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

Conal Twomey[a], Alarcos Cieza[a,c,d] and David S. Baldwin[b,e]

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. This was 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. After adjusting for covariates, a significant total costs association was yielded for functioning ($e^{\beta} = 1.02$; 95% confidence interval: 1.01–1.03), but not depressive symptom severity or 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 ($e^{\beta} = 1.04$; 95% confidence interval: 1.01–1.09), but not symptom severity. 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. ***Int Clin Psychopharmacol*** **00:000–000 Copyright © 2017 The Author(s). Published by Wolters Kluwer Health, Inc.**

International Clinical Psychopharmacology 2017, **00**:000–000

Keywords: anxiety, depression, functioning, healthcare policy

[a]School of Psychology, Faculty of Social and Human Sciences, [b]Clinical and Experimental Sciences Academic Unit, Faculty of Medicine, University of Southampton, Southampton, UK, [c]Research Unit for Biopsychosocial Health, Department of Medical Informatics, Biometry and Epidemiology – IBE, Ludwig-Maximilians-University (LMU), Munich, Germany, [d]Swiss Paraplegic Research (SPF), Nottwil, Switzerland and [e]University Department of Psychiatry and Mental Health, University of Cape Town, South Africa

Correspondence to Conal Twomey, BA, MLitt, HDipPsych, PhD, School of Psychology, Faculty of Social and Human Sciences, University of Southampton, Southampton SO17 1BJ, UK
Tel: + 44 23 8059 2612; fax: + 44 23 8059 4597; e-mail: c.twomey@soton.ac.uk

Received 23 January 2017 **Accepted** 22 March 2017

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 (Saxena *et al.*, 2003; Murray *et al.*, 2012). To meet the needs of patients, it is important that scarce mental health service resources are allocated systematically and efficiently (Essen, 2009). This could be achieved by healthcare payment systems that financially incentivize hospitals, clinics and other providers to treat as many patients as possible (Street and Maynard, 2007). Providers receive a fixed payment for every patient treated, adjusted for the 'cluster' that patients are assigned to on the basis of 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 thus 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 and Wittenbecher, 2013). However, predicting costs of care in mental disorders is complex, largely because of 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 towards a lack of progress in developing payment systems for mental health services around the world (English *et al.*, 1986; Schumacher *et al.*, 1986; Elphick and Antony, 1996; Cotterill and Thomas, 2004; Macdonald and Elphick, 2011; Mason *et al.*, 2011). It seems necessary therefore to investigate alternatives to clustering psychiatric patients according to diagnosis. The governments of Australia, New Zealand and UK have pursued a 'multidomain' 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; Macdonald and Elphick, 2012; Wang *et al.*, 2015). For example, pilot studies in the English National Health Service (NHS) have shown 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, 2006; 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: individuals with mental disorders typically experience considerable decrements in functioning and describe functional recovery as essential for remission (Zimmerman *et al.*, 2006; Lam *et al.*, 2015); poorer functioning may predict recurrence of depressive and anxiety disorders (Rodriguez *et al.*, 2005); and functioning is increasingly being recognized as a priority in the treatment and assessment of mental disorders, as reflected in the new dimensional approach of the *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed. (DSM-V), whereby it is rated alongside diagnostic severity (Gold, 2014; Lam *et al.*, 2015). The utility of functioning in predicting costs of care 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 that deployed various domain-specific operationalizations of functioning and uncosted healthcare use outcomes (Patel *et al.*, 2006; Cooper *et al.*, 2010; Twomey *et al.*, 2015a, 2015b).

We carried out a cohort study set within an NHS tertiary referral specialist clinic to investigate the utility of functioning in predicting costs of care 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 multidomain nature in an overall summary score. The PARADISE 24 differs from a previously developed ICF-based metric of multidomain functioning (i.e. the WHODAS-II) in that it includes items related to symptoms of psychological problems (e.g. anxiety and depression). We investigated the association of baseline functioning with total NHS costs at the 6-month follow-up and how functioning performed in comparison with depressive and anxiety symptom severity in cost prediction. We also investigated whether the potential predictive ability of functioning was driven by decrements related to the ICF domains of psychological health or 'activities and participation'.

