Author's response to reviews

Title: High workload and job stress were associated with lower practice performance in general practice: an observational study in 249 general practices in the Netherlands

Authors:

Pieter van den Hombergh (p.hombergh@chello.nl)
Beat Künzi (beat.kuenzi@swisspep.ch)
Michel Wensing (M.Wensing@iq.umcn.nl)
Glyn Elwyn (glyn.elwyn@btinternet.com)
Jan van Doremalen (j.vandoremalen@ives.umcn.nl)
Reinier Akkermans (r.akkermans@ives.umcn.nl)
Richard Grol (r.grol@iq.umcn.nl)

Version: 3 Date: 15 March 2009

Author's response to reviews: see over
Dear Roxane,

Thank you for reviewing our paper “High workload and job stress are associated with lower practice performance in general practice”.

It helped us to look at our paper again to improve it on details. We will react to the comment of both reviewers systematically. We picked out the comments of the reviewer and answered them straight away using inverse writing.

I First reviewer Joseph Lee

Minor Essential Revisions
  1. Minor formatting errors in Results section and headings such as "GP Workload". The heading missed a “return” and that has been corrected.

Discretionary Revisions
1. As stated in the Strengths and Limitations, this is an observational and exploratory study. As such, it is difficult to make the implication that job performance would improve as a result of increasing GP time per patient or job satisfaction. These may be true but would have to be addressed in another type of study.
   We have carefully avoided to implicate a causative relationship. Finding a linear relationship between workload and performance in an exploratory study could only open the way to look for a causative relation in a different study. We have added that in the discussion under the heading research.
   A prospective study to analyse if practice performance improves by increasing practice time per patient would be valuable for future policy on manpower.

2. This study used GP's in fee-for-service practice. There might be more discussion about the potential difference within a capitation or partial blended capitation model.
   We omitted to state in the method section that in 2006 our system changed, before 2006 we had a capitation system. It is in the Discussion under "policy".
   We added in the method section in the text:
   The study was done before The Netherlands changed from capitation payment to a mixed system.

3. This study did not make reference to GP's using other health care professionals such as nurses, nurse practitioners or others to assist in the management of patients. As there is a trend towards inter-professional primary care, this might have been addressed or discussed.
   The practice nurse was only introduced in practices from 2002 onwards. Studies into the effect of their introduction showed invariable an increase in quality of care with no reduction of the workload for GPs. Our analysis was at the practice level and looks at practice performance irrespective by whom it was realized. Yet, we should be clear about it in the paper. We added in the discussion under “strengths and limitations:
   From 2002 onwards the practice nurse was rapidly introduced in general practice in The Netherlands increasing performance in disease management. Studies into the effect of the practice nurse showed improved outcome, but no reduction in workload for the GP. Yet, future studies on the practice level should include workload and output by the practice nurse.

4. The Conclusion section is quite short and scant. One could consider commenting on some of the Discussion within the Conclusion.
   It sure was a modest conclusion, possibly not giving the study full merits. We added the following to the conclusion as recommended.
   More GP-time per patient and less job stress are related to better practice performance. The findings add weight to previous work on workload, which mostly focused on the physician rather than the practice as in our study. Workload and job stress should be monitored at the national, local and practice level using list size adjusted data. A prospective study could clarify a possible causative
relation and organisational development using feedback on workload and job stress could benefit both patients and GP.

5. Table 2 refers to Actual GP time per patient. The range of hrs/week on the job per 1000 patients is 3.1 to 62.3. This range seems unusual. I wonder whether some respondents made an error. It would be difficult to believe that a GP spends 3.1 hours per week on 1000 patients unless there is some unexplained situation such as involving a nurse practitioner. We looked at our data to find an explanation for this strangely low number of hours per 1000 patients. We found to our surprise that a number of practices with inexplicable low working hours had used the list size of the practice without correcting for having a second colleague. Also one practice used the list size of the practice of four GPs but the workload was calculated as if they were single handed. We had to omit these practices from our set except for the four single handed GPs, where we could simply correct the practice size. We repeated all the analyses and corrected the data. The graph in figure 1 became more linear. We had to adjust the results because the relation between workload and job stress and the patient’s opinion of organization of the surgery did not reach significance now. This was adjusted in the results, summary and the discussion.

