Author's response to reviews

Title: General practitioners' evaluation of community psychiatric services: responsiveness to change of the General Practitioner Experiences Questionnaire (GPEQ)

Authors:

Oyvind A Bjertnaes (oan@kunnskapssenteret.no)
Arjan Nieland (arian.nieland@sykehuset-innlandet.no)
Elisabeth Damerell (elisabeth.damerell@sykehuset-innlandet.no)
Andrew Garratt (amg@kunnskapssenteret.no)

Version: 2 Date: 4 March 2010

Author's response to reviews: see over
Dear Editor-in-Chief Norton:

**General practitioners' evaluation of community psychiatric services: responsiveness to change of the General Practitioner Experiences Questionnaire (GPEQ)**

Thank you for your careful consideration of our article about responsiveness of the General Practitioner Experiences Questionnaire.

We have considered the comments from the referees and have changed the manuscript accordingly. Our changes are described below together with where they occur in the manuscript.

In the Methods section we have confirmed that the procedures regarding informed consent, study design and data collection was approved by the Norwegian Social Science Data Services.

We hope that you find the revised manuscript satisfactory.

Yours sincerely

Oyvind A. Bjertnaes
Arjan Nieland
Elisabeth Damerell
Andrew Garratt
Comments to Referees – MS: 1864923719334082

Referee: 1

This is a concise and well-written paper about the General Practitioners Experiences Questionnaire (GPEQ), an instrument developed by the authors and validated in a previous study. The main objective of the present work is to test the GPEQ responsiveness to change, using data from a national survey involving all Norwegian GPs (however, only data in 8 municipalities in the Hamar community mental health centre catchment area are considered). In my opinion, this paper provides interesting data: there are few studies that evaluate psychometric properties of instruments for the assessment of GP’s opinions. Mostly, the assessment is made using ad-hoc scales developed for the research.

There are some Minor Essential Revisions:

1. The third paragraph of the background section is partially redundant with the second paragraph of the methods.

We agree and have shortened the third paragraph of the background section to avoid repetition. This third paragraph of the background section is now as follows:

"As is the case for patient satisfaction surveys more generally, reports of general satisfaction have limited value in quality improvement processes (14-15). Rather than simply asking GPs about their satisfaction with the community mental health centres, the GPEQ consists of concrete domains of care that are important as measures of service quality from the perspective of the GPs and hence contribute to their overall satisfaction with the centres. The GPEQ includes five scales that are supplemented with five individual items for use in repeated national surveys in Norway, as a result of the underlying development and validation work. The second national survey was conducted in 2008 (16).”

2. The description of the GPEQ is not clear; scale scores are described in two ways: 0-5 (table 2) and 0-100 (transformed, table 1, and the GPEQ paragraph in the methods). Could be useful explain this choice.

We have now included an explanation for this in the Methods section relating to GPEQ (second paragraph of GPEQ part):

"While scales are transformed to 0-100, item scores are shown in their original form as a five-point response scale 1-5. This is done to make it easier for the reader to distinguish between scales and items, and to adhere to common practice in this field”.

3. Table 3 ("a little bit better", "much better") is not consistent with text in the methods ("a little bit improved", "much improved")

We have corrected this.

4. Results, 2nd paragraph, second line: probably table 2 instead of table 1.
We have corrected this.

5. The reference n. 16 should be even translated in English

We have translated reference number 16.

Referee: 2

This is a straightforward study into the responsiveness of the already existing GP Experiences Questionnaire GPEQ, which is, by the way, a very general title for a questionnaire aimed at such a specific field as views of District Psychiatric Services. Major revision if possible:

6. However, it is also a paper about a rather specific phenomenon (the responsiveness of a questionnaire to assess GPs experiences with district psychiatric services in one small area). It seems a good idea to me, to use data from the National Surveys. This raises the question if there are not more examples of changes in psychiatric services in other locations, where the responsiveness of the questionnaire could have been tested.

We have carefully considered this proposed revision. Even though we are currently unable to gather information about quality improvement initiatives between the surveys from other centres’, we agree on the usefulness of including national results in the study. Therefore, we have decided to supplement table 1-3 with national results from 2006 and 2008 and conducted the same statistical tests for this group as the two original groups. The inclusion of national data resulted in several changes in the article:

Abstract – Methods: revised method description:

“National surveys were conducted in Norway in 2006 (n=2,415) and 2008 (n=2,209) to measure general practitioners’ evaluation of community mental health centres. GPs evaluated the centres by means of a postal questionnaire, consisting of questions focused on centre quality and cooperation with GPs. As part of the national surveys 75 GPs in 2006 and 66 GPs in 2008 evaluated Hamar community mental health centre. Between the surveys, several quality improvement initiatives were implemented which were directed at cooperation with and guidance for GPs in Stange municipality, one of eight municipalities in Hamar centre catchment area. The main outcome measures were changes in GPEQ scores from 2006 to 2008 for GPs evaluating Hamar community mental health centre from Stange municipality, and changes in scores for GPs in the other seven municipalities and nationally which were assessed for statistical significance”.

