PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Results of physician license examination and scholarship contract compliance by the graduates of regional quotas in Japanese medical schools: a nationwide cross-sectional survey
AUTHORS	Matsumoto, Masatoshi; Takeuchi, keisuke; Owaki, Tetsuhiro; Iguchi, Seitaro; Inoue, Kazuo; Kashima, Saori; Tazuma, Susumu; Maeda, Takahiro

VERSION 1 – REVIEW

REVIEWER	Yoshiharu Fukuda Teikyo University Graduate School of Public Health
REVIEW RETURNED	14-Sep-2017

GENERAL COMMENTS	The study evaluated the intermediate outcomes (passing rate of the National License Examination and the percentage of graduate who have not bought out the scholarship contract after graduation) of the subquota system and prefecture scholarship programmes. Since the shortage in rural areas and maldistribution of physicians is a serious social problem in Japan, this study showed the meaningful evidence for health policy. And the finding will provide important suggestions for other countries. The following comments will be helpful to revise the manuscript. Major point. 1. The passing rate of the subquota was compared with that of all medical school graduates. The rate differs among the schools and the rate of private medical school generally lower. Thus, the rate of the subquota should be compared with not only all medical school but such as public medical school. If the figure will be too busy, the results can be shown in the text only. Minor points 2. P3, L15: “3.25 years” could be “three years” 3. P5, L53: If the number of the medical school is specified, it should be described: “** among ** medical schools in Japan at 201*”. 4. P6, L27: The sentence of “The model of these ….” could be omitted. 5. P11, L11: “each group in each year” is correct? 6. P15, L6: The authors mentioned there is a substantial difference in the retention rates among prefecture. The difference, e.g., distribution, could be shown. 7. P15, L30: I did not understand “the risk of losing workers”. 8. P15, L35-53: Some possible reasons of lower retention rate in this study compared with in other countries will be useful for readers of the international journal.

	9. Figure 2: The legend could include explanation of “Subquota + scholarship”, “Subquota” and “Scholarship”. And figures could show N. 10. Figure 3: The scale of y axis in D is difference. The scale should be consistence for all graphs.

REVIEWER	John C. Hogenbirk Centre For Rural and Northern Health Research Laurentian University Sudbury, Ontario, Canada
REVIEW RETURNED	27-Sep-2017

GENERAL COMMENTS	Thank you for the opportunity to review “Results of physician license examination and scholarship contract compliance by the graduates of regional subquotas in Japanese medical schools”. The manuscript describes interim results from a longitudinal study of multiple cohorts of selected medical students from all of Japan’s medical schools. The study topic is important and the results described in the manuscript will be of interest to policy-makers in Japan and in other counties that use or are thinking of using similar selection and scholarship programs. The manuscript needs an intermediate amount of revision including additional information on the statistical approach and selective re-writing of the Results and Discussion sections. The following suggestions are offered for consideration by the authors. (1) “Subquota” requires clarification or replacement. Consider using “regional quota”, “prefecture quota” or similar phrase. Once defined in the manuscript, then use “quota” in the remaining text. In the context of the manuscript, quota would indicate that a minimum number of medical school seats are set aside for students from the prefecture. If it is not the case, then a better word is needed. (2) The methods section requires additional clarity on the statistical analyses. A first step would be to expand the fourth column (Subjects**) of TABLE 1 to give the number of subjects in quota with scholarship and separately for scholarship alone. (Students in quota without scholarship are provided in the last column: Subjects***). Numbers should also be provided for each bar in FIGURE 2. a.Fisher’s exact test of the passing rates. It is not clear whether this is used as a test ofindependence, as in the table below. Or was this a goodness-of-fit test using all medicalstudents (uncorrected) as the population? A test of independence seems moreappropriate. b.Kaplan-Meier test for retention. It is not clear if or how censored data were treated in the analysis conducted in support of Figure 3 D. When combining graduates from the years 2014 to 2017, those graduates in 2017 would be censored beyond the 0.25 year interval—we would not know about their retention rates beyond 0.25 years and so it is inappropriate to map out their retention up to 3.25 years without recognizing that these outcomes are censored at 0.25 years. Similar arguments apply to 2015 and 2016 graduates. It is also possible to examine retention up to 2.25 years, but only for the 2014 and 2015 cohorts. 2 (3) The choice of all medical graduates as the group for comparison is easily justified. However, without knowing how many graduates are in this group, it is not clear whether the proper group for comparison should be (a) All Graduates or (b) All Graduates minus graduates in the other groups.

