CVI Accepts, published online ahead of print on 9 June 2010
Clin. Vaccine Immunol. doi:10.1128/CVI.00079-10

Establishment of an aerosol challenge model of tuberculosis in rhesus macaques, and an evaluation of endpoints for vaccine testing

SA Sharpe[1*], H McShane[2], MJ Dennis[1], R J Basaraba[3], F Gleeson[4], G Hall[1], A McIntyre[3], K Gooch[1], S Clark[1], N E R Beveridge[2], E. Nuth[3], A White[1], A Marriott[1], S Dowall[1], AVS Hill[2], A Williams[1,] & PD Marsh[1]

[1]Health Protection Agency, Centre for Emergency Preparedness and Response, Porton Down, Salisbury, UK.

[2]The Jenner Institute, University of Oxford, Oxford, UK.

[3]Colorado State University, Fort Collins, Colorado, USA.

[4]Churchill Hospital, Headington, Oxford, UK

*Corresponding author: Tel +441980 612811; Fax +441980 612731
Email: sally.sharpe@hpa.org.uk

Key words
Tuberculosis, Vaccine, non-human primate, aerosol challenge

Running title: Aerosol challenge model of tuberculosis in rhesus macaques

Abstract

The establishment of an aerosol challenge model in non-human primates (NHPs) for the testing of vaccines against *Mycobacterium tuberculosis (M. tuberculosis)* would assist the global effort in optimising novel vaccination strategies. The endpoints used in preclinical challenge studies to identify measures of disease burden need to be accurate and sensitive enough to distinguish subtle differences and benefits afforded by different TB vaccine regimens when group sizes are inevitably small. This study sought to assess clinical and non-clinical endpoints in a challenge study using rhesus macaques as potentially sensitive measures of disease burden, using a new protocol of aerosol administration of *M. tuberculosis*. Immunological and clinical read-outs were assessed for utility in vaccine evaluation studies. This is the first example of TB vaccine evaluation in rhesus macaques where long term survival was one of the primary endpoints. However, we found that in NHP vaccine efficacy studies with maximum group sizes of six, survival did not provide a valuable endpoint. Two approaches used in human clinical trials for the evaluation of the IFN-γ response to vaccination (ELISpot, ELISA), were included in this study. The IFN-γ profiles induced following vaccination were found not to correlate with protection, nor did the level of PPD-specific proliferation. The only readout to reliably distinguish vaccinated and unvaccinated NHPs was determination of lung lesion burden using Magnetic Resonance imaging (MR) combined with stereology at the end of the study. Therefore, the currently proposed key markers were not shown to correlate with protection, and only imaging offered a potentially reliable correlate.

(249 used, 250 words max)

Introduction

Tuberculosis (TB) is a re-emerging infectious disease, and is responsible for nearly 2 million deaths and 9 million new cases each year, (37). The global TB pandemic has been exacerbated by the emergence of drug resistant strains of *M. tuberculosis* which render treatment less effective, and by the HIV epidemic, where co-infection with HIV greatly increases the risk of reactivation of latent TB and susceptibility to active TB disease.

The most effective means of controlling this global epidemic would be by prophylactic immunization. *Mycobacterium bovis* Bacille Calmette Guerin (BCG), the only licensed TB vaccine, is administered to neonates in high-risk populations as part of the WHO Expanded Programme on Immunisation. BCG consistently protects against TB meningitis and disseminated TB in childhood (28, 31) but its efficacy wanes with time and it affords only variable protection against pulmonary disease (10). A new, more effective TB vaccine is a major global health priority and is an important part of the WHO STOP TB partnership strategy.

A large international effort is underway to develop a more effective vaccine. The leading TB vaccine development strategy involves vaccinating with BCG followed by a heterologous subunit vaccine boost designed to enhance protective immunity. One such subunit vaccine is the viral-vectored subunit candidate TB vaccine developed at Oxford University, MVA85A (19) (live, replication-deficient modified vaccinia virus Ankara (MVA) (7), expressing the highly conserved, immunodominant mycobacterial antigen 85A). Enhancing BCG with systemically administered (intradermal) MVA85A has been evaluated in several pre-clinical animal models and is currently subject to ongoing evaluation in several clinical trials. This BCG-MVA85A vaccination

regime induces a high magnitude of cellular immunity in mice and cattle (19, 33) and can protect against *M. tuberculosis* in guinea pigs (34), non-human primates (NHPs) (32) and cattle (34). It is safe and highly immunogenic in healthy adults, adolescents, children and infants, and in HIV and *M. tuberculosis* infected adults (6, 20, 22), and has recently entered a large scale efficacy trial in South African infants (21).

The lack of a defined immunological correlate of protection for TB means that, in order to assess efficacy, candidate TB vaccines must enter large clinical trials involving thousands of at-risk individuals in endemic countries. There is, therefore, a need for a validated preclinical animal model that can be utilized to accurately predict the effectiveness of a candidate vaccine in humans and to aid in the identification of correlates of protection through challenge studies. Mouse models are generally used as a first screen of vaccine candidates and are very useful for studying detailed immunological responses (25). Guinea pigs are considered a more stringent model than mice to discriminate between vaccines in terms of protective efficacy, since they show a variety of pulmonary and extra-pulmonary lesion types that are similar to those observed in humans (4, 17). Although small animal models are useful, it is widely accepted that larger animals such as cattle and NHPs are potentially the most relevant model species to predict safety, immunogenicity and protective efficacy of vaccines prior to their large-scale evaluation in humans (8, 18). NHPs are naturally susceptible to infection with *M. tuberculosis* via the respiratory route and develop a disease which clinically closely mimics human disease. As with the other preclinical species, BCG vaccination of NHP provides a limited level of protection against *M. tuberculosis* that can be

quantified through a variety of clinical and non-clinical parameters (2, 3, 9, 12, 14).

