Author's response to reviews

Title: Differences in the quality of primary medical care for CVD and Diabetes across the NHS: Evidence from the Quality and Outcomes Framework

Authors:
Gary McLean (gml17y@clinmed.gla.ac.uk)
Bruce Guthrie (b.guthrie@chs.dundee.ac.uk)
Matt Sutton (m.sutton@abdn.ac.uk)

Version: 2 Date: 18 April 2007

Author's response to reviews:

Response to Reviewers' reports

Revised Manuscript: MS1117280466134288
Differences in the quality of primary medical care for CVD and Diabetes across the NHS: Evidence from the Quality and Outcomes Framework
Gary McLean, Matt Sutton, Bruce Guthrie

Dear Editor,

We thank you and the reviewers for their valuable comments. The manuscript has been revised in light of the reviewers' comments. We have included the reviewers' comments and our response (in italics) as to how we have modified the manuscript to take into account these observations in detail below.

Reviewer: Mark Strong
Reviewer's report:
Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)
A more explicit justification of which QOF indicators were categorised as "simple process", "complex process" is necessary. For example, why is checking pulses a "complex process" - this can be done in the surgery, yet measuring creatinine, requiring a blood test, is a "simple process" Due to an editing error, measuring creatinine was erroneously described as a complex process in the submitted paper. We have corrected this, and have more clearly defined simple and complex processes in the text as follows "We distinguished simple and complex processes in terms of whether they could easily be delivered opportunistically irrespective of whom the patient was seeing. A simple process is therefore one that can be delivered during routine care by any doctor or nurse, such as blood pressure measurement or blood taking. A complex process either requires referral to a specialist such as diabetic retinal screening, or is likely to be done only by particular primary care clinicians such as comprehensive diabetic foot examination." (Page 7, 1st paragraph)

It would be helpful to have an accompanying table that detailed the indicators in each of the above categories (indicators referred to as e.g. "CHD 5 - blood pressure recorded") We have added table 7 detailing indicators used and definitions

I feel that the explanation in the methods section describing how the composite scores were arrived at needs to be a little clearer. Were all the summary measures weighted for register size, for example? We have added text to methods section (Page 7, 2nd and 3rd paragraphs) to explain this more clearly.

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)
The reference to tables 3 and 4 in the results section needs to be to tables 3,4 and 5. We have changed the references to the tables
The reference to table 5 in the last para of results needs to be to table 6. Changed reference

Make clear whether the "percentage differences from England" in the tables are percentage point differences i.e. % Scotland - % England, or whether they are percentage differences i.e. (% Scotland - %England) / % England x 100%. ) Changed title to 'percentage points difference to England'

Are the mean composite means in table 2 percentages? If so a label would be helpful. Changed title to "Average percentage achievement by indicator category..."

Discretionary Revisions (which the author can choose to ignore)
I am somewhat uncomfortable with saying that x practice or y country has higher "quality" just because their QOF achievement is higher. Is it safer just to say that "recorded achievement for QOF was higher"? Or is it worth making a statement about this in the introduction? Added statement in introduction (page 4, 2nd paragraph)

Could say in last para of intro: "However, payment per point for the average practice in each country is the same" to make it clear that payment, while adjusted for prevalence, ultimately depends on achievement.. Changed to "However, as payment for the average prevalence practice achieving the same points total in each country is the same, the average per-patient rewards for quality are therefore lower in higher prevalence countries."

Could add some numbers to last table (or to the discussion) to indicate how much money a practice in Scotland or Wales will lose compared to a practice in England if they all have the same disease prevalence and list size and achieve the same "quality". We do not think this is necessary, partly because the actual payment received will also vary due to the application of the Adjusted Disease Prevalence Factor, the effects of which are unpredictable and would be difficult to quickly explain as an aside in this paper (. Guthrie B, McLean G, Sutton M. Workload and reward in the quality of outcomes framework of the 2004 general practice contract. Br J Gen Pract 2006; 56: 836-841)

Reviewer: Yvonne Engels
Reviewer's report:

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)
1. I suggest another title, referring to diabetes and CVD. Changed title by adding CVD and diabetes
2. the sentence on page 4 'QOF data ....is therefore not possible' is strange. Please adapt (I miss a verb in the part of the sentence after 'as'). Added "using patient level data"
3. Page 4-5: you mean to say that exception report differs per country, but you don't say that explicitly. Be clearer here. We now explicitly state that this may be the case in paragraph 2 on page 6.
4. Methods: add a reference to table 1 after the first sentence. Table 1 is referenced in introduction and does not use QOF data
5. Is it correct to calculate average scores without doing a factor analysis or reliability analysis? I am not sure about that. (maybe you add apples and pears). We do not believe that factor analysis is required. Our use of simple average achievement levels by indicator categories are used only to summarise the main results. Figures for each individual indicator are provided in later tables.
6. page 8: this is table 6 and not table 5. Changed table title
7. I guess that you refer to the difference between diabetes prevalence rates (28%)? It is not clear where the reader can find this 28%. Added text to show example is the difference between prevalence rates in England and Scotland (Page 10, 2nd paragraph)
8. page 10: Finally, ...... higher than IN the rest of the UK. (add 'in'). Added in
9. tables: is there a specific reason to use 99% confidentiality intervals instead of 95? I have no problems with it, as you use large numbers, but please mention it explicitly. These were used because of number of comparisons in the paper. This is now explicitly mentioned in methods (Page 7, paragraph 3).

