Author’s response to reviews

Title: The Implications of the Feminization of the Primary Care Physician Workforce on Service Supply: a Systematic Review

Authors:

Lindsay K Hedden (lhedden@alumni.ubc.ca)
Morris L Barer (mbarer@chspr.ubc.ca)
Kimberlyn M McGrail (kmcgrail@chspr.ubc.ca)
Michael R Law (mlaw@chspr.ubc.ca)
Ivy L Bourgeault (ivy.bourgeault@uottawa.ca)

Version: 2
Date: 28 February 2014

Author’s response to reviews: see over
February 28, 2014

We thank the reviewers for their very thoughtful and constructive suggestions regarding our manuscript. As discussed with the editors, we have discarded the report from Reviewer 1 and have focused here on the suggestions from Reviewers 2 and 3 (review 3 being the more recent review from the same reviewer who had earlier provided review #1). Below, we detail the actions we have taken in response to these reviews:

Reviewer 2

Point 1: I think that sometimes you need to state the year of the study when reporting an individual finding. For example page 13: "Out of Office" Section on Keane. This was written in 1991 (22 years ago). Perhaps you should say, 'In 1991, Keane et al reported'. (Note the tense correction here as well). Check other for the same issue please.

We have added years in several places throughout the manuscript when studies are initially cited and in particular for those statements referencing studies that were published before 2000 (for example, in reference to Bensing et al., page 12, paragraph 4). The tenses have been checked and, where necessary, corrected throughout the document.

Point 2: Correction: Page 13 first full para: Bensing found no difference in referrals, so I think you must mean Australian study (when you say "Like Bensing et al") in this para.

We have corrected this statement to reference the paper by Harrison et al., which is the Australian study the reviewer mentions. We thank the reviewer for picking up this error.

Comment only: I was very glad to see you recognised the limitation of the search, in that there are many studies that investigate (for example) length of consultation, patient mix, morbidity managed etc, where the sex of the GP is an outcome measure, which of course you have not picked up. I assume this is part of the lead author’s PhD, and suggest that searches on these subjects are also needed to ensure full coverage of the available literature in the thesis.

Agree. These topics are indeed part of the lead author’s Ph.D., and the intent is to cover these related areas and associated literatures in subsequent papers.

Point 3: There is one point you have failed to understand. You suggest (with some horror) that many of the studies included did not use statistics to control for confounders. Statistical methods for this process only became mainstream in the late 80s early 90s. And from memory I think the Britt paper was one of the first to apply them to the question of the sex of the GP. There is no reason to forgive lack of adjustment in more recent papers, but there is reason to accept that these were very new methods in the mid 1990s. Please add something to recognize the development of statistical methods over the period of the literature review.

In recognition of the development of statistical methods to control for the effect of confounders, we have added the following statement to page 18: “Statistical methods controlling for the effect of confounders may not yet have been accepted practice in this field when some of these earlier papers were published, which may explain their limited use.”
Point 4: Page 19 middle of first para of section 4.3: "or it may be that this has not historically been true (note you have split you infinitive here - put historically before 'this'), trends over time suggest it might become so'.

This needs to be reworded. I think if you take out 'Or it may be that' the sentence will be OK. In any case you can't start a sentence with a conjunctive (Or).

We reworded this sentence as follows: “While historically this may not have been true, trends over time suggest that it might become so in future.”

Point 5: Page 14 3.2.5 lines 1-2: sentence needs to be read twice to make sense. Try putting a bracket round the last section (rather than in small.....).

We have added the brackets as suggested.

Point 6: Practice is the verb, practice is the noun. Please correct.

We assume here that the reviewer meant to point out the UK grammatical convention where practise is a verb and practice is the noun. It is our understanding however that in the US practice can be used as either a verb or a noun. We have left this as-is but would be happy to be guided by the editors on this rather minor point.

Point 7: You 'compare to' throughout the paper. This should be "compared with". Incorrect, the term is more than or less than. This is easy to check mentally, having mentally removed the other words. For example: Last para p16, last 2 lines: ...'spent more time on unwaged childcare and other household responsibilities compared to'. Remove the bit about time and household responsibilities, and it becomes clear this should be 'than' rather than 'compared' with.

This has been corrected throughout the manuscript.

Watch your tense. In the discussion you are inclined to switch tenses. For example p 16 For example at the beginning of page 17 you use "are accounted for" when the study WAS past tense. Authors contribution: Last line. I think you mean they all gave approval of the publication of this version of the PAPER, rather than the study.

