ACRA(2012)18A

DEVELOPING THE MENTAL HEALTH FUNDING FORMULA FOR ALLOCATIONS TO GENERAL PRACTICES

Phase 3: Estimation of a formula based on person-level data

24th August 2012

Matt Sutton, University of Manchester
Søren Rud Kristensen, University of Manchester
Yiu-Shing Lau, University of Manchester
Gyles Glover, NEPHO and University of Durham
William Whittaker, University of Manchester
John Wildman, Newcastle University
Hugh Gravelle, University of York
Peter Smith, Imperial College London

Summary

This paper contains final recommendations for a person level model based on the MHMDS for the age and additional needs components of the mental health services formula.

The models control for practice-level supply variables and PCT differences and include the following sets of needs variables:

- gender and age (person level)
- mental health prevalence rates and population characteristics (practice level)
- prevalence rates of particular non-psychiatric conditions (age-sex-practice level)
- psychiatric diagnosis markers (person level)
- condition severity and enhanced care markers (person level)
- personal characteristics of service users (person level)

A model containing only the RAMP formula variables estimated using the updated data explains 32.7% of the variation in costs per capita across practices. Use of better practice-level variables (including prevalence rates from the QOF) increases this explanatory power to 42.7%. Addition of the person-level variables increases the model fit to 61.4%. Estimation of the model as a two-part model results in the highest explanatory power, at 63.6%.

We recommend use of the two-part model containing person-level variables but include a wide range of alternative models for ACRA's consideration.

1. Variable sets

The models used to underpin previous mental health formulae have included demographic variables, practice and/or area variables, and PCT indicators. Adoption of a person-based approach enables us to use a wider range of needs variables.

The sets of variables that we have considered for inclusion in the Person-based Resource Allocation for Mental health (PRAM) model are listed in Table 1. As well as the sets of variables used in previous formulae, we have included: prevalence rates of non-psychiatric conditions; markers for psychiatric diagnoses; condition severity markers and personal characteristics recorded for service users.

Descriptions of the variable sets are provided in Appendix 1.

Table 1. Sets of variables considered for inclusion in the PRAM model

Variable set	Variables
Needs variables	
Demographic group	28 gender and five-year age band interactions
Practice variables	Variables measured directly for practice populations and LSOA variables attributed to practices using the ADS
Prevalence rates of particular non-psychiatric conditions	Proportions of people in this age-sex-practice stratum diagnosed with particular conditions in previous two years of HES
Psychiatric diagnosis markers	43 flags for three-digit ICD10 codes recorded in the MHMDS in previous two years
Condition severity and care markers	8 categories of severity and an enhanced care supplement derived from care patterns in the MHMDS in previous two years
Personal characteristics of service users	Ethnicity, marital status, employment status, accommodation status and year first received psychiatric care for individual service users in the MHMDS
Supply variables	
Practice variables	Variables measured directly for practice populations and LSOA variables attributed to practices using the ADS
PCT indicators	151 binary membership indicators

2. Modelling strategy

As we have done previously, the underpinning models are estimated on all persons registered with a random 50% sample of GP practices separately for three population groups: males aged 16-64 years, females aged 16-64 years, and persons aged 65 years and over. The predictions from these models have been compared to the costs of all persons registered with the remaining 50% of practices.

We have sought to produce models that contain only statistically significant variables with plausibly-signed coefficients. We have followed a modelling strategy that allows us to examine how the coefficients and their significance change as we add new variables to the more basic models. Along the way, we show how addition of more variables adds to the fit of the models at person- and practice-levels. We focus on the model fit at practice level and as well as the R2 (as a measure of the percentage of variation explained), we provide statistics on the extent of variation across practices in predicted costs per capita and in the mean predicted values for deciles of practices grouped by their actual costs per capita. A good model fits the data well and provides a level of variation in the predictions that reflects the clustering of high cost individuals in some practices.