The establishment of an aerosol challenge model in NHPs, in which the *M. tuberculosis* challenge is delivered by the same route as occurs during natural infection would assist the global effort in optimising novel vaccination strategies. The endpoints used in preclinical challenge studies to identify measures of disease burden need to be accurate and sensitive enough to distinguish subtle differences and benefits afforded by different TB vaccine regimens, including those which enhance the protective efficacy of BCG, where partial protection is already conferred and the power to detect smaller incremental improvements in small numbers of animals is limited. The aims of this study were to establish an aerosol challenge model of TB in rhesus macaques, and to assess clinical and non-clinical end points in an NHP challenge study as potentially sensitive measures of disease burden, following vaccination with BCG or BCG boosted by MVA85A. The currently most widely used challenge model in macaques uses high dose, intra-tracheal administration of *M. tuberculosis* and it has been unclear whether low dose aerosol administration would provide a better model because although it should replicate the route of natural infection (32), larger group sizes may be necessary to compensate for inter-individual heterogeneity to a low dose challenge (8).

MATERIALS AND METHODS

Experimental animals

The animals used in this study were rhesus macaques of Indian origin obtained from an established UK breeding colony. All animals were 4 years

old at the time of challenge and naïve in terms of prior exposure to mycobacterial antigens (*M. tuberculosis* infection or environmental mycobacteria) as demonstrated by a negative tuberculin test whilst in their original breeding colony and by the IFN-γ based Primagam™ test kit (Biocor, CSL, USA) just prior to study start. Monkeys were housed according to the Home Office (UK) guidelines for the care and maintenance of primates and were sedated by intramuscular (i.m.) injection with ketamine hydrochloride (10mg/kg) (Ketaset, Fort Dodge Animal Health Ltd, Southampton, UK) for all procedures requiring removal from their cages. All procedures involving animals were approved by the Ethical Review Committee of the Centre for Emergency Preparedness and Response, Salisbury, UK. None of the animals had been used previously for experimental procedures

Vaccination

A plan of the vaccination strategy and timing of the aerosol challenge is shown in Figure 1. Twelve animals (groups A and B) were immunised intradermally in the upper left arm with 100 µl BCG, Danish strain 1331 (SSI, Copenhagen, Denmark). The viability of the BCG vaccine given to the NHPs was checked and found to be between 4.25 – 7.55 x 10^6 CFU/ml within the expected range for the vaccine batch at 2 - 8 x 10^6 CFU. Twelve weeks after immunisation with BCG, six animals (Group A) were immunised intradermally in the upper right arm with 100 µl (5 x 10^8 pfu) MVA85A. Vaccination sites were monitored and assessed for local reactions after vaccination with BCG and MVA85A.

M. tuberculosis challenge strain

The *M. tuberculosis* Erdman strain K 01 was kindly provided by Dr Amy Yang (CBER/FDA). The stocks were provided as frozen suspensions at a stated titre of 3.5 ± 1.5 x 10^8 colony forming units (CFU)/ml. On the day of challenge, five vials were thawed and diluted appropriately, as described below, in sterile distilled water. Twenty one weeks after immunisation with BCG the twelve immunised animals (Groups A and B) together with four unvaccinated animals (Group C) were challenged by the aerosol route with *M. tuberculosis* (Figure 1).

Aerosol exposure

The process to deliver a known number of viable *M. tuberculosis* bacilli in a target volume of inspired aerosol was performed as previously described (29). In brief, mono-dispersed bacteria in particles were generated using a 3-jet Collison nebuliser (BGI) and, in conjunction with a modified Henderson apparatus (11), delivered to the nose of each sedated primate via a modified veterinary anaesthesia mask. Challenge was performed on sedated animals placed within a 'head-out', plethysmography chamber (Buxco, Wilmington, North Carolina, USA) to enable the aerosol to be delivered simultaneously with the measurement of respiration rate.

Challenge regime

Aerosol exposure was performed over two consecutive days; groups of eight animals comprising two unvaccinated animals and three animals from each of the vaccine groups were challenged on each day.

A series of *in-vitro* and *in-vivo* studies were conducted in order to establish the feasibility of performing the challenge over two consecutive days, using the

same initial stock suspension. These studies are described in detail elsewhere (Clark *et al*, submitted for publication) and essentially demonstrated that:

(i) *M. tuberculosis* could be stored overnight at 4^{o}C in the dark without significant loss in viability.

(ii) The process of generating aerosols of *M. tuberculosis* using a Collison nebuliser resulted in an initial rapid loss in viability during the first 5 minutes but, after 10 minutes, the viability remained stable for approximately 20 minutes continuous aerosolisation. To avoid the initial, highly variable fluctuation in viability and to ensure that each individual exposure occurred during the period of stable viability, a strategy was developed where a 10 minute 'pre-treatment' of the challenge suspension, was performed.

