Author's response to reviews

Title: Dynamic models of pneumococcal carriage and the impact of the Heptavalent Pneumococcal Conjugate Vaccine on invasive pneumococcal disease.

Authors:

Alessia Melegaro (alessia.melegaro@unibocconi.it)
Yoon H Choi (yoon.choi@hpa.org.uk)
Robert George (robert.george@hpa.org.uk)
John W Edmunds (john.edmunds@lshtm.ac.uk)
Liz Miller (liz.miller@hpa.org.uk)
Nigel J Gay (herdimmunity@yahoo.co.uk)

Version: 2 Date: 22 December 2009

Author's response to reviews: see over
Dear Editor,

RE: Manuscript ID 1976914097304906

*Dynamic models of pneumococcal carriage and the impact of the Heptavalent Pneumococcal Conjugate Vaccine on invasive pneumococcal disease.*

Thank you for your kind consideration of our revision of this manuscript. We have made changes to the text in response to the comments from Reviewers.

The revised manuscript and our response to the comments raised by Reviewers are enclosed. The changes made to respond to reviewers’ comments are marked in bold in an Additional material file, “Melegaro 2009_cover letter_resubmission.doc”.

We look forward to hearing from you.

Yours faithfully,

Alessia Melegaro
Yoon H. Choi
W. John Edmunds
Liz Miller
Robert George
Nigel J. Gay
Answers to Reviewers report

Title: Dynamic models of pneumococcal carriage and the impact of the Heptavalent Pneumococcal Conjugate Vaccine on invasive pneumococcal disease.

Version 1. Date: 2 November 2009

Reviewer 1: Michael Haber

Major Compulsory Revisions

1. Two new papers on modeling the effects of the PCV have recently been published:

(2) S.J. Snedecor et al. – Vaccine 27:4694-4703 (2009).

The authors should detail the similarities and differences between their model and the two earlier models.

The two references have been added and the differences with the current model explained. See added sentence.

Our work is also different from other recent and more applied models which looked at the effects of infant vaccination on the epidemiology of S.pneumoniae infection, by using pre and post invasive disease surveillance data from, respectively, Australia and the US [27,28]. Neither of these two studies, however, looked at the impact of vaccination on the ecology of the bacterium and, in particular, on the possibility that non-vaccine serotypes will replace serotypes eliminated by the vaccine. This phenomenon is currently being observed in England and Wales post vaccine introduction (www.hpa.org.uk) and it is having major implications for the overall impact of the programme.

Minor Essential Revisions

Page 2 (abstract – results): ‘… infant vaccination was estimated to prevent 39,000 IPD cases in the UK …’. I could not find any mention of this finding in the Results Section.

We included this in the Results section now.

Simulation results indicated that all of the vaccination strategies considered in this paper were sufficient to eliminate VT transmission in England and Wales with base case values for vaccination coverage and other parameters. Although the model predicts some replacement of VT with NVT, the annual total of IPD cases was reduced from 6,184 to 3,765 (38% reduction) in the long-term (20 years after the PCV7 introduction) and reduction of 39,000 IPD cases accumulated over the period.

Page 7, line 19: should become .
Uninfected, unprotected (S) individuals become infected according to the age-specific force of infection of vaccine type (λVi) and non-vaccine type (λNVi) pneumococci.

Page 8, Recovery Rates: It is assumed that vaccination does not affect the length of carriage. Is this assumption justified?

We assumed that the vaccine only affects the protection against developing disease and degree against acquisition. We used the assumption no vaccine effects on the length of carriage due to lack of relevant data on this.

Page 9, bottom: The case:carrier ratios are very important parameters. They should be displayed and included in the sensitivity analysis.

We included a table of Case:Carrier ratios (Table 3) as suggested.

Pages 10-12: References to figure number are incorrect. For example page 10, line -3 ‘Fig 6’ should be ‘Fig 5’.

We corrected as suggested.

Reviewer 2:

No changes requested

Reviewer 3: Kari Auranen

Major compulsory revisions

(1) It is not clear what are the steps were actually taken and in which order when estimating the model parameters from empirical data. In particular, how many parameters were estimated from data altogether? Were all parameters estimated at the same time or in sequence? How were the six best parameter combinations identified (Table 2). As a minor remark to the last of these questions, it is not clear how the deviance per se should be interpreted here (it is always the different of deviance values rather than any particular absolute value that should count in model choice). I would suggest a better explanation of the steps in estimating the model parameters. This could be given in the Methods section or in the Appendix. Please check that the corresponding paragraph in the Results section corresponds to what is being said in the Methods.

We thank the reviewer for pointing this out. We added the following paragraphs in the Model Analysis section to clarify the steps involved in estimating the model parameters:
Model analysis

Three models were run: one pre vaccination model to identify steady state values for model parameters, one post-vaccination model to estimate key vaccine parameters, interaction between VT and NVT and the level of assortativeness of mixing pattern and finally the prediction model to assess the impact of vaccination in England and Wales.

