

This is a repository copy of *Investigating the relationship between costs and outcomes for English mental health providers : A bi-variate multi-level regression analysis*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/119695/>

Version: Accepted Version

Article:

Moran, V. and Jacobs, R. orcid.org/0000-0001-5225-6321 (2018) Investigating the relationship between costs and outcomes for English mental health providers : A bi-variate multi-level regression analysis. *European Journal of Health Economics*. pp. 709-718. ISSN 1618-7601

<https://doi.org/10.1007/s10198-017-0915-5>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Investigating the relationship between costs and outcomes for English mental health providers: A bi-variate multi-level regression analysis.

Valerie Moran and Rowena Jacobs.

Keywords: Mental health, costs, outcomes, bi-variate model, multi-level model, provider performance.

Abstract

Provider payment systems for mental health care that incentivise cost control and quality improvement have been a policy focus in a number of countries. In England, a new prospective provider payment system is being introduced to mental health that should encourage providers to control costs and improve outcomes. The aim of this research is to investigate the relationship between costs and outcomes to ascertain whether there is a trade-off between controlling costs and improving outcomes.

The main data source is the Mental Health Minimum Data Set (MHMDS) for the years 2011/12 and 2012/13. Costs are calculated using NHS Reference Cost data while outcomes are measured using the Health of the Nation Outcome Scales (HoNOS). We estimate a bivariate multi-level model with costs and outcomes simultaneously. We calculate the correlation and plot the pairwise relationship between residual costs and outcomes at the provider level.

After controlling for a range of demographic, need, social and treatment variables, residual variation in costs and outcomes remains at the provider level. The correlation between residual costs and outcomes is negative, but very small suggesting that cost-containment efforts by providers should not undermine outcome-improving efforts under the new payment system.

1. INTRODUCTION

Provider payment systems for mental health care that incentivise both cost control and quality improvement have been a policy focus in a number of countries. In the Netherlands, psychiatric care is included in the prospective activity-based payment system for inpatient and outpatient care [1]. Cost control is incentivised by nationally agreed unit prices and the system also incentivises quality improvements that reduce resource consumption [2]. In the US, psychiatric inpatient care provided under Medicare is reimbursed using a prospective per diem payment system that links payment to average cost in order to promote efficiency. The system employs variable per diem payments with higher payments at the beginning of the inpatient stay to reflect higher cost. The payment system is also designed to prevent adverse effects on quality of care by reimbursing readmissions within a short period of time at the per diem rate that was applied at the time of discharge [3].

In England, mental health services have historically been funded through block contracts that do not necessarily incentivise providers to control cost [4] nor has payment been aligned to patient outcomes [5]. The prospective activity-based system used in the acute physical health care sector – the National Tariff Payment System (NTPS) – has recently been introduced to mental health. A new classification system has been developed for mental health with a primary focus on patient need and severity, an important predictor of mental health resource use [6]. The currencies or units of activity for which payment will be made are 21 care clusters that are independent of care setting. Users of mental health care services are allocated to a care cluster by clinicians using the Mental Health Clustering Tool (MHCT). The MHCT incorporates items from the Health of the Nation Outcome Scales (HoNOS) [7] and the

Summary of Assessments of Risk and Need (SARN) [8]. HoNOS is comprised of 12 items, each scored from 0 (no problem) to 4 (severe problem) giving a total score ranging from 0 (best) to 48 (worst). Ratings are made by an individual clinician (psychiatrist, nurse, psychologist, or social worker) or are based on a team rating. The rating is made on the basis of all information available to the clinician and is based on the most severe problem that arose during the two weeks leading up to the point of rating.

After a patient is allocated to a care cluster, they should be reviewed on a regular basis to ensure that the care cluster continues to meet their needs. The time between care cluster assessments or cluster review periods (CRPs) forms the basis of contracts and prices agreed between providers and commissioners (responsible for the organisation and purchase of mental health care for their populations) [9, 10]. The aspiration is that each care cluster will have a fixed national price or tariff [10]. This will provide an incentive to control costs as providers with costs above (below) the tariff will incur financial losses (surpluses). Quality of care and patient outcomes will also be incentivised under the new payment system with the inclusion of quality and outcome measures in provider contracts and the intention to link these to prices [11]. While the care cluster currencies cover most services for working age adults and older people, some services such as children and adolescent, drug and alcohol, and specialist mental health services are not included and will be reimbursed under separate non-cluster currencies [12].

A small number of studies [13-16] have examined the relationship between costs and quality in mental health care with no clear consensus on whether a trade-off exists between the two. Previous studies investigating the relationship between costs and quality in both physical and mental health care have revealed that this is a challenging endeavour. Particular challenges relate to the availability of adequate measures of quality, small sample sizes and the endogenous relationship between costs and quality. Regarding the latter, a number of studies

[17-19] have used instrumental variables in order to consistently estimate the causal relationship. Nevertheless other studies [20, 21] have highlighted the inherent difficulty of addressing endogeneity including the limited availability of suitable instrumental variables. Given the challenge of finding suitable instrumental variables, we avoid the causal identification problem and motivated by the methodology used in studies that have assessed performance in physical health care [22, 23], we analyse costs and outcomes using two separate equations and allow for a correlation in responses. As in previous studies [14, 21] we measure quality in terms of an outcome measure – the Health of the Nation Outcome Scale (HoNOS). Following earlier studies [18, 21-23] we use multi-level modelling which allows us to examine the correlation in residual responses at provider-level to provide insight into the relationship between costs and outcomes and whether a potential trade-off exists. The use of a large, nationally representative dataset with individual-level data moves us beyond prior studies in mental health care that were constrained by aggregate data [15, 16] or small patient sample sizes [13, 14].