Abstract – Results: revised results description:

“GPs in Stange municipality rated Hamar community mental health centre significantly better on the guidance scale in 2008 than in 2006; on a 0-100 scale where 100 represents the best possible experiences the score was 26.5 in 2006 and 58.3 in 2008 (p<0.001). Apart from one item about workforce situation, none of the other scales and items showed significant changes. The control group from
the other seven municipalities gave significantly poorer rating for the emergency situation scale, the workforce situation scale and seven items in 2008 than in 2006. The national results showed small differences between 2006 and 2008, even though several scales and items were significantly different (table 1-2).

A question about changes in centre performance over the last 2-3 years showed that 82 % of GPs from Stange municipality reported that Hamar community mental health centre had improved, compared to only 36 % from the other seven municipalities and 40 % nationally which was statistically significant.”

Background (last paragraph):

“... GPEQ scores in 2006 and 2008 for GPs in Stange municipality evaluating Hamar centre were compared, in addition to comparisons between 2006 and 2008 for GPs in the control group consisting of GPs from the other seven municipalities in the catchment area, and comparisons over time at the national level”.

Methods – data collection (revised first paragraph):
(To secure comparability at the national level we have used the whole national material in 2006, including responses from GPs in the two pilot counties).

“The data reported here were based on two national surveys in Norway in 2006 and 2008 among all regular GPs in Norway. A questionnaire that included the GPEQ was mailed to 3,704 GPs in 2006 and 3,942 GPs in 2008. Both surveys included a recommendation to take part in the survey by the leader of the Norwegian Association of General Practitioners. Non-respondents were sent three postal reminders in both surveys. The 2006 survey also included telephone reminders to postal non-respondents. The response rate in the national surveys was 65.2% in 2006 (n=2,415) and 56.0% in 2008 (n=2,209). An assessment of non-response bias in the 2006 material demonstrated adequate representativeness (17).”

Methods – statistical analysis (revised paragraph):

“We constructed three groups: group 1, respondents from Stange municipality; group 2, respondents from the seven other municipalities in Hamar community mental health centres’ catchment area; group 3, all responding GPs. We present means for all groups, both for scales and single items. For the intervention group and the two control groups we tested differences in scores for scales and items in 2006 and 2008 separately by means of t-tests. Differences between GPs in Stange municipality and GPs in the two other groups on the improvement question in 2008 were tested by means of the Pearson Chi-Square test. SPSS version 15.0 was used for statistical analyses.”

Results – new third paragraph:

“There were small differences between the national results in 2006 and 2008, but several scales and items were significantly different (table 1-2)”.

Results – revised fourth paragraph:

“More than 80% of GPs in Stange municipality reported that Hamar Community Mental Health Centre had become much or a little bit better the last 2-3 years (table 3). In the control group for Hamar centre this only applied to around 36%
and nationally the percentage was 40%, both being statistically different from the intervention group (p<0.05).

Discussion – first paragraph, last sentence:

“As expected, the control group evaluating Hamar centre and the national results showed small differences on the guidance scale in 2006 and 2008”.

Discussion – new fourth paragraph:

“Poor scores for the GPEQ guidance scale in the 2006 survey led to the local quality improvement initiative on which this study was based. Our study showed a large improvement on the guidance scale in the intervention group, but also nationally the guidance scale showed largest improvement from 2006 and 2008. We currently lack information about quality improvement initiatives between the surveys from the other community mental health centres’, but the national report from this project identified 11 centres with significantly different scores in 2006 and 2008 (16). This inclusion of additional local quality improvement initiatives would have resulted in a more extensive evaluation of the GPEQ including the responsiveness of individual GPEQ scales that were targets for improvement. An electronic survey shall assess the importance and usefulness of the 2008 survey results to the community mental health centres. Data from this survey will inform further responsiveness testing of the GPEQ as part of the analysis of changes from 2008 to 2011 when the next survey is planned”.

7. Minor revisions i): Regarding the results, GPEQ seems responsive on one scale (guidance) which seems conforming the actual changes in the policy of the Hamar Community Centre. However, there are changes as well in negative direction on two other scales for the GPs form other municipalities. Is there an explanation as well? Please provide this.

We have tried to explain this in the third paragraph of the Discussion:

“The decline in scores for two scales in the control group evaluating Hamar centre was not expected in this study. However, employee data show that Hamar community mental health centre had a decline in the number of employees from 2006 to 2008 (18.2% decline). This might be an explanation for the significant decline in the workforce situation and emergency situations scale for the control group. The team working with the intervention group also had a decline in employees from 2006 to 2008 (12.5%), but did not experience significant changes in scores for these scales. One possible interpretation is that the local quality improvement project aimed at GPs from Stange municipality prevented the worsening of scores for workforce situation and emergency situations that can be expected as a result of the decline in employees. From January 2007 the team serving the intervention group also had more specialists per inhabitant than the rest of the centre, which also could be an explanation for the lack of decline in scores”.
8. Minor revisions ii): When significant changes in Table 1 and Table 2 are mentioned, it is not clear if these differences apply to differences between two locations or change in time.

We have clarified this with the following footnote to table 1 and 2:

“Tested differences between 2006 and 2008 for each group separately, independent samples t-tests”.

9. Minor revisions iii): Maybe the conclusion is a little bit too fast: there was a significant improvement where an improvement could be expected, but there were more significant downswings that were not expected. The conclusion that the instrument is responsive to change is only one side of the medal.

We have supplemented the Conclusions with a sentence about the downswings (last sentence):

“... The worsening of services for GPs in the control group evaluating Hamar centre warrants further study”.  