Author’s response to reviews

Title: Newborn screening for congenital adrenal hyperplasia in Tokyo, Japan from 1989 to 2013: a retrospective population-based study

Authors:

Atsumi Tsuji (ahchanxt@gmail.com)
Kaoru Konishi (thsa-screening@mrj.biglobe.ne.jp)
Satomi Hashimoto (thsa-screening@mrj.biglobe.ne.jp)
Akira Anazawa (thsa-screening@mrj.biglobe.ne.jp)
Toshikazu Onishi (toshikazuonishi@gmail.com)
Makoto Ono (mono-endo@umin.ac.jp)
Tomohiro Morio (tmorio.ped@tmd.ac.jp)
Teruo Kitagawa (thsa-screening@mrj.biglobe.ne.jp)
Kenichi Kashimada (kkashimada.ped@tmd.ac.jp)

Version: 6 Date: 24 July 2015

Author’s response to reviews: see over
Dear Nawsheen Boodhun

[MS: 1675870775151494]

Thank you for giving us the opportunity to revise our manuscript entitled, “Newborn screening for congenital adrenal hyperplasia in Tokyo, Japan from 1989 to 2013: a retrospective population-based study” [MS: 1675870775151494]. We have corrected our manuscript and written a point-by-point response to the referees’ comments.

During revising the manuscript, we identified a mistake for our interpretation of the data. For this point, we sincerely apologize, and please allow us to correct the data.

<Corrected point>
Total number of neonates who were tested by our screening program from 1989 to 2013 was 2,105,108. Because of this correction, we revised the number (P2L5, P7L10, P19 Table2a) and the incidence of CAH in Tokyo (P7L14, P8L16, P10L3, P19 Table2a, P23 Table4).

We believe the correction of the data will not affect the storyline of our manuscript, and I hope you will find the manuscript suitable for publication.

We are looking forward to corresponding with you further.

With best wishes

Kenichi Kashimada, PhD, MD

Assistant Professor of
Department of Pediatrics and Developmental Biology
Tokyo Medical and Dental University
1-5-45, Yushima, Bunkyo-ku, Tokyo,
Japan
Tel. +81 3 5803 5249
Fax +81 3 5803 5248
Email kkashimada.ped@tmd.ac.jp
First of all, during revising the manuscript, we identified a mistake for our interpretation of the data. For this point, we sincerely apologize, and please allow us to correct the data.

Total number of neonates who were tested by our screening program from 1989 to 2013 was 2,105,108. Because of this correction, we revised the number (P2L5, P7L10, P19 Table2a) and the incidence of CAH in Tokyo(P7L14, P8L16, P10L3, P19 Table2a, P23 Table4).

We believe the correction of the data will not affect the storyline of our manuscript.

- Discretionary Revisions
These are recommendations for improvement which the author can choose to ignore. For example clarifications, data that would be useful but not essential. : Comment 1

1. Title: It’s too long; it should be shortened. Suggestion: Newborn screening for Congenital Adrenal Hyperplasia in Tokyo, Japan from 1989 to 2013: a retrospective population based study. (Discretionary Revision)

We agree with you. According to your suggestion, we changed the title.

3-1. Methods: This section could be improved; some criteria aren’t adequately informed. Please describe the method for correcting body weight at the time of the repeated test (section Measurement of 17-OHP and criteria: 1st paragraph) and, the criteria used for diagnosis of the patients, including the type of CAH.
The algorithm and criteria of the screening is slightly complicated, and in order to clarify the point of them, we changed the description of the text (P5L22~P6L2), and replaced Table1 and Figure 1 with the new one..

3-2. Were the patients submitted to molecular testing? (section Follow-up Survey: 3rd paragraph). (Minor Essential Revision)
Genetic tests were not performed for all patients, and we added a description as below (P5L7~L8)“Diagnosis of CAH was based on the endocrinological data and physical findings[12]. Genetic tests were not carried out in all cases.”(P7L14-L15)

5. Results: The manuscript adhere to the relevant standards for reporting and data deposition; the tables and figures are clear and support the text. Table 3c could be removed, as discussed below. Despite this it’s not clear why the four SW patients with negative 17-OHP results had the tests repeated. It should be clarified (section Clinical details of CAH patients identified by the screening: 5th paragraph). (Minor Essential Revision)

The algorithm of our screening was slightly complicated, and Figure 3c is difficult to understand. In order to clarify the point, we changed the description and the figure 1 of the screening algorithm, as mentioned above.

