Reviewer’s report

Title: Feasibility, reliability and validity of a questionnaire on health care consumption and productivity loss in patients with a psychiatric disorder (TiC-P)

Version: 1 Date: 26 March 2013

Reviewer: Matthias Hunger

Reviewer’s report:

The study of Bouwmans et al. examines feasibility, reliability and validity of a self-report questionnaire on health care utilization and productivity loss for patients with mental disorders. This is an important topic, since economic evaluations to assess costs and benefits of health care intervention usually rely on self-reports and one needs to have information on the suitability of such data. In general, the manuscript is well elaborated and limitations of the study are clearly stated. However, there are some issues in the manuscript that need further clarification:

Major compulsory revisions:

1.) The authors have assessed construct validity for two specific items only: the number of contacts with psychotherapists and long term absence from work. However, at several parts of the manuscript, results for the two items are generalized to the entire TiC-P questionnaire, e.g.:

Abstract: “The construct validity was assessed by comparing patient reported data with data derived from registries”.

Abstract: “The comparison of patient-reported data and data from registrations indicate satisfactory construct validity of the TiC-P”.

Abstract: “(...) results indicate that the TiC-P is a (...) valid instrument for collecting data on medical consumption and productivity losses”.

Discussion, 1st paragraph: “Estimates on the number of contacts with healthcare providers in mental healthcare were in accordance to registries of health care providers”

Discussion, 1st paragraph: “So the TiC-P, including SF-HLQ, seems a valid instrument for measuring healthcare utilization and productivity loss”.

The authors should revise these statements making clear that their results are valid for two specific items only.

2.) Similarly, the following statement in the discussion is speculative and not supported by the data. Therefore it should be removed:

“Assuming that the reported data of contacts with psychotherapists do not differ from the reporting of other healthcare utilization, we may conclude that the TiC-P
is a valid instrument to measure healthcare utilization.”

3.) Assessment of test-retest reliability requires that the construct being measured is relatively stable over time.

As the recall periods in the authors' test retest design do not or only partly overlap, the assessment of test retest reliability assumes that the frequency of healthcare consumption is constant over time. While this may be a reasonable assumption for weekly prescribed sessions with a psychotherapist, I think it is questionable for infrequent or irregular physician visits, like e.g. those to a social worker or hospital outpatient visits.

As a consequence, reliability coefficients could indicate low agreement even if there was perfect reliability of the items. Given these concerns, the authors should think about removing the infrequent medical use categories in their reliability analyses in table 3.

Minor essential revisions:

4.) The intraclass correlation coefficient (ICC) is typically used to assess the consistency of multiple measurements on the same quantity in the case that the true score is not known. Therefore it is usually used as a measure of reliability. Is there a reason why the authors also used the ICC for the comparison between self-reports and register data (validity assessment), i.e. in a situation where the true healthcare utilization is actually known?

5.) Discussion: “However, it can be assumed that the amount of short term absence from work (e.g. absence shorter than 2-4 weeks) is much easier to recall thus resulting in accurate information from self-reports.”

I do not fully agree with this statement. In the Collaborative care study, all participants were on long term sick leave. To provide the true number of days on sick leave, these patients only have to recall the approximate day when they started their sick leave.

In contrast, for patients to correctly report their short term absence from work, it is potentially necessary to recall more than one episode of sick leave and to recall the exact number of days for each of these episodes.

The authors should remove the above statement from their manuscript unless they provide data supporting their hypothesis.

6.) Results: There is no Table 5, so, the “Insert Table 5” statement should be deleted.

7.) Results, last paragraph: “The difference between reported and registered date of absence from work was on average 1.5 days”.

Are these differences absolute differences? They should be! Why is the difference so small given that there was only an agreement of 70.9%, even after accepting a margin of 7 days?
Discretionary revisions:

8.) Table 2: This table contains little information. I think the respective numbers could also be reported in the text.

9.) Introduction, last paragraph: “(...) is widely used in the Netherlands for economic evaluations in mental health”
Please give one or two references.

10.) Methods, third paragraph: “This resulted in a number of textual changes”:
Please give a reference on work on the development process of the TiC-P.

11.) Discussion, last paragraph: “In conclusion, the results from our study indicate that the TiC-P is a relatively good alternative for collecting resource use data”.
It is unclear what “relatively good” means and what the other alternatives are.

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**
I declare that I have no competing interests.