Reviewer’s report

Title: Benefits and harms of the human papillomavirus (HPV) vaccines: systematic review with meta-analyses of trial data from clinical study reports

Version: 0 Date: 10 Jul 2018

Reviewer: Beate Wieseler

Reviewer’s report:

General remarks

The manuscript reports a systematic review based on an information retrieval using sources beyond journal publications and study registries. Therefore, in addition to the results on the benefits and harms of HPV vaccination, this review is relevant for the general discussion of sources and methods for systematic reviews.

The manuscript shows the potential of systematic reviews based on CSRs, specifically concerning the level of detail for study methods and harms data. The fact that the authors did not succeed in receiving the CSRs of all studies in the study pool underlines the need for further discussion on availability of these documents.

Overall, the reviews adds to the information available on benefits and harms of HPV vaccines. Differences in the outcomes of this review as compared to reviews based on journal publications and registry reports need to be discussed.

The reader might at some points of the manuscript benefit from additional information on the methods or on the data reported (please see comments below).

Page 4, lines 94-96, Methods, search strategy

CSRs were requested from EMA and GSK. It remains unclear why no CSRs were requested from other companies, when part of the eligible studies were sponsored by other companies. If no CSRs were requested from other companies because the CSRs received from EMA for MSD-studies were considered complete this should clearly be stated. Please clarify.

Page 4, lines 97-100, Methods, study eligibility

While the inclusion criteria are clear from the referenced review protocol, I would recommend that they should also be clear from the manuscript itself. So far it is not easy to understand from the methods section that availability of a CSR (or a comparable document) is a requirement for inclusion of a study in the analyses.
Page 5, lines 127-131

The comparison of the VigiBase harm clusters with the CSR data is difficult to understand. I understand, that you identified the 3 largest harm clusters from VigiBase, then identified the corresponding PTs in the CSR data sets, summarised them by treatment group and compared them between the treatment groups of the included studies. Thus the analyses presented in the manuscript do not include any VigiBase data. Is this correct? Please clarify.

Page 6, lines 141-147, Data synthesis and analysis

Overall it is difficult to understand from the text and tables, if and when the authors are referring to events and if and when they are referring to patients with events (also concerning the risk ratio analysis). Please clarify throughout the manuscript.

Page 6, lines 151 and 152, Characteristics of included trials

In general, there remain a number of open questions on the availability and completeness of the clinical study reports (CSRs) used in the review. First of all, it is unclear, why not all CSRs of eligibly studies were available. The authors describe that they requested CSRs of eligible studies. Have there been any reasons given by the companies why the CSRs were not provided? Furthermore, it is unclear how the pool of eligible studies relates to the pool of studies provided by / available at EMA, e.g. if EMA did not have the full study pool due to the timing or scope of their review. This relates also to the discrepancies of pages per CSR obtained from EMA and the manufacturer.

Please clarify the transmission of CSRs related to the requests and if there were any reasons given for not providing the (full) CSRs.

While this might not be the focus of this manuscript the authors might want to consider including a brief section on availability of CSRs for this project in the discussion section.

Page 6, lines 152-158, characteristics on included trials

To gain an overview of the study pool it would be helpful for the reader to have a summary listing including e.g. study ID, sponsor, vaccine, comparator, number of patients and follow-up period per study at the beginning of Additional file 2. Please consider providing such an overview.

Page 7, line 168

Please briefly clarify for the reader why an active comparator in the case of vaccination studies constitutes a high risk of bias.
Page 7, lines 175-185, Results on benefits

The meta-analyses show substantial heterogeneity. Please comment on potential sources and the consequences for interpretation of the results.

Page 7/8, lines 188-199, Results on harms

Part of the meta-analyses show substantial heterogeneity. Please comment on potential sources and the consequences for interpretation of the results. Please clarify throughout the text which numbers are based on events and which are based on patients with events.

Page 8, lines 202-203, post hoc analyses

The sentence on association of pharmacovigilance data with harms data from the study is difficult to understand. Please rephrase to clarify (please also see comment on corresponding methods section).

Page 8, lines 211-216, subgroup analyses

Is subgroup information by age available for serious neurological harms and for CRPS and POTS? Please clarify and add if possible.

Page 9, lines 223-227, Discussion

Please add information on serious neurological harms in younger participants, if possible.

