Author’s response to reviews

Title: HER2 testing on core needle biopsy specimens from primary breast cancers: interobserver reproducibility and concordance with surgically resected specimens

Authors:

Hitoshi Tsuda (hstsuda@ncc.go.jp)
Masafumi Kurosumi (mkurosumi@cancer-c.pref.saitama.jp)
Shinobu Umemura (umemura@is.icc.u-tokai.ac.jp)
Sohei Yamamoto (nrg09280@nifty.com)
Takayuki Kobayashi (dr20010@ndmc.ac.jp)
Robert Yoshiyuki Osamura (osamura@is.icc.u-tokai.ac.jp)

Version: 2 Date: 8 August 2010

Author’s response to reviews: see over
Dear Editor:

Thank you very much for your Email dated July 23, 2010. We are grateful to be informed the manuscript can be considered for publication in BMC Cancer pending revision. We have corrected the manuscript according to the referees’ comments as much as possible.

To reviewer 1 (Dr. Cathy Moleans)

1. Figures 1 and 2 have been improved as TIF files.

2. We have corrected the points of comments in Materials and Methods and Table 3 legend. According to ASCO/CAP guidelines, amplification of HER2 gene has been defined as a >2.0 ratio, and equivocal and negative were defined as a 1.8 to 2.2 ratio and a <1.8 ratio, respectively, in page 8 line 17 and in page 30 (Table 3 legend) of the revised manuscript.

3. Collaborating pathologists in the three institutes were assigned to submit almost equal number of CNB cases of each score (score 1 or 1+, 2+, and 3+), for the purpose of that almost equal number of HER2-negative, equivocal and positive cases were examined in the study. However, these institutional scores were not informed to the pathologists on the central review. Therefore, the cases were not consecutive and there was some selection bias. We have added these sentences in page 6 lines 8 to 13 in the part of tissue samples of Materials and Methods.

4. We have added standard deviations in page 9 lines 15-16, page 10 lines 1-2, page 11 lines 8-9, and footnotes of Table 4 in the revised version. Because weighed kappa is not used very popularly, we used classical kappa statistics.

Minor essential revisions:
1. We have amended the result section of Abstract in page 2 line 20, according to the comments.

2. English textual errors have been improved according to the reviewer’s comments in a, b, and c (Introduction) and d and e (Materials and Methods). The corrections are located in page 5 line 10 for a; page 5 line 22 for b; page 6 line 7 for c; and page 9 line 4 for d.

3. Yes they were. We have performed IHC and FISH assays on parallel slides of the same CNB/surgical specimens. In page 6 lines 8-13, we have added a sentence that “IHC and FISH assays were performed on parallel slides of the same CNB/surgically resected specimens”.

4. II means a roman numeral denoting two. Herceptest II is the name of the HercepTest kit for the HER2 test sold by Dako in Japan. Please leave the word as it is.

5. Three observers were experienced pathologists specialized in breast pathology (Kurosumi M, Umemura S, and Tsuda H). Because FISH was performed later than IHC judgments, they were blinded from FISH results. We have altered a sentence in page 7 lines 10-12 into “Three experienced (for >25 years) pathologists (M.K., S.U., H.T.), being blinded from institutional IHC results or the present FISH results, independently evaluated the results of IHC, and assigned a score of 0 …” in the part of immunohistochemistry in Materials and Methods.

6. The observers were blinded from IHC results and results of corresponding CNB or surgical resection specimen. We have added this sentence in page 8 lines 12-13 in the part of FISH in Materials and Methods.

7. As shown in Table 3, three (B61, B62, and B87) of 99 CNB successful cases were judged differently by two observers and the judgments were equivocal (Table 3). Therefore, they were subjected to a recount: By the recount, all three tumors were finally judged as positive because the counted data of \( \frac{\text{HER2}}{\text{CEP17}} \) ratio exceeded 2.20, although the judgment of each observer differed in two. These results have been described in page 10 lines 10-11 and 13-14 in Results. Among the 25 surgical specimens for which FISH was successful, only one (S67) was judged differently by
two observers. However, the average of the two judgments was within the range of positive (\(\text{HER2/CEP17 ratio} = 2.236\)), the case was judged as positive without re-evaluation. These results also has been described in page 10 lines 18-22 in Results in the revised version. In Table 3, we have added data of average \(\text{HER2/CEP17 ratio}\) to facilitate identifying equivocal cases.