Reviewer: Martin Roland
Reviewer's report:

General
Using the QOF as a quality benchmark to compare policy in the four countries is an interesting
approach, and of value in its own right. The broader conclusions about the different health care systems need to be treated (and presented) with some caution as they are somewhat speculative. However, there are few papers which have chosen to look at the four countries in the UK in terms of their different approaches to health reform, so the paper is welcome.

More detailed points.
The term 'payment quality' used by the authors is analogous to our 'reported achievement', and 'delivered quality' is analogous to 'population achievement' in work which we have published Doran et al NEJM 2006; 355:375-384. It would be helpful if academics could agree on these terms! The naming of payment/reported and delivered/population quality was simultaneous (Prof Roland's paper defining these terms published in October 2006, ours in November). We have changed 'delivered quality' to 'population achievement' since we agree that it better describes the nature of the data. However, we have retained our term 'payment quality' since we believe that data for payment more accurately reflects the nature of the data, and have reference the source paper where the term is defined.

The separation of types of outcome into four categories is sensible, but I have a few queries:
* Why have they left out CHD2 (referral for exercise testing), apart from the fact that it doesn't sit neatly in any of their categories?
* Why have they left out DM2 (BMI recorded) and DM13 (micro-albuminuria testing)?
The indicators that we have used are those that apply to the whole register so that a comparison of payment and delivered quality can be conducted. We have added DM2 and DM5 to the tables and changed the composite scores. The 'true' denominator for CHD2 and DM13 is not the whole register, because certain patients are automatically excluded as ineligible, rather than all exceptions being decided by practices. They are therefore not included.

* Exception reporting for intermediate outcomes has to be calculated in a different way from the other indicators. The authors haven't provide details of how they've calculated the indicators, so its difficult to comment on whether they've done this correctly, but they certainly haven't noted in the methods that they've treated intermediate outcomes indicators any differently. The method for secondary indicators is to use the denominator from a relevant primary indicator as the basis for calculating exceptions, rather than the register size for the disease in question. They rather imply that they have done the latter (page 6, line 7). We use the same way of calculating intermediate outcomes as is used for the process and treatment indicators. Thus we assume that not measured is not controlled. We have explicitly stated the method in the paper (Page 6, 3rd paragraph). The alternative is to use 'patients measured' as the denominator which assumes that those unmeasured are the same as those measured. The 'truth' is likely to lie somewhere between these assumptions - some patients without measurement will truly have controlled blood pressure etc; but the unmeasured may be less likely to be in control (based on previous research showing that those not attending care have poorer outcomes eg Griffin S. Lost to follow up: the problem of defaulters from diabetes clinics. Diabetic Medicine 15: S14-S24). Since we distinguish process and outcome, there is no element of double counting in our composites, as there would be if we were using a single composite measure of quality. The table below shows that the principal finding of differences between the four countries is unaffected.

<table>
<thead>
<tr>
<th>Intermediate outcomes</th>
<th>England</th>
<th>Scotland</th>
<th>% Difference from England</th>
<th>Wales</th>
<th>% Difference from England</th>
<th>Ireland</th>
<th>% Difference from England</th>
<th>'Population quality'</th>
<th>68.2</th>
<th>69.8</th>
<th>1.6</th>
<th>67.1</th>
<th>-1.1</th>
<th>71.5</th>
<th>3.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>'Suggested Method'</td>
<td>74.3</td>
<td>75.9</td>
<td>1.5</td>
<td>73.6</td>
<td>-0.8</td>
<td>77.3</td>
<td>3.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A few comments on the results for the countries:
Wales: It appears that Welsh GPs have to deal with a greater burden of disease, in a more deprived population (although they don't comment on the latter point), but nevertheless achieve comparable levels of quality to their English counterparts for those indicators over which they have most control. Where they fall down is on indicators that may require the cooperation of outside agencies.

Northern Ireland: NI has a relatively young population profile compare to the UK as a whole, and the slightly lower reported prevalence rates (high for CHD & stroke, low for HT and diabetes). Both these factors may affect achievement rates. There's not much the authors can do to address this problem, but they shouldn't at least acknowledge it. We have added text regarding Northern Ireland having younger population and its possible impact on achievement. (page 13, paragraph 2) and referenced Doran T, Fullwood C, Gravelle H,