Author's response to reviews

Title: Guidelines for disease combinations: a useful strategy to manage multimorbidity? A practice-based analysis of combinations of diseases in patients aged 65 or older in primary care.

Authors:

    Pauline Boeckxstaens (pauline.boeckxstaens@ugent.be)
    Wim Peersman (wim.peersman@ugent.be)
    Souhila Ghali (souhila.ghali@ugent.be)
    Gwendolyn Goubin (gwendolyn.goubin@ugent.be)
    Jan De Maeseneer (jan.demaeseneer@ugent.be)
    An De Sutter (an.desutter@ugent.be)

Version: 2 Date: 21 March 2014

Author's response to reviews: see over
Major compulsory revisions:

This manuscript reports the results of a comparably small study on combinations of diseases in patients aged 65 or older in primary care. Against the background of the many possible combinations of diseases in multimorbidity (see van den Bussche et al., 2013) the analysis of only 543 patients in two family practices in Belgium is a limitation.

The authors argue that using insurance claims data – which offer the advantage of huge data sets – may lead to an over- or underestimation of diseases, because “these studies do not necessarily identify the combinations that are relevant at the patient-provider level”. Therefore they try to identify diseases that influence clinical management at the patient level and based on these diseases they try to find specific combinations of problems that could be a relevant focus for RCTs and/or management guidelines.

However, on the way to this goal the authors performed steps that are not clear.

1) 1a) Did they apply the CIRS-CI or the CIRS? In the methods section they introduce the CIRS-CI, but in the following they state, that multimorbidity was defined as the involvement of 2 or more body systems (CIRS > 1). What does this mean exactly?

Thank you for noting this discrepancy in the paper. In fact when we state to define multimorbidity as the involvement of 2 or more body systems (CIRS > 1) this should have been CIRS-CI>1. However, as most authors currently simply refer to CIRS, we believe that the distinction earlier mentioned has become redundant within the current manuscript also because we now have put less emphasis on the CIRS (see next comment). Therefore we adjusted the manuscript to consequently use CIRS instead of CIRS-CI to describe the prevalence of multimorbidity in this study sample.

1b) Were patients included who had illnesses in at least 2 body systems according to CIRS and at least a severity score of 3 in both of the systems? Please clarify.

In fact, this paper has included all patients aged 65 or older in both practices, irrespective of their morbidity. The CIRS was merely used to provide quantitative information on the prevalence of multimorbidity within this sample and was not used as an inclusion criterion. We rephrased the methods to put less emphasis on the CIRS as we believe this was indeed distracting the reader from the inclusion criteria which were merely based on age.

As for the question on how we have defined multimorbidity within this sample we confirm that patients who had illnesses in at least 2 body systems according to CIRS and at least a severity score of 3 in both of the systems were defined to have multimorbidity. This is compatible to the scoring guidelines of Hudon et al.
2) Then, the research team constructed a list of 23 problems based on their clinical experience and reason. This selection process is not traceable and the problems are only partially comparable to problem/disease lists in other publications. Please describe in more detail.

We did not use a selection of 23 problems when including a patient but we merely performed a process of summarization in order to find a reasonable way for including all the patient’s relevant but often low prevalent problems into a sensible analysis of combinations. Some disorders, such as diabetes were analyzed at disease level, others were considered as a diagnostic group (eg cardiac rhythm disorders) or have been grouped within a body system (eg the neurological system). We revised the methods section to make this approach of summarization instead of selection more clear. Further in the paper we aimed to avoid the use of the word selection in the context of the 23 problems in order to prevent confusion on a pre-selection of chronic diseases.

3) Coming back to the authors’ statement at the beginning of the introduction: Given that these 23 problems are in fact “diseases that influence clinical management at the patient level”: to what extent do these problems differ from the problems routinely documented in the GP files? A comparison of these routine diagnosis data with these 23 problems would be of great help to substantiate the arguments in favor of this study.

The 23 problems used to analyze the prevalence of combinations are in fact not necessarily problems that directly influence the clinical management of a patient. These 23 problems were merely used to summarize the problems which had been defined by clinician-reviewers at the time of inclusion of a patient because they were of influence on the clinical management of a specific patient.

Moreover, in the Belgian context GPs currently do not report any morbidity data outwards their practice so we were unable to compare our own data to any large representative database of morbidity data from primary care in Belgium. Also within our own sample we have not systematically compared the problems included out of the charts to all the problems which had been routinely documented for these patients as this would have increased the data-load of this study considerably and this kind of analysis would have extended far beyond the assessment of multimorbidity because it would have included each and every problem the patient consulted for (e.g. also every upper airway tract infection, skin lesion or temporary knee pain). So, in order to provide the most comprehensive insight on multimorbidity at the level of clinical practice, our methodological approach explicitly aimed to include those disorders relevant for the clinical management of each specific patient. And, to our opinion, this approach required clinician chart reviews and open lists (instead of pre-selected chronic diseases). Open lists are not suited to any automated extraction of disease codes because they require a pre-selection of disease codes.

