Abstract

Background: The existing Bacillus Calmette–Guérin (BCG) vaccination provides partial protection against tuberculosis (TB). The modified vaccinia ankara virus-expressing antigen 85A (MVA85A) aims to boost BCG immunity. We evaluated the animal evidence supporting the testing of MVA85A in humans.

Methods: Our protocol included in vivo preclinical studies of the MVA85A booster with BCG compared with BCG alone, followed by a TB challenge. We used standard methods for systematic review of animal studies, and summarized mortality, measures of pathology and lung bacterial load. The comprehensive literature search was to September 2014. Two independent investigators assessed eligibility and performed data extraction. We assessed study quality and pooled bacteria load using random effect meta-analysis.

Findings: We included eight studies in 192 animals. Three experiments were in mice, two in guinea pigs, two in macaques and one in calves. Overall, study quality was low with no randomization, baseline comparability not described and blinding not reported. For animal death (including euthanasia due to severe morbidity), studies were underpowered, and overall no benefit demonstrated. No difference was shown for lung pathology measured on an ordinal scale or bacterial load. The largest mortality trial carried out in macaques had more deaths in the MVA85A vaccine group, and was published after a trial in South Africa had started recruiting children.

Conclusions: This independent assessment of the animal data does not provide evidence to support efficacy of MVA85A as a BCG booster. More rigorous conduct and reporting
of preclinical research are warranted, and we believe the results of studies should be publicly available before embarking on trials in humans, irrespective of the findings.

Key words: MVA85A, modified vaccinia virus ankara, tuberculosis, animal, review

Key Messages

- In this systematic review of animal studies evaluating MVA85A vaccine to boost BCG immunity to tuberculosis, we found eight studies in 192 animals in all. Studies were underpowered and of poor quality. No effect was demonstrated on animal death, lung pathology or bacterial load.
- The largest mortality trial in macaques had more deaths in the experimental group, and was published after a trial of the vaccine in children had started recruitment.
- Whereas it is recognized that there are problems with the predictive value of some animal models, this review indicates that animal researchers need to be more rigorous in their methods, in their reporting and in prompt publishing of the results, irrespective of the study findings.

Introduction

Modified vaccinia ankara virus-expressing antigen 85A (MVA85A) is a new tuberculosis (TB) vaccine currently undergoing clinical trials in children as a boost to Bacillus Calmette–Guérin (BCG), as BCG protection is at best partial and variable.\(^1\)\(^2\) As we set up to carry out an independent systematic review of MVA85A vaccine trials in children, there was considerable debate in the literature that highlighted the potential gains from optimizing the design, conduct and analysis of biomedical research to avoid misleading findings and wasting resources.\(^3\)\(^4\) We therefore decided to explore these issues in animal studies, using systematic review and meta-analysis as methods to carry out syntheses in animal studies.\(^5\)\(^6\) These systematic reviews are important in animal research as they ask both a scientific and moral question, because studies that are poorly designed, conducted or reported are unlikely to be reliable and the ‘animals in effect [have] been wasted’.\(^6\)\(^7\) We sought to assess the experimental design and study quality and summarize the results of studies evaluating MVA85A combined with BCG compared with BCG alone in \textit{in vivo} animals challenged with TB. This would allow us not only to independently evaluate the strength of the pre-clinical evidence, but also to assess the rigour of the design and reporting against standards in emerging animal research quality criteria in the field of vaccine development.\(^8\)

Methods

Our methods were pre-specified in a study protocol [http://www.dcn.ed.ac.uk/camarades/research.html],\(^9\) included in Supplement 1 (available as Supplementary data at IJE online). We included \textit{in vivo} controlled studies of any animal with a TB challenge, where animals were allocated to an intervention group and a control group. We defined control groups as those treated with BCG alone, and the intervention group as those treated with MVA85A vaccine given after BCG vaccination. Studies of MVA85A combined with other antigens were also included. We included studies that measured at least one of the following outcomes: (i) death, including severe morbidity that required euthanasia (termed ‘humane endpoint’); (ii) measures of lung pathology; and (iii) lung bacterial loads. We excluded parameters such as spleen bacterial loads or immunological measures as these are not considered to directly relate to functional protection against TB. These selected endpoints are defined as indicators of protection by specialists in this field in a recent review.\(^10\)