6. Table 3 is difficult to read and interpret. I would suggest consideration of reformatting the table. We have tried to present these difficult analysis in as simple a table as possible, only showing the significant results. We tried shift on both X and Y axis, but finally settled for this format. We do not see a better way to present it.

7. Overall, the study addresses an important topic using a large and representative sample.

II Second reviewer Gerard Gill

Conflict of interest.
We became befriended after finding out that we both shared a common vision on the importance of practice visits and both developed a practice visit method being geographical antipods. We were happy with his thorough and professional comments.

2. Are the methods appropriate and well described?
Discretionary Revisions
Confirm that the sampling number of patients per GP and per practice is appropriate. Michel Wensing’s work on this area is used to define the numbers of patients needed to be surveyed for accreditation in Australia. A reference to this work would strengthen the argument. Some comments on whether practice sampling is representative would also be helpful. The previous work on numbers of patients per practice referred to a specific situation, in which we looked for a reliable score per practice (reference below). However, the current study explores trends in a sample of practices. Therefore, the score per practice does not have to be highly reliable (e.g. reliability coefficient > 0.80). In fact, we can only post-hoc determine the power of our analyses and lack of statistical power is one potential explanation for not finding associations.


3. Are the data sound?
The sample appears representative of Dutch general practice apart from the oversampling of rural GPs. The Australian impression is that rural GPs are more stressed, work longer and have greater patient demands made on them compared to urban GPs. No comment has been made about this group in the results. Is another paper planned on this matter?
Rurality in The Netherlands is incomparable to Australia. Yet there are differences in service, continuity of care and overall satisfaction between Dutch urban and rural practices. It would be interesting to do a sub-analysis to explore differences between rural and urban practices, but such study is not anticipated, because of the relative meaning of Dutch rurality and the small sample (18.5 % of the total).

Is there any evidence that the sample is an overperforming group? In Australia we would look for specific training for general practice in the GPs, fellowship of the RACGP, that the practice trains medical students or GP vocational trainees as evidence of a quality practice. Patient satisfaction with aspects of GPs appears to be around the reported average internationally of 80-90%.

Practices applied for a practice visit and there has certainly been a selection of more ambitious GPs and higher quality practices. Yet practices also applied as a group thus also including the lesser motivated and lesser performing practices. Furthermore, becoming eligible for support by a practice nurse helped to apply for a visit irrespective of the practice quality. Also more than 2000 practices visits had already been done before 2003, thus taking away many early adopters. Our sample was slightly younger and female and working part time but representative of the Dutch general practice population as is shown in table 1.

Discretionary Revisions
Patient perception of GP care was low. While there is some evidence that patients are not good judges of the quality of GP technical care, is there any literature from the Netherlands about Dutch patients' ability to comment appropriately on this?

This comment refers to a debate, which exists for more than a decade. One viewpoint is that patients cannot judge technical aspects, so they should not be asked to do so (in fact, many decline to give a judgment if asked). The opposite viewpoint is that patients do judge technical aspects, from their perspective, so they can be asked to do so. We have the feeling that the second view is most predominant. Some years ago, we performed a qualitative study which examined how patients develop their evaluations of general practice (reference below). We found that they did indeed not judge technical aspects as professionals would do, but based their judgment of the more instrumental aspects of doctor-patient communication such as the clarity of information giving.


Minor Essential Revisions
The distribution of each measure of practice performance is not given. Is it normally distributed or is it skewed to the left or right.

The outcome measures based on patient questions were left skewed. The skewness for all variables was between -0.1 and -0.53 except for 'Patient opinion on accessibility/availability' (−1.2). The literature has a rule of thumb that says a variable is reasonably close to normal if its skewness and kurtosis have values between −1.0 and +1.0. Some authors also use values between -2 en 2 and given our large sample of 239 one can feel safe about it. (http://faculty.chass.ncsu.edu/garson/PA765/assumpt.htm)

4. Does the manuscript adhere to the relevant standards for reporting and data deposition?
Table 3 appears to be incomplete with a number of columns missing data. I assume that these correlations not to have reached statistical significance. This needs stating in the table. We stated in the table: "Non significant results have been omitted".