We revised both the introduction, methods and discussion section in order to make the arguments for the methodology of practice based clinician chart review more clear. Our methodological approach is not suited to provide the comparison requested by the reviewer also because the data required are not available.
In this context a discussion of the results of the study of van den Bussche et al. in the Journal of Clinical Epidemiology 66 (2013) 209e217 is missing. In this study claims data and data from GP interviews were compared with each other. Interestingly the prevalence of most diseases was lower in claims data than in GP interviews.

Thank you for this important suggestion. Indeed, within this field of research there is an important relation between the study design, the datasource used and the conclusions which can (and cannot!) be drawn from a specific study. This is especially relevant in light of estimates on the prevalence of multimorbidity which was not the primary aim of our study. This study specifically aimed to define the disease combinations which would be most relevant and useful for the development of guidelines to manage multimorbidity in primary care. Therefore, those problems were included that influence clinical management at the patient level in primary care using the methodological approach of practice-based epidemiological research (as suggested by Starfield and Valderas) by means of clinician chart review. Because problems were identified within primary care, using open lists instead if closed lists of disease codes, including all diseases, symptoms and risk factors our prevalence estimates will be relatively high compared to for instance population studies (Fortin et al). However, this practice-based individualized assessment by family doctors was the suitavle way to provide insight in the day to day presentation of multimorbidity in primary care and enabled to include those problems that were assessed as significant by clinicians for the clinical management of multimorbidity at the level of the individual patient. We have revised the discussion of the paper referring more clearly to the important work of vandenbussche et al (and fortin et al) and how this relates to the possible interpretation of our study.

Comparing the prevalences in this study with the study of van den Bussche et al. is limited. However, a few diseases allow comparison and here we see, that the prevalences are very much lower in the Belgium study than in the German:

- Hypertension: 48% in Belgium, 78 in Germany
- COPD: 14 vs. 24%
- Diabetes: 14 vs. 38%
- Ischemic heart disease: 14 vs. 31%
- Lipid disorder: 14 vs. 59% etc.

In order to discuss these huge differences we need the results from the analysis that is described in comment 3.

Because these studies concern different populations - Van den Bussche et al merely included patients defined with multimorbidity whereas our study included all patients aged 65 or older - these prevalences are not comparable.

The authors conclude that “performing trials or developing guidelines for people with specific combinations will ever be useful at the level of clinical practice.” I doubt that this
strong conclusion can be drawn from this small study that partially used subjective measures of disease relevance.

*We agree that the sample size is limited and this should be considered a limitation. Secondly, clinician chart review indeed implies a subjective assessment of disease relevance. However we are convinced that this approach is in fact a strength rather than a limitation of the study because it enabled a patient-centered and practice based assessment of multimorbidity in terms of including those problems influencing the management of a specific patient. A higher sample size may have elicited other disease combinations. However, we do doubt whether larger sample sizes would have been able to elicit combinations which would be useful at the level of clinical practice as we believe a considerable prevalence is an important consideration when assessing the usefulness of a specific guideline in primary care. Moreover most patients were defined with more than 2 problems which would in any case hamper the usefulness of guidelines merely focusing on two problems. Therefore we believe that, even despite the small sample size, we can still conclude that performing trials or developing guidelines for people with specific combinations will not be useful at the level of clinical practice.*

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

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.

Reviewer 2:

Reviewer's report:

This is an interesting and well written paper that tries to identify specific problem combinations that could be a relevant focus for RCTs or management guidelines for multimorbid patients. A strength of the approach is that not only prevalence but also clinical relevance of problems is assessed and that the authors go beyond the limitation of the ICD and also include social problems and problems from the medical history of the patients that could interfere with problem management in general practice. The authors also found a reasonable way for including relevant but low prevalent problems by using additional CIRS domains.

Major compulsory revisions:
1. The title “Multimorbidity: structure of chaos?” does not fit, because this is not a question the paper tries to answer. The authors should find a title that fits better with their analysis.

_We adjusted the title to Guidelines for disease combinations: a useful strategy to manage multimorbidity._

2. There are some interesting results – only: the discussion part is very thin. The authors state themselves that they are not the first who identified the combination “hypertension – osteoarthritis”. However, this is the only combination that is discussed regarding medical consequences. There should be lots of other combinations which have a rather high prevalence, are clinically relevant and should therefore be worth discussing.

_The other combinations with a considerable prevalence (which we have set at >5%) often include diagnostic groups or CIRS domains to which guidelines are not applicable. Therefore, we have discussed so little combinations in terms of their medical consequences. We have added a sentence in the discussion section to explain this and we have elaborated on some other combinations such as hypertension-diabetes, hypertension-lipid disorder, ... including a general discussion on the impact of combinations which include mental health problems and social problems._

3. The authors should also discuss the consequences of combinations with social problems they found for disease management by GPs, because this is one of the main strengths of this paper.

