Author’s response to reviews

Title: A predictive model of thyroid malignancy using clinical, biochemical and sonographic parameters for patients in a multi-center setting

Authors:

Jia Liu (liujia_625@163.com)
Dongmei Zheng (dmeizheng@163.com)
Qiang Li (lqiang_hrb@163.com)
Xulei Tang (xulei_tang@126.com)
Zuojie Luo (luozj_gx@163.com)
Zhongshang Yuan (yuazhongshang@sdu.edu.cn)
Ling Gao (gaoling1@medmail.com.cn)
Jiajun Zhao (jjzhao@medmail.com.cn)

Version: 3 Date: 08 Nov 2017

Author’s response to reviews:

Point-by-point Responses to the Reviewer

We would first like to thank you for carefully and patiently reviewing our manuscript and for giving positive and constructive comments and suggestions, which helped us to improve our manuscript. Specific responses to the comments by the reviewer are as follows:

Question1: Materials and methods Patients - It is unclear why the patients had total or partial thyroid surgery who had been confirmed as benign on post-operative pathology, even though these nodules were presented as benign.

Answer: Thank you very much for your professional question. As mentioned in the manuscript, this is a retrospective study enrolling patients from 2006 to 2009. The surgery indication was generally decided by the surgeons based on their comprehensive evaluation of the risk of nodules. Due to lack of effective assessment model for risk prediction and variability of surgeons when determining the nodules, some interventions were performed unnecessarily. Therefore, we
aimed to establish a practical tool for clinicians to distinguish thyroid nodules preoperatively to avoid some unnecessary interventions.

Question 2: In table 1, 'benign' category showed 2.75 ± 1.70 cm in size, which means not all nodules were excised due to large size. - Did you include patients with all dataset (clinical, laboratory, and US variables)? Or some patients had blank? How many patients were excluded due to incomplete dataset or records?

Answer: Thank you for pointing this out and we are sorry for the inclarity. In this study, we retrospectively studied consecutive 3145 patients who had complete medical records (including clinical, laboratory and US variables) and received thyroid surgery between 2006 and 2009 from four-tertiary hospitals. Patients with medullary thyroid cancer, anaplastic cancer or lymphoma were considered TSH-nonresponsive and were excluded. After the exclusions, 2984 patients were included in the analysis.

Question 3: US imaging analysis - The authors mentioned that this was a multicenter retrospective trial but details in 'US imaging analysis of Materials and methods section' were not understandable in terms of a retrospective, multi-center trial.

1) All included subjects were examined by only two skilled sonographers in only one US machine (GE LOGIQ9) in only one center? If all participant centers had US data, this description looks wrong.

Answer: Thank you for the very professional questions. We are very sorry for the confusion. Indeed, this is a multi-center trial and the subjects were examined respectively in their located hospitals by two skilled sonographers. We have re-edited the description in Page 6, line 18 “The examinations were conducted and recorded by two skilled sonographers from respective hospitals according to a standard procedure”.

Since this is a retrospective multi-center study, to limit variability of determination we contacted fifteen representative hospitals with similar levels and confirmed that US examinations were performed using US scanner GE LOGIQ9 in the four investigated hospitals (including our center) before this study. Therefore, the subjects were examined respectively in their located hospitals with the same US scanner. We have re-edited the description in Page 6, line 16 “US examinations of the four tertiary hospitals were performed using US scanner GE LOGIQ9 (USA) equipped with a 5-12-MHz linear transducer for morphological examinations and a 4.7-MHz transducer for color Doppler evaluations”. 
2) Did you check inter-observer variability when determining US findings? Usually, US findings show substantial analytic gap between observers so the authors should mention interobserver agreement of each US findings.

Answer: Your comments are quite reasonable and we agree with your opinion. First, the four participating hospitals were high-rank, representative hospitals and US examinations were performed by skilled sonographers following a standard procedure. Secondly, to check inter-observer variability on determining US findings we randomly selected 30% of targeted number of patients from the other three participating hospitals, obtained US data with US imaging findings and sent to our sonographers for re-analyzing. We didn’t find substantial analytic gap between the observers. We re-edited the description in Page 6, line 18 “The examinations were conducted and recorded by two skilled sonographers from respective hospitals according to a standard procedure and interobservers reached agreement on the results of each US findings”.

3) How did you analyze your US data? Did you obtain US data without image findings from participating hospitals? Or did you reanalyze all dataset of US findings from multi-centers?

Answer: Thank you for careful reading and insightful questions. US data from our center was analyzed by our two skilled sonographers. As mentioned in the previous question, 30% US data with US imaging findings from the three other participating hospitals were re-analyzed by our sonographers for variability check before the study. We didn’t find substantial analytic gap among different hospitals.

Question4: Laboratory variables - Four tertiary hospitals had the same free T4, free T3, and TSH kits? I wonder it would be possible to compare thyroid function test from each center.

Answer: Thank you for reminding us for paying attention to this detail. We also took this point into consideration. As mentioned in Question3, to limit measurement variability among different medical centers, we selected three out of fifteen tertiary hospitals and confirmed that the four involved hospitals (including our hospital) quantitated thyroid function using the same chemiluminescence analyzer Roche Cobas E601 (Switzerland) and the matched kit. Therefore, we believe that the results of thyroid function were comparable among these four hospitals.

We have added the description in Page 6, line 25 “The levels of serum TSH, free triiodothyronine (FT3) and free thyroxine (FT4) were determined using chemiluminescence analyzer (Switzerland) and the matched kit”.
Question 5: Results - In page 9, I am curious as to why you have categorized microcarcinoma smaller than 1 cm apart. Also, is this microcarcinoma applicable to your formula? Does the practice of microcarcinoma vary based on the study?

Answer: Thank you for the question. Microcarcinoma (malignant nodules ≤1 cm in diameter) was categorized separately upon Reviewer 1’s requirement. Since microcarcinoma was recently considered “more silent”, we also think it might be meaningful to analyze clinical, biochemical and sonographic characteristics of microcarcinoma apart. In line with thyroid nodules >1 cm in diameter, age, positive TGAb result, hypoechogenicity, microcalcification and intranodular central flow were also associated with increased risk for malignancy in microcarcinoma. Additionally, the mathematical prediction model we proposed in this manuscript was also suitable for nodules ≤1 cm in diameter because nodule size was not included in this formula.

Question 6: Discussion - Generally, discussion section is somewhat long and very redundant. Several sentences should be more concise including Page 14 line 4~11.

Answer: Exactly as you demonstrated, an excellent article should be presented in a clear and concise way, avoiding any obscure or unnecessary details. According to your suggestions, we summarized the content, removed some details and retained the necessary and important parts. We also re-edited previous Page 14 line 4-11 as you required. The revised part could be seen in Discussion (Page 11, line 8). If you think the present version still need to be modified, please let us know and we will further revise it according to your suggestions. Thanks again for your valuable comments.

Question 7: There is a question about the clinical implication of your research. After performing a risk assessment based on the formula you have proposed, if you have a high risk of malignancy, do you suggest immediate surgical resection without an FNA? If so, you are suggesting a completely different clinical approach from the current practice, and you will need a clearer basis for this.

Answer: Thank you for the valuable question. As stated in the manuscript, this is a retrospective study enrolled patients from 2006 to 2009. Unfortunately, due to skill limitation and lack of compliance FNA was not generally performed in suspicious nodules at that time. However, FNA cytology is a relatively effective and robust method for nodule discrimination. In current circumstances we would definitely recommend FNA for high risk nodules before surgical interventions.