8. We have corrected the title in Results in page 11 line 3.

9. We have corrected the sentence as was pointed out in page 11 line 19.

10. We have changed the sentence into “Possible reasons for this may have been the borderline nature of the tumors, preanalytical factors, and/or intratumor heterogeneity (e.g., No. 54)” in page 13 lines 3-4. We have added the interobserver agreement or disagreement to Table 5. Because it was clarified that interobserver disagreement regarding the HER2 score arose for either CNB or surgically resected specimens in 11 (not 10) of the 13 cases, we have corrected the number (page 13 line 2).

11. The first case was no.67 and the second case was no. 16. We have added the sentence “As mentioned above, the case was judged differently by two observers, but the average of the judgments made a score of positive (Table 3)” for the former, and the sentence “For that case, there was no disagreement in HER2 score of CNB and surgically resected specimens by IHC” for the latter in page 13 lines 11-13.

12. At present, in Japan, external quality assurance system is not introduced for the evaluation of HER2 tests. Therefore, we conducted this study to make clear the situation of relatively major institutions in Japan. Now, we are launched into the study of nationwide external quality insurance system. In the study, as a work of both Japanese Society of Pathology and the Ministry of Health, Labor, and Welfare of Japan. at first we will survey the preanalytical condition of >500 hospitals in Japan, and, then, we will evaluate if the accuracy of HER2 tests in participating hospitals will be improved after the introduction of an experimental external quality assurance system. When the data are convincing that the quality assurance system improves the accuracy, we shall establish a formal nationwide external quality assurance system. In order to this purpose, we are going to visit UK NEQAS to learn how to manage the quality control system. Therefore, at present, we do not
participate in external quality assessment. During conduction of the experimental external quality assurance system, we will compare processing protocols between all three institutes.

13. We have corrected the description according to the comment in page 14 line 6 in the revised manuscript.

14. We have changed the name of single hapten in situ hybridization into silver-enhanced in situ hybridization, and have added the name of dual-color dual-hapten methods in page 17 lines 2-3 in the revised version. We have added three novel articles as references (#31-33) in page 17 line 5 in the revised version.

15. In Discussion, page 16 line 23 to page 17 line 1, we have added a sentence that “These findings will be helpful for the decision if subsequent FISH should be performed or not. For 2+/3+ discrepancies FISH test should always be added because the percentage of HER2 amplification is high.”

16. We have added the paper of D’Alfonso et al. as the reference 29 and in page 15 line 17. We also have added the following sentences: “The paper of Moelans et al, showed that, although some studies suggested that the validity of IHC score 3+ in core biopsies is limited, reporting high rates of false positives (19.3%), there was only a slightly higher percentage of IHC 3+ positivity in biopsies compared to resections, but this did not reach statistical difference [30]. This is in line with this study (25% vs 21%)” in page 16 lines 3-7.


18. We have amended the sentence according to the referee’s comment in page 18 line 4 in the revised manuscript.
19. We have corrected the legend for Figure 1 according to the referee’s comment.

20. We have corrected the legend for Table 1 and added abbreviations such as IHC and FISH used in the table according to the comment.

21. We have corrected the legend for Table 2 and added abbreviations such as IHC and FISH used in the table according to the comment. Data of S17 FISH were not available because FISH was not successful. We have explained that in the footnote of the table. An asterisk has been deleted.

22. Case 54 showed clear intratumor heterogeneity in the invasive component (Figure 2A). To surgically resected specimen, all three observers gave a score of 3+ because >30% of cancer cells showed strong membrane staining. In contrast, in CNB specimen, the percentage of cancer cells showing strong membrane staining was around 10%, and two observers gave a score of 0, and the other gave one of 2+. We have added these sentences in page 12 lines 6-10 in Results instead of the previous sentence. As replied to the comment 13, we have added interobserver agreement/disagreement to Table 3.