	Perhaps the analysis of pass/fail should consider a 2x4 table (pass/fail X quota+ scholarship/quota/scholarship/all other grads) rather than three 2x2 tables. In SPSS one can select exact tests (or Monte Carlo option) for most contingency tables (Mehta and Patel. 2001. IBM SPSS Exact Tests). This becomes Fisher's exact test in a 2x2 table. In the 2x4 table, all groups are mutually exclusive. In SPSS, select adjusted residuals to help identify where the differences occur in tables bigger than 2x2 (e.g., Agresti 2002. Categorical data analysis. Wiley). (4) The authors present overall results, yet allude to year-by-year differences. It would be useful to include some analyses, perhaps as an online appendix, which compares years, prefectures and schools. (5) I have some suggestions and questions for the consideration of the authors in regards to the text of the manuscript. These suggestions and questions appear below. Throughout the manuscript: replace "systems" with "programs". Please clarify if you are referring to one national policy or several national policies or 1 national and several prefecture-level policies as this affects subject verb agreement in many sentences. For example, on page 3, line 23, "revision" (with or without s?). The heading on page 5, line 43 suggests 1 policy, but elsewhere the text suggests several policies. Page 2, line 15-18: replace "mostly as a chiikiwaku, a regional subquota" with "mostly as a chiikiwaku, entrants filling a regional quota" to keep "entrants" as the subject of the sentence. Page 2, line 35: additional clarity is needed here to distinguish between survey recipients and study subjects, consider "cross-sectional survey to all prefectural governments and medical schools every year from 2014 to 2017 to obtain data on medical students." Page 3, line 6 (and elsewhere): check Journal style to determine if 97.8-100% should be 97.8%–100% (add %, use en dash). Add sample sizes, p-values (test used) when describing significant differences. Page 4, line 32 (and elsewhere): replace "compared to" with "compared with". Page 4, lines 41-47: consider replacing this bullet with "This study is part of an ongoing national wide cohort study, which will also examine the geographic location of quota graduates and scholarship recipients." Page 5, line 24: it is not the disparity in the number of physicians per se but the disparity in the proportional allocation. Consider "many studies reveal that the disparity in proportional representation of urban and rural physicians has persisted." 3 Page 5, lines 32-35: consider revising as "Furthermore, additional specially training for physicians after residency training, scheduled to start in 2018, has caused concern among health care professionals…" Page 5, line 52: move the sentence from page 6, line 27 to appear here. "… Regional quota of a medical school. The model of these quota was Jichi Medical University. The quota has spread .." Page 6, line 32: consider "Apart from the scholarship coupled with quota admission…" Page 7 lines 15-21: please clarify whether interviews are required or optional as the current phrasing seems to suggest both. This has implications for other sections. Page 7, lines 26-29: the subject of the sentence seems to be switching between medical schools, students and the prefecture. Please pick one (e.g., students) and re-write.