Search strategy

We searched the following databases from inception up to 8 September 2014: MEDLINE (Pubmed); EMBASE (OVID); Science Citation Index-expanded and Science Conference Proceedings (Web of Science); and Biosis previews (Web of Science), using the following search terms in title, abstract and keywords: ‘MVA85A’ OR ‘modified vaccinia virus Ankara’ OR ‘Ag85A’ OR ‘Antigen 85A’ AND ‘tuberculosis’ OR ‘TB’ OR ‘BCG’. We did not apply any language restrictions to the searches. We also contacted experts in the field, individual animal researchers and vaccine trial groups for unpublished data. We also checked the reference lists of relevant studies.
Selection and description of studies

Two investigators independently applied the predefined inclusion criteria (R.K. and T.Y.), and extracted data from relevant studies (R.K. and E.S.). Discrepancies were discussed by the team and agreement reached with P.G. We extracted details of the vaccines used, the route of vaccine administration, the type of TB challenge strain, and the route of TB administration. We also extracted the duration between the initial BCG vaccination and MVA85A booster (BCG/MVA85A interval), the duration between the MVA85A boost and the TB challenge (MVA85A/challenge interval), and the duration between the challenge and outcome assessment.

We used aspects of the Animal Research Reporting in vivo Experiments (ARRIVE) guidelines and the survey of the quality of experimental design, statistical analysis and reporting of research using animals to assess the design quality, reporting and risk of bias of included studies. We assessed whether study objectives were stated, whether sample size calculations were reported, whether the number of animals included were clearly described and whether there was a competing interest statement. We evaluated if animals were randomized to treatment allocation and assessed for baseline comparability and whether assessors were blind to the allocated group for: (i) humane endpoint; (ii) pathology; and (iii) bacteriology assessment.

Outcome data

The numbers of animals that reached a humane endpoint in each group were recorded for each study where this was reported. We summarized our assessment of humane endpoints in a table. We also calculated the risk ratio of death in studies that reported outcome data of for macaques that reached humane endpoint, and pooled these using random effects meta-analysis. Pathological data were summarized in a table. For scores, we reported median and range values that were derived manually.

For bacterial load data, we identified individual comparisons where outcome was measured in animals receiving intervention compared with control group animals. Two authors (R.K. and E.S.) independently extracted means and corresponding variances for each experimental arm using digital ruler software from graphs. On comparison of results, in cases where there was more than a 10% difference, both authors re-extracted the data until an agreement was reached. A mean of the values extracted by both authors were used for meta-analysis. Where aggregate data were not reported but individual animal data were provided, we used Microsoft Excel to calculate the means and standard deviations. For studies that gave a standard error (SE) or confidence interval (CI) of the bacterial loads, the figures were changed to standard deviations using Review Manager 5 software. We calculated a normalized mean difference effect size for each comparison. The data were pooled using DerSimonian and Laird random effect meta-analysis.

Results

We included seven studies from 421 records identified from our search, after considering duplication of studies (Figure 1). We also found an additional study through contacting experts in the field. This additional study was the largest preclinical trial of MVA85A, carried out in monkeys, but was not identified by our electronic search or by checking references; it was not indexed under MVA85A and had a title without relevant keywords. This gives a total of eight included studies.

One publication described two similar experiments that were carried out in two different laboratories (Oxford and Berlin) with different challenge strains; we treated these as the same study but stratified the results by city. Williams et al. in the journal Tuberculosis describes four experiments, two of which met the inclusion criteria of our review. In a second publication, in the journal Infection and immunity, the authors appeared to report on one of these experiments again. We therefore included the third experiment’s data under Williams (a). We report the fourth experiment from the Tuberculosis paper, also reported in Infection and immunity, as a single experiment labelled Williams (b). Excluded studies are detailed in Figure 1 and listed in the Supplement (available as Supplementary data at IJE online).

