PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Temporal trends of physician geographical distribution and high and intermediate physician density areas and factors related to physicians' movement to low physician density areas in Japan: A longitudinal study (1996–2016)
AUTHORS	Ishikawa, Masatoshi

VERSION 1 – REVIEW

REVIEWER	Noriko Sasaki Kyoto University Graduate School of Medicine, Japan
REVIEW RETURNED	27-Jun-2020

GENERAL COMMENTS	Thank you for the opportunity to review this interesting manuscript, which reports on a topic of physician maldistribution and urban-rural flow across Japan. However, the paper fails to fully understand the hypothesis, methods, expected results, as well as how to implement the study results practically. There are some specific comments I would like to make: (1) P4 The current contents of "Strength and limitation of the study" seems to be insufficient. Readers cannot know the whole picture of the study. Please see other original articles in this journal for reference. (2) The execution of the study is limited by a lack of methods information. It would thus be helpful in the Methods to have more information about the points below: ・Could the author show a figure of data extraction process? Please clarify the inclusion and exclusion criteria. ・P7 L5-7：The author states that "In terms of geography, I categorized the 344 secondary medical areas (SMAs) in Japan in 2016 into three groups based on the combinations of population size and density". Please clarify how the author defined these three categories in detail, and explain if it is valid. In addition, how did the author cope with the frequent code changes with SMAs before the author analyzed the data? (3) If the author's interest was to observe changes over time, why did the author compare only two points (1996-2006 vs 2006-2016)? Has the author ever considered using other analytical methods such as regression discontinuity design? If so, why did the author choose the current design? (4) The observation period is 20 years long. How does the author think about the effect of economic crisis (eg Great Recession), various dramatical policy changes related to physician distribution, resident physician education system and programs, healthcare

	services, population aging, etc., to the study results? These points need to be discussed in the Discussion or in the Limitation section. (5) What is the implication of this study and how does the study results expected to be implemented in the future practically? Minor points: ・p7　Is the category 'other' appropriate? I assume this category would include physicians working at public health centers, industrial physicians, as well as unemployed physicians. For example, physicians working at public health centers may tend to remain at the same region, whereas unemployed physicians including mothers who have delivered their child may easily move with their husbands in Japan. A mixed-up category may affect the results. ・Table3: Please clarify the name of the categories '0-14/ 15-29/ 30-44/≧45' .

REVIEWER	Candice Chen George Washington University, USA
REVIEW RETURNED	20-Jul-2020

GENERAL COMMENTS	This paper is very interesting and the level of data provided on the physician workforce over time is important. The analysis of physician migration over the two-10 year periods, and the analysis of factors associated with practice in low density areas, are also important in understanding and addressing the challenge of workforce distribution to meet the needs of communities. A few specific comments: For international readers who may not know much about Japan's healthcare system, a little more background information would be helpful in the introduction to understand policy changes discussed. For example, I wasn't completely clear what "new residency training program" meant – was residency training not previously required and medical schools would send graduates directly to practice or would they previously assign graduates to residency programs? And to clarify, for the system to place students into state health care – what % of health care is state health care in Japan, does 18% successful mean 18% ended up working in state health care, and I wasn't sure what the prefectural scholarships implications were, except in context it seemed counter to the goal of state health care recruitment? As a reader, I would love to understand these issues a little better. I struggled a little in interpreting the results and tables. I think largely because sections of the methods could use more detail, including: • For the physician density categories – are the SMA categories set based on 2016 (page 8)? It would help to clarify this in the previous paragraph that discusses how the categories are set. Also, how does this compare to the categories if set in 1996? It seems by setting it in 2016, if there is improvement in physician density over the 20 year study period for some SMAs – that then change categories as a result – then this will be missed using this method. • For the physician migration, I assume each cohort includes only physicians who are present for the full 10 year time periods – for example, looking only at physicians present in 1996 and their

aggregated change over the 10 year period. This section of the methods is very short and therefore, it took me some time to figure out and interpret the tables.
• For the regressions (Table 4) it seems control variables are all based on the start of the time period? For example, 1996 age, workplace, physician density practice, type of institution, etc. Can this be clarified in the table. Also, in this case, shouldn't physician density "low" in 1996 or 2006, be included in the regression results?