The aim of this paper is to investigate the relationship between costs and outcomes for mental health providers in England to ascertain if incentives provided by the NTPS to control costs can be achieved without negatively affecting patient outcomes. We do not attempt to estimate the causal relationship between costs and outcomes; rather we estimate a multi-level bivariate model with costs and outcomes as responses. We calculate the correlation between the residual variation in costs and outcomes for providers, before and after controlling for a range of risk-adjustment covariates encompassing socio-demographic, need and treatment variables. Our method also allows us to group providers according to their performance on residual costs and outcomes.

We contribute to existing evidence in several ways. This paper is the first to use a multi-level bivariate model to examine mental health cost and outcome responses simultaneously and calculate the correlation in residual variation between two responses. We isolate the residual variation in costs and outcomes attributable to mental health providers in order to assess provider performance. Moreover, while previous studies in mental health have used data with limited geographical or provider samples, we use a nationally-representative dataset that contains data for all specialist mental health providers in England. This means that we can examine costs and quality for both admitted and non-admitted care and are not constrained to just one care setting as in previous studies. Finally, our findings provide a better understanding of the incentives introduced by the new payment system and whether a trade-off between cost containment and outcome improving efforts exists.

2. METHODS

The CRP forms the unit of observation in this analysis. As a patient can have more than one CRP we utilise a multi-level model to reflect CRPs nested within patients who are in turn nested within providers.

We estimate the following bivariate model with two response variables for CRP i in patient j in provider k : cost y_{1ijk} and outcome y_{2ijk} :

$$\begin{cases} y_{1ijk} = \alpha_1 + \beta_1 X_{1ijk} + u_{1k} + v_{1jk} + \varepsilon_{1ijk} \\ y_{2ijk} = \alpha_2 + \beta_2 X_{2ijk} + u_{2k} + v_{2jk} + \varepsilon_{2ijk} \end{cases} \quad (1)$$

$$\begin{pmatrix} u_{1k} \\ u_{2k} \end{pmatrix} \sim N(0, \Omega_u) : \begin{pmatrix} \sigma^2_{u1} & \\ \sigma_{u1u2} & \sigma^2_{u2} \end{pmatrix} \quad (2)$$

X_{1ijk} and X_{2ijk} represent vectors of risk-adjustment covariates for the cost and outcome equations respectively. The provider-level residual variance for costs and outcomes is

represented by the random effects u_{1k} and u_{2k} respectively. The patient-level residual variances for each response are denoted by v_{1jk} and v_{2jk} ; while ε_{1ijk} and ε_{2ijk} signify the error terms at the CRP level for each response. The provider-level residuals, u_{1k} and u_{2k} are both assumed to follow a bivariate normal distribution with zero mean and covariance matrix Ω_u . Our interest lies in the correlation between the residual variation in y_{1ijk} and y_{2ijk} at the provider level which can be calculated as $r_{(x,y)} = \frac{\sigma_{u_1 u_2}}{\sqrt{\sigma^2_{u_1} \sigma^2_{u_2}}}$. We calculate this correlation for responses with and without risk-adjustment to gain an insight into the extent to which our risk-adjustment variables account for correlation between outcomes.

Our cost response variable y_{1ijk} is modelled using a log-linear model and our outcome response variable y_{2ijk} using a linear model. The multilevel estimates are statistically efficient even if some observations have missing data for either response under the assumption that data is missing at random [24].

The coefficients for the cost response can be interpreted in terms of a percentage change in the geometric mean of cost. For covariates measured as dummy variables, this is the percentage change in the geometric mean resulting from a change in the variable from zero to one which can be calculated as $(\exp(\beta) - 1) * 100$. For continuous variables, the coefficient can be interpreted as the change in the geometric mean in cost resulting from a one unit change in this variable. The coefficients for the outcome response can be interpreted in terms of marginal effects; the change in this variable arising from a change in a binary independent variable from zero to one, and a one-unit change in continuous variables.

The model is estimated using restrictive iterative generalised least squares (RIGLS) which is equivalent to restricted maximum likelihood [25] in MLwiN [24] using the runmlwin command [26] in Stata 13.0 [27].

We performed a sensitivity analysis by excluding a provider that is an outlier on the outcome response.

3. DATA

3.1. Reference Cost data

Reference Cost data is submitted by publicly owned providers to the Department of Health and reflects the costs of providing mental health services. Reference Cost data for mental health care is reported as per diem costs per care cluster for admitted and non-admitted care separately for each provider. The Reference Cost data was cleaned by omitting data for outliers defined as greater than 4 times the cost reported in the previous (for 2012/13 data) or following year (for 2011/12 data) (n=102,121). This resulted in dropping one provider with consistently high costs for all clusters across both years giving a sample size of 55 providers.

To construct our cost response variable, we measured all activity during a CRP that corresponded to mental health services reimbursed under the care clusters. For each observation (CRP), we calculated the total number of days or length of stay in admitted and non-admitted care. These length of stay variables were then multiplied by the per diem unit costs for admitted and non-admitted care for the particular care cluster and provider in order to construct a variable reflecting the total cost associated with a CRP. It is important to highlight that the use of cost data reported at a provider level, albeit disaggregated by cluster and admitted and non-admitted care will conceal the true variation in cost that would be evident in data reported at the patient level. We used the 2011/12 Reference Cost data for activity between 1 April 2011 and 31 March 2012 and the 2012/13 Reference Cost data for

activity between 1 April 2012 and 31 March 2013. For activity between 1 April 2011 and 31 March 2013 we calculated a weighted average cost that reflects the number of days during a CRP in each year.

