Review Article

Replacing, Reducing and Refining the Use of Animals in Tuberculosis Vaccine Research

Rachel Tanner and Helen McShane

The Jenner Institute, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK

Summary

Tuberculosis (TB) remains a serious global health threat and an improved vaccine is urgently needed. New candidate TB vaccines are tested using preclinical animal models such as mice, guinea pigs, cattle and non-human primates. Animals are routinely infected with virulent *Mycobacterium tuberculosis (Mtb)* in challenge experiments to evaluate protective efficacy, raising ethical issues regarding the procedure of infection itself, symptoms of disease and humane end-points. We summarize the importance and limitations of animal models in TB vaccine research and review current alternatives and modifications in the context of the NC3Rs framework for replacing, reducing and refining the use of animals for scientific purposes.

Keywords: tuberculosis, vaccine, animal models, *Mtb* challenge, 3Rs

1 Introduction

Tuberculosis (TB) is the world's most deadly infectious disease, with an estimated 9.6 million new cases and 1.5 million deaths annually (WHO, 2016). Incidence of infection in endemic countries remains very high despite good coverage with BCG (Bacillus Calmette-Guérin) vaccine, the only currently available vaccine (Mahomed et al., 2006; Moyo et al., 2010). There is a desperate need for a more efficacious vaccine. New candidate TB vaccines are currently tested for safety, immunogenicity and efficacy using preclinical animal models such as mice, guinea pigs, cattle and non-human primates (NHPs). Mice are the most widely used species due to the potential for screening a high number of candidates at low cost and the availability of gene knockout strains to characterize the immune response (Apt and Kramnik, 2009). However, research in species other than mice is becoming more commonplace with increasing availability of immunological reagents (McShane and Williams, 2014). Guinea pigs have emerged as a useful model, replicating many aspects of *Mtb* infection in humans such as granuloma formation, dissemination and caseating necrosis (Clark et al., 2015). They are also considered a more stringent model in discriminating the efficacy of different vaccines due to the variety of human-like pulmonary and extrapulmonary lesions observed (Basaraba, 2008; McMurray et al., 1996). Despite the utility of small animals in early screens, larger animal models such as cattle and NHPs are considered more relevant to human TB. NHPs are naturally susceptible to infection with *Mtb*, and develop the most human-like disease with latency and reactivation (Flynn et al., 2015). BCG confers some level of protection in NHPs, which can be quantified through a variety of clinical and nonclinical parameters (Sharpe et al., 2010). In the absence of a surrogate marker of protection from TB disease, animal *Mtb* infection models remain an essential pre-requisite for novel vaccine candidates progressing to clinical trials.

To evaluate the protective efficacy of a candidate TB vaccine, animals must be infected with virulent *Mtb* in a challenge experiment following vaccination. While *M. bovis* challenge studies in cattle are classified as "mild" in severity by the UK Home Office due to a lack of clinical symptoms, *Mtb* challenge experiments in mice, guinea pigs and NHPs are generally defined as "moderate". Moderate severity indicates that the animals are likely to experience "short term moderate pain, suffering or distress or long-lasting mild pain, suffering or distress… or moderate impairment of the well-being or general condition"[1]. As TB disease progresses, animals may experience loss of body

[1] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/276014/NotesActualSeverityReporting.pdf

Received July 28, 2016;
Accepted September 22, 2016;
Epub September 26, 2016;
https://doi.org/10.14573/altex.1607281

weight, fever and respiratory distress and if left untreated will eventually die of pulmonary insufficiency (Gupta and Katoch, 2009). As this is unethical, humane euthanasia at predefined clinical endpoints, which will be discussed later in this report, is now enforced by UK Home Office legislation.

In addition to welfare concerns, the many differences between animal models and human TB bring into question the predictive value of such studies. "Protection" in animal models, as determined by the outcome of *Mtb* challenge experiments, is on a continuous spectrum and usually defined in terms of a relative improvement in a disease-related readout such as bacterial load, pathology score or long-term survival. A vaccine is considered to provide protection even if there is a measurable bacterial load or pathology in the organs or if some animals do not survive (Elias et al. 2005; McShane and Williams, 2014; Vordermeier et al., 2009). In humans, however, efficacy is binary and defined as the prevention of TB disease using clinical endpoints; any individual developing disease, however minimal, is not protected (McShane and Williams, 2014). Clearly, an artificial aerosol challenge is very different from natural transmission in humans, and the laboratory strains of *Mtb* commonly used (such as H37Rv) are genetically dissimilar to clinical isolates (Niemann and Supply, 2014), with much higher challenge doses employed (McShane and Williams, 2014). This issue has recently been addressed with new advances in ultra-low dose challenge, as discussed below. In addition to these fundamental differences in the model itself, animals are genetically distinct from humans, with several discrepancies in both innate and adaptive immunity between mice and humans (Mestas and Hughes, 2004). The widely used Balb/c and C57BL/6 mouse strains do not exhibit caseating granuloma formation following *Mtb* infection (Orme and Basaraba, 2014) and manifest a chronic phase of disease unlike latent *Mtb* infection in humans (Rhoades et al., 1997). Furthermore, responses in genetically diverse humans will be considerably more variable than in an inbred laboratory animal strain. Although outbred mice give a diversified picture of TB, which may be more representative of human disease, larger group sizes are required to offset the increase in variability (Niazi et al., 2015).

The predictive value of animal challenge models in determining TB vaccine efficacy in humans is uncertain, and will remain unclear until a successful vaccine is developed. Furthermore, there is evidence from studies of other diseases that animal models can fail to reliably predict safety in humans (Suntharalingam et al., 2006; McKenzie et al., 1995). Other difficulties include the large numbers of animals required, and the nature and slow growth of mycobacteria making experiments long and costly with the need for highly specialized Category 3 animal facilities. Given the scientific and logistical limitations of these models as well as the ethical concerns, it is imperative that potential alternatives are pursued. The principles of the 3Rs (replacement, reduction and refinement) were first proposed by Russell and Burch in the 1950s, with the aim of ensuring ethical use of animals in research (Balls, 2009). This framework, now formalized in national and international legislation, provides the basis of our discussion on the use of animals in TB vaccine research.

2 Replacement

The National Centre for the 3Rs (NC3Rs) in the UK describes replacement as "methods that avoid or replace the use of animals defined as 'protected' under the Animals (Scientific Procedures) Act 1986, amended 2012 (ASPA) in an experiment where they would have otherwise been used"[2]. Protected animals in the UK refer to all living vertebrates except humans. Alternatives include using humans, *in vitro*/cell culture models, computational/mathematical modelling, or less sentient animals. All of these have been reported in the context of TB vaccine development.

