Reviewer Response

Dear Editor,

Thank you for sending review reports for our manuscript. The reviewer’s comments were useful and have greatly improved the paper.

We enclose a revised version of the manuscript (clean and with tracked changes), which addresses the reviewer’s comments and suggestions. The reviewer’s comments are shown below with our response to each comment in italics.

Yours sincerely,

Sam Abbott

## Authors’ response to Editorial Requests

GENERAL

* In both the abstract and the main text please define all abbreviations at first use (e.g. UK; gradient boosting method – is this GBM?; 95%CI – is CI confidence interval? Credible interval?; MNAR in the discussion – is this missing not at random mechanism?). For Tables and Figures, please define all abbreviations used in a footnote below the Figure or Table.
* If possible, please use square brackets for the references with the punctuation following the reference.
* Tables should not have any empty cells as design element or because information is not available (‘NA’ (subsequently defined as not available or not applicable) can be used for example).
* For all Figure and Table titles, when possible, please indicate at the end of the title, the geographical area, period and total number of items covered by the illustration. E.g. Table 1: Summary of relevance and eligibility criteria for each cohort, United Kingdom, 2000-2010 (n = xxx individuals)
* Please check the overall text versus the tables for the name of the intervals (e.g. 95% confidence intervals vs 95% credible intervals). To give an example: In the abstract and in the main text subsection called ‘Adjusted estimates of the effects of the change in policy on school-age children’ there is ‘IRR: 0.74 (95%CI 0.61, 0.88)’ but in the table 2 there is ‘0.74 (0.61, 0.88)’ defined as IRR (95% CrI) CrI: Credible interval.

ABSTRACT

* While keeping the word count below 250 words, please add an ‘aim’ subsection, describing briefly the aim of the paper, and please specify the study type (e.g. observational study?) also in the abstract (either in the aim or the methods).

METHODS

Data sources - Does the Labour Force Survey record only the number of adults of working age (employed or unemployed) in the country, or does it record more people than these (e.g. their family members such as children)? Please describe the population and demographics recorded in the Labour Force Survey in more detail.

Statistical methods

* While it is mentioned in the methods that 95% confidence intervals were calculated around the IR values, it is not described what the intervals are around the IRRs. Please check the earlier editorial comment about the intervals that are presented in text and tables around the IRRs.

Ethical statement

* An ethical statement should be included (<https://www.eurosurveillance.org/for-authors#ethical%20statement>); if ethical approval was not needed, this should be stated in the manuscript. In this case, if possible, it should be please explained why.

RESULTS

= ‘Of notifications with a known birth status 27% (29096/106765) were UK born, in comparison to 33% (2634/8055) in cases with an imputed birth status.’ Just for editor’s understanding: does this mean that ‘Among notifications with a known birth status, 27% had the status ‘born in the UK’, while among notifications with an imputed birth status, 33% had the status ‘born in the UK’?

## Authors’ response to Reviewer #1:

This is an interesting paper in which the authors address a key TB control issue which is measuring the impact of BCG policy change in the incidence of tuberculosis in children. Whether and when targeted BCG vaccination should replace universal vaccination depends on several factors and may be a matter of discussion within national TB control programs, although in Europe most countries have a targeted BCG policy. The question is worthwhile investigating and discussing. I suggest publication in Eurosurveillance provided that the minor changes suggested are taken into account. HOWEVER, given the importance of the statistical methods in this analysis, which I do not feel confident to analyse, I recommend that a tough review of this part, and the results, be made by a statistician, which opinion is crucial to decide whether the paper is able to be published or not.

Introduction

1. Paragraph 2. In France, universal vaccination was implemented in all children before entering “community care” and primary school at the latest, and not in neonates. Methods
2. Paragraph 1. The geographical scale of TB local incidence is not specified and is an important piece of information that would be worth mentioning at this stage.
3. Paragraph 2. Not clear whether the LFS which reports “employment circumstances” of the UK population reports also data in children, who are not employed. We assume the answer is yes which should be probably clarified.
4. Definition of cohorts. Not totally clear to me. Example for comparison cohort 2: a child born in 2000 and who is diagnosed with TB at 2 years old (2002): the vaccination programme in 2002 is universal not targeted, as seen in Table 1. Maybe further clarify this crucial point.

Statistical Methods

This is a key point that should be reviewed by a statistician; I do not feel confident on giving my opinion on this.

Results

1. No comments, except that I find some parts of the results difficult to understand for a “common” public health professional" with no deep knowledge on statistics and modelling methods.

Discussion

1. BCG protects mainly against severe disease (meningitis). This point is not discussed at all. An additional analysis could have investigated the impact of policy change on the incidence of meningeal TB in children, which could have broaden the consequences of policy change in the epidemiology of TB in children. Maybe a word on this should be said in Introduction or Discussion.
2. Other interventions may decrease TB incidence, such as a better screening, a change in screening policy in migrants, a better case management…This is not discussed. Could these potential interventions have also played a role in the results presented here?
3. No figures of BCG vaccination coverage trends in targeted groups are given (except the 68% mentioned in the Discussion). This determines to a large extent the impact of vaccination. A brief description of this could be given here, or in Introduction.
4. BCG has an impact mainly in severe forms of TB, and prevents about 50% of pulmonary TB in children. But children, to my knowledge, do not frequently transmit TB. So BCG has rather poor effect on TB transmission. The authors speak in their discussion on indirect effects and ongoing transmission. Is this really the case with BCG and TB?