3.2. Mental Health Minimum Data Set (MHMDS)

The MHMDS is a patient-level data set with national coverage for England. It is mandatory for providers of specialist, including elderly, mental health services funded by the NHS to submit MHMDS data on a quarterly and annual basis. The MHMDS contains data on all the care and treatment received by a service user irrespective of setting. We use Version 4.0 of the MHMDS, which covers 2011/12 and 2012/13 and includes information pertaining to the NTPS for mental health. As patients are not potentially identifiable from the data, and are not directly involved in the research, ethical approval was not required. The MHMDS data was cleaned to remove observations that: are duplicates; have age coded as less than 18 years or greater than 110 years, and are treated by private providers. We also dropped observations with inpatient days in the 99th percentile for clusters 1 (Common mental health problems, low severity) and 2 (Common mental health problems) (n=833). Admission thresholds have increased [28-30] to the extent that patients are being admitted under the Mental Health Act (MHA) in order to access inpatient treatment [31]. Therefore, we would not expect patients in clusters 1 and 2 to receive long periods of inpatient treatment due to demand pressures on inpatient beds.

HoNOS is routinely collected as part of the MHMDS, both as part of the MHCT and as an outcome measure. We use the total HoNOS score recorded as part of the MHCT at the end of a CRP (follow-up HoNOS) as our outcome response. Risk-adjustment covariates can vary for each response. The total HoNOS score recorded at the start (baseline) of a CRP is included as a risk-adjustment variable for the outcome response only as previous studies [32-35] show

that baseline outcome is a consistent predictor of follow-up outcome. Risk-adjustment covariates for both responses cover demographic, treatment, need and social variables. Demographic variables include age, gender (with female as the reference category), ethnicity (White (reference category), Black, Asian and Other) and married/civil partner (with unmarried/no civil partner as the reference category). Age is grouped into five categories reflecting quintiles of the distribution in order to capture any non-linearities in the relationship with cost and outcomes with age 18-34 years the reference category. Variables reflecting if a patient has care co-ordinated under the Care Programme Approach (CPA) (a method of assessing, planning and reviewing the needs of a person with severe mental illness) or has been admitted to hospital under the Mental Health Act (MHA) provides information on severity and treatment. Missing values of CPA and MHA were coded as zero under the assumption that it is unlikely these observations were subject to the MHA or under CPA, given the high levels of scrutiny of these activities. We include dummy variables for the 21 care clusters to investigate the extent to which these explain variations in costs and outcomes. We use the cluster with the lowest cost (cluster 1 - common mental health problems, low severity) as the reference category. The MHMDS also contains data for a small area level geographic marker, the Lower Layer Super Output Area (LSOA) of the individual. LSOAs are a geographic hierarchy with a minimum population of 1000 and a mean of 1500 [36]. We matched LSOA codes in the MHMDS to data on the Index of Multiple Deprivation (IMD) Income Domain [37]. The IMD Income Domain measures the proportions of the population experiencing income deprivation in an area [37]. A higher score for the Income Domain indicates a greater proportion of the population in the area in which the patient lives experiences income deprivation. A dummy variable is included as a covariate for both responses in order to capture the year (2011/12 and 2012/13) in which the cluster commenced with 2011/12 as the reference category.

4. RESULTS

4.1. Response variables

Figure 1 shows our cost response variable and figure 2 our outcome response variable measured at the CRP level.

Insert Figure 1 around here

Insert Figure 2 around here

The graphs show that both variables are approximately normally distributed although the outcome response variable is slightly right skewed reflecting a smaller number of observations with high follow-up HoNOS scores (and more severe mental health problems).

4.2. Descriptive statistics

Our estimation sample is 697,022 observations treated by 55 providers of which 269,525 observations have both cost and outcome responses, 419,879 observations have the cost response only, and 7,618 have the outcome response only.

Table 1 displays the descriptive statistics for our estimation sample with reference categories in brackets.

Insert Table 1 around here

4.3. Estimation results

Table 2 displays the estimation results. All variables are statistically significant with the exception of Other ethnicity and cluster 11 (ongoing recurrent psychosis, low symptoms) for the total follow-up HoNOS dependent variable.

Insert Table 2 around here

For the log of total cost response variable, many of the cluster variables are associated with the largest effects. For example, cluster 14 (psychotic crisis) is associated with a 647% and cluster 17 (psychosis and affective disorder difficult to engage) a 555% increase in cost compared to cluster 1 (common mental health problems, low severity). Other clusters associated with psychosis (clusters 10, 13 and 15) are also associated with considerable increases of over 400% compared to cluster 1. CRPs that have an admission under the MHA variable are associated with a 98% increase in cost. In terms of demographic variables, Black ethnicity is associated with an increase of 9% in costs compared to White ethnicity while age of 63-79 years is associated with an increase in costs of 34% compared to age 18-34 years. CRPs that started in 2012/13 are associated with a 39% reduction in costs compared to CRPs that started in 2011/12. Unit costs for non-admitted care reported by providers fell from 2011/12 to 2012/13 and we believe this is driving the negative association between Total Cost and CRPs that started in 2012/13.

For the follow-up HoNOS response variable, a positive coefficient signifies a worse outcome. Covariates associated with an improved outcome include being married or having a civil partner, Asian and Black ethnicities compared to White ethnicity, and older age. Marriage/civil partnership and age 80 years or over are associated with a reduced HoNOS score of around 0.4 points while Black ethnicity is associated with a reduction of 0.3 points. The MHA and CPA variables are associated with an increase in the follow-up HoNOS score of 0.4-0.5 points. Similar to the cost response, the clusters with higher severity are associated with greater magnitudes of effects. Compared to cluster 1 (common mental health problems, low severity) cluster 15 (severe psychotic depression) and cluster 16 (dual diagnosis, substance abuse and mental illness) are associated with an increase of 3.6 points and 3.9 points respectively. Clusters for cognitive impairment or dementia are associated with an increase of 4.4 points (cluster 20 – cognitive impairment or dementia, high need) and 5.5

points (cluster 21 – cognitive impairment or dementia, high physical need) compared to cluster 1. A CRP that started in 2012/13 is associated with an increased HoNOS score of around 0.2 points.

4.4. Residual variation in provider costs and outcomes

The correlation between residual costs and outcomes at the provider level was calculated as -0.004 for unadjusted outcomes and -0.02 for risk-adjusted outcomes suggesting little evidence of a meaningful relationship between the two variables.