On the day of challenge, five vials of Erdmann strain K 01 were thawed, pooled and diluted 1:60 in sterile distilled water to a total volume of 200 ml which was split into two 100 ml aliquots. One aliquot was used immediately to challenge eight animals, the second aliquot was stored at 2-8^{o}C in the dark for challenge of the remaining eight animals on the second day. At the start of each challenge day, one 100 ml aliquot of the mycobacterial suspension was 'pre-treated' in the Collison nebuliser for 10 minutes, then divided into eight aliquots each of 12.5 ml which were stored at room temperature in the dark until use. A fresh aliquot of pre-treated Erdman was used to challenge each animal. Animals were sedated with a combination of ketamine/acepromazine/atropine (Ketaset, Fort Dodge Animal Health, Southampton, UK (x 100mg/ml ketamine), ACP, Novartis (x 10mg/ml), atropine sulphate, Martindale Pharmaceuticals, Romford, UK (x 0.6mg/ml)) in a ratio of 5:1:1 and then exposed to the aerosol of *M. tuberculosis* for

sufficient time (4 – 10 minutes) for inhalation of an acquired volume of 7 L. Macaques were challenged with estimated doses of 40-60 CFU retained in the lungs (Table 1).

Clinical assessment post challenge

Animal behaviour was observed daily throughout the study for contra-indicators such as depression, withdrawal from the group, aggression, food and water intake, changes in respiration rate, or cough. Animals were sedated every two weeks to measure weight, body temperature, blood haemoglobin levels and erythrocyte sedimentation rate, (ESR) and to collect blood samples for immunology. Blood cell haemoglobin was measured using a HaemaCue haemoglobinometer (Haemacue Ltd, Dronfield, UK) and ESR was measured using the Sediplast system (Guest Medical, Edenbridge, UK). The time of necropsy, if prior to the end of the planned study period, was determined by experienced primatology staff based on a combination of the following adverse indicators: depressed or withdrawn behaviour, abnormal respiratory rates (dyspnoea), loss of 20% of peak post-challenge weight, ESR levels elevated above normal, haemoglobin level below normal limits, increased temperature and severely abnormal chest X-ray.

Interferon-gamma (IFN-γ) ELISpot Assay

Peripheral blood mononuclear cells (PBMC) were isolated from heparin anti-coagulated blood by Ficoll-Hypaque PLUS (GE Healthcare, Buckinghamshire, UK) density gradient separation using standard procedures. Isolated cells were washed in medium (R2) consisting of RPMI 1640 (Sigma-Aldrich,

Dorset, UK) supplemented with L-glutamine (2 mM) (Sigma-Aldrich, Dorset, UK), penicillin (50 U/ml)/streptomycin (50 μg/ml) (Sigma-Aldrich, Dorset, UK), 25mM Hepes buffer (Sigma-Aldrich, Dorset, UK), 0.05mM 2-mercaptoethanol (Invitrogen, Paisley, UK) and 2% heat-inactivated foetal bovine serum (LabTech, Ringmer, England).

An IFN-γ ELISpot assay was used to estimate the numbers and IFN-γ production capacity of mycobacteria-specific T cells in PBMC using a human/monkey IFN-γ kit (MabTech, Nacka. Sweden). Cells were stimulated with PPD (10 μg/ml, SSI, Copenhagen, Denmark) or pools of overlapping 15mer peptides spanning Ag85A, or ESAT6 (Peptide Protein Research Ltd, Wickham, UK). Plates (Millipore, Watford, UK) were coated overnight at 4^{o}C with 10 μg/ml of IFN-γ antibody. 200,000, 100,000 or 5000 PBMC were plated per well in 100 μl R5 medium (RPMI 1640 supplemented with L-glutamine (2 mM), penicillin (50 U/ml)/streptomycin (50 μg/ml) and 5% heat-inactivated foetal bovine serum), with or without antigen, in triplicate, and incubated for 18 hours. Phorbol 12-myristate (Sigma-Aldrich Dorset, UK) (100 ng/ml) and ionomycin (CN Biosciences, Nottingham, UK) (1 μg/ml) were used as a positive control. After culture, plates were washed and incubated for 2 hours with biotinylated anti-IFN-γ. Spots were developed by adding streptavidin-alkaline phosphatase and freshly prepared 5-Bromo-4-Chloro-3-Indolyl Phosphate/Nitro Blue Tetrazolium (BCIP/NBT) substrate. Determinations from triplicate tests were averaged. Data were analysed by subtracting the mean number of spots in the cells and medium-only control wells from the mean counts of spots in wells with cells and antigen or peptide pools.