	Page 7, lines 41 to 60 and on to page 8, line 6: consider replacing all of these lines with ”Because of these differences, the general population and medical educators are concerned about the academic performance of quota entrants. In addition, there is a concern that many quota entrants may buy out their contract to avoid practising in the prefecture and/or in rural areas.” The Study objectives will then logically follow. Page 8, line 21: consider ”we also examine the percentages…” Page 8, line 23: if I understand correctly, graduates who have not bought out their scholarship may or may not have completed their contractual work. If this is true, then delete the last clause that says “and thus completed their contractual work.” Page 8, lines 24-29: consider revising the last sentence to read “results can be used to inform political decision-makers for future revision of the quota and scholarship programs.” Page 8, line 41: omit the sentence and move reference number 14 to the following sentence to be added after the phrase “nationwide cohort study” page 8, line 55 and on to page 9, line 6: consider ”The study includes three groups of subjects:…” page 9, lines 9-12: consider “Data on the former …. and that on the last group were collected from medical schools.” Page 9, line 32: consider “to ask which …” page 10, line 18 consider “In this study, a quota student is …” Page 10 lines 43 to 47: check the years listed as it is my understanding that data were added from years 2016 and 2017 (not 2015 and 2016). Page 10 lines 49 to 53: please clarify how Fisher’s exact test was used, in particular which data were used in comparison and whether or not the correction was made in the comparison group for the number of quota students on scholarship, etc. 4 Page 11, line 15: please clarify if and how censored data were considered in the Kaplan-Meier analysis shown in figure 3D. Page 11 and page 12: The first few sentences at the start of each sub-section within the Results can be omitted. These sentences simply refer the reader to particular table or figure without giving a summary of the results and are not needed. It should suffice to list the appropriate table or figure with the sentence that first describes the results displayed in that table or figure. In the Response Rate section, this would be the sentence that starts “Almost all prefectures and medical schools…” Pages 11-13, Results: please ensure that the sample size is included in the text, tables and figures. Pages 11-17 Results & Discussion: In the Result and Discussion sections, my comments pertain to the results as presented and may need to be changed if the authors make modifications to the statistical approach. Some of the percentages, particularly those reported for retention in years 2014-2017 may need to be revised. Page 13, lines 44 to 47: please clarify or correct the sentence because the passing rates were only highest for this group in 2017. Page 14 lines 23 to 50: This paragraph makes many points and the reader’s understanding may benefit if the points are made more distinct and described in separate paragraphs. More is needed to bring out the relevance of the information in these sentences. The last 3 sentences, beginning with “whether in a subquota or not…” seem to be a miscellany that are (as yet) not strongly related to the study outcomes.

	Page 15, line 55: By “representative value” do the authors mean “overall value” or are they making the argument that the value is truly representative? Page 15, line 6: what is the proof that there is substantial differences among prefectures? Page 15, lines 9 to 32: while there is a large difference in the amount of the scholarship, there may be other reasons for why there is differences in retention rate between Jichi students and students from other medical schools. Given that Jichi has a specific mandate to train physicians for rural practice, there may be differences in applicants and who is selected, differences in the training experiences offered to students, and there clearly is a known difference in the focus of the school. In addition, there is a difference in the years used to make the comparison: pre-2009 for Jichi vs post 2014 for the current study. All of these differences could contribute to differences in retention. Page 15 line 41: consider ”programs was reported to be 71%.” Page 15 lines 44 to 53: consider replacing the last two sentences with “while international data are favourable, it is not known if the relationships would hold for Japan. Thus further study is needed.” Page 16, lines 6-37: These are important points and this paragraph needs some re-writing to facilitate reader understanding. 5 It would be interesting to know whether there were appreciable differences in the marks of the different groups with respect to the national licensing exam and not just in the pass/fail rates. Do quota graduates have higher or lower marks than all other graduates? Page 16 lines 35 to 38: consider “is the most important outcome of these programs, and will be assessed in the near future, based on the data from the ongoing cohort study.” Page 17 line 18: consider ”regional quota of entrants to medical schools,…” Page 18 line 44: consider “to share aggregated data with stakeholders and researchers.” Table 1. Expand the fourth column (Subjects**) of TABLE 1 to give the number of subjects in quota with scholarship and separately for scholarship alone. (Students in quota without scholarship are provided in the last column: Subjects***). Table 1. If it is at all possible, please provide an estimate of the number of students (subjects) that were missed when the few prefectures did not respond to the survey. Figure 1 is not needed as this information is easily and clearly defined in the text. Figure 2: add sample size to the bars STROBE checklist. Item 12a: more information is needed on the tests as described above. Item 12b: differences among subgroups or interactions were alluded to in the text, but not examined. In the current version of the manuscript, the response to this checklist item would be “not applicable”. However, it would be of interest to see the authors explicitly address the issue of interactions and differences among years, prefectures or schools. Item 14a: Information on subjects is missing. Consider adding mean age, % females (or % males), and other selected demographic characteristics. If these demographics do not differ substantially from year to year, among groups (i.e., quota + scholarship, quota only, scholarship only, all medical students) among prefectures or among schools, then a simple description can be added to the text.