Figure 3 shows the pairwise plot in risk-adjusted residual costs and outcomes for the providers in our analysis.

Insert Figure 3 around here

The providers fit quite evenly into four groups; those associated with 1) higher costs and lower follow-up HoNOS scores (better outcome) in the top left quadrant, 2) higher costs and higher follow-up HoNOS scores (worse outcome) in the top right quadrant, 3) lower costs and higher follow-up HoNOS scores (worse outcome) in the bottom right quadrant, and 4) lower costs and lower follow-up HoNOS scores (better outcome) in the bottom left quadrant. There is an outlier provider with a residual follow-up HoNOS score of just over 4 points above the average and slightly above-average residual costs. Compared to the average, this outlier provider has higher than average baseline and follow-up HoNOS scores, and higher than average proportions of observations of White ethnicity; in the older age groups (63-79 years and 80 years and over); under CPA; in cognitive impairment or dementia clusters (clusters 18 and 19); and with lower income deprivation.

The estimates of risk-adjusted residual costs and outcomes at the provider level are normalised to have a mean of zero. The follow-up HoNOS response variable is measured on

a continuous scale from 0 (best) to 48 (worst) meaning that a positive score for the residual total follow-up HoNOS score signifies a worse outcome. The best performer in relation to outcomes is associated with a risk-adjusted residual follow-up HoNOS score of 1.36 less than the average performer while the worst performer is associated with a risk-adjusted residual follow-up HoNOS score of 4.23 greater than the average performer. The estimates of residual Total Cost can be interpreted as percentage increases or decreases in the geometric mean of Total Cost compared to the average as calculated as $(\exp(\text{EB estimate}) - 1) * 100$. The best performing provider in relation to risk-adjusted residual Total Cost is associated with a Total Cost that is 71% lower than the average performing provider while the worst performing provider in relation to risk-adjusted residual Total Cost is associated with a Total Cost that is 194% higher than the average performing provider.

4.5. Sensitivity Analysis

The exclusion of the provider with the above-average risk-adjusted residual follow-up HoNOS score of 4.23 decreased the estimation sample to 681,305 observations. The estimation results were robust to this change. The correlation between risk-adjusted residual costs and outcomes at the provider level became -0.09. Figure 4 displays the pairwise plot in residual costs and outcomes for the 54 providers in the sensitivity analysis. The risk-adjusted residual follow-up HoNOS score reduced to 1.39 for the worst performing provider on follow-up HoNOS scores compared to the average performing provider for this response. There was also a small reduction regarding risk-adjusted residual Total Cost with the worst performing provider associated with a Total Cost that is 191% higher than the average performing provider.

Insert Figure 4 around here

5. DISCUSSION

The reimbursement of mental health care providers in England is undergoing a considerable reform with the introduction of a prospective, activity-based payment system. From April 2017, all mental health providers and commissioners are required to link payment to locally agreed quality and outcome measures [11] and some local health economies have already developed outcome measures and indicators for payment purposes [38]. The new system will offer incentives for providers to deliver care more efficiently while better meeting patient needs and improving outcomes. This research has explored the relationship between mental health costs and outcomes in order to examine the scope for providers to respond to the incentives introduced by the new payment system and makes an important contribution to the limited literature for mental health care.

Prior to risk-adjustment, we find evidence of a very small negative correlation between costs and outcomes at the provider-level. After controlling for a range of demographic, need, social and treatment factors, we find that residual variation remains in both costs and outcomes at the provider level. However, the correlation between residual costs and outcomes at the provider level is miniscule, which suggests that a trade-off between cost containment and outcome improving efforts on the part of providers is not a major concern in our data. Plotting the provider-level residual costs and outcome response variables reveals that providers broadly fall into four groups with an outlier provider. Providers with higher than average residual costs and higher than average residual follow-up HoNOS scores (worse outcome) may signify poor performance but may also indicate that certain providers are treating a case-mix that our model has not fully accounted for. While patient case mix is controlled for to a certain extent by the care clusters, the clustering method does not explicitly take diagnosis into account and it is likely that the clusters are very variable in terms of diagnosis and case mix [4, 39]. It may also be the case that some patients have treatment-resistant variants of mental illness which implies that they will be consuming large amounts

of care and resources with little discernible changes in outcome scores [40]. If certain providers have a higher case-load of such patients this could well explain their unexplained higher costs and worse outcomes. If the higher costs are legitimate then these providers may warrant additional payments as defined by any outlier policy attached to the payment system. A number of providers are associated with lower residual follow-up HoNOS scores (better outcome) but also with higher residual costs. These providers in particular may face a potential trade-off between costs and outcomes and efforts to reduce costs under a national tariff may compromise outcomes if providers are induced to undertake undesirable behaviours such as skimping on patient care. A number of providers have lower than average residual costs and lower than average residual follow-up HoNOS scores (better outcome). These providers are likely to benefit financially from the new payment system if a national tariff is introduced and patient outcomes are linked to provider payment. Providers with lower than average residual costs and higher than average residual follow-up HoNOS scores (worse outcome) may have scope to make financial profits under a national tariff but these may be offset if payment is linked to outcomes. If providers are achieving lower costs at the expense of patient outcomes then they would warrant particular scrutiny by commissioners under quality and outcomes standards established in the contracting process.