Quantification of secreted IFN-γ

IFN-γ production was measured using a whole blood ELISA as described by Black *et al*., (5). In brief, heparinised blood was diluted 1 in 10 with serum-free medium (RPMI supplemented with L-glutamine, penicillin and streptomycin) and cultured with purifed protein derivative from *M. tuberculosis* (PPD; 5 µg/ml, SSI, Copenhagen, Denmark) or mitogen Phytohaemaglutinin (PHA; Sigma-Aldrich, Dorset, UK) (5µg/l), or in medium alone for 6 days. Supernatants were harvested at day 6 and stored at -20^{o}C. The quantity of IFN-γ in the supernatants was estimated using a commercially available human/monkey IFN-γ ELISA kit (MabTech, Nacka. Sweden). A purified human IFN-γ was used for the standard curve on each plate. The ELISA was developed using streptavidin and 3,3',5,5'-tetramethylbenzidine (TMB) liquid substrate system (Sigma-Aldrich, Dorset, UK) and the reaction stopped with 2M sulphuric acid (May & Baker Ltd, Dagenham, UK). ELISA plates were read at 450nm. A standard curve was plotted for each plate and used to calculate the concentrations of IFN-γ in each sample.

Intracellular cytokine staining

Freshly isolated PBMCs were stimulated with either PBS (negative control), PPD or staphylococcus enterotoxin B (SEB) (Sigma, Poole, UK) at 10 µg/ml with 0.1 µg/ml of purified αCD28 and 0.1 µg/ml of purified αCD49d (both BD Biosciences, Oxford, UK). Cells were incubated at 37^{o}C for 5 hours. Golgiplug, (BD Biosciences, Oxford, UK) was added and cells incubated for a further 12 hours at 37^{o}C. Cells were stained for expression of $CD8^{+}$, $CD3^{+}$, $CD4^{+}$ and IFN-γ using standard commercially available fluorochrome conjugated reagents (BD Biosciences, Oxford, UK and Caltag Invitrogen,

Paisley, UK) acquired on a FC500 Cytomics flow cytometer (Beckman Coulter, High Wycombe, UK) and analysed using CXP V2.1 software (Beckman Coulter, High Wycombe, UK).

Polyfunctional intracellular cytokine staining

PBMCs were thawed, washed, re-suspended in medium (R10) consisting of RPMI 1640 supplemented with L-glutamine (2 mM), penicillin (50 U/ml)/streptomycin (50 μg/ml) and 10% heat-inactivated foetal bovine serum with 1 U/ml of DNase (Sigma, Poole, UK) and incubated at 37°C. After 2 hours, the cell concentration was adjusted to 1 x 10^6 cells/ml in R10 containing 1 μg/ml of αCD28 (BD Biosciences, Oxford, UK) 1 μg/ml of αCD49d (BD Biosciences, Oxford, UK), 10 μg/ml of Brefeldin A (Sigma, Poole, UK). PBMC were stimulated with one of the following: 10 μg/ml PPD (SSI, Copenhagen, Denmark), 2 μg/ml antigen 85A complete peptide pool (66 x 15 mer peptides overlapping by 10 amino acids, 2 μg/ml final concentration for each amino acid), BCG at 1 x 10^6 CFU or 5 μg/ml of SEB (Sigma, Poole, UK). Un-stimulated PBMCs were used to assess non-specific cytokine production. Cells were incubated at 37°C with 5% CO_2 in air for 10 hours. Following stimulation, PBMCs were washed in PBS containing 0.1% BSA and 0.01% sodium azide (both Sigma, Poole, UK) by centrifuging at 400g for 5 minutes. Following a second wash step, cells were stained with the amine reactive LIVE/DEAD fixable red dead cell stain kit (Molecular Probes, Invitrogen, Paisley, UK) and labelled with the following antibodies at optimally diluted concentrations: CD4+ PE, CD8+ PE-Cy7 (BD Biosciences, Oxford, UK), CD14+ECD and CD20+ECD (Beckman Coulter, High Wycombe, UK).

After labelling for 30 minutes with light excluded, cells were washed by centrifuging before permeabilisation by incubation at room temperature for 15 minutes with Fix/Perm (BD Biosciences, Oxford, UK). PBMC were washed by centrifuging at 400g for 5 min with permwash (BD Biosciences, Oxford, UK) then labelled with CD3 APC-C7 (to account for down regulation after stimulation), and also for the following cytokines: IFN-γ FITC (both BD Biosciences, Oxford, UK), TNFα pacific blue (eBiosciences, Hatfield, UK) and IL-2 APC (Miltenyi Biotech, Bisley, UK).

Flow cytometric acquisition of polyfunctional intracellular cytokine staining samples.

Cells were analysed using a 3 laser standard LSRII (BD Biosciences, Oxford, UK). Cytokine-secreting T cells were identified using a forward scatter-height (FSC-H) versus side scatter-area (SSC-A) dot plot to identify the lymphocyte population, to which appropriate gating strategies were applied to exclude doublet events, non-viable cells, monocytes ($CD14^+$) and B cells ($CD20^+$) prior to sequential gating through $CD3^+$, $CD8^-$ and $CD4^+$ versus IFNγ, and $CD3^+$, $CD8^+$ and $CD4^-$ versus IFNγ histograms. The Boolean gate platform was used with individual function gates to create all response pattern combinations. Data for subsequent analysis were prepared using Pestle version 1.6.1 Mario Roederer, Vaccine Research Centre, NIAID, NIH), and SPICE version 4 (Mario Roederer, Vaccine Research Centre, NIAID, NIH) was used to calculate threshold and present experimental results.

Lymphoproliferation assay.