We have corrected the tense issues throughout the manuscript, and changed the wording of the authors’ contribution statement as requested.

Reviewer 3

Point 1: It is a bit unclear which particular articles are included in the study. For instance, citation 25 (Differences in income between male and female primary care physicians. J Am Med Womens Assoc 2002, 57:180–184.) is cited in the 'hours of work' section (Results from North America are similar, with female PCPs working between four and 14.5 fewer patient-care hours per week [8, 20–25]), but the study is not listed in the supplemental table nor on the summary.

We apologize for any confusion about which articles are included in the review proper, and thank the reviewer for pointing out the issue with citation 25 in particular. Only those studies that are included in the supplemental table were summarized as part of the review. We note on page 7, in section 3.1: “Of [those articles selected for full text review] 30 studies met the inclusion criteria; they are summarized in Supplement 2.”
The specific citation (25) pointed out by the reviewer was included in the results sections in error, and has been removed. The study did not meet our inclusion criteria, which is why it was not included in Supplement 2: its primary outcome was income, which we do not consider to be a direct measure of time spent working, intensity of work, scope of work, or practice characteristics. Time spent working and practice characteristics are both discussed in paper (25) and are included as model covariates, but no raw results of these outcomes are presented, which is why the study was excluded during our full-text review.

To make the inclusion criteria more explicit, we have added the following footnote to our Inclusion/Exclusion criteria Table 1: “Raw or adjusted results for one or more of these measures must have been presented. If these measures were included as covariates in a multivariate modeling exercise, the study was excluded unless raw comparisons on one of these outcomes were also presented.”

In the discussion section only, we have chosen to reference some additional literature that focuses on topics outside the scope of this review, but whose content provides some complementary perspective on our own findings. For example, we refer to studies that examine the balance between household and work responsibilities for male and female physicians. These papers did not meet the criteria for inclusion in the formal structured review, but do provide important context about why female physicians may self-report working fewer hours. Whenever we have referred to literature not covered by our review, we have been sure to clearly indicate that (e.g. page 16, section 4, paragraph 3).

Lastly, we have added the following statement to our limitations section (page 21, section 4.4, paragraph 1): “…studies that included our measures of interest as covariates, rather than as a primary outcome of interest, were excluded unless raw comparisons for these variables were also presented.”

Point 2: Table 2.

2. Citation 50, (The influence of race and gender on family physicians’ annual incomes. J Am BOARD Fam Med 2006, 19:548–556) is included in the supplement and Table, but three other similar studies from the same authors that examined pediatricians, general internists, and ob-gyns beg the question of what the authors considered to be ‘primary care.’ So my question is why, if the one was included, and general internists and pediatricians (and, in some circles) OB-GYNs are considered primary care physicians, why did the paper examining family practitioners meet the criteria but the other three did not? The definition of ‘primary care’ is not made available to readers.

We thank the reviewer for pointing out the lack of clarity in our definition of primary care. For the purposes of this review, we considered only physicians whose specialty was primary care, and so have excluded specialists who are not primary care physicians but who may be practicing, at least in part, in a primary care capacity. There were two reasons for this decision. First, we found upon reviewing the literature for these specialties, that it was impossible to determine which physicians were practicing like primary care physicians and which were practicing like specialists. Thus, we felt that any patterns observed for these specialists might not be generalizable to primary care physicians. Second, recognizing that workforce feminization is an important issue within both primary care and specialist physician workforce populations, we are currently working on a companion piece to this article, which will focus on other specialties. Those papers that focus on OB-GYNs, general internists, and primary care physicians will be included in that review, provided they meet our inclusion criteria. In particular, the three articles the reviewer flagged as missing in what was included and reviewed for the present paper, are all “in” for this second manuscript currently under development.

To clarify the definition of primary care physician used for the present paper, we have added a footnote to our inclusion criteria in Table 1: “Specialists physicians (such as paediatricians, or general internists) who may on occasion practice like primary care physicians (i.e. acting as a point of entry to the health care
system, providing person-focused care over time, and acting as a coordinator for care provided elsewhere) were not included.”

We have also added the following statement to the section on study limitations (page 21, section 4.4, paragraph 2): “Our decision to include only those studies that focused on primary care physicians (rather than also including specialists – like general internists or pediatricians – who may practice like primary care physicians under certain circumstances) may limit the generalizability of our results.”

Point 3: Another publication (How do race and sex affect the earnings of primary care physicians? Health Affairs 2009;28(2):557-566.) might provide more information on the topic.