It is important to highlight a number of limitations regarding the data we use. Concerns have been raised regarding the quality of mental health Reference Cost data in particular in relation to variations in unit costs within clusters and between providers as well as issues around missing data [41]. A further limitation of the cost data used is that it is essentially provider-level cost data that underpins our dependent cost variable. While variation in the dependent cost variable will arise for patients in different clusters with different care patterns, a greater level of variation would be observed if we had access to data on the actual costs incurred by individual patients, rather than the provider average. While the MHMDS contains variables

on primary and secondary diagnoses, poor coding of this data inhibited inclusion of diagnosis in our set of risk-adjustment variables. A further data limitation is the missing data for follow-up HoNOS scores. We assume that this data is missing at random but this assumption would not hold if, for example, patients with more severe mental health problems are more likely to drop out of care and not have a follow-up HoNOS score recorded. Due to the poor coding of diagnosis data we are unable to investigate this hypothesis. Despite these data limitations, this research provides a valuable insight into the relationship between mental health costs and outcomes that is pertinent in the context of prospective activity-based payment. Our data show no evidence of a strong relationship between costs and outcomes. This provides some reassurance that outcome improvements can be obtained without spending a lot more, or, conversely, that some savings are possible. The research will be useful to commissioners of mental health services by providing an indication of how providers perform on disparate objectives. This work also benefits policymakers in planning future refinements to the payment system.

References

1. Kobel, C., et al., DRG systems and similar patient classification systems in Europe, in *Diagnosis-Related Groups in Europe Moving towards transparency, efficiency and quality in hospitals*, R. Busse, et al., Editors. 2011, Open University Press: England.
2. Swan-Tan, S., et al., DRGs and cost accounting: Which is driving which?, in *Diagnosis-Related Groups in Europe Moving towards transparency, efficiency and quality in hospitals*, R. Busse, et al., Editors. 2011, Open University Press: England.
3. Mason, A. and M. Goddard, *Payment by results in mental health: a review of the international literature and an economic assessment of the approach in the English NHS*. CHE Research Paper 50, 2009.
4. Jacobs, R., *Payment by results for mental health services: economic considerations of case-mix funding*. *Advances in psychiatric treatment*, 2014. **20**: p. 155 - 164.
5. Monitor and NHS England, *2014/15 National Tariff Payment System*. 2013, Monitor: London.
6. Trauer, T., *Outcome measurement - applications and utility*, in *Outcome Measurement in Mental Health Theory and Practice*, T. Trauer, Editor. 2010, Cambridge University Press: UK.
7. Wing, J.K., R.M. Curtis, and A.S. Beevor, *HoNOS - Health of the Nation Outcome Scales: Report of Research and Development*. 1996, Royal College of Psychiatrists Research Unit: London.
8. Self R., Painter J., and D. R., *A Report on the Development of a Mental Health Currency Model, (Including Suggestions on the Development and Testing of PbR Tariffs)* 2008, Department of Health: London.
9. Monitor and NHS England, *Guidance on mental health currencies and payment*. 2013, Monitor: London.
10. Department of Health Payment by Results team, *Mental Health Payment by Results Guidance for 2013-14*. 2013, Department of Health: Leeds.
11. NHS England and NHS Improvement, *2017/18 and 2018/19 National Tariff Payment System: A consultation notice*. 2016, NHS Improvement: London.

12. Department of Health, Reference costs 2011-12. 2012, Department of Health: Leeds.
13. Dickey, B. and S.L. Normand, Toward a model for testing the relationship between quality of care and costs. *J Ment Health Policy Econ*, 2004. **7**(1): p. 15-21.
14. Haas, L., et al., The trade-off between costs and quality of care in the treatment of psychosomatic patients with somatoform pain disorder. *Appl Health Econ Health Policy*, 2013. **11**(4): p. 359-68.
15. Hendryx, M., State mental health funding and mental health system performance. *J Ment Health Policy Econ*, 2008. **11**(1): p. 17-25.
16. Schulz, R.I., J.R. Greenley, and R.W. Peterson, Management, cost, and quality of acute inpatient psychiatric services. *Med Care*, 1983. **21**(9): p. 911-28.
17. Carey, K. and J.F. Burgess, Jr., On measuring the hospital cost/quality trade-off. *Health Econ*, 1999. **8**(6): p. 509-20.
18. Schreyogg, J. and T. Stargardt, The trade-off between costs and outcomes: the case of acute myocardial infarction. *Health Serv Res*, 2010. **45**(6 Pt 1): p. 1585-601.
19. Weech-Maldonado, R., D. Shea, and V. Mor, The relationship between quality of care and costs in nursing homes. *Am J Med Qual*, 2006. **21**(1): p. 40-8.
20. Hvenegaard, A., et al., Exploring the relationship between costs and quality: does the joint evaluation of costs and quality alter the ranking of Danish hospital departments? *Eur J Health Econ*, 2011. **12**(6): p. 541-51.
21. Gutacker, N., et al., Truly inefficient or providing better quality of care? Analysing the relationship between risk-adjusted hospital costs and patients' health outcomes. *Health Econ*, 2013. **22**(8): p. 931-47.
22. Hauck, K. and A. Street, Performance assessment in the context of multiple objectives: a multivariate multilevel analysis. *J Health Econ*, 2006. **25**(6): p. 1029-48.
23. Gutacker, N. and A. Street, Multidimensional performance assessment using dominance criteria. *CHE Research Paper 115*, 2015.
24. Rabash, J., et al., MLwiN Version 2.1. 2009, Centre for Multilevel Modelling: University of Bristol.
25. Goldstein, H., Restricted unbiased iterative generalized least-squares estimation. *Biometrika*, 1989. **76**(3): p. 622-623.
26. Leckie, G. and C. Charlton, runmlwin: A Program to Run the MLwiN Multilevel Modeling Software from within Stata. *Journal of Statistical Software*, 2012. **52**(11).
27. StataCorp, Stata 13.0. 2014, StataCorp LP.: College Station, Texas.
28. The Commission to review the provision of acute inpatient psychiatric care for adults, Improving acute inpatient psychiatric care for adults in England. Interim report. 2015, The Commission to review the provision of acute inpatient psychiatric care for adults.
29. Sabes-Figuera, R., et al., Developing a tool for collecting and costing activity data on psychiatric inpatient wards. *Epidemiol Psychiatr Sci*, 2012. **21**(4): p. 393-9.
30. Csipke, E., et al., Inpatient care 50 years after the process of deinstitutionalisation. *Soc Psychiatry Psychiatr Epidemiol*, 2014. **49**(4): p. 665-71.
31. House of Commons Health Committee, Post-legislative scrutiny of the Mental Health Act 2007 First Report of Session 2013-14. 2013, House of Commons: London.
32. Dow, M.G., T.L. Boaz, and D. Thornton, Risk adjustment of Florida mental health outcomes data: concepts, methods, and results. *J Behav Health Serv Res*, 2001. **28**(3): p. 258-72.
33. Rosen, A.K., et al., Improving risk adjustment of self-reported mental health outcomes. *J Behav Health Serv Res*, 2010. **37**(3): p. 291-306.
34. Hendryx, M.S., D.G. Dyck, and D. Srebnik, Risk-adjusted outcome models for public mental health outpatient programs. *Health Serv Res*, 1999. **34**(1 Pt 1): p. 171-95.