A conventional tritiated thymidine proliferation assay was performed. Briefly, PBMC were cultured in triplicate at 1 x 10^5 cells/well in 96-well plates with medium alone or in the presence of PPD at a concentration range of 1, 5 and 10 μg/ml for 7 days or Concanavalin A (Sigma, 10 μg/ml) for 3 days. [^{3}H]thymidine (GE Healthcare, Little Chalfont, UK) was added at 1 μCi/well for 6 hours, cells were harvested and thymidine uptake measured using a β-counter (Tri-carb 2900TR, Liquid Scintillation Analyzer, Packard Biosciences Company, Perkin Elmer, Milano, Italy). Proliferative capacity was expressed as stimulation index (SI) (mean [^{3}H] thymidine incorporation of cells stimulated with antigen/mean incorporation in absence of stimulation i.e. medium alone).

Necropsy

Before necropsy, animals were sedated with ketamine (15 mg/ml, i.m.), weighed, photographed, chest X-rays taken, clinical data collected and ex-sanguination was effected via the heart, before termination by injection of a lethal dose of anaesthetic (Dolelethal, Vétoquinol UK Ltd,140mg/kg). A full necropsy was performed and gross pathology was scored using a relative scoring system based on the number and extent of lesions present in lungs, spleen, liver, kidney and lymph nodes as previously described (29). Samples of spleen, liver, kidneys and hilar, inguinal and axillary lymph nodes were removed, dissected on sterile trays and placed in weighed tubes for quantitative bacteriology and into formalin buffered saline for histology. The whole lungs were carefully dissected to retain their integrity. Following examination, the left bronchus of lung was clamped off using artery forceps, and the left upper and left lower lung lobes dissected away from lung. The

upper and lower left lobes were collected for quantitative bacteriology. The left bronchus was tied off with string to the left of the clamp and right hand side of lung gently infused with 10% neutral buffered formalin using a 10 ml syringe attached to a 14 CH Netalon catheter (J.A.K. Marketing, York, UK) whilst observing closely for inflation. The trachea was tied off and the lungs immersed in formalin to complete fixation.

Lung imaging

Thoracic radiographs (SP VET 3.2, Xograph Ltd) were acquired using mammography film (Xograph Imaging Systems Ltd, Tetbury, UK) before and every 2 weeks after exposure to *M. tuberculosis*. Evaluation of disease was performed by an experienced consultant thoracic radiologist blinded to the animal group and clinical status, using a pre-determined scoring system based on the amount and distribution of infiltrate (29).

Magnetic resonance imaging (MRI)

The inflated lungs that had been immersed in formalin were set in agarose for MRI scanning. Images were acquired on a 1.5 T Twinspeed HDX MRI Scanner (General Electric Healthcare, Milwaukee, WI, USA). The sample was placed in a general purpose flexible coil. T1 weighted 3D Fast Spoiled Gradient Echo (FSPGR) images were acquired with the following parameters: TE: 7.7, TE: 4.2, 16 cm field of view, 1 mm slice thickness, -0.5 mm slice gap, 192x192 matrix, 15 degree Flip Angle, 3NEX. Each acquisition took approximately 3.5 minutes.

Lesion analysis / quantification (Stereology)

Lung lesions were identified on MR images based on their signal intensity and nodular morphology relative to more normal lung parenchyma. The total lung and lesion volume relative to the fixed tissue was determined using the Cavalieri method applied to MRI image stacks, and then expressed as a ratio to provide a measure of disease burden in each animal as previously described (29). Analyses of lesion volume on magnetic resonance (MR) images were performed with the investigators reading the images blind to treatment groups.

Pathology / Histology

On return from scanning, lung lobes were removed from agar and stored in pots of 10% neutral buffered formalin. Lung lobes that had been fixed in buffered formalin were sliced at approximately 5-10 mm intervals and discrete lesions measuring 1-5 mm in diameter were counted. Lesions that had coalesced and measured more than 5 mm in diameter were counted and the dimensions of the sectioned coalesced lesions were measured and recorded. Total lesion count was the sum of the discrete and coalesced lesions. Lesions were classified according to the scheme used by Lin *et al* (16).

Bacteriology

The left lung lobes were sampled for the presence of viable M. tuberculosis post-mortem. The weights of all samples were determined. Lung lobes (1.97 to 16.25) were homogenised in 10 ml and sections of spleen (0.072 to 2.54g), liver (0.38 to 2.61g), kidney (0.011 to 2.39g) and hilar lymph nodes (0.317g) were sampled and homogenised in 2 ml of sterile water. All samples were

406 either serially diluted in sterile water prior to being plated or plated directly
407 onto Middlebrook 7H11 OADC selective agar. All plates were incubated for 3
408 weeks at 37ºC and resultant colonies were counted. The result from each
409 sample was expressed
410
411 Statistical analyses
412 To compare the immune response profiles induced in animals by vaccination
413 or challenge, the area under the curve (AUC) for each response was
414 calculated using Sigmaplot version 10 (Systat Software Inc, Hounslow, UK)
415 for each animal. The area under the curves calculated for the animals in each
416 test group were compared to those for the animals in other test groups with a
417 Mann Whitney test using Minitab, version 15 (Minitab Ltd, Coventry, UK).
418 Differences in the survival rates of animals in each test group were compared
419 with a log rank test using Minitab version 15. A Cox regression analysis was
420 performed in order to estimate the hazard ratios and 95% confidence intervals
421 using STATA Statistical Software 2007 Release 9.2. (StataCorp LP, Texas,
422 USA).
423 The Spearman correlation test was used to determine the level of correlation
424 between study parameters using GraphPad Prism, version 5.01 (GraphPad
425 Software Inc, La Jolla, California, USA).