_We agree that this is a major finding and we have elaborated on this in the discussion setting._

4. The disease associations from table 3 also need to be discussed regarding implications for general practice.

_We have elaborated on this aspect in the discussion section, especially focusing on the less expected combinations with psychiatric disorders and social problems which have been elicited in this practice-based analysis._

Minor essential revisions:

5. p.4 para 3 line 5: GPs “assessed whether or not a problem interfered with patients’ normal activity and/or whether continuous treatment was needed”. Was this an inclusion criterion for problems? If yes, this should be mentioned.

_Yes this was indeed an inclusion criterion as otherwise we would have included disorders for which patients consulted only once and which had no impact anymore. We have mentioned this now explicitly in the manuscript._

6. p.4 para 3 line 5: “social problems” and “relevant medical history” of the patients were also assessed. Please give a detailed operation definition of these issue, i.e. which facts were counted as social problems or relevant medical history.
These issues were assessed by clinician reviewers, without providing them a stringent operational definition. They were merely asked to include the problems, that influence the clinical management for that patient. As for social problems these included mostly problems which were defined by means of disease codes within the Z-chapter of ICPC (reference) and for which the additional free text was assessed to estimate the impact of the problem on the patient. As for the medical history of the patient, the clinician’s assessment would have included problems like myocardial infarction for which continuous medical treatment and follow up were required or a history of colon cancer or breast cancer which required further follow up or caused considerable fear despite remission. Problems such as kidney stones, gall bladder resection or knee prostheses with full rehabilitation of the patient were generally not included by clinicians because they were no longer relevant for the daily clinical management of that patient. To harmonize the clinical assessment, data extraction and allocation of the multimorbidity scores, 30 patient records were independently reviewed by the three family doctors who performed the assessment and results were compared and discussed in a meeting to attune the assessment approach. This has been mentioned more clearly now in the manuscript.

7. p.8 para 3 line 1: “The main strength of this study is that the morbidity estimates are derived from chart review rather than automated diagnostic codes”. In my experience data from GPs patient charts are automatically extracted and sent to health insurance companies to justify insurance claims. For this reason automated diagnostic codes are usually the same as data from chart reviews. If this should also be the case in your study, this sentence needs to be adjusted.

In Belgium there is no automatized extraction of data for health insurance companies, nor are there any separate reports for insurance companies. GPs use their morbidity data merely at patient/practice level. We did not adjust the manuscript.

8. p.8 para 3 line 4: “Moreover, the way we constructed the problem list enabled us to include every condition instead of a limited list.” If a patient had two or more problems with a low prevalence that had to be coded in one CIRS domain how did you proceed? Were they counted as only one CIRS problem? If yes, you should discuss that.

The CIRS was merely used to provide an estimate of the prevalence of multimorbidity within this sample. All problems included by the clinician reviewer out of the charts were in fact included and assessed separately. However, for the analysis of the combinations a process of summarization was performed in order to find a reasonable way for including all the relevant but often low prevalent problems. Some disorders, such as diabetes were analyzed at disease level, others were considered as a diagnostic group (eg cardiac rhythm disorders) or have been grouped within a body system (eg the neurological system). We have revised the methods section to make our methodological approach more clear.
9. p.9 para 1 line 4: “However, due to the high morbidity load and complexity of the patient population we do believe that this population was suitable to identify relevant combinations at practice level.” It should be noted as limitation of the study that – compared to the complexity of multimorbidity – a very low sample size was obtained. For this reason the study might have missed possibly relevant combinations.

We agree that the sample size is limited and we have added this more clearly in the discussion section as a limitation. However, we do doubt whether larger sample sizes would be able to elicit combinations which would be useful at the level of clinical practice as we believe a considerable prevalence is an important consideration when assessing the usefulness of a specific guideline for a GP.

10. As additional limitation it should be noted that only combinations of two problems are considered while elderly patients normally have a much larger number of problems.

Because of the low prevalence of the combinations we did not extend our analyses to the co-existence of more than 2 problems but we have mentioned the co-existence of more than two diseases as an important argument to our general conclusion as this would in any case hamper the usefulness of guidelines merely focusing on two problems. Therefore we believe that, even despite the small sample size, we can still conclude that performing trials or developing guidelines for people with specific combinations will not be useful at the level of clinical practice. We have added this as an additional limitation to our study.

There are also some typos that need to be corrected:

11. p.5 para 2 line 5: “constructed a list 23 problems” -> an “of” is missing

12. p.6 para 2 line 2: “in the study sample is f 73 years” -> the “f” should be deleted

13. p.6 para 3 line 1: “problem pairs who occur” -> “who” should be replaced by “which”

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

Quality of written English: Needs some language corrections before being published

Statistical review: No, the manuscript does not need to be seen by a statistician