PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (http://bmjopen.bmj.com/site/about/resources/checklist.pdf) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

<table>
<thead>
<tr>
<th>TITLE (PROVISIONAL)</th>
<th>Examining sufficiency and equity in the geographic distribution of physicians in Japan: a longitudinal study</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUTHORS</td>
<td>Hara, Koji; Otsubo, Tetsuya; Kunisawa, Susumu; Imanaka, Yuichi</td>
</tr>
</tbody>
</table>

VERSION 1 - REVIEW

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Matthew McGrail</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Senior Research Fellow, Monash University, Australia</td>
</tr>
</tbody>
</table>

| REVIEW RETURNED | 03-Oct-2016 |

GENERAL COMMENTS

This is a well conducted study, which provides a strong argument for improved supply actually not keeping up with demand increases. I only have minor issues.

In the abstract, why do they only state “decreases were also observed in the rural groups” without any further details? Furthermore, no details of this decrease are given in the main paper either. The -4.4% and -7.6% decreases (after adjusting for health needs) were key results for Groups 2 and 3, worthy of highlighting.

Under “Physicians and populations”, I believe that the 2nd sentence beginning as “Physician numbers…” should be “Physician ratios”.

Under the SMA heading, no detail is provided for the relative size (population and area) of this geography and particularly comparing those of urban and rural; we only know that there are 349.

Later, they split SMAs as having either ‘higher’ or ‘lower’ initial supply; however, I did not see any detail specifically how this separation is defined?

The large discrepancy between conclusions based on raw population (22.9% and 34.5% increase in G1 & G4) versus -1.3% and +3.5% after needs adjustment is very interesting. They state “…(this) indicates that it is necessary to adjust healthcare demand when analysing physician supply”. I suggest they could also add that “not doing (this) will produce misleading results”.

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>VICTORIA FAN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UNIVERSITY OF HAWAII AT MANOA (USA)</td>
</tr>
</tbody>
</table>

| REVIEW RETURNED | 09-Oct-2016 |

GENERAL COMMENTS

This paper on the growing inequality of Japan’s health workforce is a valuable contribution to the literature on health workforce studies for a country which has long demonstrated a commitment to and
leadership in health equity.

The methodology would appear to be very straightforward, except for the use of the authors’ "demand-adjusted population". I would like some greater clarification about how this adjustment was implemented, but more importantly the theoretical and empirical basis for doing such an adjustment. Perhaps I missed the relevant reference which may have been cited on this method but I did not see it.

The authors could also consider as an additional revision the use of the Theil T and Theil L inequality indices which permit decomposition by different strata (e.g. urban-rural; other geographical units).

My understanding about the policies in Japan to attract and recruit people to the rural areas was through various types of incentive programs and policies. Such discussion on those past use of policies to recruit rural doctors would be of considerable interest and lend greater policy relevance of this paper.

The authors could further develop the limitations section of their paper and discuss the limitations of the methodology of this demand-adjusted population measure that they use.

### GENERAL COMMENTS

This is an interesting and methodologically rigorous work. Nevertheless there are two aspects that were not clear to me (sorry) and I think they deserved to have been better explained:

(1) how did you calculated the expenditure per capita for each age group? Health expenditure data from Ministry of Health are disaggregated by age group?

(2) You have divided SMAs into four groups (pag. 6) based on the initial physician supply and on the degree of urbanization. You are very clear in what concerns to the "Low initial position" (line 165); and what about those classified as "high initial position"?: are all the other? are those SMAs in the fourth quartile? The same about the urbanization: have you considered the population density also categorized into quartiles? Which was the criteria used to define a SMS as urban or rural?

P.S.: sorry, if the misunderstanding is my fault.

### VERSION 1 – AUTHOR RESPONSE

Reviewer: 1 Dr. Matthew McGrail

#3 In the abstract, why do they only state “decreases were also observed in the rural groups” without any further details? Furthermore, no details of this decrease are given in the main paper either. The -4.4% and -7.6% decreases (after adjusting for health needs) were key results for Groups 2 and 3, worthy of highlighting.
Thank you for your suggestion. In accordance with the reviewer’s comment, we have added a more detailed description of the rural groups’ results in the Abstract and Discussion. (Page 2, Lines 41-43; Page 9, Lines 246-248)

#4 Under “Physicians and populations”, I believe that the 2nd sentence beginning as “Physician numbers...” should be “Physician ratios”.

Thank you for pointing this out. The reviewer is indeed correct. Although the term “physician ratios” may be easy to understand, it is not generally used in similar studies. However, the term “physician density” is used in previous studies and WHO reports, and we believe that this term would not only be technically accurate, but may also be more familiar to researchers in the same field. Therefore, we have replaced “Physician numbers...” with “Physician densities...”, and hope this is acceptable to the reviewer. We thank the reviewer for allowing us the opportunity to rethink this point. (Page 5, Line 125; Page 7, Line 180 and 182)

Reference:

#5 Under the SMA heading, no detail is provided for the relative size (population and area) of this geography and particularly comparing those of urban and rural; we only know that there are 349.

Thank you for the suggestion. We classified the SMAs into rural or urban areas using population density (the median value was set as the cut-off). In accordance with the reviewer’s suggestion, we have added a new table (Table 4 in the revised manuscript) to summarize the relative sizes of the SMAs, and provide a comparison of the rural and urban areas. Since this information would represent the baseline characteristics of the subject regions, we have included the table in the Results instead of the Methods. We think that the addition of the population and area size of each group provides crucial contextual information for interpreting our results, and thank the reviewer for their advice.

#6 Later, they split SMAs as having either ‘higher’ or ‘lower’ initial supply; however, I did not see any detail specifically how this separation is defined?

Thank you for your comment. Although the method for splitting SMAs into “higher” and “lower” initial physician supply was described in the Methods, it may have been unclear as it was described together with the separation of rural and urban areas. We have revised the sentences for clarity. (Page 7, Lines 178-182)

#7 The large discrepancy between conclusions based on raw population (22.9% and 34.5% increase in G1 & G4) versus -1.3% and +3.5% after needs adjustment is very interesting. They state “…(this) indicates that it is necessary to adjust healthcare demand when analysing physician supply”. I suggest they could also add that “not doing (this) will produce misleading results”.

We appreciate the reviewer’s comment on this point. In accordance with the reviewer’s instructions, we have added that “…failing to adjust healthcare demand will produce misleading results when examining physician supply,…” in the conclusions. (Page 2, Line 48; Page 12, Lines 335)

Reviewer: 2   Dr. Isabel M. Correia
#8 (1) how did you calculated the expenditure per capita for each age group? Health expenditure data from Ministry of Health are disaggregated by age group?

Thank you for your comment. Yes, the health expenditure data that we obtained from Japan’s Ministry of Health, Labour and Welfare are disaggregated according to patient age. We have revised the Methods section to clarify this point. (Page P5, Line 114-115)

#9 (2) You have divided SMAs into four groups (pag. 6) based on the initial physician supply and on the degree of urbanization. You are very clear in what concerns to the “Low initial position” (line 165); and what about those classified as “high initial position”? are all the other? are those SMAs in the fourth quartile?

Thank you for your comment. In examining the time-based trends in the number of SMAs with a low physician supply, we defined low supply using the first quartile of the number of physicians per 100,000 population in 2000 as the criterion value. In this part of the study, we focused only on SMAs with low physician supply to track the changes in the proportions of SMAs that fulfilled this specific criterion throughout the study period. In this sense, any SMAs outside of the first quartile may be considered as having a “higher” initial physician supply. However, there was no need to classify SMAs with a higher physician supply as we did not conduct any comparative analyses between SMAs with low and high physician supply.

#10 The same about the urbanization: have you considered the population density also categorized into quartiles? Which was the criteria used to define a SMS as urban or rural?

Thank you for your comment. In the sub-analysis on time-based trends in the number of SMAs with a low physician supply, we focused only on areas with low physician supply, regardless of the level of urbanization. In the sub-group analysis, the degree of urbanization was determined using the population density in 2000 as a proxy variable. We divided the SMAs according to the median value of the population density in 2000 to obtain two groups: SMAs with population densities above the median value were designated urban areas, whereas SMAs with population densities below the median value were designated rural areas. We have revised the Methods to clarify this point, and thank the reviewer for bringing this up. (Page 7, Lines 175-178)

Reviewer: 3 Dr. VICTORIA FAN

#11 The methodology would appear to be very straightforward, except for the use of the authors’ “demand-adjusted population”. I would like some greater clarification about how this adjustment was implemented, but more importantly the theoretical and empirical basis for doing such an adjustment. Perhaps I missed the relevant reference which may have been cited on this method but I did not see it.

We thank the reviewer for this comment. In this study, we conducted adjustment using health expenditure per capita based on Gravelle and Sutton’s findings that this factor is indicative of the degree of physician workload3. In addition, The Nuffield Trust—a charitable trust in the UK that also predicts future health expenditure—has reported a relationship between healthcare demand and health expenditure 4.

An alternative to the use of health expenditure in the adjustment of healthcare demand is the use of medical consultation rates, representing the number of times people visit healthcare providers. However, these rates have the disadvantage of not being weighted with the degree of physician workload 3.

For these reasons, we elected to use health expenditure per capita in place of medical consultation rates for adjustments. The actual adjustment method is described in the first paragraph of the
“Analysis of physician distribution” subsection.

Reference:

#12 The authors could also consider as an additional revision the use of the Theil T and Theil L inequality indices which permit decomposition by different strata (e.g. urban-rural; other geographical units).

We appreciate the reviewer's interest in additional information on decomposition using the Theil T and Theil L indices. We certainly agree that it is important to explore the structure of inequity by allowing for decomposition. However, our current study is focused on the inequity of physician supply as a whole or across the rural-urban axis, and not on exploring the structure of inequity. In addition, as our data on physician supply are generally comprehensive, the Gini coefficient would present a fairly accurate picture of inequity. Also, the differences in population sizes and structures among the SMAs throughout Japan may make it more difficult to conduct comparisons using the Thiel indices. The common use of the Gini coefficient and ease of interpretability are still (at present) advantages, and this coefficient would be more intuitively understood by a wider audience. Thus, we consider that the use of the Theil indices may not be necessary for our current study, and wish to retain the original text. However, we agree that further study is needed using the inequity indices suggested by the reviewer, and warrants consideration for downstream analyses.

#13 My understanding about the policies in Japan to attract and recruit people to the rural areas was through various types of incentive programs and policies. Such discussion on those past use of policies to recruit rural doctors would be of considerable interest and lend greater policy relevance of this paper.

Thank you for the suggestion. As pointed out, the policies about allocating physicians to rural areas are important to our paper. Although some policies on physician supply were implemented during the study period, their efficacy remains unknown. As suggested, we have added a discussion on these policies to the revised manuscript. (Page 11, Lines 285-301)

#14 The authors could further develop the limitations section of their paper and discuss the limitations of the methodology of this demand-adjusted population measure that they use.

Thank you for the comment. We think that we must allow for the possibility that individual health expenditure is not indicative of the degree of physician workload. On the other hand, at the macro level, using health expenditure is better than medical consultation rates because the latter are unadjusted and therefore completely unweighted for physician workload. As advised, we have added this as a limitation. (Page 12, Lines 320-323)

---

**VERSION 2 – REVIEW**

| REVIEWER | Matthew McGrail  
| Monash University, Australia |
| REVIEW RETURNED | 12-Nov-2016 |
| GENERAL COMMENTS | All previous comments have been addressed |