We began with simple models that contain only demographic and practice variables. For comparison purposes, we started with a model that included the practice variables used in the RAMP formula. We then sought to improve on this model by including other practice-level variables that we had found to be significant in earlier stages of this project. These include the prevalence rates for severe mental illness and dementia measured for the QOF.

At key stages in the development of the model, we provide results for variants of the model involving deletion of controversial variables or addition of potentially important variables. Before we added person-level variables, we considered in particular whether measures of the marital status composition of the local population (as measured at the 2001 Census) predicted variations in mental health care costs.

The addition of the person-level needs markers reduced the significance of the practice-level variables and some became not significant at $p<0.05$. Before we dropped these variables, we checked whether they were significant in the two-part model. We only dropped practice-level variables that were not significant in both the one-part and two-part models once we had included in the person-level variables.

Some of the condition severity and care markers might be judged to be sensitive to supply conditions and inappropriate for a needs model. To show how inclusion or exclusion of these markers affects the model fit, we considered variants that did not include them. Our preferred specification of these markers includes all but the last category, which captures those with the least severe conditions. We present variants that: include this category; also

exclude those with frequent contact with a single type of care professional; exclude the marker for CPA; and exclude hospital admission as a needs marker.[1]

Given the low proportion of the population that uses mental health services, it may be better to use a two-part model to estimate the determinants of mental health care costs. The first part of the two-part model is concerned with the predictors of the probability of using services. The second part of the two-part model focuses on the predictors of variations in costs amongst the service-using population. We would expect needs and supply variables to have different effects on these two parts of total healthcare costs. In the final stages of the modelling, we therefore re-estimated our preferred specification using a two-part model. We used logistic regression for the first part of the model and a generalised linear model with a log link for the second part of the model to ensure that the predicted values were all positive.

The final model that we estimate involves addition of the personal characteristics collected in the MHMDS. These are only available for persons that have used services and so we estimated this model only on this sample. For persons who have not used services, the prediction is calculated by multiplying the prediction from the first part of the model by the prediction from the second part of the model that does not include service user characteristics. For persons who have used services, the prediction is calculated by multiplying the prediction from the first part of the model by the prediction from the second part of the model that does include service user characteristics.

In total, we present results from 16 key models.[2] The content of the key models is summarised in Table 2. In each row, our preferred models at this stage are shown in the leftmost column. Variants at this stage of the development of the models are shown in the other columns in the same row.

[1] We have placed hospital admission as the third most severe category. When we excluded this category, we recalculated the categories into which each person was placed by the rest of the classification.
[2] Note: Due to an error in the numbering of the models, the 16 models are numbered 1-4 and 6-17. There is no model 5.

Table 2: Contents of the key models

Model	Variables	Alternatives			
M1.	Gender and age only				
M2.	M1 with RAMP variables				
M3.	M2 with PCT membership indicators				
M4.	M3 with new practice variables	M6. M4 with marital status	M7. M4 with Census LLTI	M8. M4 with HES rates	
M9.	M3 with refined practice variables				
M10.	M9 with person level needs markers				
M11.	M10 without insignificant practice variables	M12. M11 with category for ‘any historical use’	M13. M11 without frequent contact category	M14. M13 without CPA	M15. M14 without ‘admission’ category
M16.	Two-part version of M11				
M17.	M16 with more variables for service users				

3. Results

The coefficients on each of the variables and the statistics on goodness-of-fit are provided in an accompanying Excel file.

The coefficients are shown in Tables E1-E6. The contents of these tables are as follows:

- Table E1: Coefficients on the practice- and stratum-level variables for males aged 16-64 years from models 9, 11 and 17.
- Table E2: Coefficients on the practice- and stratum-level variables for females aged 16-64 years from models 9, 11 and 17.
- Table E3: Coefficients on the practice- and stratum-level variables for persons aged 65+ years from models 9, 11 and 17.
- Table E4: Coefficients on the person-level condition severity and care markers for the three age-sex groups from model 11.
- Table E5: Coefficients on the person-level psychiatric diagnostic markers for the three age-sex groups from model 11.
- Table E6: Coefficients on the service user characteristics for each of the age-sex groups from model 17.