23. Although FISH counts of case S67 was inconsistent between two observers, i.e., positive vs equivocal, we have listed that case in Table 3. Nonetheless, the average HER2/CEP17 ratio between two observers was 2.24, which exceeded the value of 2.2. Therefore, we judged we need not retest FISH.

24. We have added the word “corresponding” to the legend for Table 4, according to the referee’s comment.

25. We have amended Table 6 according to the referee’s comments. Because cut-off used in Table 7 was also 2007 ASCO/CAP guideline, we have corrected the table.

Discretionary revisions
1. The title proposed by the referee is fine. We have altered the Title in page 1 lines 1-2, according to the suggestion.

2. According to the referee’s suggestion, we have removed the indicated sentence
(indicated in page 8 lines 5-6 in the revised manuscript).

3. We have removed “levels” (indicated in page 9 line 11 in the revised version).

4. Because the indicated part was prone to misleading, we have amended the part as follows in page 14 lines 9-15 in Results: “For CNB specimens, 14 cases showed interobserver disagreement in HER2 IHC scores of between 1+ and 2+, and, of these, seven tumors each were finally scored as 1+ and 2+. Among these tumors, HER2 gene amplification was detected in 0 (0%) and 2 (29%), respectively. The other 10 cases showed interobserver disagreement in HER2 IHC scores of between 2+ and 3+, and, of these, five tumors each were finally scored as 2+ and 3+. HER2 gene amplification was detected in all of these tumors (Table 1)”.

5. We have added “between CNB and surgical specimens” in page 16 lines 12-13 in the revised version.

To reviewer 2 (Dr. Mamatha Chivukula)

1. The paper indicated had been already included as reference 25. We have added a sentence that “The present results were similar to the very first studies done by Chivukula et al. that stated as CNB a better sample [28]” according to the referee’s suggestion in page 15 lines 12-14 in Discussion.

2. Because we have adhered to ASCO/CAP 2007 guideline, we have corrected the manuscript throughout to be adherent to the guideline in this manuscript.

3. In cases of borderline FISH on CNB, it is important to select two blocks from resection and perform HER2 IHC testing (Stribel J et al.). In the present study, only one CNB block was available for three cases with equivocal HER2/CEP17 ratios. Therefore, to make data more reliable, two observers counted a total of 40 nuclei, and when the counts were equivocal, further 40 nuclei were counted.

4. As replied to reviewer 1, collaborating pathologists in the three institutes were assigned to submit almost equal number of CNB cases of each score (score 0 or 1+, 2+, and 3+), for the purpose of that almost equal number of HER2-negative, equivocal and positive cases were examined in the study. Therefore, the cases were
not consecutive and there was some selection bias. We have added these sentences in page 6 lines 8-13 in the part of tissue samples of Materials and Methods.

5. As replied to reviewer 1, case 54 showed clear intratumor heterogeneity in the invasive component (Figure 2A). To surgically resected specimen, all three observers gave a score of 3+ because >30% of cancer cells showed strong membrane staining. In contrast, in CNB specimen, the percentage of cancer cells showing strong membrane staining was around 10%, and two observers gave a score of 0, and the other gave one of 2+. We have added these sentences in page 12 lines 6-10 in Discussion instead of the previous sentence.

To reviewer 3 (Dr. Sophia Apple)

Materials and Methods
1. According to the referee’s comment, we have corrected the description from 5 micrometer to 4 micrometer in page 6 line 15.

2. We have altered description from “A total of 100 pairs of archival formalin-fixed paraffin-embedded CNB and surgically resected specimens of invasive breast carcinomas were cut into sections and subjected to HER2 testing by immunohistochemistry (IHC) and fluorescence in situ hybridization (FISH), the results being evaluated by three and two observers, respectively” to “A total of 100 pairs of archival formalin-fixed paraffin-embedded CNB and surgically resected specimens of invasive breast carcinomas were cut into sections. All 100 paired sections were subjected to HER2 testing by immunohistochemistry (IHC) and 27 paired sections were subjected to that by fluorescence in situ hybridization (FISH), the results being evaluated by three and two observers, respectively” in page 2 lines 8-12 in Abstract.