Page 9, lines 237-242, Limitations

I agree that the incomplete trial data access is a major limitation. Please provide additional information on reasons of this incompleteness. Please expand on possible consequences of this limitation. Should this prevent systematic reviews based on CSRs or should there be any action to prevent insufficient data access?

Page 9, lines 247-249, Limitations

The statement on countries implementing Gardasil 9 seems vague. Do the authors suggest any conclusions from the available data across the various vaccine types?

Page 10, lines 250-253, limitations
Please comment on the consequences of heterogeneity for the interpretation of the results.

Page 11, lines 291-292, Limitations of harms assessment

The section is difficult to follow. Please include the number of patients in the studies reporting the complete treatment period and those not reporting the complete treatment period (I assume this is the comparator). Please generally clarify which number is referring to events and which to patients (with events).

Page 11/12, lines 303-315

Not all readers might be familiar with the issue of active comparators in vaccine trials. Please briefly explain the problem of saline placebo vs. adjuvant vs. adjuvant/protein as comparators and the consequences for the interpretation of the harms (and benefits) data.

Page, 12, lines 323-327

Do the authors based on their results recommend an in-depth analyses of PTOS and CRPS based on full data (which in principle are available)?

Page 12/13, lines 330-335

Please include information about the study overlap between the 2 reviews. How many of the studies in the Cochrane review have been included in the 24 studies of the current review.

Additional point for discussion

Please comment on the approach of using CSRs for systematic reviews and on potential steps to improve availability of CSRs for systematic reviews.

Table 3

The table is an important overview of the different types of harms information available from the studies in the study pool. To fully understand the possible impact of varying harms documentation in the different studies on the review results, please add the number of patients included in the different study sub-pools.

There seem to be relevant differences in the definitions for harms documentation and analysis between GKS and Merck Sharp & Dohme. I am wondering if these differences should prevent pooling of data from reports from the two companies for "new onset diseases", because these
definitions (might) differ substantially and it is unclear what is finally described by the pooled estimates. Please see comments on Tables 6-9.

Tables 4-9

Please clarify in all tables, which data refer to events and which to patients (with events); please also clarify if the percentages refer to events or patients with events.

Given that according to Table 3 data of the different categories are available from a varying number of studies, please clarify how many patients were included in each category of AEs (e.g. the categories of new medical history and systemic adverse events seem to have been collected in different studies).

The tables are partly confusing. The percentages seem to refer to the total no. of events in a category. This might be confusing, when the number of events between groups differs. (Furthermore, I would also be interested in the number and % of patients with events.) On the other hand the risk ratios consider patient numbers. E.g. in Table 4 the data on CIN2+ show 93% vs. 92% but result in a 0,81 RR (1,87% vs. 2,32 % of patients with CIN2+)?

In general, I think that in addition to events also the proportion of patients with events is important information. Please consider adding this information to the tables.

Table 6, new onset diseases

The 2 definitions used by the 2 companies and pooled in this table seem substantially different. This is also demonstrated by the different numbers of events (per patient) in the studies conducted by GSK and MSD. E.g. there seem to be roughly 8000 MSC / 32000 patients in the vaccine groups from GSK studies and 40000 NMH / 16000 patients in the vaccine groups from MSD studies. Furthermore, the most common events from the MSD studies are among those excluded from the definition of MSC in the GSK studies. Given the number of events, the analysis is dominated by events from the MSD studies anyway, in which "new medical history" is undefined. Therefore, I am wondering if a pool of events covered by these definitions is meaningful.

I agree that in addition to fatal and serious adverse events there should be an analysis of "all adverse events (excluding local events at the vaccination side)". The general harms data from Table 7 might be more meaningful.

Figure 1, flowchart

Please include the number of eligible industry trials per company in the Box stating N = 48 eligible industry trials. So far it is unclear which part of the missing studies refers to which
vaccine/company. Please consider adding a box between the eligible studies and the obtained studies to include information on the missing studies (by company).

**Level of interest**
Please indicate how interesting you found the manuscript:

An article of importance in its field that should be highlighted to relevant networks

**Quality of written English**
Please indicate the quality of language in the manuscript:

Acceptable

**Declaration of competing interests**
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.
I agree to the open peer review policy of the journal

Were you mentored through this peer review?

No