Description of included studies

Of the eight included studies, two evaluated BCG followed by MVA85A followed by a recombinant fowlpox-expressing antigen 85A (FPAg85A) acting as a second viral vector and booster. Six studies evaluated BCG followed by MVA85A alone (Table 1). Three studies used mice, two used macaques, two used guinea pigs and one used calves. Sample sizes ranged from 4 to 14 per group and study duration ranged from 26 to 73 weeks after TB challenge. One study reported a range of animals from 9 to 12 in each group, and we used the minimum of 9 for analysis.

Intradermal administration was used in six studies and intravenous in one study. In the remaining one study, two comparisons were identified where MVA85A was administered nasally in one group of mice and parenterally in another. Three different challenge strains were used: H37Rv (five experiments), Erdman (three experiments) and
M. bovis (one experiment). The BCG/MVA85A interval ranged from 4 to 22 weeks (median 9.5 weeks); the MVA85A/challenge interval ranged from 4 to 9 weeks (median 6 weeks). The challenge to outcome assessment interval ranged from 6 to 52 weeks (median 13 weeks). One study had no reported endpoint but survival data presented did not go beyond 40 weeks.18

Five studies reported on death or animals reaching a humane endpoint; four reported on pathology; and six reported on bacterial load.

Reported study quality
Table 2 summarizes the reported study quality. None of the studies reported a sample size or power calculation. Of the eight studies, six studies described precisely the number of animals used; one reported ‘five animals in each group’ but the results showed groups of four to six in each group;14 and one reported a range of between nine and 12 animals.17

Two studies reported potential conflict of interest as some authors were ‘co-inventors of MVA85A and shareholders in the joint venture developing the vaccine’. The other six studies had no statements regarding conflict of interest although the same authors who had declared a potential conflict of interest were also co-authors of five of the studies.

One study reported random allocation of animals into treatment groups. Only one study had a baseline comparability table and reported that there was no significant difference in the body weight and age of animals between groups at the start of the experiment.19 Of the others, three reported age, sex and species; one, weight, sex and species; two, age and species; and one, species only.

We assessed the blinded assessment of outcome for each of the three endpoints separately. None of the studies
Table 1. Characteristics of included experiments