35. Kramer, T., et al., Comparing outcomes of routine care for depression: The dilemma of case-mix adjustment. *The Journal of Behavioral Health Services & Research*, 2001. **28**(3): p. 287-300.
36. Health and Social Care Information Centre, “*Lower Layer Super Output Area*”, in *NHS Data Dictionary*. 2013, Health and Social Care Information Centre: Leeds.
37. Noble, M., McLennan, D., Wilkinson, K., Whitworth, A., Barnes, H. and Dibben, C., *The English Indices of Deprivation 2007*. 2008, Communities and Local Government: London.
38. NHS England and NHS Improvement, *Linking quality and outcome measures to payment for mental health Technical Guidance*. 2016, NHS Improvement: London.
39. Yeomans, D., *Clustering in mental health payment by results: a critical summary for the clinician*. *Advances in psychiatric treatment*, 2014(20): p. 227-234.
40. Jencks, S.F., et al., *Bringing excluded psychiatric facilities under the Medicare Prospective Payment System. A review of research evidence and policy options*. *Med Care*, 1987. **25**(9 Suppl): p. S1-51.
41. PriceWaterhouseCoopers, *An Evaluation of the Reimbursement System for NHS-Funded Care: Report for Monitor*. 2012, PriceWaterhouseCoopers LLP: UK.

Table 1. Descriptive statistics

Variable	Observations	Mean	Standard Deviation
Total cost	689404	3448.076	9783.937
log of total cost	689404	6.919	1.615
Total HoNOS follow-up score	277143	11	6
Total HoNOS baseline score	277143	11	6
Married/civil partner [Unmarried/ no civil partner]	697022	0.331	0.470
[White ethnicity]	697022	0.877	0.329
Asian ethnicity	697022	0.046	0.208
Black ethnicity	697022	0.047	0.211
Other ethnicity	697022	0.031	0.173
[Age category 1 (18-34)]	697022	0.204	0.403
Age category 2 (35-46)	697022	0.191	0.393
Age category 3 (47-62)	697022	0.207	0.405
Age category 4 (63-79)	697022	0.204	0.403
Age category 5 (80+)	697022	0.194	0.395
Gender [Female]	697022	0.436	0.496
Admitted under the MHA	697022	0.088	0.283
Under CPA	697022	0.413	0.492
Cluster 0: Variance	697022	0.011	0.103
[Cluster 1: Common mental health problems, low severity]	697022	0.039	0.195
Cluster 2: Common mental health problems	697022	0.050	0.219
Cluster 3: Nonpsychotic,	697022	0.116	0.321

moderate severity			
Cluster 4: Non-psychotic, severe	697022	0.088	0.283
Cluster 5: Non-psychotic, very severe	697022	0.032	0.175
Cluster 6: Non-psychotic disorders of overvalued ideas	697022	0.017	0.128
Cluster 7: Enduring non-psychotic disorders	697022	0.039	0.193
Cluster 8: Non-psychotic chaotic and challenging disorders	697022	0.036	0.186
Cluster 10: First episode in psychosis	697022	0.027	0.163
Cluster 11: Ongoing recurrent psychosis, low symptoms	697022	0.090	0.286
Cluster 12: Ongoing or recurrent psychosis, high disability	697022	0.064	0.246
Cluster 13: Ongoing or recurrent psychosis, high symptom/disability	697022	0.046	0.209
Cluster 14: Psychotic crisis	697022	0.029	0.167
Cluster 15: Severe psychotic depression	697022	0.010	0.102
Cluster 16: Dual diagnosis, substance abuse and mental illness	697022	0.016	0.127
Cluster 17: Psychosis and affective disorder difficult to engage	697022	0.022	0.148
Cluster 18: Cognitive impairment, low need	697022	0.098	0.297
Cluster 19: Cognitive impairment or dementia, moderate need	697022	0.107	0.310
Cluster 20: Cognitive impairment or dementia, high need	697022	0.043	0.204
Cluster 21: Cognitive impairment or dementia, high physical need	697022	0.019	0.135
Income Deprivation	697022	17.988	11.785
CRP started in 2012/13 [CRP started in 2011/12]	697022	0.423	0.494