We thank the reviewer for suggesting this additional article. It was captured by our original search but excluded during full-text review. The primary outcome in this study was income. No direct results on differences in time spent working, intensity of work, scope of work, or practice characteristics were presented, although these were used as covariates in a multivariate modeling exercise. Because no direct results meeting our criteria were presented, this study was excluded.

Point 4: A more thorough search strategy would have named the different types of providers considered to be primary care; it also would have included the terms ‘gender’ and ‘sex’. And, because people who examine gender based difference in work hours and productivity are also pretty interested in annual incomes, including physician income in the search would have been a wise move.

Please see our response to Point 2 above. We were deliberate in our exclusion of specialists who may choose to practice like primary care physicians, for reasons noted there. We were also careful to include in our search a comprehensive list of terms to capture primary care physicians (three related MeSH headings and 5 keywords).

The terms gender and sex were excluded from the search strategy (though “female” and “women” were included). We found that these terms (“gender” and “sex”) tend to surface mounds of papers on topics such as gender assignment counselling, sexual dysfunctions and the like, rather than being productive in terms of surfacing anything to do with gender-based patterns of practice. Based on the advice of two University of British Columbia population health librarians, we adjusted our search strategy to include the MeSH headings and keywords for “women” and “female” in place of “sex” and “gender”. This considerably improved the precision of our search.

Terms for physician income and remuneration were also already included in our search strings under the assumption that papers that examine physician income may also include supporting data on work hours and work intensity. This proved to be the case. Several articles that focused on income as their outcome of interest were either included in the final paper (e.g. citations 22 and 26) or were excluded in the full-text review stage (because they did not present direct results for any of our outcomes of interest).

With respect to the thoroughness of our search strategy in general, it is worth noting that, in addition to using a complex carefully-crafted search string, we also conducted exhaustive forward and reverse citation searches to ensure that we picked up any remaining articles that might have been missed with our initial search. We also used multiple white and grey literature search engines to pick up any papers that would have been missed by relying on a single engine.

Point 5: In citation 50, there are some descriptors of differences in the patient population, and practice characteristics. But these boxes are not checked for that study in Figure 2.

We thank the reviewer for pointing out this oversight. It has been fixed, as has the relevant citation in Appendix 2. We corrected the corresponding in-text discussion of article themes (page 8, section 3.2).
also conducted a re-review of our abstraction tool to ensure that no other errors were made in the development of Figure 2.

Point 6: In the methods, the authors state that they use a measure of bias, but they don’t report on the outcomes of the use of that measure – they should do so in the ‘methodological issues’ section (unless the comment that none of the studies would have met the Cochrane criteria is the use of the bias instrument, in which case it would have been helpful to know why (and, I believe the Cochrane bias has to do with trials (particularly RCTs) and it’s not clear to me that one could randomize gender, so it might not be the best tool.)

Our appraisal of study quality was conducted qualitatively, and the results of that qualitative assessment are presented on page 18 section 4.2. We have modified the methods section to make this clearer, and in particular to point out that this was not done using any specific measure. We mentioned the Cochrane handbook because it extensively outlines potential quality issues (relevant for both RCTs and observational studies); however, based on the reviewer’s comment, we have removed reference (especially since we were probably incorrect to refer to it as a “tool” to begin with, as that suggests the possibility of a quality score or ranking). The final paragraph of the methods section (page 7, section 2.2) now reads, “We conducted a qualitative examination of study quality by assessing the following items: clarity of research questions and objectives; appropriateness of study design; sample size and representativeness; validity of measures; addressing possible confounders; and generalizability.”

Similarly, section 4.2 of the discussion has been reworded as follows: “As part of our qualitative assessment of study quality, we identified some significant methodological concerns with the studies included in this review…”

Point 7: While the breadth that the authors seek (across many countries) is admirable, their own acknowledgement that different systems might have different incentives and different countries have different trends in work patterns and different rates of feminization makes the combination of countries challenging. Better would be to limit to a particular country, for the purposes the authors have.

Although we agree with the reviewer’s point that different incentives and trends in work patterns complicate the discussion of feminization trends, we disagree with the suggestion that we restrict the work to an individual country. We believe that the ability to make cross-national comparisons is, in fact, a strength of our review, as we were able to comment on whether particular health care system structures tend to amplify or diminish observed differences between male and female physicians. For example, we found that certain differences between male and female physicians were smaller in countries where primary care physicians play a gate-keeping role (page 13, paragraph 2). In addition, the small number of submissions from each of the represented countries would have prevented us from being able to reach any clear conclusions, especially given the variability in outcome measures and methodologies within the selection of studies from each country.