	If there are significant and important differences, then consider adding a table on student demographics. Item 14b: need to add an estimate of how many students were missed because of the few prefectures and schools that did not reply in every year. Item 17: not (yet) applicable.

VERSION 1 – AUTHOR RESPONSE

Comments to Reviewer 1
Thank you for reviewing our manuscript and for offering such constructive feedback. The updates we made in relation to your comments have greatly improved the quality of our paper. The changes in the manuscript appear in red. The responses to each of your comments are as follows:

Major point.
1. The passing rate of the subquota was compared with that of all medical school graduates. The rate differs among the schools and the rate of private medical school generally lower. Thus, the rate of the subquota should be compared with not only all medical school but such as public medical school. If the figure will be too busy, the results can be shown in the text only.

Res: The suggestion is important. In this revised version we added the each-year passing rate of public medical schools and compared it with the rate of quota with scholarship in the “Passing rates of the National Physician License Examination” section of the Results. As the passing rate of private medical schools is usually lower than the rate of all medical schools, the comparison between quota and private schools was omitted. (Please note that we used “quota” instead of “subquota” in this revision).

Minor points
2. P3, L15: “3.25 years” could be “three years”

Res: The part has been changed accordingly.

3. P5, L53: If the number of the medical school is specified, it should be described: “** among ** medical schools in Japan at 201*”.

Res: The requested information was added.

4. P6, L27: The sentence of “The model of these ….” could be omitted.

Res: Yes, it can be omitted. But the comparison between quota and Jichi Medical University is one of the main themes in Discussion. So we consider it would be better to mention something about Jichi in Introduction. Quota is certainly not a copy of Jichi, but it still is more or less similar to Jichi and the concept of scholarship and obligation period of quotas derives from Jichi. We thus remained this sentence. According to a comment of Reviewer 2, we moved this sentence to an earlier part of this paragraph.

5. P11, L11: “each group in each year” is correct?

Res: We are sorry for the too-short description. Yes, it is correct itself, but we increased our explanation on calculation for retention rate as follows;
“The retention rate of each group was calculated in each cohort of graduation year. Also retention rate of all those who graduated between 2014 and 2017 was calculated with Kaplan-Meier survival analysis in which subjects with various observation periods can be analysed.”

6. P15, L6: The authors mentioned there is a substantial difference in the retention rates among prefecture. The difference, e.g., distribution, could be shown.

Res: It is a very important point. We added information on the range of retention rate among prefectures in the “Retention rates for contractual work (non-buying-out rates)” section of the Results. Please note that the name of prefecture and its retention rate cannot be shown because we are not permitted to do so. Also please note that license passing rate cannot be calculated on a school name basis because most of our data are from prefectures, not schools.

7. P15, L30: I did not understand “the risk of losing workers”.

Res: This part was certainly difficult to understand. So it was deleted.

8. P15, L35-53: Some possible reasons of lower retention rate in this study compared with in other countries will be useful for readers of the international journal.

Res: Actually the retention rates of Japanese quotas and scholarship programmes (92% and 90%) were higher than those of foreign programmes (71% and 67%). But it is quite difficult to compare them because observation periods of the studies and social background are so different. Thus we limited our description to simply mentioning the reported rates of the programmes.

9. Figure 2: The legend could include explanation of “Subquota + scholarship”, “Subquota” and “Scholarship”. And figures could show N.

Res: We agree. The explanation was added to the footnote of Figure 1 in which these categories first appeared in the manuscript.

10. Figure 3: The scale of y axis in D is difference. The scale should be consistence for all graphs.

Res: It was our mistake. The scale was amended. Thank you for suggesting. Unrelated to your comments, please note that, according to comments of the other reviewer, the comparison group in Figure 2 was changed from “all graduates” to “others (all graduates excluding quota and scholarship ones)”. So the results (such as p values) slightly changed.