2.1 Use of humans

Given the differences between human and animal manifestations of TB disease, one may argue that a more appropriate focus would be the target species. Controlled human challenge models have been successfully implemented for other pathogens, including those responsible for malaria and typhoid (Marwick, 1998; Sauerwein et al., 2011), and are a valuable tool for assessing vaccine efficacy. However, the safety and ethical barriers to challenging humans with live virulent mycobacteria have thus far limited the development of an *in vivo* challenge model for TB.

Already licensed for use in humans, BCG represents a potential surrogate for *Mtb* challenge, and is a safe replicating mycobacterium that causes a contained, short-term infection in immunocompetent individuals. A BCG challenge model has recently been described in which participants were challenged with intradermal (ID) BCG. Skin biopsies of the challenge site were taken 2 weeks later. BCG load was quantified by culture and quantitative polymerase chain reaction (qPCR) (Harris et al., 2014; Minassian et al., 2012). The model demonstrated the ability to detect differences in anti-mycobacterial immunity induced by BCG and MVA85A (modified vaccinia Ankara 85A, a new-generation vaccine against tuberculosis) vaccination, with a significant inverse correlation between immune signatures, particularly IFN-γ and IL-17 pathways, and BCG load detected by qPCR (Harris et al., 2014). A dose escalation study and comparison of BCG SSI and BCG TICE has also been reported (Minhinnick et al., 2016).

One criticism of intradermal challenge is that it does not mimic the natural route of infection, and to that end a clinical trial evaluating the safety and feasibility of an aerosol BCG human challenge model is currently ongoing (NCT02709278).

2.2 *In vitro* assays

The development of vaccines against other pathogens has been greatly expedited by the identification of a biomarker or immune correlate of protection (Thakur et al., 2012). Such indicators, for example antibody titer or cytokine level, may

[2] https://www.nc3rs.org.uk/the-3rs, accessed 2016

be measured using an *in vitro* assay, allowing the use of human blood or cell samples. Frustratingly, there are currently no validated correlates to reliably assess the efficacy of candidate TB vaccines. Most TB vaccine studies to date have used quantification of antigen-specific IFN-γ by ELISpot and/or intracellular cytokine staining as the primary immunological read-out, though it remains unclear whether this measure correlates with protection (Elias et al., 2005; Mittrücker et al., 2007). Although one study of BCG vaccinated infants in South Africa found no difference in frequency and extended cytokine profiles of *Mtb* specific cells between protected and non-protected infants (Kagina et al., 2010), a more recent trial indicated an association between the BCG antigen-specific IFN-γ ELISpot response and reduced risk of TB disease (Fletcher et al., 2016). This latter study also found a negative correlation between levels of *Mtb* antigen-specific IgG and risk of disease, suggesting that protective immunity may not be restricted to the T cell compartment (Fletcher et al., 2016).

An alternative to measuring predefined individual parameters is the use of mycobacterial growth inhibition assays (MGIAs), which take into account a range of immune mechanisms and their additive effects and interactions. These systems measure the ability of human or animal cells to inhibit growth of mycobacteria following *in vitro* infection. Using samples taken pre- and post-vaccination, functional efficacy may be assessed without the requirement for *in vivo Mtb* challenge or natural infection in animals. Several such MGIAs have successfully discriminated BCG vaccinated from non-vaccinated human volunteers using both whole blood or peripheral blood mononuclear cells (PBMC) (Cheng et al., 1988; Cheon et al., 2002; Fletcher et al., 2013; Hoft et al., 2002; Kampmann et al., 2004; Worku and Hoft, 2000).

Animal models provide an opportunity to test novel vaccine candidates, and an *in vitro* assay using blood or cells from vaccinated animals offers a potential surrogate of protective efficacy that may negate the need for *in vivo* challenge during early selection of vaccine candidates. MGIAs have been described using cells from mice (Cowley and Elkins, 2003; Kolibab et al., 2009; Marsay et al., 2013; Parra et al., 2009; Sada-Ovalle et al., 2008), cattle (Carpenter et al., 1997; Denis et al., 2004) and NHPs (Tanner et al., submitted). Importantly, both Parra et al. (2009) and Marsay et al. (2013) showed that differences in mycobacterial growth inhibition between groups were consistent with different levels of protection in experimentally-matched mice challenged *in vivo*, thus demonstrating the utility of animal MGIAs for biological validation. Preliminary work applying a whole blood MGIA in cynomolgus macaques has demonstrated a correlation between mycobacterial growth inhibition following vaccination and protection from BCG challenge as measured by lymph node CFU (own unpublished data).

MGIAs also allow efficacy against different strains (including hypervirulent strains) to be tested in parallel in cells from the same animal, rather than limiting to one laboratory strain, which may be unrepresentative of clinical strains affecting humans, as for *in vivo* challenge.

2.3 *In silico* modelling

The availability of genome sequences for *Mtb* and other mycobacterial species, mice and humans together with relatively recent developments in computer algorithms have facilitated the use of *in silico* bioinformatics methods for the identification of new TB vaccine candidates.

Comparative analyses of mycobacterial genomes have allowed the identification of 16 genomic regions of *Mtb* which are absent in one or more strains of BCG, known as regions of difference (RD) (Behr et al., 1999; de Jonge et al., 2005). RD proteins have been generated using recombinant methods or overlapping synthetic peptides (Mustafa, 2005) followed by testing in immune assays to identify those suitable for vaccine development. Wang et al. analyzed RD proteins *in silico* for their ability to bind to a range of HLA class I alleles and showed that a significant proportion were high-affinity binders, representing promising epitopes for inclusion in experimental TB vaccine candidates (Wang et al., 2010).

In 2011, a study by Tang et al. used novel computational search tools to identify new *Mtb* antigens activating polyfunctional $CD8^+$ T cells which were then validated in human-based assays (Tang et al., 2011). A further study scanned multiple published databases of *Mtb* gene expression to select the proteins most highly expressed in all phases of infection. The proteins were evaluated for the presence of B and T cell promiscuous epitopes and population coverage in terms of allele presentation. Sequence alignments were then used to determine identical epitopes on *M. smegmatis*, and two *M. smegmatis*-derived experimental vaccines were tested in mice to assess humoral immunogenicity and cross-reactivity with *Mtb* (Rodriguez et al., 2011).