Point 8: In some important places, I don’t think that they have adequate support for their statements. For instance, the discussion on page 16 about childbearing makes several declarations, but also state that these issues were not the primary focus and were seldom explicitly investigated. I don’t think that, based on such information, I’d make the declaration that “once family circumstances were accounted for, the gender of the physician had no significant effect on hours worked.” Without stating pretty explicitly where that statement came from and whether it was generalizable (to all countries examined, to all studies.) Similarly, it doesn’t seem right to focus on the non-work hours commitments in the general discussion when only one study in Canada evaluated that aspect of work balance. Perhaps that would work better in the ‘future research’ section.

We thank the reviewer for these very helpful comments. In response, we have reworded the section about
family circumstances and included a specific citation to support our statement (page 16, section 4, paragraph 2): “Female PCPs who had children under age 18 worked fewer hours per week and were more likely to have self-reported part-time status compared with those who did not. The dampening effect of children on work hours was twice as large for women as it was for men. One study found that once family circumstances were accounted for, the gender of the physician had no significant effect on hours worked [17].”

While family circumstances and non-work activities were not the primary focus of any studies, we make reference to them in the discussion because they were included in several articles as model covariates. In all papers that so included them, they had a substantial impact on the observed differences between male and female primary care physicians. As suggested, we have added a statement in the future research section about the impact of work balance and family circumstances: “Issues of work-life balance and child-rearing and household responsibilities are also under-researched, especially given their observed impact on full- vs. part-time job status and working hours [17, 37]. In the 2007 and 2010 Canadian National Physician Surveys, the majority of respondents identified attaining balance between personal and professional life as the most important factor for a satisfying practice [40].”

To make the paper better, I’d suggest that the authors do the following:

1. Re-do the literature search after defining primary care (or limiting to a particular primary care specialty), including the suggested mesh terms, and limiting the analysis to a single country of work.

Please see our comments in response to points 2 and point 7 above. As we note in our responses to those points, we deliberately limited the search for this paper to primary care physicians (covered exhaustively with MeSH headings and keywords), and excluded those specialties that may at times practice in a primary care capacity. These specialties will be covered in a forthcoming companion piece. Additionally, we feel that restricting our study to one country of interest would methodologically weaken the paper and would limit our ability for generalization or cross-national comparison.

2. Get at least two reviewers to determine whether a particular study addresses the particular aspects of interest.

We recognize that systematic reviews are indeed strengthened by having a secondary reviewer to determine whether individual studies meet the inclusion criteria. In this case, however, the broad scope of the review and the poor indexing of study keywords (as noted above) resulted in an extremely large pool of potentially relevant articles: 1271 abstracts and 74 full-text articles. This made it impractical (for reasons of time and cost) to add a second reviewer.

3. Use the data to actually estimate or forecast how increases in feminization might impact the workforce (since that is the focus of the effort). The theoretical framework is fine, but suggest, given an average of x% fewer hours, y% less work annually, but z% greater productivity what might that mean?

We agree with the reviewer that this type of forecasting is critically important, but would argue that one can only do so much in one paper, and that the subject matter of this paper should logically precede any such modelling efforts. Furthermore, not one of the studies we reviewed would provide sufficient data onto which to build a forecasting model. As part of the lead author’s doctoral work, we plan to build a model that will predict specific service shortage or surplus outcomes related to the trend of primary care workforce feminization in British Columbia, Canada.

4. Use the limitations (ie. Retirement differences, concept of a ‘window’ of relatively lower productivity, and, I’d suggest, the income differences) to suggest what additional work should be done and how that
work might impact their model.

As we state in response to point 3 above, we feel that the data presented in the included articles is insufficient to build predictive models at this time. In section 4.3, we have commented on how the additional work needed in this area may alter current impressions about the impact of workforce feminization. For example, we note (page 19, section 4.3) that “differences in retirement patterns between male and female physicians may partially or wholly offset other trends in service provision, when viewed over an entire life cycle.”

In closing, we wish to reiterate how valuable, refreshing, and rare it is to have the benefit of two such thoughtful and thorough sets of comments. We are also heartened by the fact that both noted that this study covers an important and under-researched topic in the field of health human resources.

Sincerely,

Lindsay Hedden, MSc, BSc (hons.)
PhD Candidate
School of Population and Public Health
University of British Columbia
201-2206 East Mall, Vancouver, BC V6T 1Z3
Tel: 604.822.5684
Email: lhedden@alumni.ubc.ca

c. Morris L. Barer
   Kimberly M. McGrail
   Michael R. Law
   Ivy L. Bourgeault