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Study</th>
<th>Duration (weeks)</th>
<th>Animal</th>
<th>MVA85A routes</th>
<th>BCG route</th>
<th>Animals in each group x groups(^a)</th>
<th>Challenge strain</th>
<th>Challenge Route</th>
<th>BCG/ MVA85A interval (weeks)</th>
<th>Boost to challenge (weeks)</th>
<th>Challenge to autopsy (weeks)</th>
<th>Funding Reference list</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCG+</td>
<td>Williams 2005</td>
<td>29</td>
<td>Guinea pigs</td>
<td>ID(^a)</td>
<td>SC</td>
<td>6 x 6</td>
<td>H37Rv</td>
<td>Aerosol</td>
<td>4</td>
<td>8(^b)</td>
<td>17</td>
<td>EU; DOH UK</td>
</tr>
<tr>
<td>MVA85a + FP(^9)Ag85a</td>
<td>Williams 2005</td>
<td>40</td>
<td>Guinea pigs</td>
<td>ID</td>
<td>SC</td>
<td>6 x 3</td>
<td>H37Rv</td>
<td>Aerosol</td>
<td>4</td>
<td>6(^b)</td>
<td>26</td>
<td>EU; Wellcome Trust; DOH UK</td>
</tr>
<tr>
<td>BCG + MVA85A</td>
<td>Goonetilleke 2003</td>
<td>32</td>
<td>Mice</td>
<td>IM/NA(^c)</td>
<td>NA</td>
<td>9–12 x 7</td>
<td>H37Rv</td>
<td>‘In lungs’</td>
<td>22</td>
<td>4</td>
<td>6</td>
<td>Welcome Trust</td>
</tr>
<tr>
<td></td>
<td>Romano 2006</td>
<td>Not stated</td>
<td>Mice</td>
<td>IV</td>
<td>IV</td>
<td>11–14 x 3</td>
<td>H37Rv</td>
<td>Intravenous</td>
<td>34.8</td>
<td>Not mentioned</td>
<td></td>
<td>FWO – Vlaanderen, Brussels capital region, Damiaanktie Belgium, European Economic community.</td>
</tr>
<tr>
<td></td>
<td>Tchilian 2008:Oxford</td>
<td>26</td>
<td>Mice</td>
<td>ID</td>
<td>SC</td>
<td>5–6 x 3</td>
<td>Erdman</td>
<td>Aerosol</td>
<td>10</td>
<td>4</td>
<td>12</td>
<td>European FP6 integrated project, NIHR Oxford Biomedical Research Centre programme</td>
</tr>
<tr>
<td></td>
<td>Tchilian 2008: Berlin</td>
<td>26</td>
<td>Mice</td>
<td>ID</td>
<td>SC</td>
<td>4–5 x 3</td>
<td>H37Rv</td>
<td>Aerosol</td>
<td>10</td>
<td>4</td>
<td>12</td>
<td>European FP6 integrated project, NIHR Oxford Biomedical Research Centre programme</td>
</tr>
<tr>
<td></td>
<td>Verrick 2009</td>
<td>34/35</td>
<td>Macaques</td>
<td>ID</td>
<td>ID</td>
<td>6 x 4</td>
<td>Erdman</td>
<td>Tracheal</td>
<td>9</td>
<td>9</td>
<td>16/17</td>
<td>Wellcome Trust; EU DOE UK; Wellcome Trust</td>
</tr>
<tr>
<td></td>
<td>Vordermeier 2009</td>
<td>28</td>
<td>Calves</td>
<td>ID</td>
<td>SC</td>
<td>10 x 4</td>
<td>M. bovis</td>
<td>Tracheal</td>
<td>8</td>
<td>6</td>
<td>14</td>
<td>Wellcome Trust; EU DOE UK; Wellcome Trust</td>
</tr>
<tr>
<td></td>
<td>Sharpe 2010</td>
<td>73</td>
<td>Macaques</td>
<td>ID</td>
<td>ID</td>
<td>4–6 x 3</td>
<td>Erdman</td>
<td>Aerosol</td>
<td>12</td>
<td>9</td>
<td>52</td>
<td>DOH UK</td>
</tr>
</tbody>
</table>

\(^a\)Source of vaccine similar to that used in Williams et al. 2005 (b).

\(^b\)The boost challenge interval shows weeks after FP-Ag85A, the second booster that was given 4 weeks after MVA85A booster compared with other studies where the TB challenge follows MVA85A, the only booster used.

\(^c\)Two intervention arms: one with MVA85A given nasally, and one arm with MVA85A given by injection.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Were sample size and power calculations given?</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Was the number of animals included in experiments clearly described?</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Was competing interest declared?</td>
<td>No COI statement</td>
<td>No COI statement</td>
<td>No COI statement</td>
<td>No COI</td>
<td>statement</td>
<td>No COI</td>
<td>statement</td>
<td>No COI</td>
</tr>
<tr>
<td>Was there randomization?</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>Was there baseline comparability?</td>
<td>All 250–300g female guinea pigs</td>
<td>Same species guinea pigs</td>
<td>All 8–10-week-old female mice</td>
<td>All 8–10-week-old female mice</td>
<td>All 6–8wk old female mice</td>
<td>Same age weight, male macaques</td>
<td>All 6-month-old cattle</td>
<td>All 4-year-old macaques</td>
</tr>
<tr>
<td>Was the outcome assessor blind?</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Yes</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>Humane endpoint</td>
<td>Not reported</td>
<td>Only lesion volume blinded</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pathology</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacteriology</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
</tbody>
</table>

