Reviewer's report

**Title:** Quality of care for major depression and its determinants: a multilevel analysis

**Version:** 3  **Date:** 23 June 2012

**Reviewer:** Isobel Cameron

**Reviewer's report:**

Review BMC Psychiatry June 2012

**General Comments**

This research set out to measure the proportion of primary care patients with MDE receiving adequate treatment using a multilevel modelling method. Multilevel modelling has been infrequently applied in investigations in this field yet it is highly appropriate given the nested nature of this data and the different levels at which quality of care could be affected. For example, factors may affect quality of treatment at the patient, GP, practice and LSN level.

A tenet of this investigation is that previous studies have assessed quality of primary care using inadequate indicators. Whilst this is the case, some of the indicators applied by the current investigators may be considered somewhat arbitrary and some are also affected by limitations inherent in methods which assume reliable retrospective self-report. Details of these concerns will be explained under the relevant sections below.

**Background**

This provides a good introduction to the investigations which follow. In the second paragraph, the authors state that numerous studies highlight an important gap in depression treatment in primary care. It would be of value to the reader if the authors specified what this gap constituted.

The authors have clearly considered the issue of measuring quality indicators for depression in primary care and where there have been limitations in previous studies.

**Methods**

Data were collected as part of the Dialogue Study which has the strength of involving participants from 67 clinics. Furthermore, the authors recruited clinics from across 15 local service networks which were selected to provide diversity.

Less than a quarter of the clinics approached participated. The authors should consider the generalisability of their findings from this sample. For example, it would be useful if the authors were able to assess whether the participating clinics differed in any meaningful and measurable way from those clinics that did not participate.

In sampling patients, the investigators sensibly recruited across randomly chosen periods of the day.
67.4% of eligible primary care patients completed a brief questionnaire in the waiting room. The researchers have done well to recruit this proportion as recruitment opportunities are brief in such a study design and individuals can be easily missed. Eligibility criteria for follow-up seem inclusive. This is appropriate given that a more thorough investigation will follow and should ensure individuals possibly eligible for the cohort sample are not missed.

Although the authors have provided detailed descriptions of the recruitment process, including a flow diagram, I found it quite confusing trying to follow what was done at what time point and the nature of exclusion at each stage. It might be helped if the subsection which describes the instruments that were used was abandoned and these methods were described within the eligibility criteria. They are after all integral to this.

I am uncertain of the reliability of the methods used to ascertain relevant diagnoses in the last 12 months. It appears that this stage relies on self-report and on participants remembering back over a year long period. Additionally, medication use relies on self-report also. The authors should acknowledge the potential limitations of such data.

The authors should state the cut-offs they applied in the HADS.

It states that the “respondent most knowledgeable about the clinic’s organisation and functioning completed a standardised questionnaire”. Is it those individuals who provide information such as “consideration of patient’s preference” etc? It is difficult to know how one representative from a clinic could respond for the practice of all doctors in the clinic. It was unclear to me what information was used for the assessment of guideline compliance regarding ADP dosage and follow up, for individual patients. Was this information gathered from the patient interview or from the organizational questionnaire? Details regarding dose and follow up appointments would most likely be most reliably recorded in case notes but there is no reference made to this source of information. It is difficult to feel confident that participants could reliably remember detailed aspects of care ranging back over a 12 month period. Did the authors make any effort to corroborate their findings, for example, checking medication type and dose against case notes? Two issues are a concern: adequacy of memory and adequacy of reporting.

The criteria for minimally adequate pharmacotherapy do not specify adequacy of dose or duration of treatment. I would consider these important indicators of quality.

The multilevel model takes into account patient and clinic level data. I wondered if an opportunity had been missed to also assess factors at LSN level given that they were selected to represent diversity.

I have not conducted multilevel modelling and therefore do not feel qualified to judge the appropriateness of the detail of the procedure followed. A statistician with experience of MLM would be best to appraise this aspect.

Results

Given that depression is a highly recurring condition and that this is not a sample
of first ever occurrence, it is not surprising that 70% of the sample reports first symptoms over 5 years ago. The authors suggest these cases may be considered “complex”. My immediate thought is that they are “typical”. The authors state that the characteristics suggest that patients were not all recruited at the same stage of their MDE. It would be very strange if they had been, given the recruitment method. They also state that a cut-off point of 8 indicates a high probability of depression. However, Zigmond and Snaith (developers of HADS) described this cut-off as indicative of “possible depression” – a lesser assertion.

The authors report the odds ratios for factors associated with more adequate treatment. Additionally they should comment on the magnitude of these. For example the effect of severity of symptoms is of low magnitude.

The illustration with the virtual patients was useful for understanding the implications.

Discussion
The authors give a detailed consideration of findings in light of existing literature. Although they acknowledge the limitations of the methods of data collection, the implications of these limitations makes the findings of limited value.

**Level of interest:** An article whose findings are important to those with closely related research interests

**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.