PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	A systematic review of the predictors of health service utilisation by adults with mental disorders in the UK
AUTHORS	Twomey, Conal; Baldwin, David; Hopfe, Maren; Cieza, Alarcos

VERSION 1 - REVIEW

REVIEWER	Panagiotis Zis Evangelismos General Hospital, Athens, Greece
REVIEW RETURNED	22-Feb-2015

GENERAL COMMENTS	Few relevant papers must be cited: Killaspy H, Zis P. Predictors of outcomes for users of mental health rehabilitation services: a 5-year retrospective cohort study in inner London, UK. Soc Psychiatry Psychiatr Epidemiol. 2013 Jun;48(6):1005-12. Trieman N, Leff J. Difficult to place patients in a psychiatric hospital closure programme: the TAPs project 24. Psychol Med. 1996 Jul;26(4):765-74. Trieman N, Leff J. The TAPS project. 36: the most difficult to place long-stay psychiatric in-patients. Outcome one year after relocation. Team for the Assessment of Psychiatric Services. Br J Psychiatry. 1996 Sep;169(3):289-92. Trieman N, Leff J. Long-term outcome of long-stay psychiatric in-patients considered unsuitable to live in the community. TAPS Project 44. Br J Psychiatry. 2002 Nov;181:428-32.

REVIEWER	Steve Gillard St George's, University of London, UK
REVIEW RETURNED	09-Mar-2015

GENERAL COMMENTS	This is a timely, innovative and well conducted study. The case for further consideration of including demographic and co morbidity variables in the case mix for the MHCT is well made. It is perhaps something of a missed opportunity that LD wasn't included as a strong case has been made elsewhere that tariffs for treatment of people with a LD co morbidity should attract a premium in order that sufficient resource exists to appropriately assess, diagnose and treat people with LD for mental and physical health conditions (see recent publications by Tuffrey-Wijne).

	The authors might have argued that the study further questions the 'symptoms, behaviour, functioning' approach adopted by the MHCT (why were impairment and risk variables not explicitly considered in this review?) given there was only some evidence that some of those variables were associated with HSU. However it is not clear whether or not the authors are suggesting that some association of some variables (in each of the domains considered) with HSU is evidence of the relevance of the domain to clustering, or of its limited usefulness. For example, there is inconsistency in that the authors conclude that the review lends support for further consideration of the role of diagnosis - alongside pathway - in clustering systems (where only personality disorders was shown to be a predictor of HSU), while not also concluding that symptoms, functioning and behaviors should likewise be given further consideration (on the basis that a single variable was also shown to be associated with HSU in each case). Specifically, why is prediction of HSU in two or more studies an indication that that variable is a sufficiently good predictor of HSU? Those studies might be in different populations, settings etc. Given the heterogeneous nature of studies - and therefore the difficulties in drawing a general conclusion about the relevance of each particular variable or domain - the observation of any predictive relevance might be an indicator that that variable warrants further study. As such, the discussion needs to point out that further research needs to be done to understand the relationships between these variables and HSU. For example, HSU might be higher in women than men because women are better at accessing treatment. As such any inflation of tariff based on gender (and similarly ethnicity) might deny providers with the funding required to better engage harder to reach populations with treatment. In the same way the relationship between intervention and reduced HSU needs to be better understood; i.e. in using an expensive but effective treatment patients might be assigned to a cluster that draws down insufficient funding to deliver the intervention, hence disincentivising the system. The authors do refer to 'gaming' in this respect, and it is acknowledged that it is not possible to explore each variable in depth in this paper, but the authors do seem to take a view - on the basis of only limited evidence of association with HSU in each case - on whether each variable (or domain) warrants further consideration. The rationale for whether a variable (or domain) 'gets the nod' for further consideration, as indicated above, does not always seem to lie in the analysis as it is presented here. A somewhat fuller discussion of the possible implications of findings is needed. As minor points, it is perhaps surprising that, given the now widespread routine collection of HSU and other data by electronic patient record, that the paper did not suggest so called 'big data' approaches to exploring the usefulness of some of the possible associations between variables and HSU indicated here. Given the huge heterogeneity in studies and the challenges in uniform operationalisation of HSU indicated by authors, it seems unlikely that there will ever be a sufficiently comparable body of research to generate suitably definitive findings using traditional methods. But what this study does very well is generate those specific questions that a health informatics approach might address. How and why were the three international studies chosen for comparison? While it is reasonable to select studies that indicate

	support for findings, some indication should be given if, for example, there are studies of diagnostic or symptomatic approaches to clustering (and its equivalents) from outside of the UK. And are there any international or other country-specific reviews similar to this one? The decision making process for including intervention studies in the review is not clear. Why has not have every intervention study that measures HSU in some way been included if intervention is a worthy domain for consideration as a predictor of future HSU?

VERSION 1 – AUTHOR RESPONSE

Reviewer Name: Panagiotis Zis
Institution and Country: Evangelismos General Hospital, Athens, Greece

1. Few relevant papers must be cited:

Killaspy H, Zis P. Predictors of outcomes for users of mental health rehabilitation services: a 5-year retrospective cohort study in inner London, UK. Soc Psychiatry Psychiatr Epidemiol. 2013 Jun;48(6):1005-12.