Comments to Reviewer: 2
Thank you for reviewing our manuscript and for offering such constructive feedback. The updates we made in relation to your comments have greatly improved the quality of our paper. The changes in the manuscript appear in red. The responses to each of your comments are as follows:

(1) “Subquota” requires clarification or replacement. Consider using “regional quota”, “prefecture quota” or similar phrase. Once defined in the manuscript, then use “quota” in the remaining text. In the context of the manuscript, quota would indicate that a minimum number of medical school seats are set aside for students from the prefecture. If it is not the case, then a better word is needed.

Res: Every “subquota” has been changed to “regional quota” or “quota”.

(2) The methods section requires additional clarity on the statistical analyses. A first step would be to expand the fourth column (Subjects**) of TABLE 1 to give the number of subjects in quota with scholarship and separately for scholarship alone. (Students in quota without scholarship are provided in the last column: Subjects***). Numbers should also be provided for each bar in FIGURE 2.

Res: Table 1 has been changed accordingly. Numbers were added to Figure 2.

a. Fisher's exact test of the passing rates. It is not clear whether this is used as a test of independence, as in the table below. Or was this a goodness-of-fit test using all medical students (uncorrected) as the population? A test of independence seems more appropriate.

Res: You are right. In the previous version we employed Fisher's exact test for all the tests shown in Figure 2. In this revised version, we used chi-square test of independence, as you suggested, for all the tests except for one in 2014 which Fisher's test fits better. As we have less than 5 subjects of quota with scholarship who failed the National License Examination in 2014, this part was examined with Fisher's test. Also according to your comment, we changed the comparison group from "all medical graduates" to "others" which excluded graduates of quota with scholarship, quota alone and scholarship alone from all medical graduates. This change has made the difference in passing rate between quota and the comparison group more obvious than before. Thank you for your suggestion.

b. Kaplan-Meier test for retention. It is not clear if or how censored data were treated in the analysis conducted in support of Figure 3 D. When combining graduates from the years 2014 to 2017, those graduates in 2017 would be censored beyond the 0.25 year interval—we would not know about their retention rates beyond 0.25 years and so it is inappropriate to map out their retention up to 3.25 years without recognizing that these outcomes are censored at 0.25 years. Similar arguments apply to 2015 and 2016 graduates. It is also possible to examine retention up to 2.25 years, but only for the 2014 and 2015 cohorts.

Res: This is an important point. We treated data on graduates in 2017, 2016, 2015 as censored data in the Kaplan-Meier's survival analysis. Total number of subjects (including censored ones) at each time point is shown in the following table (for reviewer only). We increased explanation on this point in the second paragraph of "Statistical analyses" section of Results.

Table for reviewer
type=1 (Quota with scholarship)
time, N with risk, N with event, survival std.err, lower 95% CI, upper 95% CI
0.25 1556 24 0.985 0.00312 0.978 0.991
1.25 934 7 0.977 0.00416 0.969 0.985
2.25 441 10 0.955 0.00803 0.939 0.971
3.25 144 5 0.922 0.01651 0.89 0.955

type=2 (Scholarship alone)
time n.risk n.event survival std.err lower 95% CI upper 95% CI
0.25 1429 11 0.992 0.00231 0.988 0.997
1.25 1047 21 0.972 0.00486 0.963 0.982
2.25 702 29 0.932 0.00866 0.915 0.949
3.25 307 11 0.899 0.01294 0.874 0.925

(3) The choice of all medical graduates as the group for comparison is easily justified. However, without knowing how many graduates are in this group, it is not clear whether the proper group for comparison should be (a) All Graduates or (b) All Graduates minus graduates in the other groups. Perhaps the analysis of pass/fail should consider a 2x4 table (pass/fail X quota+scholarship/quota/scholarship/all other grads) rather than three 2x2 tables. In SPSS one can select exact tests (or Monte Carlo option) for most contingency tables (Mehta and Patel. 2001. IBM SPSS Exact Tests). This becomes Fisher's exact test in a 2x2 table. In the 2x4 table, all groups are mutually exclusive. In SPSS, select adjusted residuals to help identify where the differences occur in tables bigger than 2x2 (e.g., Agresti 2002. Categorical data analysis. Wiley).