The pre-vaccination model was programmed in Excel to estimate the age-specific force of infection for, respectively, VT and NVT using the carriage data and, consequently, the case:carrier ratios by fitting the new infections to the UK pre-vaccination IPD data by minimising the Poisson deviance using the SOLVER. This model also generated the fully assortative and fully proportionate mixing matrices between six age-groups.

A transmission dynamic model was programmed in Berkeley Madonna (R. I. Macey & G. F. Oster, UC Berkeley, CA, USA) and fitted to the pre- and post- vaccination IPD data from the US to estimate degree and duration of protection of the vaccine against invasive disease, the competition parameter (cN), and the mixing parameter (ε).

Once the parameters were generated, the epidemiological model was also used to produce post vaccination predictions for England and Wales. The system was solved using the Euler method to integrate ordinary differential equations with fixed time steps of 0.001 years: The model simulated 50 years (five years pre-vaccination and 45 years after vaccination).

(2) The notation is not always consistent. For example, the sub-index denoting the individual being vaccinated is sometimes “V” and sometimes “ν” (see e.g. the equations in the Appendix I). There are more comments on this below (see ‘Minor comments’).

It was not $\nu$ but $V$. The italic presented a bit strangely. Notations have been corrected.

(3) The use of the term “validation” is not clear. For example, it is not clear whether parameter estimation, as explained in the first paragraph of the Results section, should be called validation at all. I understand a set of crucial model parameters were estimated, based on the US pre- and post-vaccination data on IPD and the E&W data on carriage. But clearly this is not validation, rather “estimation” of “model calibration”.

The title of the paragraph has been changed into model calibration

(4) How sensitive are the results to the assumption that serotype 6A behaves like a vaccine type? No references are given about the effect of PCV on 6A carriage (references [34] and [35], p.7, line 5). I’m not sure if the latter of these actually discusses carriage at all (but AOM).
We have included additional references to the sentence which show that serotype 6A cross reacts. We have also included an immunological study which shows good antibody response.

Though 6A is not explicitly included in PCV7, carriage of this serotype is assimilated to the VT one as PCV7 was shown to elicit strong opsonic capacity against serotype 6A [36] and a significant reduction of 6A carriage was observed following vaccination [13,37-39]. In the model, carriage of any of the remaining pneumococcal serotypes is considered a NVT carriage episode.

(5) In the Discussion it is said that the cause of the difference between the model predictions and the recent data about the extent of replacement in IPD in the UK needs to be studied further. What do you mean by saying that this could be “due to vaccine introduction with or without a natural increase in NVT”? Do you refer to calendar time trends as a possible explanation here?

The sentence has been changed as follows

Based on the UK surveillance after the PCV7 introduction, the replacement was much bigger than the replacement in the US (http://www.hpa.org.uk) and the cause of this difference is currently being investigated.

(6) p. 14, line 1-2. It is well known that differences in the mixing pattern may have much more dramatic effects on model predictions than any statistical parameter uncertainty. Do you actually mean that the strength of the analysis was to estimate the one weighing parameter from empirical data (and not refer to statistical uncertainty)?

The sentence has been changed as follows

Here, although we make a strong assumption regarding the mixing pattern (reducing it to a single parameter, $\varepsilon$), we reduce the number of unknowns by using post vaccination herd immunity effects to inform the level of assortativeness of mixing. Alternative approaches to measure age-specific contact patterns directly through diary-based methods are currently being explored using recently collected contact pattern data [45-48].

Minor revisions

Abstract

• What time frame does the decrease of 39,000 cases refer to? How do these figures relate to those given in the Results (page 11, the last paragraph)?

In the first twenty years. We deleted a word 'and' to clarify that.

• p2, line -6, the last sentence is a little obscure. In what sense was ‘uncertainty reduced’ (e.g. with respect to what). Also, the claim about improved validity seems to be a comment which
is more suitable to the discussion part of the article. This does not mean that I objected the idea to use empirical data to estimate model parameters!

We believe that using post-PCV7 introduction data from the US allowed the estimation of key parameters which otherwise should have been assumed. However we see the reviewer’s point on the fact that this is more of a comment and thus we have deleted the sentence from the abstract.

• p3, line 4, there is something wrong in this sentence (“… and to then being able…”)

We changed the sentence accordingly

The techniques developed here can be used to assess the implementation of vaccination programmes in developing countries and to evaluate the cost-effectiveness of alternative scenarios.

Materials and methods

• Legend to Figure 1: notation ‘VT’ should be explained here.

We changed the last sentence as the following to explain VT:

One serotype, 6A, is included in the vaccine serotype group due to evidence of cross-protection.