Grey shading - outcome not evaluated in study.
COI, conflicts of interest
reporting humane endpoints or bacteriology reported blinding their assessment. For pathology, two of the four studies reported that the assessors were blind to treatment allocation.\(^1,\)\(^19\)

**Mortality, including euthanasia for severe morbidity**

Five studies with a total of 107 animals in the relevant arms assessed mortality (Table 3). Two used macaques, two used guinea pigs and one used mice. Two of the five studies that assessed mortality/euthanasia did not provide data on the number of animals euthanized. Williams \textit{et al.} (a) reported that the results ‘did not allow any distinction to be made between MVA85A boosting and BCG alone’, but no data were reported.\(^15\) Williams \textit{et al.} (b) reported a ‘statistically different increase in survival (P = 0.018)’.\(^16\)

In the remaining three studies, results were variable but the numbers are small (Figure 2). In one study with the longest follow-up, five out of the six macaques died in the MVA85A group, compared with two out of six in the BCG group.\(^13\) The other two studies reported no deaths in the MVA85A groups, compared with four and one in the BCG groups, respectively.\(^16,\)\(^19\) Another study reported median survival rather than the number of events at a single time, so we were unable to include these data in the meta-analysis. In this study, median survival time of animals in the MVA85A group (18.5 weeks) was reported as not statistically different from the BCG group (19 weeks).\(^18\)

**Lung pathology**

Four studies with a total of 100 animals in the comparison arms reported pathological changes after TB challenges were given, reported on ordinal scales. The text in the papers made inferences that were not evident from the data; for example, one study implied benefit by comparing the MVA85A group with controls rather than with BCG,\(^1\) but overall there was no effect obvious in comparisons between MVA85A with BCG compared with BCG alone (Table 4).

**Bacterial load**

Six studies with a total of 137 animals measured bacterial loads in animals after a TB challenge (Table 5). Of these six studies, we excluded the study by Williams \textit{et al.} from the meta-analysis as they did not report the data required to perform meta-analysis. The results in this study were reported as being significantly or not significantly better when compared with the BCG alone group, without outcome data.\(^15\) From the five remaining studies, we extracted

<table>
<thead>
<tr>
<th>Study I.D</th>
<th>Reference</th>
<th>Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Williams 2005 (a)</td>
<td>Experiment 3(^6)</td>
<td>BCG + MVA85A + FP9Ag85a</td>
</tr>
<tr>
<td>Williams 2005 (b)(^16)</td>
<td>Guinea pigs</td>
<td>BCG alone</td>
</tr>
<tr>
<td>Romano 2006 (^{18})</td>
<td>Mice</td>
<td>MVA85A boost</td>
</tr>
<tr>
<td>Verrick 2009 (^{19})</td>
<td>Macaques</td>
<td>BCG</td>
</tr>
<tr>
<td>Sharpe 2010 (^{13})</td>
<td>Macaques</td>
<td>MVA85A</td>
</tr>
</tbody>
</table>

\(^a\)Study did not report number of deaths but that median survival time of animals in the MVA85A group (18.5 weeks) was statistically different from the BCG group (19 weeks).

\(^{ND, no data reported.}\)
seven comparisons of MVA85A boosting vs BCG alone that we included in meta-analysis. Overall, MVA85A boosting showed no reduction in bacterial loads (3.28%, 95% CI 3.5 to 9.8, \( P = 0.267 \)) (Figure 3).