I found the negative relationship between rural practice and physician density interesting. However, I wonder if this relates to smaller populations in rural communities – driving the physician to population density up due to the denominator rather than the numerator. This might be worth exploring further since the study finds rural practice is negatively associated with working with low physician density areas and this might be interpreted as a less than desirable outcome.

I wasn't quite sure about the finding that increases in physicians working in high density areas was less than increases in low density areas. Based on table 1, in 2016 only 11.7% of all physician were working in low density areas vs. 12.1% in 1996. In rural areas, it goes from 7.7% in 1996 to 6.0% in 2016. This seems worth mentioning - understanding that underlying this is an overall increase in the physician workforce, so low density and rural areas may be better off as the author discusses due to diffusion to rural/low density areas. However, it also brings up an interesting question of the implications of increasing physician numbers in high density areas and whether this will be an area of concern in the future (for example, over-use). Is this a concern in Japan?

VERSION 1 – AUTHOR RESPONSE

Response to Reviewer Comments

Reviewer: 1

(1) P4 The current contents of "Strength and limitation of the study" seems to be insufficient. Readers cannot know the whole picture of the study. Please see other original articles in this journal for reference.

Response: Thank you for this suggestion. This section has been revised based on your suggestions as follows.

- *This study longitudinally examines the geographic distribution of physicians in Japan focusing on physician density by secondary medical districts using individual physician data with permission from the national government.*
- *To improve the uneven distribution of physicians, especially for less-experienced physicians, a new system that considers these factors would create opportunities for younger physicians to work in these low-density areas.*
- *This study only focused on correlations and was unable to determine causality. Future studies could use interviews and questionnaires to facilitate more comprehensive research for physician migration.*

- *The observation period is 20 years. The effects of various environmental changes, such as the global economic crisis, policy changes for physician maldistribution, and population aging, were not considered.*

(2) The execution of the study is limited by a lack of methods information. It would thus be helpful in the Methods to have more information about the points below:

・Could the author show a figure of data extraction process? Please clarify the inclusion and exclusion criteria.

Response: Thank you for this comment. The methods section was revised based on your suggestions as follows.

"When creating the cohort dataset, we analyzed the physicians who answered in both years. Additionally, in the original data obtained from the Ministry of Health, Labor and Welfare, there were no incomplete or missing data."

・P7　L5-7：The author states that "In terms of geography, I categorized the 344 secondary medical areas (SMAs) in Japan in 2016 into three groups based on the combinations of population size and density". Please clarify how the author defined these three categories in detail, and explain if it is valid. In addition, how did the author cope with the frequent code changes with SMAs before the author analyzed the data?

Response: Thank you for pointing this out. The methods section was revised based on your suggestions as follows.

"Based on the classification used by MHLW, the first group (urban) consists of areas with a population of at least 1 million or a population density of at least 2000 people/km2. The second group (intermediate) consists of areas with a population of at least 100,000 or a population density of at least 200 people/km2. The third group (rural) consists of areas that do not belong to the first or second groups. The municipality borders that were altered because of mergers were adjusted based on the borders used in 2016."

(3) If the author's interest was to observe changes over time, why did the author compare only two points (1996-2006 vs 2006-2016)? Has the author ever considered using other analytical methods such as regression discontinuity design? If so, why did the author choose the current design?

Response: Thank you for this comment. As you pointed out, various research design methods are possible. As shown in Table 1, there was no significant change in the ratio of the number of doctors in each physical density area category for every two years with regard to physical demographics and professional characteristics from 1996 to 2016, so a regression discontinuous design was not adopted. Two points were set every 10 years for the past 20 years, and a multivariable logistic regression analysis was performed. The limitations section was corrected based on your suggestions as follows.

"Table 2 and Table 4 analyze the whereabouts of physicians at two points, 1996 and 2006, or 2006 and 2016, and do not consider changes during the period."

(4) The observation period is 20 years long. How does the author think about the effect of economic crisis (eg Great Recession), various dramatical policy changes related to physician distribution, resident physician education system and programs, healthcare services, population aging, etc., to the study results? These points need to be discussed in the Discussion or in the Limitation section.

Response: Thank you for pointing this out. The limitations section has been revised based on your suggestions as follows.

"The observation period was 20 years. The effects of various environmental changes such as the global economic crisis, policy changes for physician maldistribution, and population aging, were not considered."