More recently, Monterrubio-López et al. (2015) identified potential vaccine targets using NERVE (New Enhanced Reverse Vaccinology Environment) prediction analysis of the *Mtb* H37Rv proteome. Proteins were further down-selected based on VaxiJen-predicted antigenicity and amino acid sequence alignments, with 6 novel candidates finally selected. Bowman et al. described the incorporation of the machine learning approach support vector machine (SVM) classification, which resulted in superior accuracy in discriminating protective antigens from non-antigens (Bowman et al., 2011). In addition to the 3Rs benefits, reverse vaccinology offers several advantages over conventional methods including speed, reduced cost and ability to identify all the putative protective antigens rather than just the most abundant (Bowman et al., 2011).

2.4 Less sentient animals

As opposed to absolute replacement of animal models with *in vitro* or inanimate systems, another 3Rs approach involves the replacement of more sentient vertebrates with animals thought to have a lower potential for pain perception. The amoeba *Dictyostelium discoideum* and the fruit fly *Drosophila melanogaster*, though useful in understanding host-pathogen interactions and innate immune responses during mycobacterial infection, have limited applicability for the study of vaccines due to their lack of adaptive immunity (Dionne et al., 2003; Hagedorn et

al., 2007, 2009). Zebrafish, however, have an immune system similar to that of humans with a fully developed adaptive arm in adults, and represent a popular model organism for various pathogens (Meijer and Spaink, 2011). It has been suggested that the course of mycobacterial infection in zebrafish has some parallels to that of human TB, with high-dose infection leading to progressive disease resembling acute TB, and low-dose infection leading to spontaneous latency with reactivation following immunosuppression (Parikka et al., 2012; Swaim et al., 2006). Importantly, many virulence factors, host genes and immune cell types involved in human *Mtb* pathogenesis have conserved functions in the zebrafish-*M. marinum* model (Cronan and Tobin, 2014). Zebrafish have already proven useful in elucidating the early events of a mycobacterial infection, the role of the innate immune system in resistance and understanding the mechanisms of granuloma formation and its role in controlling infection, with the limitation that zebrafish do not have lungs (Clay et al., 2007; Cronan and Tobin, 2014; Swaim et al., 2006). It is also a potentially promising model for aiding the development of TB therapeutics and vaccines for preventing reactivation of latent TB. In a study by Oksanen et al., both BCG and a DNA-based vaccine protected fish from mycobacterial infection, reducing mortality and bacterial burden following infection with a lethal dose (Oksanen et al., 2013).

3 Reduction

Reduction is defined as "methods that minimise the number of animals used per experiment or study, either by enabling researchers to obtain comparable levels of information from fewer animals, or to obtain more information from the same number of animals, thereby avoiding further animal use"[1]. Examples include reducing replication by increased data sharing, improved experimental design and technologies enabling longitudinal studies in the same animals.

3.1 Reducing replication

Publication bias arises when negative or non-confirmatory findings are suppressed, either by researchers themselves choosing not to submit for publication or lack of journal acceptance (Song et al., 2013). Emphasis is frequently placed on impact rather than quality of research or reproducibility, leading to what has been described as a "crisis of false positives" in biomedical research where many published results are false or exaggerated, with an estimated 85% of resources wasted (Macleod et al., 2014). Importantly, failure to share negative findings results in needless repetition of animal experiments. One potential solution is the prospective registration of preclinical studies similar to that of clinical human trials such as the BMJ AllTrials campaign, aiming for "all trials registered, all results reported"[3]. In recent years, some measures have been taken to encourage the publication of negative findings by provision of such repositories as the BioMedicine Journal of Negative Results[4] and a recent PLOS ONE collection of negative, null and inconclusive results. Furthermore, BMC Research Notes was produced with the specific objective of publishing repeat studies and negative results[5].

3.2 Experimental design

The NC3Rs state that "appropriate experimental design and statistical analysis techniques are key means of minimising the use of animals in research"[1]. One critical consideration is sample size, which should not be so large as to use an unnecessary number of animals. However, under-powering an experiment with too few animals to provide a biologically meaningful result is equally wasteful (Festing and Altman, 2002). Rigorous statistical calculation such as power analysis should be performed to identify an appropriate sample size; methods for this have been described by Dell et al. (2002). However, Williams et al. (2009) highlighted the issues involved in powering TB challenge experiments based on survival using guinea pigs and larger animals. The authors note that while survival studies can be extremely informative in establishing that a new candidate vaccine can confer equivalent protection to BCG and that this protection is sustained, demonstrating significant improvement over BCG is more challenging. Due to the binomial nature of survival data, statistical power is extremely low, and a simulation exercise revealed that for a substantial increase in mean survival time (from 150 to 250 days), a prohibitively large group size of 74 would be required to reach a significant p-value of 0.05 (Williams et al., 2009). This provides further support for the use of predefined fixed endpoints, as alternative measures such as bacterial load in target organs offer superior statistical and discriminative power (Williams et al., 2009).

Minimizing variation (for example by controlling for confounding variables such as age, weight and genetics) also improves power, allowing the same effect to be detectable with a smaller number of animals (Festing and Altman, 2002). Experiments should be unbiased with random allocation of animals to treatment groups and blinding of researchers, and it is encouraging that many of these factors are now taken into consideration when granting ethical permission to conduct animal trials. Furthermore, multiple questions may be answered, and therefore numbers of experiments and animals reduced, by applying adequately powered factorial designs (Festing and Altman, 2002). The NC3Rs recently launched an online Experimental Design Assistant (EDA) to guide researchers in the design of experiments and ensure the minimum number of animals is used to achieve the scientific objectives[1].

3.3 *In vivo* imaging

In vivo imaging techniques enable longitudinal studies of the same animals through the course of infection, reducing the number of groups required for assessment at sequential timepoints and therefore variation. As opposed to CFU quantifica-

[3] http://www.alltrials.net/
[4] http://jnrbm.biomedcentral.com/
[4] http://bmcresnotes.biomedcentral.com/

tion in the lungs, which necessitates euthanasia, bioluminescent or fluorescently-tagged mycobacteria can be tracked in live animals for real-time assessment of vaccine efficacy (Zelmer et al., 2012). Such non-invasive techniques also fall into the "refinement" category.

Zhang et al. used an autoluminescent strain of *Mtb* as a surrogate marker to replace CFU counts. Relative light units (RLU) *in vivo* paralleled CFU counts *in vitro* during the active phase of bacterial growth, and the ability of a recombinant BCG vaccine to limit bacterial growth was demonstrated. Although the modest sensitivity of the system necessitates a greater bacterial burden, leading to more widespread dissemination of infection, the authors suggest methods by which the degree of luminescence may be improved (Zhang et al., 2012).

In vivo imaging may also be used to visualize clinical symptoms of TB disease, again reducing the need for endpoint measures such as CFU. Lewinsohn et al. (2006) described the use of computed tomography (CT) scanning in macaques, demonstrating a strong correlation with pathohistologic findings at necropsy.

