the community. *British Journal of Psychiatry, 181*, 428-432.

Tuffrey-Wijne, I., Goulding, L., Giatras, N., Abraham, E., Gillard, S., White, S., . . . Hollins, S. (2014). The barriers to and enablers of providing reasonably adjusted health services to people with intellectual disabilities in acute hospitals: evidence from a mixed-methods study. *BMJ Open, 4*, e004606. doi:10.1136/bmjopen-2013004606

Tulloch, A. D. (2012). Care clusters and mental health Payment by Results. *British Journal of Psychiatry, 200*(2), 161-161. doi:10.1192/bjp.200.2.161

Tulloch, A. D., David, A. S., & Thornicroft, G. (2015). Exploring the predictors of early readmission to psychiatric hospital. *Epidemiology and Psychiatric Sciences, [Epub ahead of print]*, 1-13. doi:10.1017/s2045796015000128

Tulloch, A. D., Khondoker, M. R., Fearon, P., & David, A. S. (2012). Associations of homelessness and residential mobility with length of stay after acute psychiatric admission. *BMC Psychiatry, 12*, 121. doi:10.1186/1471-244X-12-121

Tulloch, A. D., Khondoker, M. R., Thornicroft, G., & David, A. S. (2014). Home treatment teams and facilitated discharge from psychiatric hospital. *Epidemiological and Psychiatric Sciences*, 1-13. doi:10.1017/S2045796014000304

Twomey, C., Prina, A. M., Baldwin, D. S., Das-Munshi, J., Kingdon, D., Koeser, L., . . . Cieza, A. (2016). Utility of the Health of the Nation Outcome Scales (HoNOS) in Predicting Mental Health Service Costs for Patients with Common Mental Health Problems: Historical Cohort Study. *PLoS One, 11*(11), e0167103. doi:10.1371/journal.pone.0167103

Twomey, C. D., Baldwin, D. S., Hopfe, M., & Cieza, A. (2015). A systematic review of the predictors of health service utilisation by adults with mental disorders in the UK. *BMJ Open, 5*(7), e007575. doi:10.1136/bmjopen-2015-007575

Twomey, C. D., Prince, M., Cieza, A., Baldwin, D. S., & Prina, A. M. (2015). Cross-sectional associations of depressive symptom severity and functioning with health service use by older people in low-and-middle income countries. *International Journal of Environmental Research and Public Health, 12*(4), 3774-3792. doi:10.3390/ijerph120403774

Ullrich, S., & Coid, J. (2009). Antisocial personality disorder: Co-morbid Axis I mental disorders and health service use among a national household population. *Personality and Mental Health, 3*(3), 151-164. doi:10.1002/pmh.70

Vasiliadis, H.-M., Lesage, A., Adair, C. E., Wang, P. S., & Kessler, R. C. (2007). Do Canada and the United States Differ in Prevalence of Depression and Utilization of Services? *Psychiatric Services, 58*(1), 63-71.

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gotzsche, P. C., Vandenbroucke, J. P., & Initiative, S. (2008). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Journal of Clinical Epidemiology, 61*(4), 344-349. doi:10.1016/j.jclinepi.2007.11.008

Wade, A. G., Saragoussi, D., Despiegel, N., Francois, C., Guelfucci, F., & Toumi, M. (2010). Healthcare expenditure in severely depressed patients treated with escitalopram, generic SSRIs or venlafaxine in the UK. *Current Medical Research And Opinion, 26*(5), 1161-1170.

Walters, K., Buszewicz, M., Weich, S., & King, M. (2011). Mixed anxiety and depressive disorder outcomes: Prospective cohort study in primary care. *British Journal of Psychiatry, 198*(6), 472-478.

Wang, R., Shaw, I., & Middleton, H. (2015). Delaying the implementation of Payment by Results in mental health: the application of standardisation. *Mental Health Review Journal, 20*(3), 156-165. doi:10.1108/MHRJ-03-2014-0008

Wanless, D. (2006). Securing good care for older people: taking a long-term view. *King's Fund, London*.

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine, 30*(4), 377-399. doi:10.1002/sim.4067

WHO. (2001). *World Health Organization: International Classification of Functioning, Disability and Health: ICF.* Geneva.

Wing, J. K., Beevor, A. S., Curtis, R. H., Park, S. B., Hadden, S., & Burns, A. (1998). Health of the Nation Outcome Scales (HoNOS). Research and development. *British Journal of Psychiatry, 172*(1), 11-18. doi:10.1192/bjp.172.1.11

Woods, R. T., Bruce, E., Edwards, R. T., Elvish, R., Hoare, Z., Hounsome, B., . . . Russell, I. T. (2012). REMCARE: reminiscence groups for people with dementia and their family caregivers - effectiveness and cost-effectiveness pragmatic multicentre randomised trial. *Health Technology Assessment (Winchester, England), 16*(48), v. doi:10.3310/hta16480

Wright, S., Gournay, K., Glorney, E., & Thornicroft, G. (2000). Dual diagnosis in the suburbs: Prevalence, need, and in-patient service use. *Social Psychiatry and Psychiatric Epidemiology, 35*(7), 297-304. doi:10.1007/s001270050242

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica, 67*(6), 361-370.

Zimmerman, M., McGlinchey, J. B., Posternak, M. A., Friedman, M., Attiullah, N., & Boerescu, D. (2006). How should remission from depression be defined? The depressed patient's perspective. *Amercian Journal of Psychiatry, 163*(1), 148-150. doi:10.1176/appi.ajp.163.1.148