(5) What is the implication of this study and how does the study results expected to be implemented in the future practically?

Response: Thank you for pointing this out. The discussion section has been revised based on your suggestions as follows.

"As I mentioned in the background section, the Medical Care Act revision draft took effective measures for geographical maldistribution. It requires hospital directors to procure those with work experience in low physician density areas for a certain period.[11] This study's results would support this policy's effects."

・ p 7　Is the category 'other' appropriate? I assume this category would include physicians working at public health centers, industrial physicians, as well as unemployed physicians. For example, physicians working at public health centers may tend to remain at the same region, whereas unemployed physicians including mothers who have delivered their child may easily move with their husbands in Japan. A mixed-up category may affect the results.

Response: Thank you for this suggestion; the limitations section has been revised accordingly as follows.

"the "other" physician category includes public health centers, industrial physicians, and unemployed physicians. A heterogenous category may affect the results."

・ Table 3: Please clarify the name of the categories '0-14/ 15-29/ 30-44/≧45' .

Response: Thank you for noting this point. Table 3 has been revised.

Reviewer: 2

For international readers who may not know much about Japan's healthcare system, a little more background information would be helpful in the introduction to understand policy changes discussed. For example, I wasn't completely clear what "new residency training program" meant – was residency training not previously required and medical schools would send graduates directly to practice or would they previously assign graduates to residency programs? And to clarify, for the system to place

students into state health care – what % of health care is state health care in Japan, does 18% successful mean 18% ended up working in state health care, and I wasn't sure what the prefectural scholarships implications were, except in context it seemed counter to the goal of state health care recruitment? As a reader, I would love to understand these issues a little better.

Response: Thank you for these suggestions. We have revised the background section accordingly as follows.

"Residency training was not previously mandatory, and medical schools would send graduates directly to practice."

"Second, a system for selecting students was developed with the primary objective of recruiting physicians, mainly in rural areas. This system included 1,674 medical students in 2017, 18% of the capacity of medical schools."

I struggled a little in interpreting the results and tables. I think largely because sections of the methods could use more detail, including:

• For the physician density categories – are the SMA categories set based on 2016 (page 8)? It would help to clarify this in the previous paragraph that discusses how the categories are set. Also, how does this compare to the categories if set in 1996? It seems by setting it in 2016, if there is improvement in physician density over the 20 year study period for some SMAs – that then change categories as a result – then this will be missed using this method.

Response: Thank you for making this point. Physician density was set in 1996, 2006, and 2016 based on the MHLW physician density classifications in those years. The methods section has been revised according to your suggestions as follows.

"Based on the classification used by MHLW, the first group (urban) consists of areas with a population of at least 1 million or a population density of at least 2000 people/km2. The second group (intermediate) consists of areas with a population of at least 100,000 or a population density of at least 200 people/km2. The third group (rural) consists of areas that do not belong to the first or second groups. The municipality borders that were altered because of mergers were adjusted based on the borders used in 2016.

• For the physician migration, I assume each cohort includes only physicians who are present for the full 10 year time periods – for example, looking only at physicians present in 1996 and their aggregated change over the 10 year period. This section of the methods is very short and therefore, it took me some time to figure out and interpret the tables.

Response: Thank you for this comment. The methods and limitations sections have been revised based on your suggestions as follows.

"Regarding the number of physicians per SMA in 1996, 2006, and 2016, the top 33.3% were classified as areas with many physicians and the bottom 33.3% with fewer physicians based on the MHLW physician density classifications in 1996, 2006, and 2016."

I analyzed the whereabouts of physicians at two points, 1996 and 2006, or 2006 and 2016, and did not consider changes during the period.

6

• For the regressions (Table 4) it seems control variables are all based on the start of the time period? For example, 1996 age, workplace, physician density practice, type of institution, etc. Can this be clarified in the table. Also, in this case, shouldn't physician density "low" in 1996 or 2006, be included in the regression results?

Response: Thank you for pointing this out. Regarding the first comment, Table 4 has been revised accordingly. For the latter point, I identified the factors that predicted whether physicians would begin working in low physician density areas after 10 years from 1996 and 2006. Hence, for the two cohorts, logistic regression analysis was performed by excluding doctors who were already in the low physician density areas.