Further, we also changed the description of the text that explains Table 3b and 3c (P8L25~P9L6).

6-1. The discussion is extensive, and the authors present some study limitations. However one point which brings concern is the fact that NC21OHD clinical features could not be recognized until adulthood. As the authors’ suggest it is also inappropriate to predict the form of CAH according to the value of 17-OHP. According to these points, the incidence reported for NC21OHD patients identified by the screening could be lower than the real one, and this point should be discussed and clarified (4th paragraph).
We agree with the reviewer’s suggestion, and according to it, we re-wrote the 5th paragraph of Discussion (we have added a new paragraph in Discussion and formerly 4th paragraph is now the 5th paragraph). (P11L20-P12L3)

6-2. Also, the reasons for the normal 17-OHP levels of the four SW patients who were negative on the first test should be discussed, false negative patients (2th paragraph). (Minor Essential Revision)

Actually, the 17-OHP levels on the first tests of the four SW patients were slightly elevated, requiring retest in our algorithm.

In order to clarify it, we replaced Figure1 with new one, and changed and simplified an explanation of the results (P8L25~P9L3).

“On the first test, most SW patients (94.5%) showed remarkably elevated levels of 17-OHP, and were referred to hospitals (Table 3c). While, four SW patients (No. 53, 84, 99, 101) showed mildly elevated levels of 17-OHP on the fist test (Table 3b) and required the repeated tests. Those suggest that mildly elevated 17-OHP does not exclude the possibilities of classical 21-OHD.” (P8L25-P9L3)
First of all, during revising the manuscript, we identified a mistake for our interpretation of the data. For this point, we sincerely apologize, and please allow us to correct the data.

Total number of neonates who were tested by our screening program from 1989 to 2013 was 2,105,108. Because of this correction, we revised the number (P2L5, P7L10, P19 Table2a) and the incidence of CAH in Tokyo(P7L14, P8L16, P10L3, P19 Table2a, P23 Table4).

We believe the correction of the data will not affect the storyline of our manuscript.

Background:
Minor essential revisions:

1. In the background section, the authors refer appropriately to a previous paper describing newborn screening programmes in Japan (ref 3), however it would be helpful if there were more information about the Japanese screening programme for CAH in this paper, as well as a statement about whether the Tokyo programme is typical of newborn screening elsewhere in Japan or East Asia.

Unfortunately, there is no published data that shows the details of the screening system of every lab in Japan, and difficult to describe whether Tokyo program is typical of MS in Japan or not. However, we think that our system could be unique because we used different cut off criteria for preterm and low birth weight infants.
In order to clarify it, we cited previous reports of CAH screening from Japan and Asian countries and added a more detailed description to the third paragraph in Background(P3L24-P4L2), as below
“on the basis of our pilot study, the screening system in Tokyo had the different cut-off criteria for preterm or low birth weight infants from that for term infants, because false positive of the preterm infants is one of the major concerns of the newborn screening for CAH.”

Methods:

Major Compulsory Revisions:

i. The authors describe the 17OHP measurement in sufficient detail however they should provide more detail on the age at newborn screening as this may have an important impact on the effectiveness of the screening programme and its ability to identify babies at risk prior to collapse, as well as its comparability with other programmes.

Basically the age of 4 to 7 days were recommended for blood sampling in Japan including Tokyo, and in our analyses, the most children were tested at 4 or 5 days after birth.

In order to clarify those issues, we changed the description as below,

“Basically we recommended to collect the blood sampling from the age of 4 to 7 days” (P5L5)

“The median age at first screening for them was the age of 5 days (range 0-62), consistent with our recommendation.” (P7L12-13)

ii. Table 1 suggests that screening may be undertaken at 5-10 days of age, however in the body of the text, the authors state only that blood sampling was performed at ‘5 days or after’; there is no information about the spread of ages at screening. Such data should be added.

It was simple mistake, and the description of Table 1 was incorrect.

We changed the description of Table 1, as below.

“before the age of 7”

Further, we added the information of age at screening as follow.