Res: As shown above, we changed the comparison group from all graduates (including quota and scholarship ones) to others (excluding quota and scholarship) (Figure 2).
Also according to your comment, we conducted and compared the following two tests;
1: 2x2 table analysis that we originally employed
2: 2x4 table analysis
The results of the second one are shown in a separate Figure for Reviewer attached to this letter. P value not in parenthesis is for difference in passing rate among the four groups, and p value in parenthesis is the one based on the adjusted residual of each group. Both of the tests showed the similar thing, i.e. the rate of quota with scholarship was higher and that of the "others" was lower. In this paper, however, we want to focus on the difference between quota/scholarship and "others" because this is the point educators and policy makers in Japan want to know. The difference is more clearly and easily shown with the first test. So we retained the first test in this revised version even though we understand the merits of the second one.

(4) The authors present overall results, yet allude to year-by-year differences. It would be useful to include some analyses, perhaps as an online appendix, which compares years, prefectures and schools.

Res: Inter-prefecture and inter-school comparisons are meaningful. But we are not permitted, when obtaining data, to show the prefecture-level results. So in this revised version we showed the rage of retention rate among prefectures (2nd paragraph of "Passing rates of the National Physician License Examination" of the Results). Inter-school comparison cannot be conducted because we don't have school-based data except for quota without scholarship.

(5) I have some suggestions and questions for the consideration of the authors in regards to the text of the manuscript. These suggestions and questions appear below.
Throughout the manuscript: replace "systems" with "programs".

Res: Replaced.

Please clarify if you are referring to one national policy or several national policies or 1 national and several prefecture-level policies as this affects subject verb agreement in many sentences. For example, on page 3, line 23, "revision" (with or without s?). The heading on page 5, line 43 suggests 1 policy, but elsewhere the text suggests several policies.

Res: It is a "revision." Quota and scholarship are based in a national policy to increase rural doctors, and prefecture governments provide budgets for the scholarship according to the national policy. We tried to make this point clear in this revised version changing "policies" to "policy" and added a sentence in "Chiikiwaku (regional quota) and prefecture scholarship as a national policy" section of Introduction.

Page 2, line 15-18: replace “mostly as a chiikiwaku, a regional subquota” with “mostly as a chiikiwaku, entrants filling a regional quota” to keep “entrants” as the subject of the sentence.

Res: Replaced.

Page 2, line 35: additional clarity is needed here to distinguish between survey recipients and study subjects, consider “cross-sectional survey to all prefectural governments and medical schools every year from 2014 to 2017 to obtain data on medical students.”

Res: Changed accordingly.

Page 3, line 6 (and elsewhere): check Journal style to determine if 97.8-100% should be 97.8%–100% (add %, use en dash). Add sample sizes, p-values (test used) when describing significant differences.

Res: We checked the author instruction of BMJ Open but did not find any style with regard to this point. So we employed your advice. Because of the limitation of word number the journal imposes on us, we cannot add p values and sample size in the Abstract. Also we checked recent articles published in BMJ Open and found one in which such values were not included in Abstract. So the values were not added in this revised version.

Page 4, line 32 (and elsewhere): replace “compared to” with “compared with”.

Res: Replaced.

Page 4, lines 41-47: consider replacing this bullet with “This study is part of an ongoing national wide cohort study, which will also examine the geographic location of quota graduates and scholarship recipients.”

Res: Replaced.

Page 5, line 24: it is not the disparity in the number of physicians per se but the disparity in the proportional allocation. Consider ”many studies reveal that the disparity in proportional representation of urban and rural physicians has persisted.”

Res: Changed accordingly.

Page 5, lines 32-35: consider revising as “Furthermore, additional specially training for physicians after residency training, scheduled to start in 2018, has caused concern among health care professionals…”

Res: Changed accordingly.

Page 5, line 52: move the sentence from page 6, line 27 to appear here. “… Regional quota of a medical school. The model of these quota was Jichi Medical University. The quota has spread ..”

Res: Changed accordingly.

Page 6, line 32: consider “Apart from the scholarship coupled with quota admission…”

Res: Changed accordingly.