Table 2. Estimation Results

	Observations per group			
	Number of observations	Minimum	Average	Maximum
Level 3: Provider	55	33	12673	54090
Level 2: Person	414092	1	1.7	43
Level 1: CRP	697022			
Log likelihood	-2065741			
	Log of total cost		Total follow-up HoNOS	
	Coefficient	Standard Error	Coefficient	Standard Error
Total HoNOS baseline score			0.388	0.002***
Married/civil partner	0.009	0.004*	-0.378	0.025***
Asian ethnicity	0.026	0.009**	-0.121	0.057*
Black ethnicity	0.083	0.010***	-0.302	0.063***
Other ethnicity	0.031	0.011**	-0.050	0.070
Age category 2 (35-46)	0.086	0.006***	0.191	0.037***
Age category 3 (47-62)	0.147	0.006***	0.169	0.036***
Age category 4 (63-79)	0.295	0.007***	-0.338	0.041***
Age category 5 (80+)	0.181	0.008***	-0.401	0.048***
Gender	0.011	0.004**	0.236	0.023***
Admitted under the MHA	0.681	0.008***	0.484	0.042***
Under CPA	0.231	0.005***	0.407	0.026***
Cluster 0: Variance	0.286	0.019***	1.094	0.128***
Cluster 2: Common mental health problems	0.378	0.012***	0.539	0.084***
Cluster 3: Nonpsychotic, moderate severity	0.686	0.010***	1.262	0.075***
Cluster 4: Non-psychotic, severe	1.019	0.011***	2.232	0.076***
Cluster 5: Non-psychotic, very severe	1.323	0.013***	3.129	0.087***
Cluster 6: Non-psychotic disorders of	1.284	0.016***	3.003	0.102***

overvalued ideas				
Cluster 7: Enduring non-psychotic disorders	1.280	0.013***	2.995	0.085***
Cluster 8: Non-psychotic chaotic and challenging disorders	1.347	0.013***	3.390	0.087***
Cluster 10: First episode in psychosis	1.684	0.014***	1.550	0.092***
Cluster 11: Ongoing recurrent psychosis, low symptoms	1.035	0.011***	0.125	0.076
Cluster 12: Ongoing or recurrent psychosis, high disability	1.468	0.012***	1.833	0.078***
Cluster 13: Ongoing or recurrent psychosis, high symptom/disability	1.720	0.013***	3.020	0.083***
Cluster 14: Psychotic crisis	2.011	0.014***	3.412	0.089***
Cluster 15: Severe psychotic depression	1.626	0.020***	3.562	0.118***
Cluster 16: Dual diagnosis, substance abuse and mental illness	1.528	0.017***	3.871	0.103***
Cluster 17: Psychosis and affective disorder difficult to engage	1.880	0.015***	3.373	0.095***
Cluster 18: Cognitive impairment, low need	0.186	0.011***	0.385	0.079***
Cluster 19: Cognitive impairment or dementia, moderate need	0.550	0.011***	2.310	0.079***
Cluster 20: Cognitive impairment or dementia, high need	0.808	0.013***	4.354	0.089***
Cluster 21: Cognitive impairment or dementia, high physical need	0.681	0.016***	5.462	0.113***
Income Deprivation	0.000	0.000*	0.014	0.001***
CRP started in 2012/13	-0.494	0.004***	0.195	0.020***
Constant	5.934	0.057***	4.648	0.148***
Random Effects Parameters	Estimate	Standard Error	95% Confidence Interval	
Level 3: Provider				
Variance: Log of total cost	0.170	0.033	0.106	0.234
Variance: Follow-up total HoNOS	0.748	0.052	-0.109	0.095
Covariance: Log of total cost, Follow-up total HoNOS	-0.007	0.159	0.436	1.060
Level 2: Person				
Variance: Log of total cost	0.291	0.004	0.284	0.298
Variance: Follow-up total HoNOS	5.407	0.013	0.038	0.089
Covariance: Log of total cost, Follow-up total HoNOS	0.063	0.073	5.263	5.551
Level 1: CRP				
Variance: Log of total cost	1.768	0.004	1.760	1.776
Variance: Follow-up total HoNOS	21.000	0.015	-0.078	-0.017
Covariance: Log of total cost, Follow-up total HoNOS	0.047	0.078	20.848	21.153