“The median age at first screening for them was the age of 5 days (range 0-62), consistent with our recommendation.” (P7L12-13)
iii. The authors should clearly state whether the Tokyo screening programme is representative of newborn CAH screening throughout Japan and East Asia, and therefore whether the findings of the study are generalisable.

In our manuscript, we did not say that our study represents throughout Japan and East Asia. In order to avoid misunderstanding, we rewrote the third paragraph in Background(P3L20-L24).

iv. The authors provide very little information about the sources of data, data variables and quality of the data collected in the follow-up survey. Were these data complete for every child and were the sources of data reliable?

Regarding sex, body weight and gestational age, we have added a description (P5L15-P5L17).

To collect more clinical information, we performed the survey, however, it was only for the children who were referred to hospitals due to the positive results of the screening. The survey was performed for the physicians who directly treated the patients, so the data and information are reliable.

According to the reviewer’s suggestion, we changed the description as below.

“We performed follow-up survey of the patients who were referred to hospitals due to the positive results. We collected the clinical information of the patients from the physicians of the hospitals, and collected information was the diagnosis of the patients including the type of CAH, laboratory data before the start of the treatment (17-OHP, Na, K), and the brief clinical courses during early infantile period.” (P6L5-L9)

Results:

Minor essential revisions:

i. The authors’ presentation of their results is somewhat confusing. Detailed information is present within the tables but an improvement would be to provide additional summary text
interpreting the tables and figures, and thus to draw the reader’s attention to the key points presented.

In order to assist reader’s understanding, we have added a description that addresses the composition of our analysis (P7L3-L6), as below.

“Firstly, we comprehensively analysed our data, including the incidence and the positive predictive value of the screening. Subsequently, we had examined the clinical details of the CAH patients who were identified by our screening, and finally, one of the purpose of the screening, sex assignment issue, was analysed.”

Major Compulsory Revisions:

ii. To more clearly present their findings, I believe the authors should begin by describing the characteristics of the children in the population undergoing screening (e.g. % term/preterm, median birth weight, median age at screening) to provide a baseline understanding of the population screened. They should then describe the number diagnosed with CAH and the characteristics of those with positive screen results and an eventual CAH diagnosis, e.g. proportion/number with 21OHD or other forms of CAH, sex ratio, etc. A comparison could then be more clearly drawn between the background population and those with positive screen results and/or with CAH diagnoses, in particular with regard to birth weight and gestation.

We agree with the reviewer’s suggestion, however, unfortunately, we do not have the data of the characteristics of the children in the population undergoing screening. Instead, we have the data of Japanese newborn from Ministry of Health and Welfare in Japan, and we have added the information to the text (P8L4-L6).

iii. The authors report the number of babies screened but should state whether screening coverage was complete,
We estimated the coverage of the screening, and added a description (P7L10-L11)

“Coverage of the screening was 93% of newborn babies registered in Vital Statistics of Japan.”

as it is unclear whether the incidence of 1:19,934 refers to 1 diagnosed case per 19,934 children screened or 1 diagnosed case per 19,934 births in Tokyo.

As we describe in the text (P7L14) and Table 2a, the incidence is 1:19,859 refers to 1 diagnosed case per 19,859 children screened.

iv. The authors provide positive predictive value and false positive rates but should also describe in the text more detail of the repeat testing required (which appears considerable).

Our retrospective data is based on our survey which was performed for neonates with the positive results on the screening, and unfortunately, we do not have the data the reviewer is requiring.

Major Compulsory Revisions:

i. The authors make several statements in their discussion that do not appear to be well-supported by the data presented earlier in the paper. In the first paragraph of the discussion, they state that the population screened were homogeneous in race/ethnicity but they should provide evidence to support this.

We removed the description.

ii. The authors further state a primary aim of screening was to prevent adrenal crises, so they should present any data relating to adrenal crises within their population.
Unfortunately we do not have any data that shows adrenal crisis of our population because our screening was not designed to identify the false negative cases, as described before. Instead, we cited literatures showing that the number of false negative would be quite low (P10L15-L22).

iii. The authors appear to assume that there is no migration in or out of Tokyo (such that newborns may have presented clinically with CAH to hospitals outside of Tokyo), so should provide evidence to suggest that the population is stable in the newborn period and that they have good population coverage.