The goodness-of-fit statistics at person- and practice- levels for each of the models are shown in Tables E7 and E8. The remaining tables in the Excel file provide summary statistics for the variables.

As we have shown in previous stages of the work, the model that contains only the gender and age interactions (Model 1) has virtually no explanatory power (R2=0.36%) at practice level. Addition of the practice variables from the RAMP formula in Model 2 increases the R2 to 21.02%. Addition of the PCT membership indicators in Model 3 increases the percentage of variation explained to 32.69%.

Use of the new set of practice-level variables in Model 4 increases the goodness-of-fit of the model to 42.67%. The new sets of practice needs and supply variables are listed in Table 3.

Table 3: Practice needs and supply variables included in Model 4

Needs variables	Supply variables
Persons aged 16-64 years	
QOF Severe Mental Illness prevalence Proportion receiving IB/SDA for a mental health condition Student practice (negative sign) Proportion Black Caribbean or Black African (in the model for males only)	PCT membership indicators
Persons aged 65+	
QOF Dementia prevalence Standardised Mortality Ratio where mental illness excluding dementia is indicated Proportion of population receiving pension credit as a single person Proportion reporting longstanding health problem in GP Patient Survey	PCT membership indicators Proportion providing more than 19hours unpaid care per week (negative sign) Distance to Integrated Community Mental Health Team base (negative sign)

In Model 6, we experiment with including the following marital status composition variables from the 2001 Census:

- Proportion of the population who are single (never married)
- Proportion of the population who are separated (but still legally married)
- Proportion of the population who are divorced
- Proportion of the population who are widowed

We find that none of these variables is statistically significant for persons aged 65+. For both males and females, the proportion who are widowed and the proportion who are single are both significant and positively signed. For females, the proportion who are divorced is also significant and positively signed.

In Model 7 we checked whether the proportion reporting a longstanding health problem in the GP Patient Survey should be replaced with the same measure from the Census in the older people's model. The Census measure was not statistically significant.

In Model 8, we included the prevalence rates of non-psychiatric conditions from Hospital Episode Statistics. The statistically significant rates for each group are shown in Table 4. Although the prevalence rate of "unknown and unspecified causes of morbidity" is significant in all of the models, we have not included it in the preferred models as we suspect it reflects variations in supply rather than need.

Table 4: Significance of prevalence rates of non-psychiatric conditions from HES

Diagnostic group	Males, 16-64	Females, 16-64	Persons, 65+
B15-B19 - Viral hepatitis	+		
R40-R46 - Symptoms & signs involving cognition, perception etc.		+	
R69 - Unknown & unspecified causes of morbidity	+	+	+
T00-T07 - Injuries involving multiple body regions			
T08-T14 - Injuries to unspecified part of trunk limb or body			
T15-T19 - Effects of foreign body entering through natural orifice			
T36-T50 - Poisonings by drugs & biological substances	+	+	

Our preferred model at this stage, prior to the introduction of person-level needs variables, is Model 9. As well as the variables in Model 4, we include the significant marital status composition measures and non-psychiatric prevalence rates. The explanatory power of this model is only a marginal improvement over Model 4 (42.75% versus 42.67%).

The addition of the person-level psychiatric diagnoses and condition severity markers in Model 10 results in a much larger increase in the R2 to 61.41%. The proportion of the population from Black Caribbean or Black African ethnic groups becomes insignificant in the model for males aged 16-64 years. We checked that this variable was also not significant in the two-part model and removed it to arrive at Model 11. The remaining practice-level variables remained significant or were significant in the two-part model.

The psychiatric diagnoses with the largest and most significant coefficients are Schizophrenia (F20), Schizoaffective Disorders (F25) and Bipolar Affective Disorder (F31). The coefficients on the diagnoses of disorders due to use of addictive substances tend to have negative coefficients. This is expected, because our costs exclude addiction services. In general, few of the coefficients on the diagnosis markers are statistically significant because, as we have previously noted, they are rarely recorded and recorded only if an individual has had contact with a psychiatrist. Note also that these estimated effects are in addition to the effects of the severity categories, including hospitalisation. They become more significant if the severity categories are removed.