***p<0.001, **p<0.01, *p<0.05

Figure 1. Log of Total Cost

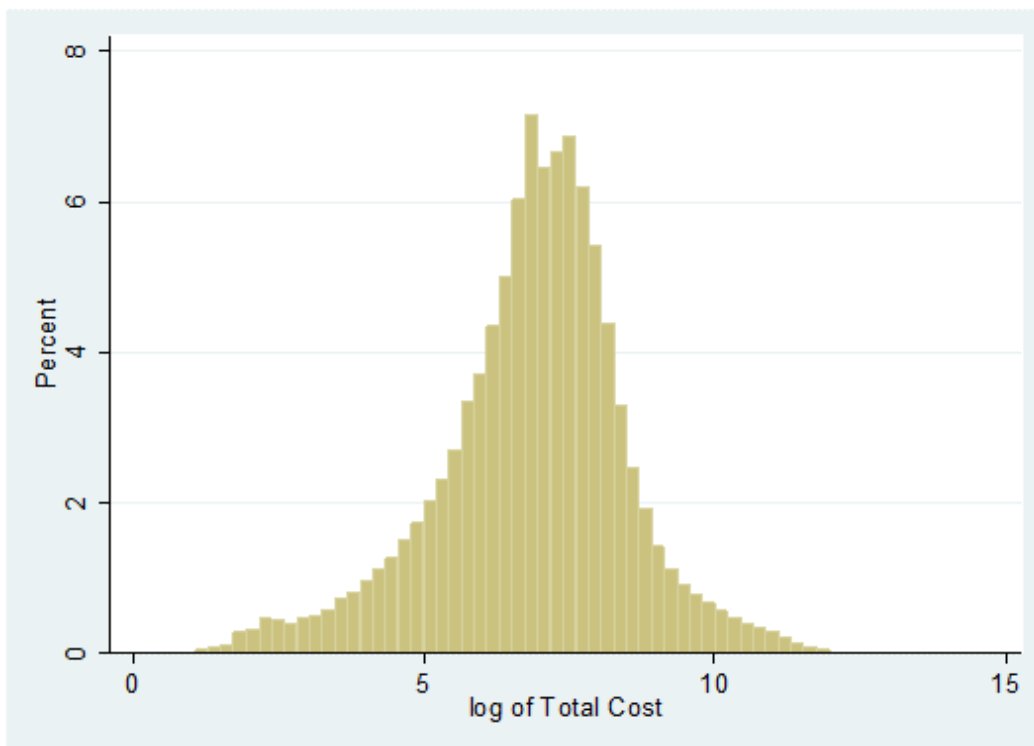


Figure 2. Total follow-up HoNOS score

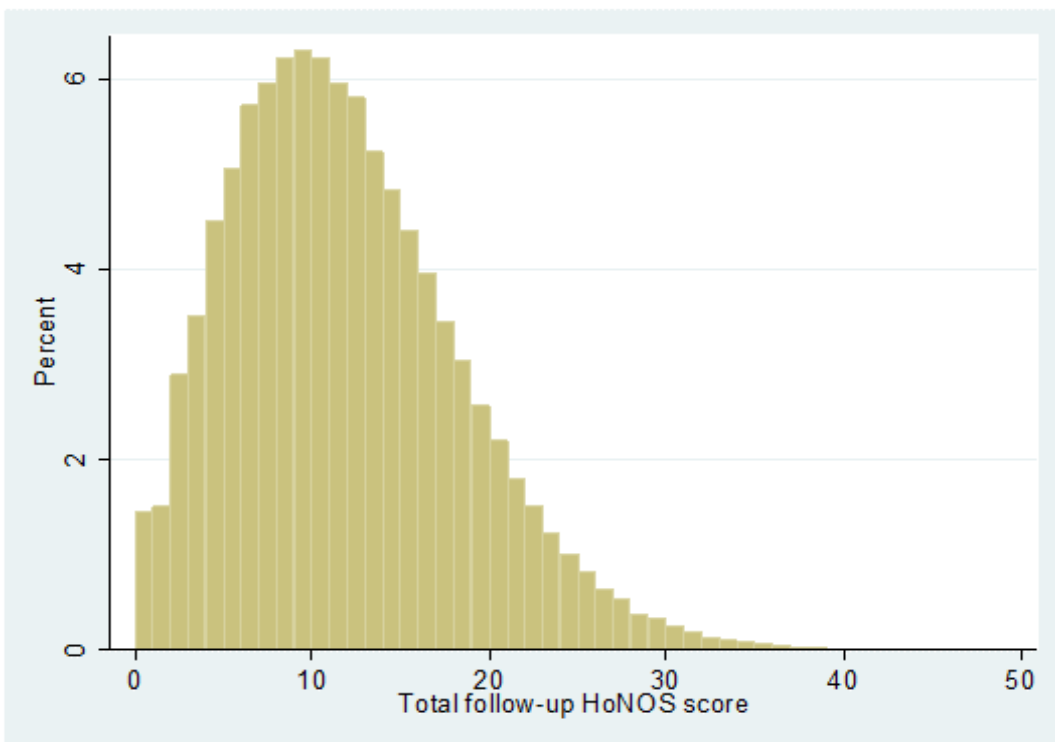


Figure 3. Pairwise plot of residual costs and outcomes for providers

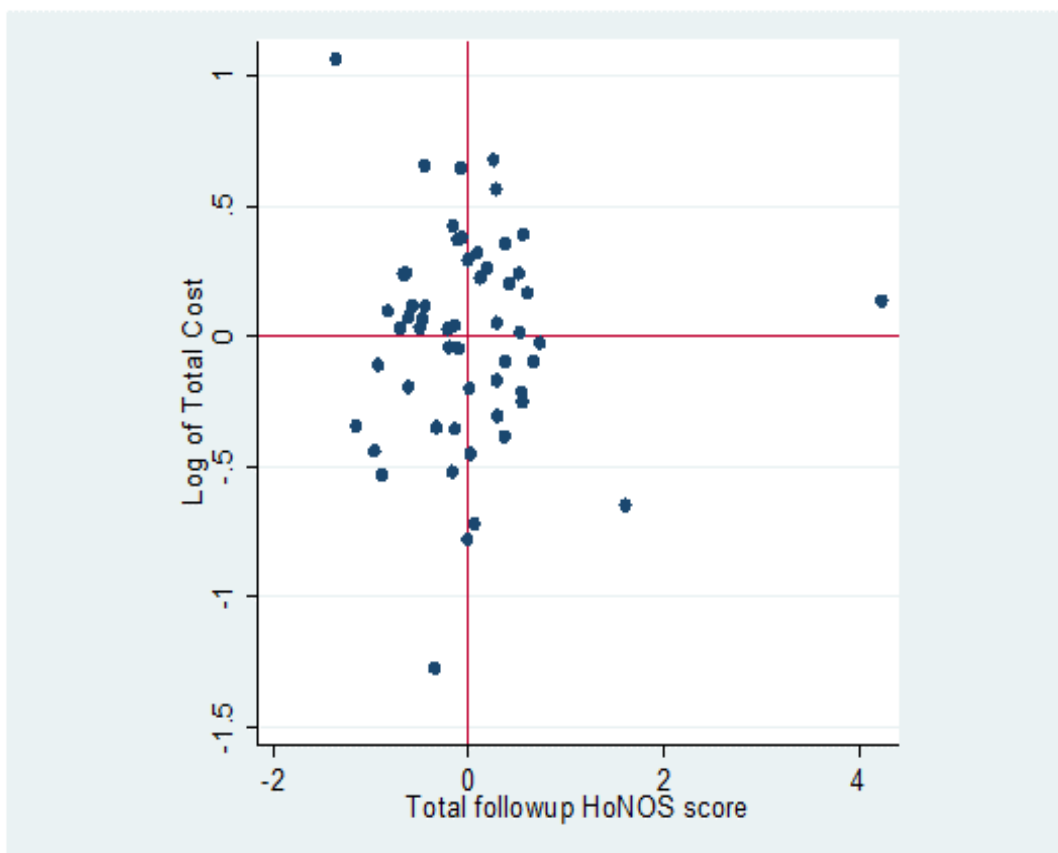


Figure 4. Pairwise plot of residual costs and outcomes for providers for sensitivity analysis