Participants and methods

Participants

The study involved a convenience sample of adult patients attending an 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 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 *et al.*, 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 an accompanying information sheet. Participation involved completing questionnaires in a baseline data collection meeting, providing information on recent healthcare use by telephone at follow-up, and agreeing that author C.T. 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, sex, ethnicity, marital status, educational, employment status, Index of Multiple Deprivation score (Noble *et al.*, 2006), general health comorbidity (Self-Administered Comorbidity Questionnaire) (Sangha *et al.*, 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 used widely in clinical populations (Zigmond and Snaith, 1983; Bjelland *et al.*, 2002). Functioning was measured using PARADISE 24, a metric developed using the 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 (0.7–1.3) 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 α) 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); and 2 (A lot). The raw score ranges from 0 to 48 before transformation into a more intuitive scale ranging from 0 to 100 (Cieza *et al.*, 2015).

Outcome: total NHS costs at the 6-month follow-up

The primary outcome was NHS total costs, with secondary analyses carried out on the subcategories of mental health service costs and general health service costs. Costs were calculated in two stages. First, we counted the number of contacts that 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 and Knapp, 2001) administered by author C.T. at 3-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 (Department of Health, 2015). As all required unit costs were not available in this source, some were based on 2015 costs provided by the Personal Social Service Research Unit (Curtis and Burns, 2015) and internal financial records. Table 1 details the unit costs used in this study.

Statistical analysis

Main analyses

Analyses were carried out using STATA 13 (StataCorp LP, College Station, Texas, USA). Descriptive statistics were used for sample characteristics. The separate associations of baseline depressive symptom severity, anxiety symptom severity and functioning with costs at the 6-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 illustrated by the presence of markedly higher mean costs (£3899) than median costs (£1595); GLM-log-gamma is widely considered 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 the 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, whereas a coefficient of 1.10 indicates a 10% increase.

We sought to investigate how functioning performed in comparison with 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 into 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

On the basis of previous research showing their associations with mental health service costs (Durbin *et al.*, 2015; Twomey *et al.*, 2015a), initial adjustments were made for age, sex, marital status, ethnicity, employment status, area-level deprivation, general health comorbidity, psychiatric comorbidity, clinician-rated severity of illness, NHS costs incurred in the 3 months before baseline, functioning, depressive symptom severity and anxiety symptom severity. To safeguard statistical power, we subsequently removed several covariates that (a) were not associated with costs in exploratory analysis and (b) yielded P values more than 0.20 in this association. These variables were sex, marital status, ethnicity, employment status, area-level deprivation and clinician-rated severity of illness.

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[a]	£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[b]	£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 min)	£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

A&E, Accident & Emergency; 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.
[a]The mean of the care clusters unit cost per occupied bed day was calculated and used.
[b]The mean cost of different types of A&E attendances was calculated and used.

Missing data

Missing cost data (three participants, 2.9%) arose because of 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 and Simpson, 2014). Imputation models included all predictive variables entered into the GLM. A total of 100 imputed datasets were created, resulting in the introduction of minimal SE, as per guidelines (White *et al.*, 2011). Checks between imputed and original values produced no anomalies. Estimates were combined using Rubin's rules (White *et al.*, 2011).

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 the sample characteristics ($n = 102$). The mean age was 50.6 years. Sixty-one per cent of the population were women 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. Scores on measures of functioning, depressive symptoms and anxiety symptoms were normally distributed. About half of the participants were deemed to be 'moderately ill' according to Clinical Global Impression Scale scores. The mean of the total costs accrued during follow-up was £3899 (SD = 7997), with a median of £1595.

Internal consistency of predictor variables

Internal consistency for the three main predictor variables was high: PARADISE 24 ($\alpha = 0.93$); HADS-depression ($\alpha = 0.86$); and HADS-anxiety ($\alpha = 0.84$).

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

Table 3 summarizes the unadjusted and adjusted associations. In unadjusted models, significant total costs associations were yielded for functioning ($e^{\beta} = 1.05$; 95% CI: 1.03–1.07) and depressive symptom severity ($e^{\beta} = 1.10$; 95% CI: 1.02–1.18), but not for anxiety symptom severity ($e^{\beta} = 1.08$; 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 ($e^{\beta} = 1.02$; 95% CI: 1.01–1.03), but not

Table 2 Sample characteristics ($N = 102$)