## Authors’ response to Reviewer #2:

Thank you for providing an excellent overview on the effects of the policy changes in BCG vaccination in England.

## Authors’ response to Reviewer #3:

This study is of high relevance, not just for the England, but for the global tuberculosis community: The UK (England and Wales notably) had been one of the few countries with a universal vaccination of adolescents and a clinical trial demonstrating high efficacy for at least 20 years in the age vaccinated groups and beyond to include young adults during follow-up.

From the switch from universal adolescent vaccination to selective (largely the descendant newborn of foreign-born ancestors) one would thus hypothetically expect an increase in the incidence rate among the former and a decrease in the incidence rate among the latter (among the respective cohorts in both groups), if confounding resulting from secular trends could be properly controlled for. This study aimed to do precisely that.

The authors created 8 cohorts, each for a 5-year follow-up, using tuberculosis surveillance data for the numerator and Labor Force Survey data for the population denominator. These procedures are all clearly explained. As it gets substantially steeper and difficult to follow for the more general user’s consumption is when the models are elaborated in the same paragraph. Perhaps it would help to split the elements of describing the basic epidemiology of incidence rates and ratios from the latter into two separate paragraphs, so that those unfamiliar with “negative binomial models”, “penalised by model complexity”, and the intricacies of Bayesian estimation with Markov Chain Monte Carlo, etc. This would clarify that here is a basic methodology, and here is a complex modeling description, the former for every epidemiologist, the latter only for the seasoned expert in statistical modeling using special techniques. I concur that it should and has to be there, but the more basic working techniques of calculation might be preferably “visually” separated from complex modeling techniques, notably when unexplained acronyms such as “GBM” are interspersed that do not help in easily following what the modeler actually did. Nor does it become that clear why exactly these models were chosen and not something else. The reader should probably have some inkling what the motivation for model selection is.

In any case, it becomes again clear that the procedures resulted in obtaining adjusted estimates of the effects of the policy change from universal adolescent to selective neonatal vaccination. The effect was measurable but seemingly modest. Apparently, the model required adjustments by including the change in policy, age, incidence rates in the non-UK born, and year of eligibility as a random effect. To the non-initiated, all these elements make intuitively sense and reflect the complexity of interplaying variables that all played a substantial role in what was finally observable. Such complexity may also contribute to variability and thus uncertainty about the unbiased estimation of the magnitude of the observable effect. However, the authors later also state that all models (does this refer only to the non-UK born?) that included at least age in the adjustment resulted in comparable effects. As parsimony is king in all modeling, perhaps it could be stated more explicitly that always the model with the minimum required variables was used and perhaps why recourse was taken to retaining more complex models if a more simple one gave actually comparable results. Perhaps I am missing here something, but then it would indeed be useful to review the sum of the modeling exercises to the point.

In table 3, among the other variables, is also the year of study eligibility (please correct spelling). The point estimate incidence rate ratio varies from 0.83 to 1.12 (UK born) with no discernible trend. Is this therefore an expression of the intrinsic uncertainty of the effect of the model? If so, then perhaps the authors would have to be more explicit that there remains substantial uncertainty in the modeled effect as small as it may be. Thus perhaps the variability in the estimated magnitude of effect for cases prevented in the following paragraph which is substantial is therefore not all that meaningful after all (CrIs not uncommonly overlapping zero).

Thus, in the end in the discussion, the authors speak - statistically probably correctly - of “evidence of an association”, but perhaps disappointingly it is not as clear-cut as one may have hypothesized (including this reviewer’s pre-reading subjective gut feeling). Perhaps there is more not easily definable confounding that couldn’t be addressed in the model. Because it seems indisputable that at least the BCG / Vole bacillus UK trial had given unusually solid long-term protection against pulmonary tuberculosis among young adults and it remains baffling that stopping this scheme of adolescent vaccination should have left such little effect. Would the reasons for this phenomenon deserve a bit more extensive addressing in the discussion? That BCG vaccination has given various levels of protection (say, nothing in Chingleput) seems entirely irrelevant for the UK, here it gave a non-varying high-level of protection for about ten to 12 years before it precipitously dropped towards zero. This fits with what we know from effector cell immunity and it is not clear why this should not also be reflected more substantially in the period following policy change unless the models couldn’t really account for all factors that matter, but it is surely not variability in vaccine protection elsewhere that is an issue here, this is only about the UK after all.

I appreciate the work of the authors with these thought-provoking study results!

Minor comments:

Several references have capitalization vs lower case issue be in names or e.g. “bcg”, and also whether words in the title should be capitalized or not, etc.

References 7 and 8 are duplicates

Reference 11 is a bit of a screw-up (2018 and January-2011)