3. We have added a sentence that “to 100 CNB and 100 paired surgically resected tumor sections,” in page 7 lines 7-8.

4. In the present study, for 100 surgically resected specimens, we performed HER IHC to all the specimens, but retest of FISH was performed if the \( \text{HER2/CEP17} \) ratio was equivocal, or the diagnosis of two observers differed together. When HER2 IHC score was 2+ but judgments by three observers were concordant and the consensus IHC scores were also concordant between CNB and surgically resected specimens
in the same tumor, FISH was no performed.

5. We studied 40 additional nuclei in another tumor area on invasion. We have added the sentence in page 8 line 18 in Materials and Method.

Results

1. As replied to reviewers 1 and 2, case 54 showed clear intratumor heterogeneity in the invasive component (Figure 2A). To surgically resected specimen, all three observers gave a score of 3+ because >30% of cancer cells showed strong membrane staining. In contrast, in CNB specimen, the percentage of cancer cells showing strong membrane staining was around 10%, and two observers gave a score of 0, and the other gave one of 2+. We have added these sentences in page 12 lines 6-10 in Discussion instead of the previous sentence.

2. The borderline nature means that the state of HER2 expression was borderline between 1+ and 2+ or between 2+ and 3+, namely, the conditions that it was difficult to judge whether he entirely circumscribing membrane immunoreactivity of the HER2 was moderate (2+) or strong (3+), or whether the weak membrane HER2 immunoreactivity wad entirely (2+) or incompletely (1+) circumscribing the membrane. We have added these sentences between page 11 line21 to page 12 line 4.

3. We only counted HER2/CEP17 ratio. SO, we have no data on polysomy 17.

4. In case 92, we performed FISH on the part including both invasive carcinoma and DCIS components. We described that in page 12 lines 14-15.

Abstract

1. Conclusion has been reworded according to the reviewer has recommended as indicated in page 3 line 4-9.

2. We have amended the error that was pointed out by the reviewer in page 2 line 20 in the revised version.

3. We have changed the sentence as “we have clarified that CNB specimens showed
almost equal reliability to surgically resected specimens for testing of HER2 expression in terms of interobserver agreement levels and concordance with FISH results” in the first sentence of Conclusion in page 18 lines 1-3.

Additional comments
1. We have included a specific % of positive cells used as 1+, 2+, and 3+ in page 7 lines 12-17 as follows: a score of 0 (no staining), 1+ (weak, incomplete membrane staining in any proportion of tumor cells), 2+ (complete membrane staining that is either nonuniform or weak in intensity but with obvious circumferential distribution in at least 10% of tumor cells, or invasive tumors show intense, complete membrane staining of 30% or fewer tumor cells), or 3+ (uniform, intense membrane staining of 30% of invasive tumor cells).

2. We have presented a poster entitled “Influence of fixation period on the detection of HER2/neu protein and HER2/neu gene signals” by Hiroi S, Tsuda H, Oya T, Hayama T, and Kawai T, in the 92th annual meeting of Japanese Society of Pathology in April 2003 in Fukuoka, Japan. This manuscript is now in preparation. Therefore, we have added the sentence that “it is shown that a prolonged formalin fixation could lose FISH amplification and/or yield to unsuccessful test (Hiroi S, Tsuda H et al, manuscript in preparation)” in page 17 lines 11-13.

3. We have included number of years of experiences of three pathologists in page 7 line 10.

4. I also believe that ideally FISH should have been done in all 100 paired surgically resected specimens with CNB specimens. However, because of the limitation of grant and the purpose of the present study can be achieved by the FISH analysis of 27 specimens, we did not perform FISH to all 100 surgically resected specimens.

We believe the corrections we have made are satisfactory to you. If there are insufficient points in the correction, please let me know. We are looking forward to hearing from you soon.

Sincerely yours,

Hitoshi Tsuda