CT and MRI scanning have since been applied in a number of NHP TB drug and vaccine studies to determine number, structure and distribution of pulmonary lesions across the lung nodes following *Mtb* challenge (Lin et al., 2013; Rayner et al., 2013; Sharpe et al., 2016). Importantly, whereas routine readouts such as bacterial burden and gross pathology necessitate the use of high doses of *Mtb*, sensitive imaging techniques permit much lower challenge doses of *Mtb*, as discussed in the following section (Rayner et al., 2013; Sharpe et al., 2016).

4 Refinement

Refinement refers to "methods that minimise the pain, suffering, distress or lasting harm that may be experienced by the animals"[1]. This applies to all aspects of animal use, including housing and husbandry. Wolfensohn et al. (2015) recently described a quantitative system for assessment of lifetime experience, assigning a combined welfare score (CWAS) to various permutations of the macaque model of TB. This system scored four parameters (physical, psychological, environmental and procedural [experimental and/or clinical]), which contributed to a level of combined severity. This measure will be discussed in the context of TB challenge models; improvements in more general areas such as importation and living conditions are beyond the scope of this review.

4.1 BCG challenge

Replacing virulent *Mtb* with attenuated BCG would not only reduce the severity of animal challenge experiments, but also offer the opportunity to conduct challenge experiments in humans as a tool for assessing and prioritizing candidate vaccines at an early stage of development as described.

As described above, a BCG challenge model for use in humans has recently been developed (Harris et al., 2014; Minassian et al., 2012; Minhinnick et al., 2016). BCG challenge in animal models similarly represents an alternative to the use of virulent *Mtb*, reducing the severity of pathology; with the caveat that reduced pathology makes for a less realistic and perhaps less sensitive model. The impact on lifetime experience that could be achieved through the use of a BCG (rather than *Mtb*) challenge model in macaques was evaluated through comparison of the CWAS during the post-challenge phase, which was shown to be considerably greater (i.e., improved welfare) at 2 weeks post-BCG challenge compared with 16, 26 or 52 weeks post-*Mtb* challenge (Wolfensohn et al., 2015). Using a BCG challenge model in cynomolgus macaques, significantly lower levels of BCG were detected in the axillary lymph nodes draining the site of challenge in BCG-vaccinated compared with naïve animals. Furthermore, higher *ex vivo* PPD-specific IFN-γ ELISpot responses and enhanced *in vitro* mycobacterial growth inhibition were associated with lower CFU counts in the draining lymph node, suggesting utility of this model in identifying correlates of immunity (own unpublished data). A similar model was employed in cattle, demonstrating that BCG vaccinated animals had lower BCG CFU counts than naïve animals following intranodal challenge with BCG (Villarreal-Ramos et al., 2014). It has been shown that BCG-vaccinated mice later challenged with intradermal BCG had reduced mycobacterial growth, and this protection was predictive of BCG efficacy against aerosol *Mtb* challenge (Minassian et al., 2011).

4.2 Ultra-low-dose challenge

High doses of *Mtb* are typically required to induce meaningful changes in clinical parameters and pathology that permit the measurement of vaccine efficacy. However, a lower challenge dose would not only more closely resemble natural infection, but also reduce disease burden and therefore symptoms. Whereas NHP challenge models typically use inoculum sizes of between 50 and 3000 CFU of *Mtb* (Langermans et al., 2001; Lewinsohn et al., 2006; Lin et al., 2012; Verreck et al., 2009), an ultra-low dose challenge recently described by Sharpe et al. (2016) exposed macaques to less than 10 CFU. Macaques did not exhibit abnormal behaviors or marked clinical signs, unlike with normal high dose challenge. Furthermore, comparison of the CWAS score for unvaccinated animals during the first 16 weeks after challenge with high and low dose *Mtb* shows that there was a beneficial effect on welfare of using a reduced dose (Wolfensohn et al., 2015). However, such a model does require more sensitive approaches to evaluate disease burden such as FHG PET-CT *in vivo* imaging, though these offer the added advantage of serial assessment as described above. Using these methods, the authors were able to discriminate between rhesus and cynomolgus macaques in terms of disease burden and progression, reflecting previously described differences in disease

outcome (Sharpe et al., 2016). Concerns that ultra-low doses would lead to increased variability or fail to reliably infect all of the challenged animals did not appear to be founded (Sharpe et al., 2016).

4.3 Humane end-points

The NC3Rs defines humane end-points as "clear, predictable and irreversible criteria which substitute for more severe experimental outcomes such as advanced pathology or death"[1]. Unfortunately, such criteria often remain poorly defined, and the long duration typical of animal experiments involving TB and other chronic progressive infections provides greater potential for ambiguity. If no measures are taken to treat *Mtb*-infected mice, they will succumb to infection and die before their average life-span (Medina and North, 1998), and in a systematic review of endpoints implemented in 80 murine TB studies published in 2009, 47% of the studies were classified as "lethal" (not terminated before animals reached advanced stages of disease, which would rapidly progress towards spontaneous death if no other endpoints were applied). 66% of these were categorized at the highest severity level, meaning that animals were allowed to die spontaneously or reach a moribund state (Franco et al., 2012).

In addition to the welfare concern, survival may not necessarily be the most controlled or statistically powerful measure, as described above. In a TB vaccine study of long-term survival in rhesus macaques, lung lesion burden using MR imaging and stereology, but not survival time, was able to distinguish naïve and vaccinated NHPs (Sharpe et al., 2010). Such "in-life" imaging may represent a more humane readout of vaccine efficacy than survival.

In most TB candidate vaccine studies reported in recent years, animals were euthanized at a fixed time-point following *Mtb* challenge and alternative measures of disease severity such as CFU counts in lungs and spleen were assessed (Gillis et al., 2014; Stylianou et al., 2015). Various measures have been described as a cut-off parameter for euthanasia including non-transient hypothermia and, more commonly, change in body weight. However, although weight itself is quantitative and objective, this measure is confounded by natural fluctuations and the definition of the