RESULTS

IFN-γ response to vaccination

The mycobacterium-specific IFN-γ response was measured by ELISpot (Figure 2 A and B), whole blood ELISA (Figure 2C) and intracellular cytokine staining (Figure 2D). All three assays (Figure 2A, 2C, 2D) revealed similar response profiles with low levels of PPD-specific IFN-γ responses after vaccination with BCG. Statistical analysis using AUC and Mann Whitney tests confirmed significant differences in PPD-specific responses between BCG vaccinated animals in groups A and B and unvaccinated animals in Group C (ELISpot: Group A versus Group C: $p = 0.01$ Group B versus Group C: $p = 0.01$; ELISA: Group A versus Group C: $p = 0.025$, Group B versus Group C: $p = 0.013$).

Although an increase in the PPD-specific response was seen in two of the six animals (K65, K86) which received the MVA85A booster vaccine, one week after immunisation, comparison of the AUC of the PPD response between week 12 (time of MVA85A vaccination) and week 20 (just prior to challenge) using Mann Whitney tests, did not reveal significant differences in the immune profiles induced in the whole group of animals which received BCG alone (Group B), or those which received BCG followed by MVA85A (Group A) (Figure 2A and C). However, the level of response in both Group A and Group B were significantly greater than in the unvaccinated control group (ELISpot: Group A versus Group C: $p = 0.04$, Group B versus Group C: $p = 0.04$; ELISA: Group A versus Group C: $p = 0.011$, Group B versus Group c: $p = 0.014$), by performing Mann Whitney tests on AUC analysis of the same time period.

There were no significant differences in the response profiles made by vaccinated and unvaccinated animals to Ag85A after vaccination with BCG (Figure 2A). After vaccination with MVA85A, the response profile in Group A which received the booster vaccination was found to be significantly different from profiles in both the unvaccinated group, Group C ($p = 0.04$) and the group vaccinated with BCG alone, Group B ($p = 0.02$), using AUC analysis and Mann Whitney tests from week 12 to 20.

Polyfunctional T cell responses following vaccination.

The cytokine responses made by antigen-specific T cell populations isolated from animals in the period following vaccination were further investigated using multi-parameter flow cytometry. The $CD4^+$ and $CD8^+$ T cell populations were evaluated for their capability to produce IFN-γ, IL-2 and TNFα, either individually (single positive cells) or in combination (double and triple positive cells) prior to vaccination with BCG, on the day of MVA85A vaccination, then one, two and eight weeks after MVA85A vaccination.

Ag85A-specific single, double and triple positive CD4+ T cells were only detected in animals that received vaccination with MVA85A (Figure 3C). There was considerable heterogeneity between animals within each group. Raised levels of single, double and triple positive IFN-γ secreting cells were seen in K65 and K86 one week after vaccination with MVA85A. Ag85A-specific IFN-γ secreting $CD4^+$ T cells were not detected in animals K20, K61 or K79 after vaccination with MVA85A (Figure 3A). $IL\text{-}2^+IFN\gamma^+$ and $IL\text{-}2^+TNF\alpha^+$ double positive Ag85A-specific CD4+T cell responses were not seen at levels above those measured prior to vaccination in any of the animals in the study (data not shown). Through all experimental groups, production of IL-2 was either low relative to IFN-γ or TNFα, or not detected.

476 A number of PPD-specific cytokine producing $CD4^+$ T cell populations were
477 detected in vaccinated animals at levels above those seen prior to vaccination
478 in the period between MVA85A vaccination and challenge. However, the
479 functional profiles did not differentiate between animals vaccinated with BCG
480 alone, or the BCG - MVA85A regimen. Single positive PPD-specific, TNFα-
481 secreting, PPD-specific $CD4^+$ T cells were seen at levels above those seen
482 before immunisation, during the period after vaccination with MVA85A (study
483 weeks 13 to 20) in three animals that received the BCG-MVA85A vaccine
484 regimen (K79, K61, K86) and two animals that received the BCG vaccine
485 alone (K52, K43) but not in any of the unvaccinated animals (Figure 3E).
486 Similarly, an increase in PPD-specific IFN-γ producing $CD4^+$ T cells was
487 detected in K69, K86, and K79 from the BCG - MVA85A vaccinated group and
488 K43 of the BCG-only vaccinated group (Figure 3D). An increase in levels of
489 double positive ($IL\text{-}2^+TNF\alpha^+$) PPD-specific $CD4^+$ T cells was detected in two
490 animals (K79, K86) one weeks after MVA85A vaccination and in one animal
491 (K43) from the group vaccinated with BCG alone during the same period
492 (Figure 3G) at week 20. Increases in double positive $IFN\text{-}\gamma^+TNF\alpha^+$ (Figure 3E)
493 and triple positive PPD-specific $CD4^+$ T cells (Figure 3H) were seen one week
494 after MVA vaccination in three animals (K65, K79, K86). Increased levels of
495 single positive $IL\text{-}2^+$, and double positive $IFN\text{-}\gamma^+IL\text{-}2^+$ secreting PPD-specific
496 $CD4^+$ T cell responses did not increase in either vaccinated or unvaccinated
497 animals (data not shown). Neither double, nor triple positive PPD, nor Ag85A-
498 specific $CD8^+$ T cell were detected in animals in this study although some IFN-
499 γ^+ and $IL\text{-}2^+$ single positive responses were seen (IFN-γ: BCG-MVA group,

K86, K79, K61; BCG group K43; IL-2: BCG-MVA group, K20; BCG group K43, K44, K59, Unvaccinated group K62 (Figure 3I and J).