Trieman N, Leff J. Difficult to place patients in a psychiatric hospital closure programme: the TAPs project 24. Psychol Med. 1996 Jul;26(4):765-74.

Trieman N, Leff J. The TAPS project. 36: the most difficult to place long-stay psychiatric in-patients. Outcome one year after relocation. Team for the Assessment of Psychiatric Services. Br J Psychiatry. 1996 Sep;169(3):289-92.

Trieman N, Leff J. Long-term outcome of long-stay psychiatric in-patients considered unsuitable to live in the community. TAPS Project 44. Br J Psychiatry. 2002 Nov;181:428-32.

• As per below, the first and last of these papers are now cited in the methodological considerations section (p13). The rationale for excluding these studies from the review them is also explained there.
• The second and third of the above papers provide earlier data from the final Trieman and Leff (2002) study that we have now cited. They also were undertaken before the year 2000 (it was an inclusion criterion for studies to be published after the year 2000). Therefore, we think it is not necessary to cite these papers.

There is relevant research relating to HSU by people with mental disorders not included in this review. This was for various methodological reasons, for example, differing conceptualisations of HSU in investigations by Killapsy and Zi [22] and Trieman and Leff. [23] These studies focused on the stability of HSU over time and were excluded because they do not address our study question which concerns identifying predictive variables contributing to an increase or decrease in HSU.

Reviewer Name Steve Gillard
Institution and Country St George's, University of London, UK
Please state any competing interests or state ‘None declared’: None declared

1. This is a timely, innovative and well conducted study. The case for further consideration of including demographic and co morbidity variables in the case mix for the MHCT is well made. It is

perhaps something of a missed opportunity that LD wasn't included as a strong case has been made elsewhere that tariffs for treatment of people with a LD co morbidity should attract a premium in order that sufficient resource exists to appropriately assess, diagnose and treat people with LD for mental and physical health conditions (see recent publications by Tuffrey-Wijne).

• The decision to exclude people with intellectual disabilities has now been explained more clearly in the methods section (p6; entry 1 below), but we understand the view of this being a missed opportunity. Therefore, to address this important issue, we have added the second entry below as a future research direction (p14):
(For the purposes of this review, mental disorders included adults experiencing elevated symptoms of mental disorders, or adults formally diagnosed with a mental disorder. Studies with participants with iIntellectual disability was not classed as a mental disorderwere excluded due to the specific additional needs of this population which have to be met beyond the healthcare system (e.g. in the education or labour systems);

Third, the HSU of people with intellectual disabilities were not examined in this review due to the specific additional needs of this population which have to be met beyond the healthcare system. However, it is an important area of research since UK-based studies have highlighted the widespread failure of health services to make required additional accommodations (e.g. extended appointment hours) for this patient group, with no additional funding currently allocated for these purposes to NHS acute trusts.[24] Determining how the inadequate provision of additional accommodations impacts upon the HSU of people with intellectual disabilities could inform future decisions surrounding allocation of resources.

2. The authors might have argued that the study further questions the 'symptoms, behaviour, functioning' approach adopted by the MHCT (why were impairment and risk variables not explicitly considered in this review?) given there was only some evidence that some of those variables were associated with HSU. However it is not clear whether or not the authors are suggesting that some association of some variables (in each of the domains considered) with HSU is evidence of the relevance of the domain to clustering, or of its limited usefulness. For example, there is inconsistency in that the authors conclude that the review lends support for further consideration of the role of diagnosis - alongside pathway - in clustering systems (where only personality disorders was shown to be a predictor of HSU), while not also concluding that symptoms, functioning and behaviors should likewise be given further consideration (on the basis that a single variable was also shown to be associated with HSU in each case).

• We agree with this point and there was some inconsistency in our conclusions in this regard. To clarify this issue, we have added that the domains of the MHCT should also be considered further in future research. (Conclusions section, p15):

The findings support the need to investigate to determine the association of the MHCT (and its domains of behaviour, symptoms, impairment, social functioning and risk factors) with HSU, the need to investigate whether combining broad diagnoses with care pathways is an effective alternative method for mental health clustering, and the need for research to further examine the association between existing mental health clusters and HSU. Overall, this review has highlighted important unresolved issues related to the Mental Health Payment by Results system.

3. Specifically, why is prediction of HSU in two or more studies an indication that that variable is a

sufficiently good predictor of HSU? Those studies might be in different populations, settings etc. Given the heterogeneous nature of studies - and therefore the difficulties in drawing a general conclusion about the relevance of each particular variable or domain - the observation of any predictive relevance might be an indicator that that variable warrants further study.

• Being a predictor of HSU in two or more studies was not the sole indication of a variable being a sufficiently good predictor of HSU. Also taken into account were study quality and the proportion of times a variables was not predictive of HSU. Moreover, heterogeneity of studies was discussed as a methodological limitation. However, in line with this suggestion, variables predictive of HSU in just one study could be considered worthy of further research. Therefore, the following entry has been added to the future research section (p14).