Variables	*n* (%)	Mean (SD)	Median (25th–75th centile)
Age[a]	–	50.6 (13.5)	52.0 (42.8–60.3)
Sex			
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[b]	–	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)	–	–
HADS-depression score (range: 0–21)	–	10.0 (4.9)	10.0 (7.0–13.0)
HADS-anxiety score (range: 0–21)	–	11.6 (4.7)	12.0 (8.0–15.0)
PARADISE 24 functioning score (range: 0–100)	–	63.1 (15.9)	65.0 (52.0–73.5)
Psychiatric comorbidity			
Yes	50 (49.0)	–	–
No	52 (51.0)	–	–
Clinical Global Impression score	–	3.7 (0.8)	4.0 (3.0–4.0)

HADS, Hospital Anxiety and Depression Scales; NHS, National Health Service; SCQ, Self-Administered Comorbidity Questionnaire.
[a]Age range is 18–79 years.
[b]The lower the decile score, the higher the relative deprivation in the area.

depressive symptom severity ($e^{\beta} = 1.03$; 95% CI: 0.98–1.07) or anxiety symptom severity ($e^{\beta} = 1.03$; 95% CI: 0.98–1.07). In the final model that 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 ($e^{\beta} = 1.04$; 95% CI: 1.01–1.09), but not depressive symptom severity ($e^{\beta} = 0.99$; 95% CI: 0.94–1.05) or anxiety symptom severity ($e^{\beta} = 0.98$; 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 divided 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 or anxiety symptom severity. Interpreting the magnitude of the significant association, for every one point increase in the mean PARADISE 24 score, there was a 2% increase in costs. As the PARADISE 24 has 100 points, this can be considered as a relatively strong association. In the final model that 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 individuals with mental disorders using an ICF-based measure of functioning. Sample representativeness was strengthened by a high participation rate (90%), a low dropout rate (3%) and use of multiple imputation. The normal distribution of scores on measures of functioning, depressive symptoms and anxiety symptoms safeguarded the validity of statistical analyses in terms of the prediction of costs. However, statistical power was limited by the sample size of 102, which reduced the number of covariates that could be included in statistical models. The study was set within a tertiary referral specialist clinic; thus, 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 sex and White ethnicity. Furthermore, the service is led by author D.S.B., who administers treatment on the basis of clinical judgement: different treatment decisions may be made in other services, which could limit the generalizability 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 were collected using the CSRI, which may have been subject to recall errors. The precision and applicability of unit costs data are limited because of 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

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

	e^{β} (95% CI)		
Predictor variable and covariates	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[a]	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[a]	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[a]	1.03 (1.01–1.05)	1.00 (0.98–1.01)	1.02 (1.01–1.03)
Functioning (PARADISE 14)[b,c]			
Unadjusted	1.14 (1.07–10.20)	1.03 (0.98–1.08)	1.11 (1.06–1.16)
Age, baseline total costs, comorbidity[a]	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)

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

applicable to certain NHS services; and the data had to be extracted from two different sources. The 6-month time frame for the analysis of costs does not take into account possible seasonal effects on resource use to the same degree as studies lasting over 1 year. The duration of our time frame 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 operationalizations of functioning and uncosted healthcare use outcomes (Patel *et al.*, 2006; Cooper *et al.*, 2010; Twomey *et al.*, 2015a, 2015b). Domain-specific operationalizations of functioning may be less representative of care needs than the multi-domain PARADISE 24, whereas uncosted healthcare use outcomes do not provide a weighted summary of resource consumption and are therefore less precise than costed outcomes. Nevertheless, our findings that 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 general practitioner attendances by individuals with 'common mental disorders'(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 patients with schizophrenia (Patel *et al.*, 2006). The utility of functioning in predicting costs of care 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 is in agreement with evidence from numerous studies involving large-scale and national datasets (English *et al.*, 1986; Schumacher *et al.*, 1986; Elphick and Antony, 1996; Cotterill and Thomas, 2004; Macdonald and Elphick, 2011; Prina *et al.*, 2015; Twomey *et al.*, 2015a).

Potential implications

Our findings support the utility of functioning in predicting costs of care 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, however, 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 related 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 shown 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 (Self *et al.*, 2008; Lovaglio and Monzani, 2011, 2012; Speak and Muncer, 2015; Wang *et al.*, 2015) and a future study directly comparing its utility with that of the PARADISE

24 metric in costs prediction might inform policy debates in UK 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 that payments should be closely linked to agreed patient outcome standards rather than costs, to incentivize quality of care (Monitor and NHS England, 2015).

