29-Oct-2016

CC: Yang, Dawei; Zhang, Xiaoju; Powell, Charles; Ni, Jun; Wang, Bin; Zhang, Jianya; Zhang, Yafei; Wang, Lijie; Xu, Zhi-hong; Zhang, Li; Wu, Guoming; Song, Yong; Hu, Jiaan; Zhang, Yong; Song, Yuanlin; Zhou, Jian; Bai, Chunxue

Dear Prof. Chunxue Bai:

We are sorry to notify you that your manuscript # CHEST-16-2258 entitled "Probability of Cancer in High Risk Patients Predicted by the Protein-Based Lung Cancer Biomarker Panel in China: LCBP Study" has not been accepted for publication in CHEST. The decision to decline your paper was based not only on the evaluation of the reviewers, but also on an assessment of the manuscript’s content, scope, priority, and interest to CHEST readers.

We are enclosing any copies of the reviewers comments for your records.

While we were not able to accept this paper, we encourage you to keep CHEST in your plans for future submissions.

Sincerely,

Richard S. Irwin, MD, Master FCCP

Editor in Chief, CHEST

Reviewer(s)' Comments to Author:

Reviewer: 1

Comments to Author

Overview:

The authors aim to develop a blood-based test comprised of four proteins to improve the ability to identify which patients have lung cancer.  This study is limited by several confounding variables including smoking histories, incompletely described methods and training/validation sets, high potential for overfitting of the clinical, nodule, and biomarker risk models, lack of comparison to existing well validate clinical risk models and nodule risk models.

Major comments:

- The previous study referenced in the introduction in which the protein panel biomarker was derived and evaluated it in a single center does not appear to be a peer-reviewed publication and is available only as an abstract.

We provided a Pubmed link for this article: http://doi.wiley.com/10.1002/cncr.29551

Dawei Yang, Yong Zhang, Qunying Hong, Jie Hu, Chun Li, BaishenPan,Qun Wang, Feihong Ding, JiaxianOu, Fanglei Liu, Dan Zhang, Jiebai Zhou, Yuanlin Song, Chunxue Bai\*. Role of a Serum-Based Biomarker Panel in the Early Diagnosis of Lung Cancer for a Cohort of High-Risk Patients. Cancer, 121(17):3113-3121, 2015. (IF = 5.068)

- What is the rationale for selecting these specific four proteins for this biomarker, and not others?

We chose the biomarkers based on the current consensus for early diagnosis of primary lung cancer in China (in Chinese, Chinese Journal of Tuberculosis and Respiratory Diseases, 2014,37(3), <http://d.g.wanfangdata.com.cn/Periodical_zhjhhhx201403005.aspx)>.

- What is the overlap of samples included in this study, and those included in the prior single center analysis (ref 17)?

None. The Ref 17 cases were enrolled between July 2011 and February 2012, while in this article the cases were enrolled between October 2012 and February 2014.

- This study’s definition of high risk varies considerably from other clinical risk models and from the inclusion criteria used for the NLST, Nelson study, and DANTE studies.  What is the rationale for this?

We used the definition of high risk varies based on the current consensus for early diagnosis of primary lung cancer in China (in Chinese, Chinese Journal of Tuberculosis and Respiratory Diseases, 2014,37(3), <http://d.g.wanfangdata.com.cn/Periodical_zhjhhhx201403005.aspx)>.

- One of the study’s inclusion criteria was an indeterminate pulmonary nodule.  What was considered to be an indeterminate pulmonary nodule?

Pulmonary nodule was occasionally found, which was less then 3cm in diameter and without previously diagnosis based on pathological biopsy or CT surveillance.

- What were the exclusion criteria?  In what context and at which point in their clinical care were these patients enrolled, evaluated and samples collected?  For example, the abstract and introduction describe a biomarker of risk, but the CT scan and models appear to have been performed in a diagnostic setting?

In ‘Patient Selection’ part we have declared the exclusion criteria were (1) absence of histopathological diagnosis; and (2) prior treatment by chemotherapy or surgery. In the patient risk model, we evaluated the patients by blood test rather than CT scan. While in the nodule risk model, we both evaluated the patients by blood test and CT scan.

- How did the authors determine which patients/samples were part of the training set, and which were part of the validation set?

The training data set included participants enrolled from three medical centers (Peking Union Medical College Hospital, Henan Provincial People's Hospital and Jinling Hospital, Nanjing University School of Medicine). While the rest of the participants from the other two medical centers (Zhongshan Hospital Fudan University and Third Military Medical University) were enrolled in the validation set.