The severity categories predict costs as expected. Table 5 shows the numbers of persons (in the 50% estimation sample) in each category and the estimated coefficients when all are entered into a model (Model 12). Persons who have had some medium secure care in the previous two years have costs in 2010/11 that are approximately £12,000 more than those who have not. Persons who have had detained days but no days of medium secure care cost approximately £9,000 more. A history of hospital admission of at least two nights, with no history of medium secure care or detention, increases future costs by approximately £5,000. High, medium and low levels of multi-disciplinary care, with no history of hospitalisation, also increase future costs. Frequent contact with a single type of care professional increases future costs by approximately £2,500. Persons in none of the above categories, but in contact with mental health care services in the previous two years, have on average £300 (£600 if aged 65 years or over) higher costs in the next year. Finally, enhanced CPA increases costs by between £300 and £700 in the next year, depending on the person's gender and age.

Table 5: Frequencies of, and coefficients on, condition severity and care markers

	Number of persons	Coefficients (Model 12) Males (16-64)	Females (16-64)	65 and over
Category				
Medium secure care	1,280	£12,588	£13,801	£11,223
Detained under the Mental Health Act	39,128	£9,800	£8,787	£8,762
At least two nights in a mental health hospital	48,578	£4,967	£5,386	£7,271
At least 4 contacts with each of 4 different professionals	4,784	£5,580	£5,281	£5,436
At least 3 contacts with each of 3 different professionals	26,577	£3,313	£3,282	£2,886
At least 2 contacts with each of 2 different professionals	120,790	£1,639	£1,605	£1,605
Frequent care - at least 30 contacts with 1 professional	9,426	£2,756	£2,571	£2,356
Any contact with mental health care services	552,147	£331	£312	£611
Supplement				
Enhanced CPA	305,126	£680	£500	£286

Model 12, which includes the any historical marker category, has an R2 that is slightly higher than Model 11 (61.55% versus 61.42%). Model 13 shows the effect of dropping the frequent contact category from Model 11. There is only a small reduction in the goodness of fit at practice level (from 61.42% to 61.38%). In Model 14 we also drop the enhanced CPA marker. This results in a higher R2 compared to Model 13 and to Model 11 at 61.48%. Omission of the hospital admission category in Model 15 reduces the R2 to 59.28%.

Model 16 is the two-part version of Model 11. Use of the two-part approach increases the goodness of fit from 61.42% to 62.54%. As we would expect, the practice-level needs variables exert a stronger influence on the proportions of persons using services than on the costs of those who do use services. The person-level needs variables are more significant in explaining variations in costs amongst service users.

Finally, in Model 17, we use the personal characteristics from the MHMDS in the second part of the two-part model for service users. Compared to Model 16, the R2 increases from 62.54% to 63.62%. The coefficients (in Table E6) show that single, divorced and widowed marital status, non-settled accommodation, unemployment, Caribbean and African ethnicity, and earlier first contact with psychiatric care (amongst those aged 16-64) are all associated with higher costs.

Table E7 shows the mean predicted costs for persons with zero costs and deciles of costs for persons with non-zero costs from each of the models. The predictions from the models that do not contain person-level needs markers do not increase monotonically across the actual cost deciles. With the exception of decile 3, the models containing historical markers perform much better. Model 17 explains 18.2% of the variation in costs at person level.

Table E8 shows the mean predicted costs per capita for practices grouped into deciles of their actual costs per capita. The mean cost per capita is £89. The 10% of practices with the lowest actual costs per capita have a mean cost per capita of £22. The 10% of practices with the highest actual costs per capita have a mean cost per capita of £213. The standard deviations of predicted costs per capita, and the steepness of the predicted values across the actual cost deciles, follows the pattern of the practice R2-statistics. Model 17 has the

widest range of predictions, from £34 per capita in the lowest-cost practices to £172 per capita in the highest-cost practices.