Page 7 lines 15-21: please clarify whether interviews are required or optional as the current phrasing seems to suggest both. This has implications for other sections.

Res: Clarified.

Page 7, lines 26-29: the subject of the sentence seems to be switching between medical schools, students and the prefecture. Please pick one (e.g., students) and re-write.

Res: Changed accordingly.

Page 7, lines 41 to 60 and on to page 8, line 6: consider replacing all of these lines with ”Because of these differences, the general population and medical educators are concerned about the academic performance of quota entrants. In addition, there is a concern that many quota entrants may buy out their contract to avoid practising in the prefecture and/or in rural areas.” The Study objectives will then logically follow.

Res: Replaced.

Page 8, line 21: consider ”we also examine the percentages…”

Res: Replaced.

Page 8, line 23: if I understand correctly, graduates who have not bought out their scholarship may or may not have completed their contractual work. If this is true, then delete the last clause that says “and thus completed their contractual work.”

Res: Your understanding is correct. Deleted.

Page 8, lines 24-29: consider revising the last sentence to read “results can be used to inform political decision-makers for future revision of the quota and scholarship programs.”

Res: Replaced.

Page 8, line 41: omit the sentence and move reference number 14 to the following sentence to be added after the phrase “nationwide cohort study”

Res: Changed accordingly. Please note that the reference 14 is now reference 15.

page 8, line 55 and on to page 9, line 6: consider ”The study includes three groups of subjects:…”

Res: Changed.

page 9, lines 9-12: consider “Data on the former …. and that on the last group were collected from medical schools.”

Res: Changed.

Page 9, line 32: consider “to ask which …”

Res: Changed.

page 10, line 18 consider “In this study, a quota student is …”

Res: Changed.

Page 10 lines 43 to 47: check the years listed as it is my understanding that data were added from years 2016 and 2017 (not 2015 and 2016).

Res: You are right. Corrected.

Page 10 lines 49 to 53: please clarify how Fisher’s exact test was used, in particular which data were used in comparison and whether or not the correction was made in the comparison group for the number of quota students on scholarship, etc.

Res: Clarified in the first paragraph of “Statistical analyses” section of Methods.

Page 11, line 15: please clarify if and how censored data were considered in the Kaplan-Meier analysis shown in figure 3D.

Res: Clarified in the above response. Also clarified in the second paragraph of “Statistical analyses” section of Methods.

Page 11 and page 12: The first few sentences at the start of each sub-section within the Results can be omitted. These sentences simply refer the reader to particular table or figure without giving a summary of the results and are not needed. It should suffice to list the appropriate table or figure with the sentence that first describes the results displayed in that table or figure. In the Response Rate section, this would be the sentence that starts “Almost all prefectures and medical schools…”

Res: These sentences were deleted.

Pages 11-13, Results: please ensure that the sample size is included in the text, tables and figures.

Res: All the sample sizes are now included in the table and figures. Because sample size is different between groups and between years (and thus the information overwhelms readers), we did not include them in the text.

Pages 11-17 Results & Discussion: In the Result and Discussion sections, my comments pertain to the results as presented and may need to be changed if the authors make modifications to the statistical approach. Some of the percentages, particularly those reported for retention in years 2014-2017 may need to be revised.

Res: Please refer to above responses.

Page 13, lines 44 to 47: please clarify or correct the sentence because the passing rates were only highest for this group in 2017.

Res: That part was certainly misleading, and thus was deleted.

Page 14 lines 23 to 50: This paragraph makes many points and the reader’s understanding may benefit if the points are made more distinct and described in separate paragraphs. More is needed to bring out the relevance of the information in these sentences. The last 3 sentences, beginning with

“whether in a subquota or not…” seem to be a miscellany that are (as yet) not strongly related to the study outcomes.

Res: The miscellany was deleted. And then the paragraph was rewritten.

Page 15, line 55: By “representative value” do the authors mean “overall value” or are they making the argument that the value is truly representative?

Res: Yes, it means overall value. The word was replaced.

Page 15, line 6: what is the proof that there is substantial differences among prefectures?

Res: We added the data on the difference in the second paragraph of “Passing rates of the National Physician License Examination” section of Results.