Acknowledgments

The authors thank Magda Nowak for her dedicated work in carrying out the study.

All authors collaborated in the conception of the research paper. C.T. and D.S.B. undertook various aspects of data collection in collaboration; C.T. interviewed all participants; C.T. carried out data management, statistical analyses and drafted versions of the manuscript; A.C. and D.S.B. provided regular consultations and specific feedback on statistical analyses, interpretation of results and manuscript write-up, until final draft completion. All authors read and approved the final manuscript.

The research is funded by the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007–2013/(REA grant agreement No. 316795). Alarcos Cieza is a member of the 'Psychosocial fActors Relevant to BrAin DISorders in Europe' (PARADISE) consortium that developed the PARADISE 24 metric. She was not involved in data collection.

Conflicts of interest

There are no conflicts of interest.

References

Appleby J, Harrison T, Hawkins L, Dixon A (2012). *Payment by results – how can payment systems help deliver better care?*. London, UK: The King's Fund.

Beecham J, Knapp M (2001). Costing psychiatric interventions. In: Thornicroft G, Brewin C, Wing JK, editors. *Measuring mental health needs*. London, UK: Gaskell. pp. 163–183.

Bjelland I, Dahl AA, Haug TT, Neckelmann D (2002). The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *J Psychosom Res* **52**:69–77.

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

Busse R, Schreyögg J, Smith PC (2006). Editorial: hospital case payment systems in Europe. *Health Care Manag Sci* **9**:211–213.

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.

Cieza A, Sabariego C, Anczewska M, Ballert C, Bickenbach J, Cabello M, *et al.* (2015). PARADISE 24: a measure to assess the impact of brain disorders on people's lives. *PLoS One* **10**:e0132410.

Cooper C, Bebbington P, McManus S, Meltzer H, Stewart R, Farrell M, *et al.* (2010). The treatment of common mental disorders across age groups: results from the 2007 Adult Psychiatric Morbidity Survey. *J Affect Disord* **127**:96–101.

Cotterill PG, Thomas FG (2004). Prospective payment for Medicare inpatient psychiatric care: assessing the alternatives. *Health Care Financ Rev* **26**:85–101.

Curtis L, Burns A (2015). *Unit costs of health & social care*. University of Kent; UK: Personal Social Service Research Unit.

Department of Health (2015). *NHS reference costs 2014–2015*. London; UK: Department of Health.

Durbin A, Moineddin R, Lin E, Steele LS, Glazier RH (2015). Examining the relationship between neighbourhood deprivation and mental health service use of immigrants in Ontario, Canada: a cross-sectional study. *BMJ Open* **5**: e006690.

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**:172–177.

Elphick M, Antony P (1996). Casemix groupings for psychiatry: strengths and weaknesses of 'Version 2.0 Healthcare Resource Groups' (HRGs). *J Ment Health* **5**:443–450.

English JT, Sharfstein SS, Scherl DJ, Astrachan B, Muszynski IL (1986). Diagnosis-related groups and general hospital psychiatry: the APA Study. *Am J Psychiatry* **143**:131–139.

Essen AMV (2009). New hospital payment systems: comparing medical strategies in The Netherlands, Germany and England. *J Health Organ Manag* **23**:304–318.

Faul F, Erdfelder E, Lang AG, Buchner A (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* **39**:175–191.

Garcia Alvarez L, Aylin P, Tian J, King C, Catchpole M, Hassall S, *et al.* (2011). Data linkage between existing healthcare databases to support hospital epidemiology. *J Hosp Infect* **79**:231–235.

Gold LH (2014). DSM-5 and the assessment of functioning: the World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0). *J Am Acad Psychiatry Law* **42**:173–181.

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. *Int J Qual Health Care* **23**:331–341.

Guy W (1976). The Clinical Global Impression Scale. In: Guy W, editor. *ECDEU assessment manual for psychopharmacology-revised*. Rockville, MD: US Department of Health, Education and Welfare, ADAMHA, MIMH Psychopharmacology Research Branch. 218–222.

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

Hopfe M, Stucki G, Marshall R, Twomey CD, Üstün BT, Prodinger B (2016). Capturing patients' needs in casemix: a systematic literature review on the value of adding functioning information in reimbursement systems. *BMC Health Serv Res* **16**:1–17.

