Reviewer's report

Title: Using a practical molecular capsular serotype prediction strategy to investigate Streptococcus pneumoniae serotype distribution and antimicrobial resistance in Chinese local hospitalized children

Version: Date: 26 August 2015

Reviewer: Hans-Christian Slotved

Reviewer's report:

Major Compulsory Revisions

1. In this study, the authors mention three different types of data. 1. They have developed a new approach for serotype identification, 2. They describe serotype distribution and antibiotic susceptibility based on 193 isolates, 3. The collected isolates are mainly from children with different levels of pneumonia. I think the authors need to focus much more on the molecular identification approach, a this study uses isolates (both carrier and IPD isolates) from sick children with different levels of pneumonia, bacteraemia and meningitis, which I believe are not representative for the region (Supplementary Table S3). I think the authors need to reduce the focus on the serotype distribution in the manuscript, and focus more on the molecular typing strategy. Furthermore, I think the authors need to add information on how representative the isolates are to the region, if they intend to compare these isolates. I therefore recommend that the manuscript (particularly the discussion) is reduced with more focus on the molecular approach.

2. Regarding the susceptibility data, you also have to specify the limitation of the collected isolates, see comment 1.

3. The collected material is mainly from children with pneumonia. Why do the authors believe that these data can be used to describe the pneumococcal serotype distribution in the region?

4. In this study, the authors base their serotype identification solely on molecular techniques. This means that they show that the isolates possess the capsular genes. However, by using only molecular techniques, the numbers of NT isolates will be very low when compared to a phenotypical method. I acknowledge that the authors state that their presented method is only to be used for predicting the serotype distribution. However, I think the authors need to add a section in the discussion, where they discuss the differences between a molecular method and a phenotypical method, particularly in relation to possible isolates, which do not express the capsular gene (often referred to as NT isolates or unknown serotypes). Remember that PCV vaccines only protect against pneumococcal isolates that express their capsules!

5. How would the authors handle a PCV-13 vaccinated child, who a year after vaccination was brought to the hospital with an IPD due to pneumococcal
infection, the authors tested the isolates with their molecular techniques and the test results showed serotype 19F?. Would this be enough information to go to the Vaccine Company and claim that their vaccine does not protect against 19F, or would an additional test be required to show if the capsular gene was expressed?

6. My English is not sufficient to evaluate the quality of the overall written English; however, there are several examples where the sentences do not make sense! One example is on page 6 line 109 to 111. I therefore suggest a thorough editing of the language.

Minor Essential Revisions
7. Page 5, line 78: what does “first-tier” mean?
8. Page 11, line 224: what does sequetyped mean?

Discretionary Revisions
9. On page 12 line 226, the authors mention the possibility of new serotypes. What do the authors find that defines a new pneumococcal serotype?

**Level of interest:** An article of importance in its field

**Quality of written English:** Not suitable for publication unless extensively edited

**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