Page 15, lines 9 to 32: while there is a large difference in the amount of the scholarship, there may be other reasons for why there is differences in retention rate between Jichi students and students from other medical schools. Given that Jichi has a specific mandate to train physicians for rural practice, there may be differences in applicants and who is selected, differences in the training experiences offered to students, and there clearly is a known difference in the focus of the school. In addition, there is a difference in the years used to make the comparison: pre-2009 for Jichi vs post 2014 for the current study. All of these differences could contribute to differences in retention.

Res: You are right. The reasons you mentioned were added.

Page 15 line 41: consider ”programs was reported to be 71%.”

Res: Changed.

Page 15 lines 44 to 53: consider replacing the last two sentences with “while international data are favourable, it is not known if the relationships would hold for Japan. Thus further study is needed.”

Res: Replaced.

Page 16, lines 6-37: These are important points and this paragraph needs some re-writing to facilitate reader understanding. It would be interesting to know whether there were appreciable differences in the marks of the different groups with respect to the national licensing exam and not just in the pass/fail rates. Do quota graduates have higher or lower marks than all other graduates?

Res: It would be great if we knew the marks of the exam, but we don’t have such data.

Page 16 lines 35 to 38: consider “is the most important outcome of these programs, and will be assessed in the near future, based on the data from the ongoing cohort study.”

Res: Changed.

Page 17 line 18: consider ”regional quota of entrants to medical schools,…”

Res: Changed.

Page 18 line 44: consider “to share aggregated data with stakeholders and researchers.”

Res: Changed.

Table 1. Expand the fourth column (Subjects**) of TABLE 1 to give the number of subjects in quota with scholarship and separately for scholarship alone. (Students in quota without scholarship are provided in the last column: Subjects***).

Res: Expanded accordingly.

Table 1. If it is at all possible, please provide an estimate of the number of students (subjects) that were missed when the few prefectures did not respond to the survey.

Res: It is impossible unfortunately. We cannot know how many subjects were missed when a prefecture didn't respond. We suppose the proportion of the missed subjects among all potential subjects was very small though.

Figure 1 is not needed as this information is easily and clearly defined in the text.

Res: We consider Figure 1 is needed by readers, particularly those who didn't read our previous paper and don't know the basic structure of our cohort study. So we retained this figure in the revised version.

Figure 2: add sample size to the bars

Res: Added.

STROBE checklist.
Item 12a: more information is needed on the tests as described above.

Res: Please refer to the above responses.

Item 12b: differences among subgroups or interactions were alluded to in the text, but not examined. In the current version of the manuscript, the response to this checklist item would be "not applicable". However, it would be of interest to see the authors explicitly address the issue of interactions and differences among years, prefectures or schools.

Res: As mentioned above inter-prefecture difference in retention rate was added in the text. Other differences and interactions cannot be analysed in this study. The answer was changed to "not applicable".

Item 14a: Information on subjects is missing. Consider adding mean age, % females (or % males), and other selected demographic characteristics. If these demographics do not differ substantially from year to year, among groups (i.e., quota + scholarship, quota only, scholarship only, all medical students) among prefectures or among schools, then a simple description can be added to the text. If there are significant and important differences, then consider adding a table on student demographics.

Res: We don't have information on demographics of subjects unfortunately. In the on-going cohort study (not in this cross-sectional study) we will have demographic and some other characteristics of subjects, which will be shown in the future paper.

Item 14b: need to add an estimate of how many students were missed because of the few prefectures and schools that did not reply in every year.

Res: As mentioned above it is impossible to know how many were missed.

Item 17: not (yet) applicable.

Res: Changed to “not applicable”.

VERSION 2 – REVIEW

REVIEWER	Yoshiharu Fukuda Teikyo University Graduate School of Public Health
REVIEW RETURNED	30-Oct-2017

GENERAL COMMENTS	The manuscript has been revised according to our comments.

REVIEWER	John C. Hogenbirk Centre for Rural and Northern Health Research Laurentian University Canada
REVIEW RETURNED	09-Nov-2017

GENERAL COMMENTS	The revised version reads well, the authors have addressed previous concerns and have added material to address anticipated questions.