Article details: 2020-0037	
Title	Calculating physician supply using a service day approach: a descriptive analysis
Authors	Terrence McDonald MD MSc, Brendan Cord Lethebe MSc, Lee A. Green MD MPH
Reviewer 1	Mathews, Maria
Institution	Division of Community Health and Humanities Memorial University of Newfoundland, St. John's, NL
General comments (author response in bold)	1. The service day method included a number of judgements (e.g. minimum 10 visits, 3 service days per week, 46 weeks per year) in its calculations. A rationale for why these cut-offs were used and how they were determined would be helpful. Sensitivity analyses for 3 days a week and 46 weeks per year cut-offs (to demonstrate that they are reasonable) would be helpful. **We have added to our sensitivity analysis and provide further rationale for the 3 day cut-off through out the text.** 2. More information on the Health Quality Council of Alberta primary physician method (at least in terms of how it might differ from other well-known methods) Please lines 142-48 for further details. 3. In general, the results are difficult to follow. I'm not sure whether Table 1 has been cut off or if it is presented in its entirety in the manuscript (it looks off centre on the page). Given that the research objective was to compare the two methods of calculating FTE and PTE, the presentation of results does not quite meet expectations. For example, which physicians were identified consistently as FTE or as PTE using either method (and what are their characteristics)? Which physicians are inconsistently classified by the two methods and what are the characteristics of the physicians who were inconsistently classified? I'm not sure this is the same as the 85% reported for physicians who were classified as either FTE or PTE by either method as reported in table 1. I'd like to see % agreement based on how an individual physician is categorized by the two methods rather than a summary statistic based on final counts. **Please see new tables 1 and 2 and kappa values in the sensitivity analysis for clarity.** 4. The text in the manuscript does not present the data in the same order as it is presented in the table. For example, the text discusses weekend days worked before days worked (which is in the opposite order as the table). The figure "4.95" in line 161 should be 4.93. The "Zone" variable in the table needs more description as the zones would not be understandable to anyone outside Alberta. **The text has been edited accordingly to now follow the results.** 5. Appendix 1 (a detailed histogram of average service day by GP cohort) is missing. Figures 1 and 2 appear to present the same data (and it's unclear why only 3 of the 4 groups are included in the figures). The text and figures should also clarify that these are billings not income. **The histograms and Figures have been removed. Table 1 and 2, as well the Appendix Tables A1-5 have been edited and created to add clarity and highlight the main points of the manuscript.** 6. In table 2, the total number of physicians grouped by the two methods in a single year differ (e.g. in 2015 there are 3560 in the service method and 3572 in

	the income method, in 2014, there are 3419 in the service method and 3427 in the income method). **Should these not be the same? See above response in 5.** 7. The conclusion does not reflect what was presented in the manuscript but rather what the next steps in the research will be. **The conclusion has been edited accordingly.** 8. Overall, I think the intention of the article is important for physician workforce planning. However, while the results present a generally convincing case (aside from the issues described above) that methods provide similar results, I'm more interested for whom (which physicians) they produce identical classification and for which they produce inconsistent classifications. This would help me better understand the implications of using one method over another. **Please see Table 2 and Table A5 for clarification.**
Reviewer 2	M. Lavergne
Institution	UBC Centre for Health Services and Policy Research
General comments (author response in bold)	1. Analysis describes alignment of categories based on pre-determined cut points (FT, defined as providing 3 or more service days per week over 46 weeks of the year is compared to income percentiles method used for CIHI FTE). Substantial space is devoted to describing the characteristics of the various cohorts in the text. Appendix 1 histograms for average service day by GP cohort. However, there is no overarching histogram to show the distribution of service days across all GPs. I think this is important information to contextualize cut points. **The histograms and Figures have been removed. Table 1 and 2, as well the Appendix Tables A1-5 have been edited and created to add clarity and highlight the main points of the manuscript, as well as next steps to assess comprehensive versus more focused practices.** 2. Sensitivity analysis explores descriptive statistics for each fiscal year, as well as varying the definition of a "service day" but I feel like sensitivity analysis using multiple cut-points for service days would be more important to feature in main results. **A sensitivity analysis has been added including kappa values.** 3. I found Figures 1 and 2 to be hard to follow. I'm not clear why it is relevant to compare across income quintiles within cohorts. I would suggest clarifying the description of these features, or possibly including alternate tables/figures in the main results suggested in points 1 and 2. **The histograms and Figures have been removed. Table 1 and 2, as well the Appendix Tables A1-5 have been edited and created to add clarity and highlight the main points of the manuscript, as well as next steps to assess comprehensive versus more focused practices.** 4. Patient-level continuity was calculated based on primary provider assigned based on a published algorithm. If I understand correctly, patients with one visit would have had 100% continuity - is it possible that continuity is over-estimated for walk-in clinic type physicians? Might it be helpful to present a physician-level measure that is independent of the panel algorithm (for example, of the unique patients a physician saw within a year, what proportion of all those patients' contacts were with the index physician)?

	Thank you for this very helpful comment. Though beyond the scope of this current work, we plan to explore the continuity of patent care beyond the algorithm of panel calculations in our future work. 5. (At least until recently) Alberta has had higher payments to physicians than average. Are there any reasons to think this method might not apply in other provinces? **We have consulted with colleagues throughout Canada, and have provided some specific information on how this novel method might be applied. See Interpretation lines: 261-69. Based upon reviewer comments below we have revised the methods section to provide additional clarity for reproducibility.** 1. Line 29 – rephrase for clarity The text has been adjusted/edited. 2. Lines 54-56, by definition the income threshold approach does classify high-billing physicians as full time. **We have provided some clarification on this throughout the manuscript.** 3. Line 77 – can you explain what service delivery sites could include? How was this used in analysis? **For billing purposes, GPs register one main physical address where care is provided, and fee-for-service claims are submitted and paid to the care provided at that location. See line 110-12.** 4. Presumably study cohort includes GPs/Family Medicine physicians. It may be helpful to define this in methods or use more inclusive language (e.g. primary care physicians). "Registered as Family Physician" is a variable in Table 1. What does this mean? **The "%" registered family physicians has been removed and "GP" has been used inclusively within the text, for those who (presumably) practice primary care, it includes both GP's and Family Physicians. The registration of a 'Family Physician' is collected and made available by the Alberta College of Physicians and Surgeons.** 5. Please explain urban/rural classification in more detail See Methods and lines: 151-54. **The rural urban continuum are divided into 7 different geographic areas based on on the aggregation of Local Geographic Areas also known as LGAs. Numerous components were used to classify these area including population density and distances to both health and non-health services in both rural and urban centres to name a few. As example, metro centres include population > 500,000 people such as Alberta's two largest cities: Edmonton and Calgary. Whereas urban includes centres with > 25,000 but less than 500,000 and include places such as Red Deer, Lethbridge or Medicine Hat. For further details please see reference** **Alberta Health Services and Alberta Health (2017), Official Standard Geographic Areas, Alberta, Canada.**