Reviewer's report

Title: Meta-analysis of quantitative diffusion-weighted MR imaging in the differential diagnosis of breast lesions.

Version: 6 Date: 1 November 2010

Reviewer: Shannon Agner

Reviewer's report:

In this paper, the authors present a meta-analysis of diffusion weighted imaging (DWI) for diagnosis of breast cancer. It is shown that DWI may have better specificity than DCE-MRI in distinguishing between benign and malignant lesions. The manuscript is improved from the prior submission. Although the manuscript is recommended for acceptance after minor essential revisions, the first three revisions suggested in this review are very important to solidify the manuscript. The manuscript also still has many spelling and grammatical errors, which are addressed in minor essential reviews.

Minor essential revisions:
1. In Discussion, fifth paragraph, sentence 9: This sentence is too bold a conclusion to be drawn from meta-analysis results. I suggest changing “should” to “could.”

2. In the Conclusion, a comment should be made on why a threshold for benign/malignant lesions classification could not be made based on this study and how this could be rectified.

3. In Methods, Selection of Articles, how is “sufficient data” to calculate TP, FP, TN, and FN defined?

4. In the Abstract, Conclusions: “Quantitative DWI has a higher specificity to differentiate between benign and malignant breast lesions…”

5. In Background, first paragraph, second sentence: “…application of mammography and ultrasound, differentiation between benign and malignant breast lesions remains a difficult diagnosis problem, especially in dense fibroglandular breasts.”

6. In Background, third paragraph, third sentence, “Comparison of the diagnostic performance of breast quantitative DWI among the studies may have been compromised by differences in the patient characteristics, MR imaging techniques, and diagnostic criteria for malignancy in the studies.”

7. In Background, third paragraph, last sentence, the authors should state the purpose of the study, not the conclusions. For example, “In this study, we pool a number of DWI studies of the breast to evaluate the diagnostic performance of DWI in breast lesion characterization.”

8. In Methods, Literature Search, third sentence: “We limited our search to publications in English and Chinese languages…” (leave word languages in).

10. In Methods, Selection of Articles, first sentence: “Inclusion criteria were: varied pathology within the dataset; total number of lesions # 30, with the number of both malignant and benign lesions each # 10; histopathologic analysis (performed at surgery and biopsy) and follow-up by ultrasound, mammography, or MRI used as the reference standards…”

11. In Methods, Selection of Articles, third paragraph, first sentence: remove “to verify that stability and reliability of pooled analysis”

12. In Meta-Analysis, Threshold effect, third sentence: log (SEN) instead of “log sensitivity” and log(1-SPE) instead of log 1-specificity

13. In Meta-Analysis, Publication bias analysis, last sentence, change “nil” to “zero.”

14. In Results, first paragraph, third sentence: “The objective of studies was…”

15. In Results, second paragraph, first sentence: “were included from 13 studies.”

16. In Results, second paragraph, sentence 5: “The abstracted data of these individual studies are summarized in Tables 1 and 2.”

17. In Results, second paragraph, sentence 6: “used to perform breast DWI in Baltzer et al. [40]; Marini et al. [18] used…in Rubesova et al. [8].

18. In Results, fourth paragraph, second sentence: “A homogeneity test…”

19. In Results, fourth paragraph, sentences 6-7: “A Spearman rank correlation is performed as a further test for threshold effect. The Spearman correlation coefficient was equal to 0.097 (P=0.66) and indicates…”

20. In Results, sixth paragraph, first sentence: “Having found notable heterogeneity beyond threshold…” (remove “the”).

21. In Results, fifth paragraph, first sentence: “Heterogeneity, the pooled weighted sensitivity and specificity, and AUC are analyzed again when [46] and [47] were excluded…”

22. In Discussion, second paragraph, first sentence, “in a ROC space, and Spearman correlation coefficient is computed between log (SEN) and log (1-SPE).”

23. In Discussion, second paragraph, remove “Comprehensive literature search may reduce publication biases.” In next sentence: “is supplemented by checking references of relevant studies in order to reduce publication bias [25].”

24. In Discussion, third paragraph, “Because b values used in the included studies varied, we needed to explore whether b values were the source of heterogeneity.”

25. In Discussion, third paragraph, “Furthermore, sensitivity analysis is performed in the subgroup of b = 1000 s/mm2. When the studies…”

26. In Discussion, combine fourth and fifth paragraphs.
27. In Discussion, fifth paragraph: “The significance of Peter et al. [6] lies in…”
28. In Discussion, fifth paragraph, sentence 9: “Because MR is one of the important methods for breast cancer diagnosis, DWI, having the advantages of a short examination time and…”
29. In Discussion, the sixth paragraph is very long. Consider breaking it into 2 paragraphs.
30. In Discussion, sixth paragraph, first sentence: remove “considered to be combinable”
31. In Discussion, sixth paragraph, sentence 9: “As the index test, DWI was always performed first, and interpretation of the results of the DWI was usually done without knowledge…”
32. In Discussion, sixth paragraph, sentence 11: “Test accuracy studies may produce various number of uninterpretable results.” This sentence does not make sense, and I am unsure what you are trying to convey.
33. In Discussion, sixth paragraph, sentence 21: “Therefore a systematic data reporting method such as the standards for reporting diagnostic accuracy (STARD)…”
34. In Discussion, seventh paragraph, sentence 3: “Even in the subgroup of studies using maximum b=1000 s/mm2, the minimum and maximum…”
35. In Discussion, seventh paragraph, sentence 5: “… benign breast lesions because selection of the threshold value should be…”
36. In Discussion, seventh paragraph, sentence 6: “For example, a relatively higher threshold value may be recommended to minimize false positives in breast cancer screening.”
37. In Discussion, seventh paragraph, sentence 7: “If DWI is appended..., a relatively lower threshold value may be recommended to reduce false positive results.”
38. In Reference 6, remove [0] from authors section.

Level of interest: An article of importance in its field

Quality of written English: Needs some language corrections before being published

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:
I declare that I have no competing interests.