We cannot catch the point of the suggestion. We have never described that there is no migration in or out of Tokyo in our text. Of course, there are, but we did collect the survey from all patients who referred to hospitals for further examination (P6L8-L9), and the locations of the hospitals did not matter for our survey.

Describing limitations of the study:

Major Compulsory Revisions:

i. The authors acknowledge the lack of follow-up of negative screens to identify false negative results as a limitation, however they should describe why they made no attempt to identify false negative cases (‘missed’ diagnoses amongst living or dead children) within the screened population. The lack of follow-up data for negative cases means they cannot assess screening programme performance fully as they cannot report sensitivity nor specificity.

As most other screening programs, our screening program is not designed to follow the false negative cases, and that is the reason.

Of course, we are aware of the problem, so we have not mentioned the specificity and sensitivity of our screening system in our first manuscript, instead, we mentioned positive predictive value which does not need false negative data.
ii. The authors should discuss the large number of repeat tests required as an additional limitation to screening.

We have added a description of the limitation as the new 7th paragraph in Discussion. (P12L17-L20)

What are the psychological and cost implications of repeating screening in this population?

We think those points would be important for newborn screening.

But, we do not have exact data of those issues.

Further, we think those issues are beyond our scope and would make our manuscript more complicated.

iii. The authors state that no deaths were reported but should provide evidence that the passive reporting by paediatric endocrinologists was sufficiently well-conducted to have accurately identified such cases. As studies in other countries suggest that deaths may occur before diagnosis or without diagnosis, it seems unlikely that passive reporting of deaths would have completely identified deaths. The authors could discuss this as a limitation or provide evidence that demonstrates that reporting of deaths was accurate and complete.

As the reviewer suggested, our screening does not have a system to identify false negative cases. So we added a description and cited literatures showing that the number of false negative would be quite low (P10L15-L22), and further, we also added a new paragraph (7th) in Discussion describing that lacking the systems to identify false negative cases is limitation of our screening system (P12L17-L20).

The authors make appropriate reference to previous studies in the East Asia region and elsewhere.
According to the reviewer's suggestion, we cited previous reports from Asian countries (P3L22).
First of all, during revising the manuscript, we identified a mistake for our interpretation of the data. For this point, we sincerely apologize, and please allow us to correct the data.

Total number of neonates who were tested by our screening program from 1989 to 2013 was 2,105,108. Because of this correction, we revised the number (P2L5, P7L10, P19 Table2a) and the incidence of CAH in Tokyo(P7L14, P8L16, P10L3, P19 Table2a, P23 Table4).

We believe the correction of the data will not affect the storyline of our manuscript.

1-1. The second paragraph of the Background section indicates that a primary purpose of the paper is to address a lack of knowledge about the incidence and epidemiology of CAH in Japan. However, the incidence of CAH in Japan is known to vary between 1 in 18,000 (Suwa 1994) and 1 in 21,000 (Morikawa et al. 2014), the same as in other countries. The authors should provide a context for their study in relation to the previous literature.

Suwa’s report is meta-analysis of independent 51 screening labs in Japan that had different screening algorithm, and that was the reason why we did not mention the literature in our first manuscript. However, the scale of the study is extremely large and we agree with the reviewer’s comment. According to the reviewer’s suggestion, we totally re-wrote the second paragraph in Background (P3L13-P4L2) and 1st paragraph in Discussion (P10L3-L8). We mentioned Suwa’s report and Morikawa’s report in the new 1st paragraph in Discussion. We also added the data of Nagasaki’s study and Morikawa’s study to Table 4.
1-2. They should indicate which aspects of their study are novel and which confirm previously published findings. In particular, it is the screening algorithm and its high positive predictive value (PPV) that appears novel. To explain why this is novel, it is necessary for the authors to summarize the relevant CAH screening literature and the lower PPV and higher rates of false positives that have been reported from other screening programs.

As the reviewer suggested, one of the novel point and the most important point of our study is higher PPV.

In order to clarify the point, we rewrote the 2nd and the 3rd paragraph in Background (P3L13–P4L2), emphasizing the point.

Further, we separated the 3rd paragraph in Discussion into two paragraphs, now new 3rd and 4th paragraphs in Discussion (P11L1–L18), discussing the issue more precisely and clearly.

1-3. Similarly, statements in the Discussion such as “This is the first large scale retrospective study of CAH newborn screening in East Asia” are incorrect.