Tab. 1: Summary of 3Rs approaches to tuberculosis vaccine research

3Rs area	Method	Examples
Replacement	Human challenge	BCG challenge as a surrogate for *Mtb* (Harris et al., 2014; Minassian et al., 2012; Minhinnick et al., 2016)
	In vitro assays	IFN-γ ELISpot and IgG ELISA as correlates of risk (Fletcher et al., 2016) Mycobacterial growth inhibition assays (Cheng et al., 1988; Cheon et al., 2002; Fletcher et al., 2013; Hoft et al., 2002; Kampmann et al., 2004; Worku and Hoft, 2000)
	In silico modelling	Identification and selection of epitopes for inclusion in novel TB vaccine candidates (Monterrubio-López et al., 2015; Mustafa, 2005; Rodriguez et al., 2011; Tang et al., 2011; Wang et al., 2010)
	Less sentient animals	Zebrafish (Clay et al., 2007; Oksanen et al., 2013)
Reduction	Reducing replication	Improved sharing of negative results/trial registering
	Experimental design	Sample size and power (Dell et al., 2002; Festing and Altman, 2002; Williams et al., 2009) Controlling variation, factorial design, blinding (Festing and Altman, 2002)
	In vivo imaging	Fluorescently-tagged mycobacteria (Zelmer et al., 2012; Zhang et al., 2012) Visualising clinical symptoms (Lewinsohn et al., 2006; Lin et al., 2013; Rayner et al., 2013; Sharpe et al., 2016)
Refinement	BCG challenge	BCG challenge in non-human primates and cattle (Minassian et al., 2011; Wolfensohn et al., 2015; Villarreal-Ramos et al., 2014)
	Ultra-low-dose challenge	Non-human primates (Wolfensohn et al., 2015; Sharpe et al., 2016)
	Humane end-points	Fixed endpoints (Gillis et al., 2014; Wolfensohn et al., 2015; Stylianou et al., 2015) Prevention of infection as an endpoint (McShane and Williams, 2014)

upper boundary varies considerably across studies, ranging from 10 to 30% weight loss (Franco et al., 2012; Williams et al., 2009). As described, survival studies remain informative in certain circumstances (for example in demonstrating sustained protection by a vaccine candidate), but even then "death" is now defined as the time at which moribund animals are humanely euthanized (Franco et al., 2012).

Wolfensohn et al. (2015) compared CWAS scores to quantify differences in lifetime experience when a survival endpoint (52-week post-challenge follow-up) was used as opposed to a fixed end-point (16 or 26 weeks post-challenge) for the evaluation of TB vaccine efficacy. There was a considerable reduction in welfare "cost" with decreasing time post-challenge. Although alternative read-outs such as bacterial load usually support survival data, further work is required to develop more accurate predictors of death or survival; indeed parameters such as blood glucose homeostasis have been shown to correlate with degree of severity in other disease models (Yu et al., 2011).

Alternatively, McShane and Williams suggest that the field may consider a movement from reduction of disease to prevention of *Mtb* infection as the endpoint of animal challenge experiments (McShane and Williams, 2014). Though larger numbers of animals would be required, such a model would eliminate many of the issues associated with TB disease endpoints and better align with the definition of vaccine efficacy in humans.

5 Conclusion

Animal models have been, and continue to be, invaluable in the TB vaccine field. As well as enhancing our understanding of TB disease and immunology, they are essential for assessing immunogenicity, efficacy and safety of vaccine candidates prior to advancement into clinical trials. UK legislation dictates that research must be performed in accordance with the 3Rs framework to ensure ethical rigor. We have reviewed examples from the TB vaccine field of replacing, reducing and refining animal experiments, summarized in Table 1. Replacements include the use of humans for BCG challenge experiments, *in vitro* assays such as IFN-γ ELISpot and MGIAs, *in silico* techniques and the use of less sentient animals. The number of animals used may be reduced by avoiding data replication, improving experimental design, and technologies such as *in vivo* imaging enabling longitudinal studies in the same animals. Refinements specific to TB challenge models include the use of lower doses of *Mtb* or less pathogenic mycobacterial strains, and clearly-defined humane fixed end-points. Recent technological advances have opened the door to alternatives and improvements which must now be taken forward to ensure impact. The approaches described not only advance animal welfare, but also demonstrate the ability of projects with a 3Rs justification to drive forward the science in the development of improved TB vaccines.

References

Apt, A. and Kramnik, I. (2009). Man and mouse TB: Contradictions and solutions. *Tuberculosis (Edinb) 89*, 195-198. https://doi.org/10.1016/j.tube.2009.02.002

Balls, M. (2009). The origins and early days of the Three Rs concept. *Altern Lab Anim 37*, 255-265.

Basaraba, R. J. (2008). Experimental tuberculosis: The role of comparative pathology in the discovery of improved tuberculosis treatment strategies. *Tuberculosis (Edinb) 88, Suppl 1*, S35-47. https://doi.org/10.1016/S1472-9792(08)70035-0

Behr, M. A., Wilson, M. A., Gill, W. P. et al. (1999). Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science 284*, 1520-1523. https://doi.org/10.1126/science.284.5419.1520

Bowman, B. N., McAdam, P. R., Vivona, S. et al. (2011). Improving reverse vaccinology with a machine learning approach. *Vaccine 29*, 8156-8164. https://doi.org/10.1016/j.vaccine.2011.07.142

Carpenter, E., Fray, L. and Gormley, E. (1997). Cellular responses and Mycobacterium bovis BCG growth inhibition by bovine lymphocytes. *Immunol Cell Biol 75*, 554-560. https://doi.org/10.1038/icb.1997.86

Cheng, S. H., Walker, L., Poole, J. et al. (1988). Demonstration of increased anti-mycobacterial activity in peripheral blood monocytes after BCG vaccination in British school children. *Clin Exp Immunol 74*, 20-25.

Cheon, S. H., Kampmann, B., Hise, A. G. et al. (2002). Bactericidal activity in whole blood as a potential surrogate marker of immunity after vaccination against tuberculosis. *Clin Diagn Lab Immunol 9*, 901-907. https://doi.org/10.1128/cdli.9.4.901-907.2002

Clark, S., Hall, Y. and Williams, A. (2015). Animal models of tuberculosis: Guinea pigs. *Cold Spring Harb Perspect Med 5*, a018572. https://doi.org/10.1101/cshperspect.a018572

Clay, H., Davis, J. M., Beery, D. et al. (2007). Dichotomous role of the macrophage in early Mycobacterium marinum infection of the zebrafish. *Cell Host Microbe 2*, 29-39. https://doi.org/10.1016/j.chom.2007.06.004

Cowley, S. C. and Elkins, K. L. (2003). $CD4^+$ T cells mediate IFN-gamma-independent control of Mycobacterium tuberculosis infection both in vitro and in vivo. *J Immunol 171*, 4689-4699. https://doi.org/10.4049/jimmunol.171.9.4689

Cronan, M. R. and Tobin, D. M. (2014). Fit for consumption: Zebrafish as a model for tuberculosis. *Dis Model Mech 7,* 777-784. https://doi.org/10.1242/dmm.016089

de Jonge, M. I., Brosch, R., Brodin, P. et al. (2005). Tuberculosis: From genome to vaccine. *Expert Rev Vaccines 4*, 541-551. https://doi.org/10.1586/14760584.4.4.541