I found the negative relationship between rural practice and physician density interesting. However, I wonder if this relates to smaller populations in rural communities – driving the physician to population density up due to the denominator rather than the numerator. This might be worth exploring further since the study finds rural practice is negatively associated with working with low physician density areas and this might be interpreted as a less than desirable outcome.

Response: Thank you for the suggestion. The discussion section has been revised accordingly as follows.

"The results show that rural practice is negatively associated with low physician density areas. This is related to the smaller populations in rural communities, which drives the physician-to-population density up due to the denominator rather than the numerator. This might be worth exploring further."

I wasn't quite sure about the finding that increases in physicians working in high density areas was less than increases in low density areas. Based on table 1, in 2016 only 11.7% of all physician were working in low density areas vs. 12.1% in 1996. In rural areas, it goes from 7.7% in 1996 to 6.0% in 2016. This seems worth mentioning - understanding that underlying this is an overall increase in the physician workforce, so low density and rural areas may be better off as the author discusses due to diffusion to rural/low density areas. However, it also brings up an interesting question of the implications of increasing physician numbers in high density areas and whether this will be an area of concern in the future (for example, over-use). Is this a concern in Japan?

Response: Thank you for making this point; we have revised the discussion section based on your suggestions as follows.

"This study revealed that the increases in physicians per population in high-density areas were less than the increases in low-density areas, suggesting that the physician geographical imbalance has improved."

Please ensure that you have fully described the methodological limitations of the study in the strengths and limitations section.

Response: Thank you for this comment. The limitations section has been expanded as follows.

- This study only focused on correlations and was unable to determine causality. Future studies could use interviews and questionnaires to facilitate more comprehensive research for physician migration.

7

- The observation period is 20 years. The effects of various environmental changes, such as the global economic crisis, policy changes for physician maldistribution, and population aging, were not considered.

VERSION 2 – REVIEW

REVIEWER	Noriko Sasaki Kyoto University Graduate school of Medicine
REVIEW RETURNED	24-Sep-2020

GENERAL COMMENTS	I have no further comment to address.

REVIEWER	Candice Chen George Washington University
REVIEW RETURNED	25-Sep-2020

GENERAL COMMENTS	Overall, a comprehensive, thoughtful study of the physician workforce and migration over time. The data and methods are strong, and ongoing study and understanding of the workforce - particularly changes and migration over time - are important to track, particularly as new policies are implemented and additional policies are considered to ensure health care access, as described in the manuscript. A few specific comments: This sentence is unclear to me: "Despite these issues, a certain amount of evidence supports the policy." It's not clear why this is "Despite these issues" – What are the issues with the described policies? Should it be: Despite these policies, maldistribution persists? It's actually the US Office of Management and Budget – although the definition of rural isn't always consistent in the US either. Table 3 – This table is confusing. Can the row categories be clarified? What is "change in area of practice" and what is included in the "started working in low density area"? Which individuals are or are not included in these categories? What is the estimated annual retention rate? I did not understand the method of square rooting the biannual retention rate and would rather have the retention over 2 years reported. For Table 4 and related text (including the abstract), it would help to clarify the analysis examines "physicians in high and intermediate areas and factors related their movement to low physician density areas," if I'm understanding the methods correctly. It currently reads as if the analysis is examining overall practice decisions to work in low density areas when physicians who start in low density areas are explicitly excluded from the analysis. The conclusion that geographical imbalances are improving based on greater increases in physician per population in low density areas appears based on percentage improvements over time, but looking at Figure 1, it appears that the increase in the physician

	per population ratio for high density areas is greater than for low density areas (nearly 75 per 100,000 vs. less than 50 per 100,000). It's definitely arguable though that low density areas saw an increase in physicians per population over time. Along the same lines, for the migration, it might be noted that the absolute numbers of physicians moving from high to low areas is actually greater than the absolute number moving from low to high (for both 1996 to 2006 and 2006 to 2016). In the discussion, what does the sentence "The number of physicians across all categories (low, intermediate, high) tended to decrease" mean? Didn't the number of physicians increase over time, in all categories? The author acknowledges the low, intermediate, high physician density areas are set at three different points of time (1996, 2006, 2016), additional information on the number of SMAs that change classification between those time periods would help to understand if/how much of an issue this might be.

VERSION 2 – AUTHOR RESPONSE

Response to Reviewers Comments

Reviewer: 2

1. This sentence is unclear to me: "Despite these issues, a certain amount of evidence supports the policy." It's not clear why this is "Despite these issues" – What are the issues with the described policies? Should it be: Despite these policies, maldistribution persists?