Proliferative responses following vaccination

T cell recognition and response to PPD stimulation was investigated using a standard tritiated thymidine incorporation assay. In the group that received the BCG-MVA85A vaccination regimen, all animals made low PPD-specific proliferative responses (SI< 20) in the period after BCG vaccination and before vaccination with MVA85A (weeks 0 to 12). Response levels increased after vaccination with MVA85A (SI range = 40 – 60) in two of the six animals (K65 and K69). All six animals in the group that only received the BCG vaccine made PPD-specific proliferative responses during the period before challenge. The largest responses were seen in K44 (SI range = 20 - 80); in contrast responses in the other five animals were much lower (SI <10), (data not shown). Statistical analysis using AUC and Mann Whitney tests did not demonstrate a significant difference in PPD-specific responses made by animals vaccinated with BCG and MVA85A (Group A) and those vaccinated with BCG alone (Group B), either after vaccination with BCG (weeks 0-12), or during the period after MVA85A vaccination and challenge (weeks 12 to 20).

Clinical status following vaccination

All animals in the study showed the weight gain profiles expected in normal healthy animals during the period prior to challenge. Temperature, erythrocyte sedimentation rate (ESR) and haemoglobin levels all remained within the normal range during the vaccination phase of the study. Local skin reactions at the site of BCG immunisation were seen in all vaccinated animals.

Reactions appeared 2 to 4 weeks after BCG vaccination in 10 of 12 immunised animals, but resolved in all animals 22 weeks after immunisation. A local skin reaction at the site of immunisation was seen in all six animals vaccinated with MVA85A. Reactions peaked one week after vaccination and had resolved eight weeks after immunisation in all animals. Vaccination with MVA85A also induced a increase in size of axillary lymph nodes draining the site of vaccination in all animals following vaccination with MVA85A. Maximum lymph node size was seen one to two weeks after immunisation and returned to normal in all animals six to eight weeks after vaccination.

Changes in clinical parameters following aerosol challenge with *M. tuberculosis*

Animals in all three study groups showed changes in behaviour and clinical parameters consistent with progression of tuberculosis-induced disease and were euthanized when set humane endpoints were reached ahead of the planned end of the study. Disease progressed to reach humane end-points in all four animals in the naïve control group C (K32, K47, K54, K62) , two of six BCG-vaccinated animals in group B (K44, K52) and five of six BCG-MVA85A vaccinated animals in group A (K20, K61, K69, K79, K86) (Figure 4).

All animals that met humane endpoints showed a range of clinical signs in the five days before euthanasia. These included depression, and withdrawal from the group. Abnormal rapid respiration rates and shallow, laboured breathing were also observed in this group, and coughing by all animals except K32. Ruffled fur was seen in four animals (K20, K32, K54, K61). In the period prior to termination, all the animals which met humane endpoint criteria showed progressive weight loss and elevated ESR levels (Figure 5F and G). At

termination, all the unvaccinated animals (K32, K47, K54, K62), one of six BCG-vaccinated animal (K52) and two of six BCG-MVA85A animals (K61, K69) showed erythrocyte haemoglobin concentrations below the normal range. No adverse clinical, or behavioural indicators were exhibited by BCG-vaccinated animals K43, K50, K59, or BCG-MVA85A-vaccinated animal K65 at the time of termination 52 weeks after challenge.

Pulmonary abnormalities were identified on X-radiographs from all animals after challenge and remained until the end of the study period. Abnormalities were not detected until 4 weeks after challenge in any of the animals, at which point, infiltrates were detected on both sides of the lung (score 3) in all except two, K86 and K59, where detection of bilateral infiltration was delayed until week 12 and 28, respectively. Infiltrates were extensive (score 4) by the time of euthanasia in three of the four control animals (K32, K47, K62,), two BCG-vaccinated animals (K52, K80) and four of the BCG-MVA85A vaccinated animals (K20, K61, K69, K79) (Figure 5E). Calcified nodules were observed on radiographs taken from week 24 after challenge onwards in all the animals which remained clinically healthy during the study. Calcified nodules were first identified at week 24 in K65, week 30 in K43 and K80, week 34 in K50, and week 38 in K59.