- What was the duration of follow up for patients before receiving a diagnosis of lung cancer?

We could retrospect this data by checking CRF record from each centers, but it might take one-month period to recall the data. We have collected blood draw data and the diagnosis data of lung cancer, so we could calculate the time period between the enrollment and the diagnosis time.

- The smoking histories of the participants are incompletely described in the results and tables.  What proportion of cancer/control samples in the training and validation sets were current smokers, former smokers, and nonsmokers?  What were the pack years?  How many years since quitting?

We could retrospect this data by checking CRF record from each centers, but it might take one-month period to recall the data. We have questionaries for smoke status, pack years, and quit year.

- What is the rationale for developing a new patient risk model?  How does this compare to existing and well validated models for risk of lung cancer?  Why were other factors such as pack years, occupational exposures, family history, etc not included or evaluated?

The current patient risk model is based on western population’s data, we developed the new patient risk model based on Chinese population. We will compare the new model with PLCOm2012 model (Tammemägi MC1, Katki HA, Hocking WG, Church TR, Caporaso N, Kvale PA, Chaturvedi AK, Silvestri GA, Riley TL, Commins J, Berg CD. Selection criteria for lung-cancer screening. N Engl J Med. 2013 Feb 21;368(8):728-36. doi: 10.1056/NEJMoa1211776.)

We have no enough information about smoke history and smoke status to apply PLCO m2012 model.

- What is the rationale for developing a new nodule risk model?  How does this compare to existing and well validated models incorporating nodule characteristics in the prediction of risk of lung cancer?  Why were other factors, such as nodule density, total number of nodules, etc not included or evaluated?

The current nodule risk model is based on western population’s data, we developed the new patient risk model based on Chinese population. In Result part, we compared the nodule risk model and ACCP model for the prediction of lung cancer diagnosis in the 163 participants with nodules in the training data set. In this population, the AUC of ACCP model was 0.8360, which was less than that of the nodule risk model (P=0.001) (Figure 5). We have not collected the nodule density and total number of nodules information in the study. But we could retrospect this data by checking PACS record from each centers, but it might take one-month period to recall the data.

- The authors describe statistically significant differences in the levels of three of the four proteins measured in blood.  What are the results when corrected for confounding variables and for multiple testing?

All samples in the training set were run in duplicate. For the validation study, all assays were run on automated analyzers per the manufacturer’s instructions.

The results are based on a multiple variables logistic regression method, which was corrected for confounding variables.

- What is the rationale for including all four proteins in the risk models, when only three were significantly different between the groups?

The four proteins measurement results were treated as an entity?(The vendor want to treat the four variable as an entity and we keep them in the variable list when construct a logistic regression model.)

- The initial analysis of these protein levels was in the total population (training plus validation), raising the concern for overfitting in the subsequent models.  The large decrease in AUC, sensitivity, and specificity between the training and validation sets for these models also suggests this.  What is the performance of the clinical and nodule models combined without the serum protein measurements?

In Table 5 we showed that the performance of the clinical and nodule models combined without the serum protein measurements is less efficient than with the serum protein measurements both in accuracy and sensitivity.

- How do these markers perform without the use of clinical or imaging variables in predicting lung cancer?

In Table 5 we showed that the performance of these markers performed without the use of clinical or imaging variables in predicting lung cancer is more specific than only use of clinical and imaging variables, but less sensitive.

- What is the incremental change in risk prediction with the addition of these four serum markers to clinical and/or imaging variables?

In Result part, we compared the nodule risk model and ACCP model (which is combined clinical and imaging variables) for the prediction of lung cancer diagnosis in the 163 participants with nodules in the training data set. In this population, the AUC of ACCP model was 0.8360, which was less than that of the nodule risk model (P=0.001) (Figure 5).

# minor comments

- The manuscript needs to be reviewed / corrected for grammar

We have reviewed and corrected for grammar.

Reviewer: 2

Comments to Author

Thank you for the opportunity to review your work.

General Comments: This manuscript presents an attempt at discovering models of serum protein biomarkers capable of identifying people with lung cancer from a general population as well as in characterizing lung nodules. This is an area in need of further development and a sizeable number of patients are included. The interpretation of the reported results appears flawed.

Specific Comments:

1. The abstract states that this is a clinical utility study. This is not correct. Models are being built then tested from the same population. Thus this is a discovery level study. Validation of the findings is necessary before pursuing a clinical utility study. Clinical utility studies assess the impact of the findings on clinical decisions.