Lam RW, Parikh SV, Michalak EE, Dewa CS, Kennedy SH (2015). Canadian Network for Mood and Anxiety Treatments (CANMAT) consensus recommendations for functional outcomes in major depressive disorder. *Ann Clin Psychiatry* **27**:142–149.

Lee KJ, Simpson JA (2014). Introduction to multiple imputation for dealing with missing data. *Respirology* **19**:162–167.

Lovaglio PG, Monzani E (2011). Validation aspects of the health of the nation outcome scales. *Int J Ment Health Syst* **5**:20.

Lovaglio PG, Monzani E (2012). Health of the nation outcome scales evaluation in a community setting population. *Qual Life Res* **21**:1643–1653.

Macdonald AJ, Elphick M (2011). Combining routine outcomes measurement and 'Payment by Results': will it work and is it worth it? *Br J Psychiatry* **199**:178–179.

Macdonald A, Elphick M (2012). Care clusters and mental health Payment by Results (author's reply). *Br J Psychiatry* **200**:163.

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. *J Ment Health* **20**:234–248.

Mathauer I, Wittenbecher F (2013). Hospital payment systems based on diagnosis-related groups: experiences in low- and middle-income countries. *Bull World Health Organ* **91**:746–756A.

Monitor and NHS England. Masters A, Strathdee G, Baumann P (2015). Proposed changes to local payment rules covering mental health services in the NHS. *2016/17 national tariff proposals*. London; UK: Monitor and NHS England. pp. 1–9.

Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, *et al.* (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. *Lancet* **380**:2197–2223.

Noble M, Wright G, Smith G, Dibben C (2006). Measuring multiple deprivation at the small-area level. *Environ Plan A* **38**:169–185.

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

Prina AM, Cosco TD, 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. *J Psychosom Res* **78**:25–33.

Rodriguez BF, Bruce SE, Pagano ME, Keller MB (2005). Relationships among psychosocial functioning, diagnostic comorbidity, and the recurrence of generalized anxiety disorder, panic disorder, and major depression. *J Anxiety Disord* **19**:752–766.

Sangha O, Stucki G, Liang MH, Fossel AH, Katz JN (2003). The Self-Administered Comorbidity Questionnaire: a new method to assess comorbidity for clinical and health services research. *Arthritis Rheum* **49**:156–163.

Saxena S, Sharan P, Saraceno B (2003). Budget and financing of mental health services: baseline information on 89 countries from WHO's project atlas. *J Ment Health Policy Econ* **6**:135–143.

Schumacher DN, Namerow MJ, Parker B, Fox P, Kofie V (1986). Prospective payment for psychiatry-feasibility and impact. *N Engl J Med* **315**:1331–1336.

Self R, Rigby A, Leggett C, Paxton R (2008). Clinical decision support tool: a rational needs-based approach to making clinical decisions. *J Ment Health* **17**:33–48.

Speak B, Muncer S (2015). The structure and reliability of the Health of the Nation Outcome Scales. *Australas Psychiatry* **23**:66–68.

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

Street A, Maynard A (2007). Activity based financing in England: the need for continual refinement of payment by results. *Health Econ Policy Law* **2**:419–427.

Tulloch AD (2012). Care clusters and mental health Payment by Results. *Br J Psychiatry* **200**:161–162.

Twomey CD, Baldwin DS, Hopfe M, Cieza A (2015a). A systematic review of the predictors of health service utilisation by adults with mental disorders in the UK. *BMJ Open* **5**:e007575.

Twomey CD, Prince M, Cieza A, Baldwin DS, Prina AM (2015b). Cross-sectional associations of depressive symptom severity and functioning with health service use by older people in low-and-middle income countries. *Int J Environ Res Public Health* **12**:3774–3792.

Wang R, Shaw I, Middleton H (2015). Delaying the implementation of Payment by Results in mental health: the application of standardisation. *Ment Health Rev J* **20**:156–165.

White IR, Royston P, Wood AM (2011). Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* **30**:377–399.

Zigmond AS, Snaith RP (1983). The hospital anxiety and depression scale. *Acta Psychiatr Scand* **67**:361–370.

Zimmerman M, McGlinchey JB, Posternak MA, Friedman M, Attiullah N, Boerescu D (2006). How should remission from depression be defined? The depressed patient's perspective. *Am J Psychiatry* **163**:148–150.