Reviewer's report

Title: Morbidity and doctor characteristics only partly explain the substantial healthcare expenditures of frequent attenders - A linkage study between patient data and reimbursements data -

Version: 1 Date: 18 July 2013

Reviewer: Richard Morriss

Reviewer's report:

Comments for authors.
This study is an extremely important one in the field and this is an interesting hypothesis to explore. The methods are broadly of high quality and the discussion covers most points. There is a shortage of information on key points so my main comments are a set of questions seeking more information.

Major compulsory revisions.
My main concern is that I cannot follow the analysis to determine how the main hypothesis was tested. The data is clustered by practice and GP who of course see multiple patients so this data cannot be handled in the same manner as individual patient characteristics. It is unclear how this clustering was statistically handled. Some problems seem to be considered arbitrarily in the analysis as possible explanatory variables based on a rather limited view of the literature. Some other factors from a rather limited list of possible variables are considered as cofounders. I found the distinction arbitrary and contentious e.g. locomotor problems are also found often among FAs and FAs in some studies die more often from cancer. The list of variables considered is not exhaustive.

Even more concerning is the statement that each problem had to be likely to be present for 6 months - how can the authors make such judgements only from notes on such a scale and who made them? If it s not possible to assess the severity of the problem then how can its likely duration be assessed and won't the duration also depend on the management that is proposed? How was the validity and reliability of each judgment or a proportion of them tested? It is possible that all of these issues have been addressed in the research but the manuscript does not provide enough detail to assure the reader that the hypothesis was adequately explored.

Abstract. Results do not seem to address the hypothesis that additional costs can be explained by patient and primary care physician's characteristics.

Minor essential revisions.
Page 5. Problem of only using data from one insurer (so 55% data) need to be acknowledged. Possible selection bias and possible problems of generalisability.
Page 5. More detail on how linkage was achieved and verified i.e. checks to
ensure that the data on patients and GPs were correctly linked and results of those checks. The worry is that the reason no association is found is because the data has not been linked properly. What checks were employed by the independent organisation to ensure that the data sets were linked accurately?

Page 5. Was the same record system used throughout the three year period?

Page 5. Were all secondary care and emergency care costs covered by the insurers? What costs would the insurance system cover and what would it not cover? Could coverage vary for different groups of patients e.g. residents without EU passports? Were there some contacts e.g. emergency less reliably recorded than others?

Pages 8 and 9. Why not use a systematic classification e.g. see Barnett and Guthrie Lancet 2012; 380:37-43 who identified over 40 different medical categories that could be established from primary care records?

Pages 8 and 9. Having 2 models in the analysis does not in any way test residual confounding when the distinction between an explanatory variable and a confounding variable is so arbitrary and tenuous. Residual confounding is inevitably a problem with this type of study.

Table 1. How were depression, anxiety and medically unexplained symptoms defined and recorded? Primary care physicians tend to use an enormous and sometimes idiosyncratic set of terms for these problems. Judgements about medically unexplained symptoms are particularly contentious. Were other diagnostic information e.g. diabetes mellitus recorded in a systematic and/or standardised way?

Discretionary revision

Pages 8 and 9. I am unclear what the questionnaire sent to practices and GPs revealed but it seems to have elucidated little e.g. nothing about consultation style, stress perceived by GP, aim to provide personal continuity of care or greater emphasis on providing immediate access to care. The manuscript would be clearer if it outlined the information requested in the questionnaire sent to the GP.

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.

Declaration of competing interests:

I declare that I have no competing interests