Author’s response to reviews

Title: Host genetic variability and pneumococcal disease - A systematic review and meta-analysis

Authors:

Anne Kloek (a.t.kloek@amc.nl)

Matthijs Brouwer (m.c.brouwer@amc.nl)

Diederik Van De Beek (d.vandebeek@amc.uva.nl)

Version: 1 Date: 25 Jun 2019

Author’s response to reviews:

Thomas H. Hampton, Ph.D (Reviewer 1): I really enjoyed reading this. This is well thought out, executed, relevant and well articulated.

Please address the following concerns:

The authors conclude:

"Studies have identified several host genetics factors influencing risk of pneumococcal disease, but many result in non-reproducible findings due to methodological limitations."

"In conclusion, several host genetic polymorphisms have been identified to influence susceptibility and outcome of pneumococcal disease, but most of these studies are hampered by methodological flaws or were not reproduced (yet). Carefully designed whole-genome association and replication studies are needed with detailed clinical meta-data to further clarify and confirm the genetic basis of pneumococcal disease."

Essentially, the authors say that yes, some of the variability in pneumococcal disease is explained by host genetic factors but that, unfortunately, methodological limitations are "getting in the way" of our ability to clarify and and confirm the roles played by host genetic factors.

Indeed, something like 60 studies known to the authors have studied this question. Looking any them together through meta-analysis, the authors show that the vast majority of the supposed connections between pneumococcal disease and host genetics are false.

Importantly, early research into the question of what role host genetic variability might play began with intense study of the genes that make the most sense to explain either susceptibility or outcomes. And these studies failed to find a consistent relationship.
While it is possible that 60 studies share design and analysis features that led them to false conclusions, an even simpler explanation is that individual SNPs explain almost nothing about pneumococcal disease susceptibility or outcomes because this is a genetically complex phenotype.


Would we predict that susceptibility/outcomes in pneumococcal disease are controlled by single genes, and that further research will identify a series of SNPs that explain most individual differences? I don't think so.

Please briefly address the possibility that factors other than those studied in this meta analysis drive much of the variation in susceptibility or outcome. I think you fine analysis is pointing that way.

We agree pneumococcal disease is a complex trait and outcome and susceptibility are probably not controlled by single SNPs or genes. We included this in our discussion (line 375-377) – with the reference as suggested.

Line 88: Citation for Review Manager

We have included the citation.

Line 89: Were both fixed and random effect models used on the same meta-analysis? Does this choice matter affect conclusions?

No we applied the model that fitted the results of the meta-analysis. Random effect models were used if effects between studies were too heterogeneous.

Line 146: I believe it is more conventional to say "neither study showed any effect" than "both studies showed no effect."

We changed the sentence of line 146 as suggested.

Line 330: I believe you mean "Few findings were replicated in independent cohorts

That is correct. We changed the sentence in line 330.

Line 351: I believe you mean "The results of our meta-analyses should be interpreted with caution because many included methodologically flawed studies

We changed line 351.
Line 355: "Many studies had a retrospective inclusion design which poses a risk for selection bias." Design is the subject here, so we need "poses".

We corrected line 355.

Line 374: "In the last years many loci are identified by GWAS." Try "In recent years, many loci have been identified by GWAS" to match tense in the rest of the sentence.

We changed line 374 as suggested.

Mahamadou Diakite, PharmD, DPhil (Reviewer 2): This is a good review paper on susceptibility genetic to pneumococcal disease. It's well written with appropriate references.

Mohammad Reza Jabal Ameli Forooshani, PhD (Reviewer 3): Kloek et al. provided a comprehensive and well-written meta-analysis of host genetic variability in the context of pneumococcal meningitis. Through an in-depth analysis of published literature during the past 35 years (1983-2018), they have demonstrated that the majority of identified genetic variants are found in genes involved in host immune response to microbes. Based on their meta-analysis, they identified that variants on CD14 and MBL2 genes significantly associate with pneumococcal disease.

The manuscript follows a structured and detailed review of association signals involved in individual components of the host immune system. The method section thoroughly addresses the potential concerns regarding the aggregation of data from diverse resources and Importantly, the limitations of their meta-analysis are adequately discussed in the final section. Nevertheless, there are few comments that if adequately addressed will increase the readability of the paper and addresses minor scientific flaws.

Comments:

1. Figure 1 resolution does not meet the standard requirement of a high-quality paper. Please consider providing a high-resolution figure according to the journal instructions.

In the re-submission we included our high-resolution PDF version.

2. In multiple occasion, authors referred to association studies that carried out on white patients (for example page 4, line 46). Although colloquially correct, the term "White" does not appropriately indicate a specific ethnicity. Please consider replacing it with more appropriate terms such as "Caucasians" or "European descendant" patients.
3. PCR is a molecular technique for amplification of target sequences and it does not reveal the genotype unless coupled with other molecular methods such as RFLP. Hence the use of PCR for determining genotypes (Page 5, line 12 & page 13, line 22) is not correct. Provide the correct method used for genotyping of the patients.

Thank you. The methods for determining genotypes were heterogeneous so we summarized the genotyping as “PCR followed by various methods of allelic discrimination”. (lines 120 and 366)

4. Tables in figure 2 (Forest plots derived from meta-analysis) are not consistent. Some tables have an extra column indicating the year of the study while the year column from other tables missing.

We changed the tables.

5. Please consider using figure sub-legends for figure 2 and reference to the relevant plot in the "candidate gene approach" section.

The suggestion of the reviewer is not completely clear for us.

6. Page 6- line 46 "In the meta-analysis…": Please consider using appropriate punctuations in a compound/complex sentence.

We corrected the punctuations (line 168-170).

7. Figure 3- Please add the takeaway message from the funnel plot before listing abbreviations.

We added the takeaway message from the funnel plot.

8. Please add HGVS representation of the SNPs to the figure legends: for example; rs1800450 (NM_000242.2(MBL2):c.161G>A (p.Gly54Asp).

We added the HGVS representation of the SNPs in figure 2 and 3. We could not find all HGVS representations for all SNPs. We suggest to remove the HGVS representations in figure 3 to make the lay-out more uniform.

9. Please correct the typo in the spelling of word "independent" in line 12, page 13.

Corrected.
10. A significant association signal does not imply a causal relationship and in fact, due to LD, the causal allele might be up to few Kbp apart from the association signal. Identification of the causal allele requires the application of fine-mapping methods. Please consider paraphrasing the last sentence in line 48 of page 12 to reflect the correct scientific inference from the mentioned results.

We rephrased the last sentence of page 12 (line 348).

11. Please consider including a paragraph in the discussion section of the paper reflecting on the availability of low-cost high-throughput sequencing in the clinic and its implication on the susceptibility gene finding and improved diagnosis of patients at risk.

We feel that that this is beyond the scope of this ms, but are happy to add if needed.

Overall, I believe the current manuscript is a timely and well-written review of susceptibility genes in the context of pneumococcal disease and will attract the attention of experts in the field. Provided that authors adequately address the aforementioned comments, I recommend this paper for publication.