• p.6, England & Wales IPD data: the total number of isolates should be mentioned

We added the following sentence:

The average annual number of IPD cases during the pre-vaccination period (between 2003/04 and 2204/05) in the England and Wales was 6,088.

• Legend to Figure 2: notations ‘VT’ and ‘NVT’ should be explained in the legend. The same applies to other figure legends, i.e., some of the notations should perhaps be written out, despite they had been defined in the main text.

Changes as follows:

Serotype distribution of invasive pneumococcal disease (IPD) isolates reported to the national enhanced surveillance system of invasive pneumococcal disease for England and Wales. IPD cases in the vaccine serotype group are shown in black bars and IPD cases in the non-vaccine serotype group in white bars. Data shown refer to the years 2004 and 2005.

• p. 6, the last paragraph (‘Population’): What is the mortality rate, i.e. how does it depend on age (class)? What is the age structure of the population?
We added the following sentence at the end of the paragraph.

The population sizes for each annual age cohort from the National Statistics Office were used to calculate the IPD cases in each age cohort.

- Figure 4
  o Both arrows between compartments $V_i$ and $B_i$ point to the same direction towards $B_i$. The other of these should be reversed.
  
  It is now reversed.

  o Subindex ‘i’ is missing in the clearance rates of the non-vaccine types, i.e., should it not read $r_{[N_i]}$? There is also inconsistency in this notation in the main text (check page 7)

  They are corrected as suggested.

- p. 7, line -11, ‘V’ and ‘N’ are obviously sub-indices to $\lambda$?

  They are corrected as suggested.

- p 7, line -10, there is misplaced comma between $c_N$ and the ‘is’

  It is deleted as pointed.

- p 7, line -6, the clearance rate $r_i$ should be italicised. Similar mistakes occur also later (at least on p. 8, line -7).

  They are corrected as suggested.

- p.7, line-2: The percentage reduction in the rate of acquisition of carriage, due to the individual begin vaccinated, is called ‘degree of protection’ here. In the legend to Figure 4, I understand the same parameter is called ‘vaccine efficacy against VT carriage acquisition’. Could either of the two used consistently?

  We will use as "degree".

- p. 8, line -10: Does it mean that the inverse of the rate of clearance (“the mean duration”) was 72 days in children 0-1 years of age etc?

  Yes, it is.

- p. 8, line -8. I understand the work ‘infection’ here, but maybe it is clearer to use ‘carriage’ or ‘carriage episode’ also here.

  Changed to "carriage" as suggested.

- p. 8, line -4, index ‘j’ should be italicised

  Changed as suggested.

- p.8, the equation on the last line, should read $\lambda_{V_i}$ (i.e. subindex ‘i’ to the sub-index $V$)
Changed as suggested.

- I understand the parameterisation and estimation of the parameters are explained in the one section starting on page 8 (in Material and Methods) and then in the first section of the Results. While I like the approach to explain which data sets carried information about which parameters, I don’t find it easy to follow how different data sets were used in combination or successively.

  o For example, I gather that the force of infection was based on the E&W data and the mixing matrix on the US data, but I think this is not clearly stated in the main text. Were there parameters estimated simultaneously?

  o How was the E&W data used in the estimation of the case:carrier ratios? How about the US data? Was either of the two IPD data sets used for validation or both for estimation of the parameters?

  o Were the case:carrier ratios estimated simultaneously with the “beta” matrix and other carriage-related parameters?

  o In summary: How many parameters were estimated altogether. Were all parameters estimated in one step? How were the two data sets used in combination?

Detailed explanations are now in the "Model analysis" section before "Results".

- p. 9, the middle paragraph: What do you refer with 'Poisson deviance' to? Is this simply a deviance, based on Poisson likelihood for observed IPD counts?

  Yes.

Results

- p 10, line -3, it should obviously read ‘Fig. 5’ (instead of Fig. 6)

  Changed as suggested.

- p. 10, line -2. What does it mean that the sensitivity analysis was conducted to “study uncertainty of different number of PCV/ dose effects?” How was this done, otherwise than changing the PCV/ coverage (which seems a different thing to me)?

The following sentences have been changed and new text included.

Graphical comparison between the observed and predicted VT and NVT IPD cases by age groups and for the 8 counties where the Active Bacterial Core Surveillance data collected the serotyped IPD data between 1998 and 2004 is presented in Fig. 5 where the baseline value of the annual PCV7 coverage was set at 86% with 1+ doses. Sensitivity analysis of the annual PCV7 coverage level was conducted to study the uncertainty around the degree of vaccine protection and the number of doses given. In particular, the effect of increasing the vaccine coverage level to also include those individuals that were only partially vaccinated and that, for this reason, were not included in the base case scenario (being 86% the vaccine coverage for fully vaccinated infants) was considered as well as reductions in vaccine coverage to explore the effects of lower protection levels.
• “Figure 6” should read “Figure 5”. In general, the enumeration of figures is shifted from this figure onwards. Please correct.