Fourth, the review identified a number of variables (i.e. attending a community outreach service, attending a psychiatric liaison service, unspecified ICD-10 diagnosis, insomnia symptoms, self-harming behaviour) examined in relation to HSU in just one study yet predictive of HSU. Therefore, the associations of these variables with HSU could be explored in future research.

4. As such, the discussion needs to point out that further research needs to be done to understand the relationships between these variables and HSU. For example, HSU might be higher in women than men because women are better at accessing treatment. As such any inflation of tariff based on gender (and similarly ethnicity) might deny providers with the funding required to better engage harder to reach populations with treatment. In the same way the relationship between intervention and reduced HSU needs to be better understood; i.e. in using an expensive but effective treatment patients might be assigned to a cluster that draws down insufficient funding to deliver the intervention, hence disincentivising the system. The authors do refer to 'gaming' in this respect, and it is acknowledged that it is not possible to explore each variable in depth in this paper, but the authors do seem to take a view - on the basis of only limited evidence of association with HSU in each case - on whether each variable (or domain) warrants further consideration. The rationale for whether a variable (or domain) 'gets the nod' for further consideration, as indicated above, does not always seem to lie in the analysis as it is presented here. A somewhat fuller discussion of the possible implications of findings is needed.

• To address this point, the below entry has been added to the discussion section (p13). As suggested, we now discuss the potential inappropriateness of certain demographic variables for clustering and suggest. We already discuss intervention variables in this regard but have added some wording to make our point clearer.
• The latter part of this suggestion was addressed in response to item 3- we have expanded our conclusions on the selection of variables for future research.

Regarding additional variables worth considering in the clustering process, various demographic (i.e. comorbidity, age, female gender, marital status, non-white ethnicity, high previous HSU) and intervention (i.e. IAPT, medication) variables with good preliminary evidence relating to their ability to predict HSU were identified. Future research could investigate if adding these variables into the ‘case mix’ of the MHCT adds to the economic validity and reliability of mental health clusters. However, it is worth noting that variables that are predictive of HSU are not always suitable for clustering and resource allocation purposes. For example, concerning demographic variables, it could be argued that it would be unfair to distribute resources on the basis of increased HSU by females (relative to

males). Similar arguments could be made regarding other population groupings with contrasting HSU patterns (e.g. certain ethnic groups). Moreover, the benefit of using intervention variables for clustering purposes is may be somewhat limited because it is relatively easy for providers to use these variables to 'game' the system (i.e. when patients are inappropriately and deliberately allocated to clusters that attract higher fixed payments) in order to generate additional revenue. [7]

5. As minor points, it is perhaps surprising that, given the now widespread routine collection of HSU and other data by electronic patient record, that the paper did not suggest so called 'big data' approaches to exploring the usefulness of some of the possible associations between variables and HSU indicated here. Given the huge heterogeneity in studies and the challenges in uniform operationalisation of HSU indicated by authors, it seems unlikely that there will ever be a sufficiently comparable body of research to generate suitably definitive findings using traditional methods. But what this study does very well is generate those specific questions that a health informatics approach might address.

• This is a good point and we agree. To address this, the following entry has been added to the future research section (p13):
Finally, further large-scale case register studies (including participants from shared service catchment areas) would address the study heterogeneity found in this review and provide more robust evidence on the predictors of HSU by people with mental disorders in the UK.

6. How and why were the three international studies chosen for comparison? While it is reasonable to select studies that indicate support for findings, some indication should be given if, for example, there are studies of diagnostic or symptomatic approaches to clustering (and its equivalents) from outside of the UK. And are there any international or other country-specific reviews similar to this one?

• The reason for choosing these international studies is provided (p10). However, extra wording has been added to clarify this (as below: entry 1).
• The international-based reviews are referred to in the previous section ('Comparison of main findings with other reviews'; p10) and there are no country-specific reviews similar to this one previously conducted.
• We agree it is a good idea to refer to clustering approaches outside of the UK. Therefore, the second entry below has been added to the methodological considerations section (p13).

As the review was limited to UK studies only, it is informative to compare the findings with those from international studies of HSU by adults with mental disorders. Three recent international studies were chosen for comparative purposes because with of their large samples comprising adults with a range of mental health problems were selected for comparative purposes.[18-20]

Third, the review was limited to UK studies only, meaning the list of identified variables is not exhaustive, and the findings may not be applicable to services in other countries. Indeed, this

applicability is particularly limited given that only a few other countries (e.g. Australia, New Zealand, Canada, the Netherlands, Norway, USA) have made progress implementing mental health payment systems, using heterogeneous clustering and resource distribution methodologies. [24]

7. The decision making process for including intervention studies in the review is not clear. Why has not have every intervention study that measures HSU in some way been included if intervention is a worthy domain for consideration as a predictor of future HSU?

• We did not place limits on the types of variables to examine, meaning that all identified (UK-based) intervention studies that predicted HSU were included. Nevertheless, in line with this point, we have made clearer the types of studies we included (p6).

Only the following types of studies were included in the review: (1) observational and intervention studies that predicted HSU by adults with mental disorders.

We trust we have attended satisfactorily to the comments and queries raised by the three assessors, and believe that the manuscript has been improved as a result.