Table 6 provides the practice-level R2 statistics for each of the key models discussed in the paper.

Table 6: Goodness-of-fit at practice level of the key models

Model	Variables	Alternatives			
M1.	Gender and age only (0.36%)				
M2.	M1 with RAMP variables (21.02%)				
M3.	M2 with PCT membership indicators (32.69%)				
M4.	M3 with new practice variables (42.67%)	M6. M4 with marital status (42.70%)	M7. M4 with Census LLTI (43.09%)	M8. M4 with HES rates (42.82%)	
M9.	M3 with refined practice variables (42.75%)				
M10.	M9 with person level needs markers (61.41%)				
M11.	M10 without insignificant practice variables (61.42%)	M12. M11 with category for 'any historical use' (61.55%)	M13. M11 without frequent contact category (61.38%)	M14. M13 without CPA (61.48%)	M15. M14 without 'admission' category (59.28%)
M16.	Two-part version of M11 (62.54%)				
M17.	M16 with more variables for service users (63.62%)				

4. Recommendations

For the three gender and age groups, the models with the highest explanatory power contain the following needs variables:

Males aged 16-64 years	**Females aged 16-64 years**	**Persons aged 65+**
Five-year age bands	Five-year age bands	Five-year age bands
QOF SMI prevalence IB/SDA for a mental health condition Student practice (negative sign) Proportion of population widowed Proportion of population single	QOF SMI prevalence IB/SDA for a mental health condition Student practice (negative sign) Proportion of population widowed Proportion of population single Proportion of population divorced	QOF dementia prevalence SMR for mental illness Single person pension credit Rate of long-term health problems
Prevalence rates in previous 2 years: • Viral hepatitis • Poisonings by drugs	Prevalence rates in previous 2 years: • Cognition/perception symptoms • Poisonings by drugs	
43 types of psychiatric diagnosis in previous two years	43 types of psychiatric diagnosis in previous two years	43 types of psychiatric diagnosis in previous two years
8 categories of condition severity and mental health care patterns in previous two years	8 categories of condition severity and mental health care patterns in previous two years	8 categories of condition severity and mental health care patterns in previous two years
For service users only: ethnicity, marital status, employment status, accommodation status and year first received psychiatric care	*For service users only*: ethnicity, marital status, employment status, accommodation status and year first received psychiatric care	*For service users only*: ethnicity, marital status, employment status, accommodation status and year first received psychiatric care

It is possible to exclude one category of mental health care patterns (frequent contact with a single type of care professional) and the supplement for enhanced CPA without a marked effect on the goodness-of-fit at practice level. We have no strong recommendation on these variables but have included them for completeness. Omission of the 'hospital admission for at least two nights' category reduces the fit more substantially and we recommend that this variable be included.

Adoption of the two-part specification (in which the differential effects of particular needs variables on the probability of using services and the average costs of those using services can be examined) leads to a more substantial increase in the goodness-of-fit at practice level. We recommend this approach. The predictions become closer to the distribution of actual costs when the personal characteristics of service users are also taken into account.

As these gains in the predictive power of the model at practice level can be attained with plausible effects of a wide range of needs variables, we recommend these person-based models for the mental health resource allocation formula.

Appendix 1: Sets of variables considered for inclusion in the models

Prevalence rates of non-psychiatric conditions
In earlier stages of this work, we considered using activity data from Hospital Episode Statistics (HES) to estimate the mental health model. We found that the MHMDS was preferable as it included mental health service activity provided in the community. However, HES has the advantage of including non-psychiatric admissions and therefore a wider range of diagnoses.

In the earlier work, we found seven non-psychiatric diagnostic predictors of psychiatric care costs. We calculated prevalence rates by strata (defined by age, sex and practice) for each of these conditions based on HES data for 2008/9 and 2009/10 and matched them to our modelling dataset. We then considered these as possible predictors of person-level costs in the MHMDS in 2010/11.