Changed as suggested.

• The legend to this figure uses abbreviation ABC, which might not be obvious to all readers, despite it being defined in the main text.

Explanation is now included in the legend as suggested.

• It is not clear how many and which parameters were estimated at a time, and what level of the others were assumed (see my respective comments to the Methods section above)

Detailed explanations are now in the "Model analysis" section before "Results".

• What were the actual case:carrier ratios? This would be an interesting piece of information, and also provide some information about how the model accounts for pneumococcal epidemiology.

We have added the case:carrier ratio in Table 3.

• Related to the question above, were the model validated against the pre-vaccination IDP data from E&W or how was this piece of data used?

In the prediction model the initial steady state was validated against pre-vaccination IPD date for the UK. The E&W pre-vaccination data was also used to estimate the CC ratios for the E&W.

• p. 11, line 10, it should obviously read ‘Fig. 6’ (instead of Fig. 7)

Legend to Figure 6: Were there only four parameters (\(\gamma, \epsilon, c_N\) and \(c_V\)) that define the joint likelihood for the number of IPD cases? How about the waning rate for protection (cf. Table 2), or other parameters?

We only fitted four parameters (\(\gamma, \epsilon, c_N\) and waning (not \(c_V\)). We fixed \(c_V\) to 0.5 as it did not show any significant difference from two extreme values (0 and 1).

• p 10, line 11, confidence interval area? should be confidence area?

Changed as suggested.

• p. 11, line 16: Do you mean that “…predictions were insensitive to the assumptions”?

The paragraph is now changed to the following:

Seven pairs of the degree and duration of PCV7 vaccine protection were chosen within the 95% confidence area (including the best fitted pair with the baseline US PCV7 coverage) and showed that the UK transmission model predictions of IPD changes are insensitive to the assumptions. The fitting results indicated that the higher the degree is the shorter the duration of protection. The baseline parameters used for the following simulations of the England and Wales predictions are the ones estimated when fixing the coverage level to 86% (Table 1). The Case:Carrier ratios for the UK (Table 3) show that the chance of developing IPD with NVT can be twice as high as the one for VT.
• p. 11, the same paragraph: It is not clear why you say “thus the baseline parameters…”.

It is also changed to:

The baseline parameters used for the following simulations of the E&W predictions are the ones estimated when fixing the coverage level to 86% (Table 1).

• p. 11, line -7, Figure 8 should be Figure 7

Yes. Corrected as suggested.

• p. 11, line -4, does the 63% refer to reduction in the accumulated number or the annual incidence after the “system” has settled to a new equilibrium?

Yes it reached to the equilibrium point after twenty years from the PCV7 introduction.

• p. 11, bottom line, Figure 9 should be Figure 8

It is an accumulated IPD cases in every five years. The legend in the Figure 8 is now changed to clarify that.

Estimated number of accumulated IPD cases in every five years of 20 years after PCV7 introduction using the three alternative vaccination strategies.

• p. 12, line 6, Figure 10 should be Figure 9

Changed as suggested.

Discussion

• p 13, line 1, It is said that the reduction in the annual IPD number if 3,700 in the long-term. How does this relate to the 2,300 prevented cases in the long-term (see p. 11, line -2)?

6000 and 3700 are the number of cases, respectively, pre and post vaccination. The reduction is around 2300.

Appendix 1

• The subscript for the vaccinated susceptible is $\nu$. In Figure 4, $V$ is used.

It was not $\nu$ but $V$. The italic presented a bit strangely. The Figure 4 now has all in italic.

Please be consistent in this notation.

• The term $\pi_i(t)$ is clearly a rate, not just ‘function’

The term $\pi_i(t)$ is actually a time-dependent variable since the vaccine coverage changes from Year 1 to Year 2.

• The $\beta_{ij}$ is now explained, without reference to the force of infection. A reference to the main text would be appropriate here.

The beta is now removed from the Figure 4 since they are shown in the equations.
• Subscripts $\nu$ and $V$ are used interchangeably to denote vaccination. See also p. 8, the equation on the last line.

It was not $\nu$ but $V$. The italic presented a bit weirdly. The Figure 4 now has all in italic.

Appendix 2

• It seems that the competition parameters were estimated first. How was this done without knowledge of the beta parameters? Shouldn’t they be estimated simultaneously when applying the model to the US data?

It is very difficult to estimate the vaccine related parameters and a competition parameter together. We first fix a competition parameter to estimate the force of infection parameters, CC ratios and beta matrices then used this information to estimate the best fitting set of the vaccine related parameters. We continue to do this with other competition parameter values to find the best fitting competition parameter to explain the US surveillance data (a grid search with a full range of competition parameter, cN, from 0 to 1 with a 0.01 increment).