Reviewer's report

Title: Compliance with referrals to medical specialist care: patient and general practice determinants: a cross-sectional study

Version: 0 Date: 28 Sep 2015

Reviewer: Andreas Sonnichsen

Reviewer's report:

This is an interesting study on referral compliance and factors affecting it. The authors are using a large primary care database and a claims database. Patient data between the two data bases are linked by date of birth, postal code and sex. The study is based on a very large number of data sets.

The study appears to be methodologically sound but with limited validity due to incompleteness of data and possible selection bias. I have a few remarks and suggestions:

Methods

For readers not familiar with the Dutch health care system the authors should clarify the usage of the data bases. Who is participating in the NIVEL database? Which percentage of GPs? Is there a difference between NIVEL-participants and non-participants? Who is using Vektis? Are there other possibilities for specialists to claim their services?

Practices or "practice years" were excluded from the study depending on data quality. How was this assessed? Which criteria were applied? In which case was "data on care episodes" or "data on contact" or "data on prescriptions" etc. judged to be "incomplete"? What do the numbers behind these exclusion criteria mean (±25%, ±30%, ±30%, ±50%)? What does the "±" mean? Why don't the numbers add up to 100%? Why don't the authors give the exact figures?

How can the data on capitation fees be used to judge whether a patient comes from a deprived urban area?

At the end of the "study design and population" section the authors claim that the included practices were representative of Dutch general practices with respect to urbanisation and region. How was this measured/assured? Practice characteristics regarding urban/suburban/rural/deprives/region are not shown in table 1. The practice characteristics of all "Dutch general practices" should be shown in comparison. Equally the patients' characteristics of all patients or at least all patients in the NIVEL database could be shown.
Results

Data in the text don't match with table 2 (text 92.1% eye problems; table ear problems). Are any of these differences statistically significant (using Chi²)? - I imagine, not. Even less so the difference between diagnosis and diseases (87.4% versus 86.0%). It should be stated that these differences are insignificant.

What does the second column of table 3 stand for? Most of the numbers are identical between the first and second column. Sometimes there are only numbers in the first, sometimes only in the second column. This needs to be explained.

The "significant" odds ratio of 1.01 for the distance to the nearest medical specialist needs some further explanation. "km" is a continuous variable. How was this variable categorized for the calculation of an odds ratio? Even though this variable may be "significantly" associated with referral compliance due to a hardly credible very small confidence interval, an odds ratio of 1.01 of course is nothing of any relevance. Please check this result. Please omit the sentence in the discussion that "those who lived further away from medical specialist care facility were more compliant". This result should also be deleted in the abstract, or the odds ratio of 1.01 must be reported there as well.

Also, face-to-face contacts and guideline adherence of referrals were defined as continuous variables in the methods section. How were these variables categorized for the calculation of odds ratios?

Discussion

The limitation section should discuss the problem of multiple testing in an explorative study leading to significant results (e.g. odds ratios with 95% CIs completely below o above 1) just by chance.

In the discussion the authors refer to a higher compliance in patients with chronic conditions, but the data are not shown. Why? If this is really the case, it appears to be more important than the odds ratio of 1.01 for the distance to the nearest specialist which is reported and discussed extensively.

The passage in the discussion about the relationship between referral compliance and the distance to a medical specialist should be reworded, as the odds ratio of 1.01 is really nothing - even though "significant" - and is - as a number - in concordance with the literature that there is no correlation between referral compliance and distance to medical facilities.

I would like to suggest for the discussion that each point in the discussion is started with a new paragraph (lines 15/16; 33/34; 54/55).
The fact that general practice characteristics were not associated with compliance does not allow the conclusion that this may also be true in all practices in the Netherlands. It might be just the other way around. Practices participating in NIVEL and practices with high quality documentation in the EMR (i.e. the practices included in the study) may very well have higher referral compliance than the rest.

In the present sample, referral compliance was very high. This may in part be due to the methodology applied and the selection of practices used.

**Are the methods appropriate and well described?**
If not, please specify what is required in your comments to the authors.

No

**Does the work include the necessary controls?**
If not, please specify which controls are required in your comments to the authors.

Yes

**Are the conclusions drawn adequately supported by the data shown?**
If not, please explain in your comments to the authors.

No

**Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?**
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I recommend additional statistical review

**Quality of written English**
Please indicate the quality of language in the manuscript:

Needs some language corrections before being published
Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below.
If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal