Author's response to reviews

**Title:** Community characteristics that attract physicians in Japan: A cross-sectional analysis of community demographic and economic factors

**Authors:**

Masatoshi Matsumoto (matmo10@jb3.so-net.ne.jp)
Kazuo Inoue (inouek@m.u-tokyo.ac.jp)
Satomi Noguchi (snoqu-ty@umin.ac.jp)
Satoshi Toyokawa (t-satoshi@umin.ac.jp)
Eiji Kajii (kajii@jichi.ac.jp)

**Version:** 2  **Date:** 6 December 2008

**Author's response to reviews:** see over
Response to Reviewer 1 (Dr Nanako Tamiya)

We appreciate the Reviewer 1’s thorough reading and analysis of the manuscript. The feedback was very constructive. In light of the Reviewer’s comments, we have rewritten the manuscript to improve clarity. All altered or inserted parts are underlined in the manuscript, so that they can be easily located. Response to each comment follows in the order of Reviewer 1’s concerns.

Major Revisions
1) Reviewer 1 raised a concern that differences in correlation coefficients described in original Table 3 might be caused entirely by chance, and thus should be tested. Statistically, however, the difference cannot and should not be tested because the coefficients in this study were derived from the total population (total municipalities) rather than from subgroups of the total population. A statistical test for detecting significant difference in two values is possible only when the two values are from different subgroups of the total population (such as two groups of municipalities from two different prefectures). If both values are from the total population (i.e. total municipalities as in our study), estimation intervals of the values cannot be calculated, and thus statistical tests cannot be done. For the same reason, the difference between two Gini indices also cannot be tested. When the two values are different, the values are interpreted as ‘different’ without any p value in a study using the total population.

With regard to Table 4 (Table 3 in current version), we did not include ‘population density’ in the multiple regression analysis because ‘population density’ has a strong collinearity with ‘daytime population density’. In a simple correlation analysis as that shown in Table 3 (Table 2 now), the existence of collinearity is not a problem. So we can enter whatever variables we like to the model. But in the case of multiple regression analysis (like current Table 3), the situation is different. If we enter concurrently two variables into the model that strongly correlate with each other, the variables interfere each other and their correlation coefficients lose statistical significance. For reference, we show results of the regression analysis to which ‘population density’ was added.
As can be seen in these data, coefficients of both population density and daytime population density lost their significance and VIFs for both variables became remarkably high. VIF is a parameter of collinearity. In order to conduct multiple regression analysis properly, it is commonly said that VIF should be less than 10. Thus, in the above example, either population density or daytime population density must be deleted from the model. In the paper we deleted population density because its correlation with physician-to-population ratio was lower than that of daytime population density. An explanation on the issue of collinearity was added to the text (page 13, line 4-6).

2) Some citations were added to the sentence on page 4 (the last 2 lines). We thank the Reviewer for the suggestion.

3) The Reviewer's comment on night or weekend clinics is quite important. According to the statistics from Ministry of Health, Labour and Welfare [1], a small proportion of hospitals officially open at night. Please refer to the following table.

<table>
<thead>
<tr>
<th>Hospitals open at night and on weekend</th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thu</th>
<th>Fri</th>
<th>Sat</th>
<th>Sun</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of hospitals officially open after 18 o'clock</td>
<td>12.4</td>
<td>11.0</td>
<td>11.4</td>
<td>10.3</td>
<td>12.2</td>
<td>2.4</td>
<td>3.5</td>
</tr>
<tr>
<td>% of hospitals officially open at any time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As for clinics (institutions with less than 20 inpatient beds), the proportion is even higher.

<table>
<thead>
<tr>
<th>Clinics open at night and on weekend</th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thu</th>
<th>Fri</th>
<th>Sat</th>
<th>Sun</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of clinics officially open after 18 o'clock</td>
<td>26.3</td>
<td>23.7</td>
<td>19.7</td>
<td>13.7</td>
<td>26.1</td>
<td>3.7</td>
<td>4.5</td>
</tr>
<tr>
<td>% of clinics officially open at any time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

So, the existence of night and weekend clinics and hospitals can potentially be a bias of
the results, but still out statement ‘most medical services are provided during business hours’ holds true. We inserted sentences explaining this issue (page 16, line 8-11).

Minor Revisions
1) and 2) We changed the parts accordingly.
3) ‘Urbanization’ and ‘urbanity were uniformed to ‘urbanity’
4) We corrected the misspelling.

Reference

Responses to Reviewer 2 (Dr Hideto Takahashi)

We appreciate the Reviewer 2’s thorough reading and analysis of the manuscript. The feedback was very constructive We took the comments seriously and rewrote the manuscript almost exactly according to the suggestions. All of the altered or inserted parts were underlined in the manuscript so that they can be easily located. Our responses to comments follow.

Major Comments

We again asked a native and professional proof-reader to make the English clear. We believe the revised version is clearer and easier to read than the previous version.

According to Reviewer comment, the original Table 2 was deleted. Also means and interquartile ranges (due to skewed distribution of most of the variables) were added to Table 3 (Table 2 in the current version).

Minor Comments
In Table 4 (currently Table 3), ‘B’ was replaced by ‘coefficient’ and 95% confidence interval of each coefficient was shown. The column of ‘tolerance’ was deleted.

An explanation on VIF (variance inflation factor) was inserted into the Methods (page 10, line 9-12).
An explanation on what p value indicates was added to Table 3 and 4 (footnotes of current Table 2 and 3).

As for the Figure of Lorenz curve, the suggestion that it is not necessary is understandable. But we were concerned that most of the readers of Human Resources for Health (i.e. non-economic readers) may not understand what a Gini index indicates unless its graphic presentation is shown supplementary to the verbal explanation. We thus retained the figure of the Lorenz curve in the manuscript. After rereading, if it is still considered superfluous, please do not hesitate to contact me again.