Dell, R. B., Holleran, S. and Ramakrishnan, R. (2002). Sample size determination. *ILAR J 43*, 207-213. https://doi.org/10.1093/ilar.43.4.207

Denis, M., Wedlock, D. N. and Buddle, B. M. (2004). Ability of T cell subsets and their soluble mediators to modulate the replication of Mycobacterium bovis in bovine macro-

phages. *Cell Immunol 232*, 1-8. https://doi.org/10.1016/j.cellimm.2005.01.003

Dionne, M. S., Ghori, N. and Schneider, D. S. (2003). Drosophila melanogaster is a genetically tractable model host for Mycobacterium marinum. *Infect Immun 71*, 3540-3550. https://doi.org/10.1128/IAI.71.6.3540-3550.2003

Elias, D., Akuffo, H. and Britton, S. (2005). PPD induced in vitro interferon gamma production is not a reliable correlate of protection against Mycobacterium tuberculosis. *Trans R Soc Trop Med Hyg 99*, 363-368. https://doi.org/10.1016/j.trstmh.2004.08.006

Festing, M. F. and Altman, D. G. (2002). Guidelines for the design and statistical analysis of experiments using laboratory animals. *ILAR J 43*, 244-258. https://doi.org/10.1093/ilar.43.4.244

Fletcher, H. A., Tanner, R., Wallis, R. S. et al. (2013). Inhibition of mycobacterial growth in vitro following primary but not secondary vaccination with Mycobacterium bovis BCG. *Clin Vaccine Immunol 20*, 1683-1689. https://doi.org/10.1128/CVI.00427-13

Fletcher, H. A., Snowden, M. A., Landry, B. et al. (2016). T-cell activation is an immune correlate of risk in BCG vaccinated infants. *Nat Commun 7*, 11290. https://doi.org/10.1038/ncomms11290

Flynn, J. L., Gideon, H. P., Mattila, J. T. and Lin, P. L. (2015). Immunology studies in non-human primate models of tuberculosis. *Immunol Rev 264*, 60-73. https://doi.org/10.1111/imr.12258

Franco, N. H., Correia-Neves, M. and Olsson, I. A. (2012). How "humane" is your endpoint? Refining the science-driven approach for termination of animal studies of chronic infection. *PLoS Pathog 8*, e1002399. https://doi.org/10.1371/journal.ppat.1002399

Gillis, T. P., Tullius, M. V. and Horwitz, M. A. (2014). rBCG30-induced immunity and cross-protection against Mycobacterium leprae challenge are enhanced by boosting with the Mycobacterium tuberculosis 30-kilodalton antigen 85B. *Infect Immun 82*, 3900-3909. https://doi.org/10.1128/IAI.01499-13

Gupta, U. D. and Katoch, V. M. (2009). Animal models of tuberculosis for vaccine development. *Indian J Med Res 129*, 11-18.

Hagedorn, M. and Soldati, T. (2007). Flotillin and RacH modulate the intracellular immunity of dictyostelium to Mycobacterium marinum infection. *Cell Microbiol 9,* 2716-2733. https://doi.org/10.1111/j.1462-5822.2007.00993.x

Hagedorn, M., Rohde, K. H., Russell, D. G. and Soldati, T. (2009). Infection by tubercular mycobacteria is spread by nonlytic ejection from their amoeba hosts. *Science 323*, 1729-1733. https://doi.org/10.1126/science.1169381

Harris, S. A., Meyer, J., Satti, I. et al. (2014). Evaluation of a human BCG challenge model to assess antimycobacterial immunity induced by BCG and a candidate tuberculosis vaccine, MVA85A, alone and in combination. *J Infect Dis 209*, 1259-1268. https://doi.org/10.1093/infdis/jit647

Hoft, D. F., Worku, S., Kampmann, B. et al. (2002). Investigation of the relationships between immune-mediated inhibition of mycobacterial growth and other potential surrogate markers of protective Mycobacterium tuberculosis immunity. *J Infect Dis 186*, 1448-1457. https://doi.org/10.1086/344359

Kagina, B. M., Abel, B., Scriba, T. J. et al. (2010). Specific T cell frequency and cytokine expression profile do not correlate with protection against tuberculosis after bacillus Calmette-Guérin vaccination of newborns. *Am J Respir Crit Care Med 182*, 1073-1079. https://doi.org/10.1164/rccm.201003-0334OC

Kampmann, B., Tena, G. N., Mzazi, S. et al. (2004). Novel human in vitro system for evaluating antimycobacterial vaccines. *Infect Immun 72*, 6401-6407. https://doi.org/10.1128/IAI.72.11.6401-6407.2004

Kolibab, K., Parra, M., Yang, A. L. et al. (2009). A practical in vitro growth inhibition assay for the evaluation of TB vaccines. *Vaccine 28*, 317-322. https://doi.org/10.1016/j.vaccine.2009.10.047

Langermans, J. A., Andersen, P., van Soolingen, D. et al. (2001). Divergent effect of bacillus Calmette-Guérin (BCG) vaccination on Mycobacterium tuberculosis infection in highly related macaque species: Implications for primate models in tuberculosis vaccine research. *Proc Natl Acad Sci U S A 98*, 11497-11502. https://doi.org/10.1073/pnas.201404898

Lewinsohn, D. M., Tydeman, I. S., Frieder, M. et al. (2006). High resolution radiographic and fine immunologic definition of TB disease progression in the rhesus macaque. *Microbes Infect 8*, 2587-2598. https://doi.org/10.1016/j.micinf.2006.07.007

Lin, P. L., Dietrich, J., Tan, E. et al. (2012). The multistage vaccine H56 boosts the effects of BCG to protect cynomolgus macaques against active tuberculosis and reactivation of latent Mycobacterium tuberculosis infection. *J Clin Invest 122*, 303-314. https://doi.org/10.1172/JCI46252

Lin, P. L., Coleman, T., Carney, J. P. et al. (2013). Radiologic responses in cynomolgous macaques for assessing tuberculosis chemotherapy regimens. *Antimicrob Agents Chemother 57*, 4237-4244. https://doi.org/10.1128/AAC.00277-13

Macleod, M. R., Michie, S., Roberts, I. et al. (2014). Biomedical research: Increasing value, reducing waste. *Lancet 383*, 101-104. https://doi.org/10.1016/S0140-6736(13)62329-6

Mahomed, H., Kibel, M., Hawkridge, T. et al. (2006). The impact of a change in bacille Calmette-Guérin vaccine policy on tuberculosis incidence in children in Cape Town, South Africa. *Pediatr Infect Dis J 25*, 1167-1172. https://doi.org/10.1097/01.inf.0000243765.33880.54