**Discussion**

We confined this review to functional parameters of protection in animal models. Indeed, the vaccine scientists state the decisions to move to trials in humans are defined by animal studies with ‘statistically significant improvements in disease compared to control groups as measured by bacterial load, severity of pathology, and time to death’.10 We used standard methods for systematic reviews and meta-analysis in animal studies.5,6

Our meta-analysis suggests an apparent lack of evidence of efficacy in animals in data collected before the start of the recent phase 11b trial in South African children that enrolled children between 15 July 2009 and 4 May 2011.20 We acknowledge that the decision to progress to clinical trial is not solely based on evidence derived from preclinical efficacy studies, but these studies are an important component of the TB vaccine development paradigm.10 Selection of adequately powered endpoints in preclinical studies, such as the 60% improvement in clinical efficacy that is required for licensing a vaccine, may result in similar magnitude of improvement in animals and thus be more predictive of human trial efficacy.10 Indeed, the recent clinical trial testing MVA85A in 2797 South African infants did not demonstrate a beneficial effect of MVA85A.20 In subsequent papers, the authors of this trial have explained that this trial did not have an effect due to species differences, clinical trial settings, *M. tuberculosis* strain and exposure, magnitude of efficacy, definition of protection and whether to use reduction of disease or prevention of disease as an endpoint.10 The principal investigator of a number of the animal studies was also the senior principal investigator in the South Africa infant trial.

We believe that, as with any research endeavour, validity depends on experiments conducted and reported with rigour. None of the studies in our review report a sample size or power calculation; few studies were either blinded or randomized, casting doubt on the internal validity of the studies. There were too few studies to statistically assess for selective reporting of outcomes and a priori protocols of the analysis were not available. We know the shortcomings in the conduct and reporting of animal research may lead to over- or under-estimations of treatment effects,21,22,25 and this indicates critical, systematic and independent appraisal of animals studies is a potentially important component of translational research. The limitations of the animal model in tuberculosis also causes the researchers themselves to be concerned about the predictive abilities.23

We appreciate that the effects of a boost vaccine to BCG may be more modest, with requirements for larger samples of animals.23 Indeed, research that is inadequately powered could be regarded as unethical, as the design is insufficient to answer the question the research is trying to address. It is important that the ethics in animal research and power are more fully debated. In addition, it is important that if the decision by researchers to proceed from animal studies to children is not based on ‘statistically significant improvements in disease compared to control groups’, but on other evidence, this evidence needs to be systematically summarized, appraised, checked for completeness and documented to allow transparency in the translational process.10 There are gating criteria for TB vaccines published, that include safety, immunogenicity and animal efficacy.24

Adherence to reporting guidelines will allow consumers of animal research to draw informed conclusions from results presented. Journal editors, peers and granting bodies should drive this improvement. In addition, a priori protocols, where investigators provide details of appropriate experimental design and statistical analysis, are important to this process.25,26 Finally, it may be that publication bias further confounds our conclusion, but we identified too few studies to allow us to assess for this using standard techniques.

However, there did seem to be some evidence of a delay of 2 years in publication of one study, which concerned us. The longest follow-up study in our dataset, containing almost half (16/34) of the data testing this vaccine in

![Figure 2](http://ije.oxfordjournals.org/)

**Figure 2.** MVA85A combined with BCG compared with BCG alone. Death, including euthanasia for severe morbidity.
Table 4. Experiments MVA85A boosting vs BCG alone or naive, outcome 2: pathology changes