The training data set included participants enrolled from three medical centers (Peking Union Medical College Hospital, Henan Provincial People's Hospital and Jinling Hospital, Nanjing University School of Medicine) (Table 1). While the rest of the participants from the other two medical centers (Zhongshan Hospital Fudan University and Third Military Medical University) were enrolled in the validation set (Table 2).

2. The “ACCP model” is better called the “Mayo model”. When comparing the biomarker results to the Mayo model it is important that you use the validation set results (AUC 0.5836), not the training set. The Mayo model appears to outperform your biomarker model. This should also be reflected in Figure 5.

We have change the ‘ACCP model’ into ‘Mayo model’. We also test Mayo model in our validation set and updated the result into Figure 5. Apply LCBP ACCP model in our validation set and the AUC of them are very close, and the p value of Delong test of AUC is 0.7971, show that there is no statistical significant difference between LCBP and ACCP model in validation set.



3. It is difficult to understand the patient groups based on the information provided. You state a high risk group was recruited but stated patients could be as young as 18 and could be never smokers if women. It is not clear how you insured they were high risk. You have a high percentage of non-smokers – were these former or never smokers? Similarly, for the nodule section, were they symptomatic patients or not. In the nodule group did you only include stage I cancers that presented as nodules? It appears that you have very few stage I cancers. If not then it is hard to claim the model relates to nodule management.

We have excluded this young group of patients and analyzed this subgroups as special high risk population. We would like to figure out the special risk factors, which cause the younger lung cancer incidence. We could retrospect this data by checking CRF record from each centers, but it might take one-month period to recall the data. We have questionnaires for smoke status, pack years, and quit year. We also have questionnaires for symptoms and the reason for visiting hospital. We divided previous nodule group into new nodule groups (lesion <=3cm) and mass group (lesion >3cm).

4. You state all patients had CT scans of the thorax. Were these performed as screening or for symptom evaluation?

We could retrospect this data by checking CRF record from each centers, but it might take one-month period to recall the data. We have questionnaires for the aim for visiting hospital.

5. It is difficult to reconcile the results that are provided. For the patient risk model with the AUCs provided at around 0.7, it is impossible to have a sensitivity and specificity > 90%. In examining the ROC curve of the validation set, at a sensitivity of 90% the specificity looks to be around 25%. I have similar concerns about the listed accuracies for the nodule models. It is hard to reconcile accuracies reported to be 69% with a sensitivity of 92.9 and specificity of 99.5.

A patient with a calculated cancer probability of 0.65 (cut-off value 1) or higher was considered to have cancer (high-risk), and one with a probability less than 0.21 (cut-off value 2) was considered not to have cancer (low-risk). So we chose the different threshold for picking up the optimized sensitivity and specificity.

1. Cut-off point 1: with sensitivity as high as 95% and relative lower specificity. With high sensitivity, this cut-off point can be used as a screening method to figure out the one with potential risk of lung cancer and trigger further investigation of them.
2. Cut-off point 2: with specificity as high as 95% and lower sensitivity. This cut-off point can be used as …(David, would you please fill in the usage of this cut-off point here?)

6. You describe a cutoff whereby you consider someone to have cancer and another whereby you consider someone not to have cancer. What do you do with all of the results in between these cutoffs? This question applies to both sets of models.

A patient with a calculated cancer probability of 0.65 (cut-off value 1) or higher was considered to have cancer (high-risk), and one with a probability less than 0.21 (cut-off value 2) was considered not to have cancer (low-risk).

7. In the methods, list the features of the nodule that were recorded (e.g. spiculation).

We recorded the features of the nodule, such as diameter, speculation and lobe location from CT scan.

8. The discussion mis-states (over-estimates) the actual accuracy reported.

We also test Mayo model in our validation set and updated the result in discussion part.

9. The conclusion section is far too long. It appears to be more of an extension of the discussion.

We have shortened the conclusion section.

10. Figure 1 does not make sense. Should the lower arrows be pointing down?

We have changed the arrow direction in Figure 1.

Note: Because all authors as well as CHEST staff are being copied on decision letters, please be advised, if you hit "Reply All" to this e-mail, [EditorialOffice@chestnet.org](mailto:EditorialOffice@chestnet.org) will receive a copy of your reply and any confidential comments may be shared with the CHEST editorial staff.

(DL-13)