Reviewer’s report

Title: Co- and multimorbidity patterns in primary care based on episodes of care: results from the German CONTENT project

Version: Date: 7 June 2007

Reviewer: Helena Britt

Reviewer’s report:

General
This Paper describes some results from the CONTENT project in Germany which involved episode based data collection in 17 general practices, for 42,456 patients and 90,400 episodes. The patient sample size is large, but (as is recognized by the authors) the GP sample is small. The paper is interesting but is written in a vacuum. There is no recognition of the many other studies that have attempted to estimate prevalence and establish the relationship between patient characteristics such as age, sex, multimorbidity and visit frequency. There is also no recognition, or discussion of the issues already raised by others about the completeness (or lack of it) in EHRs and other issues of reliability that have been highlighted for this method of data collection.

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

Background: The background is light
Needs some brief description of the German medical system i.e. gate-keeper role, what proportion of the population see a GP, is it user pays, or covered by universal medical insurance system, or a mixture etc. This is needed to gain a better insight into what this study actually represents.
Ref 6: the electronic version was not available in 1987. You need a different reference here. (By the way it is only available free of charge for research or testing. If it is to be rolled out in EHRs it needs a national licence to be gained from Wonca.)

The background merely describes the CONTENT project, and is in large a repeat of the background from your previous paper in Informatics in Primary Care, yet there is a mile of literature about estimating prevalence from GP records and about relationships between visit frequency, and multi morbidity. Currently you are writing this in isolation from the rest of the world. Create a cone in the Background from broad concept to why you are doing this study and what it aims to do. The reverse cone should occur in the Discussion – how do your results relate to other experience or findings?

Last paragraph: it is unusual not to limit the study to chronic conditions. This means that multiple episodes of such morbidities as URTI, etc, will be included, and that surely is not the subject of interest. Chronic conditions have been defined according to ICPC-2 by O’Halloran J, Miller GC, Britt H 2004. Defining chronic conditions for primary care with ICPC-2. Fam Pract 21(4):381-386.) and this definition is easily applied to analyses.

Methods:
The methods are far too brief and some are included in the Results section instead of the Methods section (e.g. the first four lines of para 4 in the results. The lack of detail in the methods makes it almost impossible to form a fair judgment on this paper.

1. You have a cluster sample study design of 17 practices with an average cluster of about 2,500 patients (the unit of analysis). Did you adjust for the cluster? If not you have a huge design effect in your analysis. The statistical power of the study would be far reduced after adjusting for the cluster.

2. I assume you tested for all the listed variables (age, sex, number of ICPC codes)? Or were there others you tested which were not included in the model.

3. Where you have multiple codes (e.g. K74-76) ad L89-91) in Table 4, has this been counted on the ‘at least one’ basis? I can only assume so, because you call it ‘prevalence’ For example: patient has IHD with angina and then has a Myocardial Infarct. I assume the Infarct is
recorded as the ICPC code within the IHD episode record, and then …. The question is, is this patient counted twice, for the K74-76 IHD group?

4. I think you included all ICPC-2 codes, irrespective of whether they were RFEs or problem labels – there are three different terms used in the paper – in Table 3 you speak of multiple ICPC-2 codes. In Results you speak of ‘number of different codes documents in ICPC (symptoms and diagnoses) (para 2). In para 3 you refer to the number of patients’ different “ICPC Diagnoses”. Figure 2 speaks of ICPC-diagnoses. In Table 3 you refer to them as ‘codes” – what IS included please?

The need for definitions
There are no definitions: please provided them. Is a ‘different prescription” = to the number of different brands? To the number of different generics? To the number of different occasions of prescribing?

Is the average number of referrals only those to specialists? Is it new referrals for that patient for that problem, or does it include repeat referrals which (in some countries) are required on an annual basis? Is a ‘consultation’ equivalent to a contact? Since contacts can be indirect, I don’t know if the label in Table 3 should be contacts, or whether you have reduced the analysis for this variable to direct contacts only (i.e. consultations).

What is a ‘different ICPC- diagnoses? For example if the GP had an episode labeled “Osteoarthritis, and within the episode over a couple of consultations s/he recorded L89 and in another L91 (different sites), is this counted as 2 occurrence and therefore added twice to the ‘prevalence’ estimate? I think I am having a lot of trouble absorbing this paper because I have no understanding of the structure of the data. For example, if the patient with an ongoing episode of OA of the hip, presents with a RFE of pain in hip (code L13), is this one of the codes that would be in the episode?

Classification and grouping issues
I am interested in your grouping of OA, to include L89-91. It is incomplete: L84 includes (in its published inclusion criteria) OA of the back – although it also includes many other concepts; and L83 includes osteoarthritis of the neck. L92 includes OA of the shoulder.

Yet you use L84 as a comorbidity, which comes out as the number 1 co-morbidity of OA – hardly surprising, as it should be included in the count for OA

I understand your problems with their inclusion, because it means you would be including lots of other shoulder problems, and back problems that are not OA – but if you are going to leave them out you have to give your reasons, and discuss the effect in the Limitations section of the paper. The problem you face in the analysis of OA is overcome in Australia by coding at a more specific terminology level, and regrouping the OA terms from all the above rubrics into OA.

Discussion:
Currently the discussion is very light.
As stated above you need to relate your results to something. Perhaps you could consider them in comparison to any national health survey data collected from the community?

Perhaps you could discuss them compared with other published data in Europe. Then, when they differ you can highlight the limitations of the study in terms of GP sample size; size of cluster; classification issues etc.

Alternatively you could discuss your findings in terms of the limitations already identified by others in data collection from EHRs.

Tables and Figures. There are far too many. I prefer the graphics in Figure 1, to the Tables, in which the same data are repeated. You could add the SEs to the graphics and eliminate Table 1.

Figure 2 is a repeat of Table 2. In any case, both of these should be considered in terms of sex specific rates. Currently they are sex distribution of rate type 1 (i.e. per cent males who have 0 or 1. Since females made up about 60% then a 40-60 split on “number of ICPC diagnoses) is a one to one relationship. Then decide whether to present as a Table or as a figure.

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author
can be trusted to correct)
Replace "parallelly" with 'Parallel'

Discretionary Revisions (which the author can choose to ignore)

**What next?:** Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

**Level of interest:** An article of limited interest

**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.

**Declaration of competing interests:**

I declare that I have no competing interests