<table>
<thead>
<tr>
<th>Study</th>
<th>Model</th>
<th>Outcome</th>
<th>Measured</th>
<th>Summary value</th>
<th>MVA85A</th>
<th>BCG alone</th>
<th>No vaccine</th>
<th>MVA85A vs BCG alone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Williams 2005 (b)</td>
<td>Guinea pigs</td>
<td>Lung consolidation</td>
<td>%</td>
<td>ND\textsuperscript{a}</td>
<td>ND\textsuperscript{a}</td>
<td>Unclear</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Williams 2005 (b)</td>
<td>Guinea pigs</td>
<td>Lung consolidation</td>
<td>%</td>
<td>ND\textsuperscript{a}</td>
<td>ND\textsuperscript{a}</td>
<td>Unclear</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verrick 2009\textsuperscript{c}</td>
<td>Macaques</td>
<td>Lung lesions</td>
<td>No. of animals</td>
<td>n/N</td>
<td>3/6</td>
<td>5/6</td>
<td>6/6</td>
<td></td>
</tr>
<tr>
<td>Vordermeier 2009\textsuperscript{1}</td>
<td>Calves</td>
<td>Lung TB lesions</td>
<td>No. of animals</td>
<td>n/N</td>
<td>5/10\textsuperscript{b}</td>
<td>7/10</td>
<td>10/10</td>
<td></td>
</tr>
<tr>
<td>Vordermeier 2009\textsuperscript{1}</td>
<td>Calves</td>
<td>Lymph node involvement</td>
<td>No. of animals</td>
<td>n/N</td>
<td>6/10</td>
<td>7/10</td>
<td>9/10</td>
<td></td>
</tr>
<tr>
<td>Sharpe 2010\textsuperscript{13}</td>
<td>Macaques</td>
<td>Lesions in lungs\textsuperscript{d}</td>
<td>No. of lesions/animal</td>
<td>Median, range</td>
<td>153.13, 360 to 360</td>
<td>211.97, 160 to 280\textsuperscript{b}</td>
<td>98.44, 80 to 200</td>
<td></td>
</tr>
<tr>
<td>Sharpe 2010\textsuperscript{13}</td>
<td>Macaques</td>
<td>Extra-pulmonary granulomas</td>
<td>No. of animals</td>
<td>n/N</td>
<td>4/6</td>
<td>3/6</td>
<td>4/4</td>
<td></td>
</tr>
<tr>
<td>Sharpe 2010\textsuperscript{13}</td>
<td>Macaques</td>
<td>Lung pathology score\textsuperscript{d}</td>
<td>Median, range</td>
<td>20.625, 10.25 to 23.5</td>
<td>19, 5.75 to 21.75</td>
<td>26, 18.5 to 25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sharpe 2010\textsuperscript{13}</td>
<td>Macaques</td>
<td>Total pathology score\textsuperscript{d}</td>
<td>Median, range</td>
<td>25.3, 14.8–37.4</td>
<td>22.9, 18–26\textsuperscript{c}</td>
<td>30.96, 25.6–37.4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{a}Vaccinated significantly less pathology than naive group.
\textsuperscript{b}Vaccinated significantly more pathology than naive group.
\textsuperscript{c}No difference in all other pathology parameters used including lung congestion, consolidation, fibrous pleuritis, lung oedema and emphysema.
\textsuperscript{d}Results estimated from extrapolation on graphs.
\textsuperscript{e}No data given.
\textsuperscript{f}BCG vaccination reduced the number of infected lobes significantly ($P < 0.05$) and MVA85A reduced these even further which is reflected in the higher level of statistical significance observed ($P < 0.001$)."
monkeys, was published in June 2010, almost a year after the South African trial had started testing the vaccine in South African children and 2 years after this trial in monkeys had been completed. This trial reported that five out of the six monkeys given the experimental vaccine with BCG died or were so ill they needed to be euthanized, compared with only two out of six in the BCG control group. In addition, even when published, Sharpe 2010 is not detected on a standard MEDLINE search using MVA85A because of MVA85A is not mentioned in the title or abstract.

Our review presents a useful summary of the preclinical data on MVA85A but there are limitations to our approach. We are confident that our search strategy was robust, but there were only eight studies that met our inclusion criteria. Indeed, two of the eight studies included FPag85A in addition to MVA85A; these were included as overall data were limited, and if effects were seen this could be discussed in terms of confounding. The data in these two studies were similar to the others. Few studies using different experimental designs and species may be considered akin to lumping oranges with apples, which may mask subtle but relevant differences in efficacies. However, we used a normalized mean difference effect size in our pooling of bacteriology; this is presented as a percentage improvement relative to the magnitude of effect in normal healthy animals.6

We acknowledge that our findings are at variance with a narrative review by the academic groups responsible for MVA85A. In this review, they discuss six of the eight studies included in this review and concluded from individual study data from three of the studies that MVA85A had a ‘variable and modest level of efficacy in animals that failed to predict efficacy in BCG-vaccinated infants to a level required for progression of the vaccine development’.10