As mentioned above, Suwa’s report is meta-analysis of different screening program in Japan. So, we believe our study is the largest study in East Asia to date.

In order to clarify the point and to avoid misunderstanding,

We changed the description of 3rd paragraph in Background, citing previous reports of screening for CAH from Asian countries. (P3L21-L22)

“our study is the largest retrospective analysis of CAH new born screening by using a single screening program in East Asia.”

We also described that Suwa’s report is meta-analysis (P10L6), implying that the study is not the analysis based on a single screening system.

2. The Methods section needs to explain the relationship of the newborn screening program in Tokyo to the screening program in other parts of Japan.
2-1. Are the screening cutoffs or algorithms different in Tokyo?

Yes, it is different, and we added a description of the issue to the 1st paragraph in Discussion.(P10L3-L6)

2-2. Who makes those decisions?

Maybe, a person of the government of each district makes them, but details are not clear for us.

2-3. Is there a designated screening laboratory for Tokyo hospitals?

Our lab (Tokyo Health Service Association, Newborn Screening, Tokyo, Japan) is in charge of the screening in Tokyo, and all the samples from Tokyo were sent to our lab.

2-4. Has Tokyo always used gestational age cutoffs for CAH screening?

Yes, and in order to clarify it, we added a description to the second paragraph in Methods (P5L20-L22).

“The criteria for preterm and low birth weight infants were used since the introduction of the screening in Tokyo.”

2-5. Who made that decision?

A pediatric endocrinologist in charge at the time when the screening was introduced in Tokyo did it, His name is Dr. Shimozawa.

2-6. Was it influenced by the experience of other countries such as Switzerland that have long used gestational age-specific 17-OHP cutoffs?

We determined to use the criteria just based on our pilot study. Unfortunately, the data of our pilot studies were not published in English language, then, we described just “our pilot study” (P3L24-L25, P5L19-20).
3. The Results section should present information that can be compared with previously published CAH newborn screening results from Japan. In particular, it is important to report the recall rate for this study, which appears to be 0.02%. That is extremely low compared with other screening programs. In Sapporo City, which does not use gestational age or birth weight cutoffs, the recall rate is reported to be at least 10 times higher (Morikawa et al. 2014). Also, just 25% of the infants in Tokyo who were recalled were born prior to 37 weeks, compared with 75% in Sapporo.

We agree with the review’s suggest. Our unique cut off criteria for preterm neonates reduced the recall rate of preterm neonates.

According to the reviewer’s suggest, we have added the data of Sapporo and Niigata in Table 4. The two screening systems do not have criteria for preterm infants.

Further we also calculated the recall rate and added the data to Table 4.

In order to emphasize the point the reviewer suggested, we totally re-wrote the 3rd paragraph in Discussion, making the new 3rd and 4th paragraphs (P11L1-L18).

4. The statement, “All but four (94.5%) of the SW patients were positive on the first test” requires further explanation. The Methods section states that only infants who have abnormal results on a first screen undergo a second screen. How were these 4 infants detected if they were not positive on the first screening test? Was the screening algorithm different from that described in the Methods section?

The algorithm of our screening was slightly complicated, and not described well in our first text. In order to clarify the issue, we added a description of our criteria and algorithm (P5L21~P6L2), and for further understanding, we changed the Table 1 and Figure1 showing the algorithm of our program more precisely.
5. This study notes in the Discussion that no fatalities were reported among 106 CAH cases identified by screening, which belongs in the Results section. However, the Discussion should indicate that does not rule out fatal adrenal crises in this population. Fatal cases occurring in at least 1 in 80 cases in Japanese infants with CAH detected by NBS (Ogawa et al. 2003). In other countries, about 1% of infants who have classical CAH detected by screening die in infancy (Grosse and Van Vliet 2007). The fact that no deaths were reported could be due to chance, with small numbers of cases, or to a possible lack of thorough reporting of deaths in the follow-up survey.

As the reviewer suggested, our screening does not have a system to identify false negative cases. So we discussed the issue in 7th paragraph in Discussion as the limitation of our study (P12L17-L20). We also cited literatures showing that the number of false negative would be quite low (P10L15-L22), because in Ogawa’s report, there were no death cases that were missed by the screening.