Psychiatric diagnosis markers
Up to 12 diagnosis can be recorded for each mental health care spell in the MHMDS. We identified the 62 three-digit ICD-10 diagnoses that were recorded in at least 1,000 spells in the 2008/9 and 2009/10 MHMDS. We omitted the diagnoses that were for physical conditions and most of the Z codes, and created 43 person-level diagnostic markers using the same methodology as used in PBRA3.

Condition severity and care markers
Within the MHMDS, person-level needs markers can be created from historic patterns of service use as well as diagnostic markers. The measures of service use in the MHMDS are:

- Number of medium secure bed-days
- Number of days detained under the Mental Health Act
- Number of intensive psychiatric inpatient bed-days
- Number of ordinary psychiatric inpatient bed-days
- Number of outpatient attendances
- Number of Community Psychiatric Nurse contacts
- Number of Clinical Psychologist contacts
- Number of Occupational Therapist contacts
- Number of Physiotherapist contacts
- Number of Consultant Psychotherapist contacts
- Number of Social Worker contacts

In previous work we have concentrated on nine putative markers of severe mental illness (SMI) that one of us (Glover, 2010) devised for a report to the Care Quality Commission.[3] The purpose of that work, using data from 2006/7 to 2008/9, was to derive prevalence counts of persons with SMI from the MHMDS that could be compared to the prevalence counts reported by GP practices in the Quality and Outcomes Framework.

[3] Glover G. *QOF Mental Health registers and the Mental Health Minimum Dataset.* Report to the CQC, July 2010.

For a needs-based resource allocation formula, we require person-level markers that reflect both the prevalence and intensity of need for mental health services. These markers should be as independent as possible of variations in service configuration so that they are comprehensive and comparable measures of person-level needs. In addition, we require markers that do not contain perverse incentives; i.e. it should not be possible for commissioners to manipulate these measures to increase the budgets that they receive and the specification of the measures should not seem to reward unnecessary or inappropriate care.

The combined use of the markers derived by Glover (2010) does not meet these criteria as this could reflect volumes as well as types of care. We have therefore derived eight mutually exclusive and exhaustive categories of care provision. The intention is that these categories should reflect decreasing levels of condition severity. The eight categories are:

1. Medium secure care
2. Detained under the Mental Health Act
3. Admitted to a mental health hospital for at least two nights[4]
4. High multi-disciplinary care – At least 4 contacts with each of at least 4 different types of care professional
5. Medium multi-disciplinary care - At least 3 contacts with each of at least 3 different types of care professional
6. Low multi-disciplinary care - At least 2 contacts with each of at least 2 different types of care professional
7. Frequent uni-disciplinary care - at least 30 contacts with a single type of care professional
8. Any contact with mental health care services

In addition, we consider the inclusion of a further marker indicating whether the individual is recorded as having had an enhanced Care Programme Approach (CPA) to their care. Individuals that need enhanced CPA are those who need: multi-agency support; active engagement; intense intervention; support with dual diagnoses; and who are at higher risk. Examples of aspects that suggest a need for enhanced CPA include; current or significant history of severe distress/instability or disengagement; significant reliance on carer(s) or having significant caring responsibilities; unsettled accommodation/housing issues; and employment issues when mentally ill.

Personal characteristics of service users

The MHMDS contains personal characteristics on service users that are likely to predict their costs. We have considered the following personal characteristics:

- Ethnic group
- Marital status
- Employment status
- Accommodation status
- Year first received psychiatric care

[4] The restriction to only include persons admitted for at least two nights excludes those individuals admitted for further assessment only.

The variables on employment and accommodation status were collected as a National Indicator in the PSA Delivery Agreement Number 16 on social exclusion. We considered the year in which an individual first received psychiatric care as a marker of an enduring mental health condition.

These variables are not recorded, not known or missing for some individuals. To avoid loss of data, we create separate categories for missing values on these variables.