**Table 5. Experiments in MVA85A-boosted animals vs BCG-vaccinated or naïve animals: bacterial loads, reported as log10 CFU**

<table>
<thead>
<tr>
<th>Intervention Model</th>
<th>Study ID</th>
<th>Study 2005 (a)</th>
<th>Study 2005 (b)</th>
<th>Study 2005 (c)</th>
<th>Study 2008 (Berlin)14</th>
<th>Study 2008 (Oxford)14</th>
<th>Study 200919</th>
<th>Study 201013</th>
</tr>
</thead>
<tbody>
<tr>
<td>MVA85A booster</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>5.5 ± 0.13</td>
<td>5.5 ± 0.18</td>
<td>6.2 ± 0.03</td>
<td>6.0 ± 0.03</td>
</tr>
<tr>
<td>No vaccine</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>5.5 ± 0.13</td>
<td>5.5 ± 0.18</td>
<td>6.2 ± 0.03</td>
<td>6.0 ± 0.03</td>
</tr>
<tr>
<td><strong>Difference detected?</strong></td>
<td>Mean SE</td>
<td>Mean SE</td>
<td>Mean SE</td>
<td>Mean SE</td>
<td>Mean SE</td>
<td>Mean SE</td>
<td>Mean SE</td>
<td>Mean SE</td>
</tr>
<tr>
<td><strong>MVA85A/BCG vs BCG</strong></td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Conclusions

Our review raises a question about the robustness of claims that MVA85A animal studies provide evidence of protection against TB challenge. We also found that there was a need to attend to methodological standards in the design, execution and reporting of pre-clinical animal studies. Many were inadequately powered and little attention was given to potential sources of biases. We would echo a recent publication that stated research needs to improve ‘reproducibility practices, more appropriate (usually more stringent) statistical thresholds; and implement in study design standards’11 and researchers of animal studies should publish their results promptly, irrespective of the findings.
Supplementary Data
Supplementary data are available at IJE online.

Funding
This project has been supported by the US President’s Emergency Plan for AIDS relief (PEPFAR) through HRSA under the terms of T84HA21652 and via a bursary from the Stellenbosch University Rural Medical Education Partnership Initiative (SURMEPI). This paper is a research activity within the Cochrane Infectious Diseases Group funded by DfID for research synthesis aimed to benefit the poor in developing countries. This study was funded by the Department for International Development UK for research synthesis aimed to benefit the poor in developing countries. The contribution of E.S. was funded by the UK National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs) infrastructure award: ivSyRM – the CAMARADES–NC3Rs in vivo systematic review and meta-analysis facility. The study sponsors had no role in design, collection, analysis or interpretation of data, writing of the report or decision to submit the manuscript for publication. DfID funded Open Access through a grant to CIDG. R.K., the corresponding author, had full access to the data and made the final decision to publish.

Acknowledgements
To Vittoria Lutje for carrying out the search and providing full texts, and to Malcolm Macleod for advice and suggestions on the methods and the interpretation.
P.G. conceptualized the study, brought the author team together and provided advice and input at all stages. R.K. wrote the protocol and carried out the data collection, analysis and interpretation, and writing of the first draft. T.Y. contributed to conception and design, eligibility assessment, data interpretation, reviewing the manuscript and resolving differences. E.S. provided critical input to the methods and interpretation, carried out double data entry and conducted the meta-analysis and reviewing the manuscript. All authors contributed to the writing of the paper and the submitted manuscript.

Conflict of interest: The authors declare they have no competing interests.

References

Figure 3. Meta-analysis using normalized mean difference for MVA85A vs BCG alone: bacterial loads after TB challenge given to animals.


18. Romano M, D’Souza S, Adnet PY et al. Priming but not boosting with plasmid DNA encoding mycolyl-transferase Ag85a from Mycobacterium tuberculosis increases the survival time of Mycobacterium bovis BCG vaccinated mice against low dose intravenous challenge with M. Tuberculosis H37Rv. *Vaccine* 2006;24:3353–64.