Response: Thank you for this comment. The introduction section was revised based on your suggestions as follows.

"Despite these policies, maldistribution persists."

2. It's actually the US Office of Management and Budget – although the definition of rural isn't always consistent in the US either.

Response: Thank you for this comment. The methods section was revised based on your suggestions as follows.

"In Japan, as in the US Office of Management and Budget, the definition of rural is not always consistent."

3. Table 3 – This table is confusing. Can the row categories be clarified? What is "change in area of practice" and what is included in the "started working in low density area"? Which individuals are or are not included in these categories? What is the estimated annual retention rate? I did not understand the method of square rooting the biannual retention rate and would rather have the retention over 2 years reported.

9

Response: Thank you for this comment. The result section was revised based on your suggestions as follows.

"* Those who were engaged in a low physician density area at baseline and were still engaged in the low physician density area two years after the baseline. Those who did not respond two years after baseline are not counted.

** Those who were engaged in a low physician density area at baseline but were engaged in a high or intermediate physician density area two years after baseline. Those who did not respond two years after baseline are not counted.

*** Those who were engaged in a low physician density area at baseline and were engaged in a high or intermediate physician density area at the time of the survey conducted two years ago."

Regarding the calculation of retention rate, I stopped using the method of taking the square root of the biannual retention rate and posted the retention rate over the two periods reported. I have also corrected the related text and abstract description.

4. For Table 4 and related text (including the abstract), it would help to clarify the analysis examines "physicians in high and intermediate areas and factors related their movement to low physician density areas," if I'm understanding the methods correctly. It currently reads as if the analysis is examining overall practice decisions to work in low density areas when physicians who start in low density areas are explicitly excluded from the analysis.

Response: Thank you for this comment. The results and related text were revised based on your suggestions as follows.

Result

"Table 4 first shows the logistic regression results that served to identify the variables in 1996 that predicted physicians being engaged in high and intermediate areas and the factors related to their movement to low physician density areas in 2006."

5. The conclusion that geographical imbalances are improving based on greater increases in physician per population in low density areas appears based on percentage improvements over time, but looking at Figure 1, it appears that the increase in the physician per population ratio for high density areas is greater than for low density areas (nearly 75 per 100,000 vs. less than 50 per 100,000). It's definitely arguable though that low density areas saw an increase in physicians per population over time.

Response: Thank you for this comment. The discussion section was revised based on your suggestions as follows.

"This study revealed that the increases in physicians per population in high-density areas were less than the increases in low-density areas, suggesting that the physician geographical imbalance has improved based on percentage improvements over time, although the actual increase in the physician-per-population ratio for high-density areas is greater than that for low-density areas."

6. Along the same lines, for the migration, it might be noted that the absolute numbers of physicians moving from high to low areas is actually greater than the absolute number moving from low to high (for both 1996 to 2006 and 2006 to 2016).

Response: Thank you for this comment. The discussion section was revised based on your suggestions as follows.

"With regard to migration, the absolute number of physicians moving from high to low areas is actually greater than the absolute number moving from low to high areas (for both 1996 to 2006 and 2006 to 2016."

7. In the discussion, what does the sentence "The number of physicians across all categories (low, intermediate, high) tended to decrease" mean? Didn't the number of physicians increase over time, in all categories?

Response: Thank you for this comment. The discussion section was revised based on your suggestions as follows.

"The number of physicians across all categories (low, intermediate, high) tended to increase."

8. The author acknowledges the low, intermediate, high physician density areas are set at three different points of time (1996, 2006, 2016), additional information on the number of SMAs that change classification between those time periods would help to understand if/how much of an issue this might be.

Response: Thank you for this comment. The discussion section was revised based on your suggestions as follows.

"There were 51 SMAs (15%) whose classification changed during the period between 1996 and 2006 and 62 (18%) between 2006 and 2016, as shown in Table 5."

VERSION 3 – REVIEW

REVIEWER	Candice Chen George Washington University, USA
REVIEW RETURNED	30-Oct-2020

GENERAL COMMENTS	This version is clear and the research is important. Very minor note: The second set of columns in Table 5, I believe should be labeled as "Physician density in 2016" (not 2006).