5. Are the discussion and conclusions well balanced and adequately supported by the data?
I am a little concerned about the conclusion implied in Figure 1. We know from Australian consultation analyses and the BEACH study that women GPs who work less hours are more likely to provide longer consultations, more prevention, and address psychological issues. One of the worries patients express in Australia is GP availability after hours but does that translate into hard health
outcomes. Australia has the second highest reported health adjusted life expectancy in the world. Meeting patient preferences may not result in better patient outcomes!

*It is true that better practice performance let alone more time per patient warrants better hard outcomes like life expectancy and/or QUALY’s. However, recent studies in UK, The Netherlands and Germany could relate better organized practice care with a practice nurse in chronic diseases (disease management) to better outcome. Since our study did not have data on outcome, we used patient reports and delegation as a proxy for performance. So it is beyond the scope of our study to discuss these consequences.*

**Discretionary Revisions**

More could be made of the need for non patient contact work such as teaching, CME and I assume practice audit reflecting and instituting practice management techniques as prerequisites for good practice performance. No comment was made of the fact that this survey covered a period in Holland where governmental policies created some stress for GPs with a proposal to claw back some of the payments to GPs.

*It is true that recent years (2003-2006) have been turbulent, but this may be the political presentation of what has probably been a successful period for GPs. There were no signs of increased workload (out of hours services had become arranged in large GP-out of office-posts.) the practice nurse was introduced and net income improved. There is no clear picture on the true impact of all the upheaval. We did not want to confuse the reader with it.*

6. **Are limitations of the work clearly stated?**

Yes ,comment was made on the cross sectional nature of the survey and capitation nature of Dutch general practice funding. A factor neglected by the authors which has generalisability impact is that of practice size. Internationally the trend has been to larger sized practices with more practice manager and some aspects of care being delivered by other health professionals. The Dutch model of a general practice workforce of only GPs and practice attendants working in 1-3 GP practices is radically different to Australian, New Zealand, British or North American general/family practice.

*Dutch General Practices hover somewhere in between Southern European GPs, who work virtually without assistance and the Australian comprehensive primary care centers. This certainly limits generalisibility. We have added that in the text:* Generalisability of the results to other countries is limited due the size of Dutch General Practice, which can be positioned between French and Belgian GPs, who work virtually without assistance on the one hand and British, Scandinavian or Australian comprehensive primary care centers on the other.

7. **Do the authors clearly acknowledge any work upon which they are building, both published and unpublished?**

No difficulties.

**Discretionary Revisions**

A better references for Australian and NZ GP contact times is 2. The contact time includes not only the individual consultation time but the number of GP contact hours per year.

*The number of contacts per year would be an interesting indicator for the time spent on an average patient. Yet, it does not reflect physician’s workload. We did not collect these data. We hope this indicator will be collected in our future visit method.*

8. **Do the title and abstract accurately convey what has been found?**

I have some reservations about the use of practice performance to describe what has been used as indicators of quality care. Four of the five measures are derived from patient surveys. I consider the markers utilised indicate quality care as such markers are used in Australia in the RACGP practice accreditation standards. To my view practice performance is a wider concept and is much more clinically based on markers such as disease management for example diabetic control using glycosated haemoglobin and BP levels or prevention.

*Our indicators are a proxy of practice performance with a focus on disease management. When chronic diseases are delegated, it is a change in organization (process) known to increase*
performance and/or quality of care. Because it is process we prefer to use performance, because quality would require data on outcome.

9. Is the writing acceptable?
Some paragraphs were a little stilted for a native English speaker. For example, “Lowering only the hours of direct care provided by GPs may not help and may prove counter productive. Providing feedback, coaching and supervision are also recommended solutions.” might be better written as Lowering only the hours of direct care provided by GPs may not help burn out and stress and may prove counter productive. What experts recommend to deal with these problems are feedback, coaching and supervision. 
We asked Glyn Elwyn as a native speaker to look again for ‘stilted English’ phrases. However, he had already corrected the paper for it previously. He went over it again, scrutinizing it for English and for better formulations. This phrase mentioned above was definitely an improvement and we took over the phrasing of dr Gill in the paper.

This hopefully addresses the points made by the reviewers.
Considering the above we like to ask if you are willing to review the paper again.

On behalf of the authors

Pieter van den Hombergh