Marsay, L., Matsumiya, M., Tanner, R. et al. (2013). Mycobacterial growth inhibition in murine splenocytes as a surrogate for protection against Mycobacterium tuberculosis (M. tb). *Tuberculosis (Edinb) 93*, 551-557. https://doi.org/10.1016/j.tube.2013.04.007

Marwick, C. (1998). Volunteers in typhoid infection study will aid future vaccine development. *JAMA 279*, 1423-1424. https://doi.org/10.1001/jama.279.18.1423

McKenzie, R., Fried, M. W., Sallie, R. et al. (1995). Hepatic failure and lactic acidosis due to fialuridine (FIAU), an investigational nucleoside analogue for chronic hepatitis B. *N Engl J Med 333*, 1099-1105. https://doi.org/10.1056/NEJM199510263331702

McMurray, D. N., Collins, F. M., Dannenberg, A. M. and Smith, D. W. (1996). Pathogenesis of experimental tuberculosis in animal models. *Curr Top Microbiol Immunol 215*, 157-179.

McShane, H. and Williams, A. (2014). A review of preclinical animal models utilised for TB vaccine evaluation in the context of recent human efficacy data. *Tuberculosis (Edinb) 94*, 105-110. https://doi.org/10.1016/j.tube.2013.11.003

Medina, E. and North, R. J. (1998). Resistance ranking of some common inbred mouse strains to Mycobacterium tuberculosis and relationship to major histocompatibility complex haplotype and Nramp1 genotype. *Immunology 93*, 270-274. https://doi.org/10.1046/j.1365-2567.1998.00419.x

Meijer, A. H. and Spaink, H. P. (2011). Host-pathogen interactions made transparent with the zebrafish model. *Curr Drug Targets 12*, 1000-1017. https://doi.org/10.2174/138945011795677809

Mestas, J. and Hughes, C. C. (2004). Of mice and not men: Differences between mouse and human immunology. *J Immunol 172*, 2731-2738. https://doi.org/10.4049/jimmunol.172.5.2731

Minassian, A. M., Ronan, E. O., Poyntz, H. et al. (2011). Preclinical development of an in vivo BCG challenge model for testing candidate TB vaccine efficacy. *PLoS One 6*, e19840. https://doi.org/10.1371/journal.pone.0019840

Minassian, A. M., Satti, I., Poulton, I. D. et al. (2012). A human challenge model for Mycobacterium tuberculosis using Mycobacterium bovis bacille Calmette-Guerin. *J Infect Dis 205*, 1035-1042. https://doi.org/10.1093/infdis/jis012

Minhinnick, A., Harris, S., Wilkie, M. et al. (2016). Optimization of a human bacille Calmette-Guérin challenge model: A tool to evaluate antimycobacterial immunity. *J Infect Dis 213*, 824-830. https://doi.org/10.1093/infdis/jiv482

Mittrücker, H. W., Steinhoff, U., Köhler, A. et al. (2007). Poor correlation between BCG vaccination-induced T cell responses and protection against tuberculosis. *Proc Natl Acad Sci U S A 104*, 12434-12439. https://doi.org/10.1073/pnas.0703510104

Monterrubio-López, G. P., González-Y-Merchand, J. A. and Ribas-Aparicio, R. M. (2015). Identification of novel potential vaccine candidates against tuberculosis based on reverse vaccinology. *Biomed Res Int 2015*, 483150. https://doi.org/10.1155/2015/483150

Moyo, S., Verver, S., Mahomed, H. et al. (2010). Age-related tuberculosis incidence and severity in children under 5 years of age in Cape Town, South Africa. *Int J Tuberc Lung Dis 14*, 149-154.

Mustafa, A. S. (2005). Mycobacterial gene cloning and expression, comparative genomics, bioinformatics and proteomics in relation to the development of new vaccines and diagnostic reagents. *Med Princ Pract 14, Suppl 1*, 27-34. https://doi.org/10.1159/000086182

Niazi, M. K., Dhulekar, N., Schmidt, D. et al. (2015). Lung necrosis and neutrophils reflect common pathways of susceptibility to Mycobacterium tuberculosis in genetically diverse, immune-competent mice. *Dis Model Mech 8*, 1141-1153. https://doi.org/10.1242/dmm.020867

Niemann, S. and Supply, P. (2014). Diversity and evolution of Mycobacterium tuberculosis: Moving to whole-genome-based approaches. *Cold Spring Harb Perspect Med 4*, a021188. https://doi.org/10.1101/cshperspect.a021188

Oksanen, K. E., Halfpenny, N. J., Sherwood, E. et al. (2013). An adult zebrafish model for preclinical tuberculosis vaccine development. *Vaccine 31*, 5202-5209. https://doi.org/10.1016/j.vaccine.2013.08.093

Orme, I. M. and Basaraba, R. J. (2014). The formation of the granuloma in tuberculosis infection. *Semin Immunol 26*, 601-609. https://doi.org/10.1016/j.smim.2014.09.009

Parikka, M., Hammarén, M. M., Harjula, S. K. et al. (2012). Mycobacterium marinum causes a latent infection that can be reactivated by gamma irradiation in adult zebrafish. *PLoS Pathog 8*, e1002944. https://doi.org/10.1371/journal.ppat.1002944

Parra, M., Yang, A. L., Lim, J. et al. (2009). Development of a murine mycobacterial growth inhibition assay for evaluating vaccines against Mycobacterium tuberculosis. *Clin Vaccine Immunol 16*, 1025-1032. https://doi.org/10.1128/CVI.00067-09

Rayner, E. L., Pearson, G. R., Hall, G. A. et al. (2013). Early lesions following aerosol infection of rhesus macaques (Macaca mulatta) with Mycobacterium tuberculosis strain H37RV. *J Comp Pathol 149*, 475-485. https://doi.org/10.1016/j.jcpa.2013.05.005

Rhoades, E. R., Frank, A. A. and Orme, I. M. (1997). Progression of chronic pulmonary tuberculosis in mice aerogenically infected with virulent Mycobacterium tuberculosis. *Tuber Lung Dis 78*, 57-66. https://doi.org/10.1016/S0962-8479(97)90016-2

Rodriguez, L., Tirado, Y., Reyes, F. et al. (2011). Proteoliposomes from Mycobacterium smegmatis induce immune cross-reactivity against Mycobacterium tuberculosis antigens in mice. *Vaccine 29*, 6236-6241. https://doi.org/10.1016/j.vaccine.2011.06.077

Sada-Ovalle, I., Chiba, A., Gonzales, A. et al. (2008). Innate invariant NKT cells recognize Mycobacterium tuberculosis-infected macrophages, produce interferon-gamma, and kill intracellular bacteria. *PLoS Pathog 4*, e1000239. https://doi.org/10.1371/journal.ppat.1000239

Sauerwein, R. W., Roestenberg, M. and Moorthy, V. S. (2011). Experimental human challenge infections can accelerate clinical malaria vaccine development. *Nat Rev Immunol 11*, 57-64. https://doi.org/10.1038/nri2902

Sharpe, S. A., McShane, H., Dennis, M. J. et al. (2010). Establishment of an aerosol challenge model of tuberculosis in rhesus macaques and an evaluation of endpoints for vaccine testing. *Clin Vaccine Immunol 17*, 1170-1182. https://doi.org/10.1128/CVI.00079-10

Sharpe, S., White, A., Gleeson, F. et al. (2016). Ultra low dose aerosol challenge with Mycobacterium tuberculosis leads to divergent outcomes in rhesus and cynomolgus macaques. *Tuberculosis (Edinb) 96*, 1-12. https://doi.org/10.1016/j.tube.2015.10.004

Song, F., Hooper, L. and Loke, Y. (2013). Publication bias: What is it? How do we measure it? How do we avoid it? *Open Access J Clin Trials 5*, 71-81. https://doi.org/10.2147/OAJCT.S34419

Stylianou, E., Griffiths, K. L., Poyntz, H. C. et al. (2015). Improvement of BCG protective efficacy with a novel chimpanzee adenovirus and a modified vaccinia Ankara virus both expressing Ag85A. *Vaccine 33*, 6800-6808. https://doi.org/10.1016/j.vaccine.2015.10.017

Suntharalingam, G., Perry, M. R., Ward, S. et al. (2006). Cytokine storm in a phase 1 trial of the anti-CD28 monoclonal antibody TGN1412. *N Engl J Med 355*, 1018-1028. https://doi.org/10.1056/NEJMoa063842

Swaim, L. E., Connolly, L. E., Volkman, H. E. et al. (2006). Mycobacterium marinum infection of adult zebrafish causes caseating granulomatous tuberculosis and is moderated by adaptive immunity. *Infect Immun 74*, 6108-6117. https://doi.org/10.1128/IAI.00887-06

Tang, S. T., van Meijgaarden, K. E., Caccamo, N. et al. (2011). Genome-based in silico identification of new Mycobacterium tuberculosis antigens activating polyfunctional $CD8^+$ T cells in human tuberculosis. *J Immunol 186*, 1068-1080. https://doi.org/10.4049/jimmunol.1002212

Tanner, R., O'Shea, M. K., White, A. D. et al. (submitted). The influence of haemoglobin and iron on in vitro mycobacterial growth inhibition assays.

Thakur, A., Pedersen, L. E. and Jungersen, G. (2012). Immune markers and correlates of protection for vaccine induced immune responses. *Vaccine 30*, 4907-4920. https://doi.org/10.1016/j.vaccine.2012.05.049

Verreck, F. A., Vervenne, R. A., Kondova, I. et al. (2009). MVA.85A boosting of BCG and an attenuated, phoP deficient M. tuberculosis vaccine both show protective efficacy against tuberculosis in rhesus macaques. *PLoS One 4*, e5264. https://doi.org/10.1371/journal.pone.0005264

Villarreal-Ramos, B., Berg, S., Chamberlain, L. et al. (2014). Development of a BCG challenge model for the testing of vaccine candidates against tuberculosis in cattle. *Vaccine 32*, 5645-5649. https://doi.org/10.1016/j.vaccine.2014.08.009

Vordermeier, H. M., Villarreal-Ramos, B., Cockle, P. J. et al. (2009). Viral booster vaccines improve Mycobacterium bovis BCG-induced protection against bovine tuberculosis. *Infect Immun 77*, 3364-3373. https://doi.org/10.1128/IAI.00287-09

Wang, J., Zhang, H. and Wang, H. (2010). Analysis of predicted $CD8^+$ T cell epitopes from proteins encoded by the specific RD regions of Mycobacterium tuberculosis for vaccine development and specific diagnosis. *Mol Biol Rep 37*, 1793-1799. https://doi.org/10.1007/s11033-009-9613-4

WHO – World Health Organisation (2016). Tuberculosis Fact Sheet. Reviewed March 2016. http://www.who.int/mediacentre/factsheets/fs104/en/, accessed 22.9.2016.

Williams, A., Hall, Y. and Orme, I. M. (2009). Evaluation of new vaccines for tuberculosis in the guinea pig model. *Tuberculosis (Edinb) 89*, 389-397. https://doi.org/10.1016/j.tube.2009.08.004

Wolfensohn, S., Sharpe, S., Hall, I. et al. (2015). Refinement of welfare through development of a quantitative system for assessment of lifetime experience. *Animal Welfare 24*, 139-149. https://doi.org/10.7120/09627286.24.2.139

Worku, S. and Hoft, D. F. (2000). In vitro measurement of protective mycobacterial immunity: Antigen-specific expansion of T cells capable of inhibiting intracellular growth of bacille Calmette-Guérin. *Clin Infect Dis 30, Suppl 3*, S257-261. https://doi.org/10.1086/313887

Yu, Y., Clippinger, A. J. and Alwine, J. C. (2011). Viral effects on metabolism: Changes in glucose and glutamine utilization during human cytomegalovirus infection. *Trends Microbiol 19*, 360-367. https://doi.org/10.1016/j.tim.2011.04.002

Zelmer, A., Carroll, P., Andreu, N. et al. (2012). A new in vivo model to test anti-tuberculosis drugs using fluorescence imaging. *J Antimicrob Chemother 67*, 1948-1960. https://doi.org/10.1093/jac/dks161

Zhang, T., Li, S. Y. and Nuermberger, E. L. (2012). Autoluminescent Mycobacterium tuberculosis for rapid, real-time, non-invasive assessment of drug and vaccine efficacy. *PLoS One 7*, e29774. https://doi.org/10.1371/journal.pone.0029774

Conflict of interest

The authors have no conflicts of interest or copyright issues to declare.

Acknowledgements

We are grateful to UFAW (Universities Federation for Animal Welfare) for a grant supporting this work. HMcS is a Wellcome Trust Senior Clinical Research Fellow.

Correspondence to

Rachel Tanner, PhD
The Jenner Institute
Nuffield Department of Clinical Medicine
University of Oxford
Oxford, UK
Phone: +44 1865 617617
e-mail: rachel.tanner@ndm.ox.ac.uk