Survival post-challenge

The length of time that animals in each group survived before disease progressed to reach clinical end-points was compared using two statistical approaches, the Log Rank Test and Cox regression analysis. Both

approaches showed no overall difference in survival time post-challenge between the 3 groups (Log rank test, p = 0.13, Cox regression, p = 0.093). Survival times were then compared between each of the vaccination groups A and B and the unvaccinated control group C. Survival of BCG-vaccinated animals was found to be significantly different to that of the unvaccinated control animals using the Kaplan Meier Log Rank test (p = 0.048), but not significantly different using Cox regression (p = 0.07). No difference in the survival time was seen between the group vaccinated with BCG alone and the group vaccinated with BCG followed by MVA85A by either test (Log rank test, p = 0.1, Cox regression, p = 0.17).

Measures of tuberculosis-induced pulmonary and extra pulmonary disease burden

At the end of the study the level of tuberculosis-induced pulmonary and clinical disease was investigated using quantitative and qualitative approaches (Figure 5). Pulmonary disease burden was quantified using stereology on MR images to determine the total lesion volume and the total lung volume, then expressed as a ratio of lesion to lung volume. A significant difference in lesion volume to lung volume ratio was seen between the unvaccinated control animals and animals vaccinated with BCG alone, or BCG boosted with MVA85A (Figure 5A; Table 2), with both groups of vaccinated animals having a lower ratio than the controls. The lesion volume to lung volume ratio was not significantly different between BCG and BCG-MVA85A groups. Lesion volume to lung volume ratio was lower in animals which did not show adverse clinical indicators at termination (K65, K43, K50, K59).

Pulmonary disease burden was also quantified by serial sectioning and manual counting of both discrete and coalesced lesions (Figure 5B). Lesion numbers were not significantly different between the unvaccinated and the BCG-MVA85A vaccinated animals, or the BCG vaccinated and the BCG-MVA85A vaccinated groups. However, a significantly higher number of lesions were seen in the BCG vaccinated animals compared to the unvaccinated animals (Table 2).

The distribution and type of pulmonary granulomas (classified according to the scheme of Lin *et al.*, 2006 (16) identified in sections of lung did not distinguish between the different groups, nor did this predict survival; neither did disease spread to extra-pulmonary tissues measured by lesion presence in hilar lymph node, liver, kidney and spleen. Histological examination revealed granulomas in extra-pulmonary tissues in all four of the unvaccinated animals, three of the six BCG vaccinated animals (K43, K50, K59) all of which remained clinically well until the end of study, and four of the six BCG-MVA85A-vaccinated animals (K86, K61, K69, K20) all of which succumbed to disease during the study.

Pulmonary disease burden was also measured using qualitative approaches conventionally used for the evaluation of TB vaccine efficacy in NHPs including pulmonary pathology score (Figure 5C), total gross pathology score (Figure 5D), and chest X-ray (Figure 5E). The level of pulmonary disease measured by each of these approaches generally agreed and reflected the trends identified using the quantitative measures. The gross pathology scoring system was able to distinguish between groups of BCG vaccinated and unvaccinated animals ($p = 0.017$) (Table 2), but there was no difference

between BCG-MVA85A vaccinated and BCG vaccinated and unvaccinated controls using this measure. The pathology scoring system did not predict survival in the BCG vaccinated group. No significant differences between study groups were seen in terms of chest X-ray score, lung pathology score, bacterial load, or clinical disease measures, weight loss, or ESR (Table 2). The level of pulmonary disease by all scoring systems also agreed with the level of changes determined in the clinical measures of disease severity at termination, namely, weight loss, decreased erythrocyte haemoglobin concentration and increase in ESR, with most changes occurring in the animals with the greatest level of pulmonary disease. Across all animals in the study regardless of treatment grouping, lesion volume to lung volume ratio, ESR and mean cell haemoglobin concentration for animals correlated with survival time after challenge (lesion:lung volume ratio : $p = 0.0267$; ESR: $p = 0.0038$, haemoglobin concentration: $p = 0.0120$).

Lung bacterial load tended to be lower in animals which did not show adverse clinical indicators at termination (K65, K43, K50, K59) (Figure 5H). Extra-pulmonary dissemination occurred in all animals in the study and the pattern of spread to spleen, kidneys and / or liver did not distinguish between test groups.

IFN-γ response following aerosol challenge with *M. tuberculosis*

The *Mycobacterium*-specific IFN-γ response was measured by ELISpot (Figure 6A, B, C) and ELISA (Figure 6D). Comparison of the responses made by animals in each test group during the first six weeks after challenge suggested a trend for an increased frequency of Ag85A-specific and PPD-specific IFN-γ producing cells in BCG-MVA85A vaccinated animals (Group A)

than in the BCG vaccinated (Group B) and the unvaccinated animals (Group C) (Figure 6, B, C). The difference in the PPD-specific response was found to be statistically significant ($p = 0.03$) between the BCG-MVA85A vaccinated and the unvaccinated animals, but not statistically different between the BCG and the BCG-MVA85A animals. The frequency of ESAT6-specific IFN-γ producing cells and the quantity of IFN-γ secreted following stimulation with PPD was similar in all the test groups during this initial six week period. Comparison of responses made by animals in the different test groups was only valid in the early phase after infection while all animals remained in the study.

Proliferative response following aerosol challenge with *M. tuberculosis*

PPD-specific proliferative responses were detected in five of the six BCG vaccinated animals (K43, K44, K50, K59, K80), five of the six BCG-MVA85A vaccinated animals (K61, K65, K69, K79, K86) and one of unvaccinated animals (K32) following infection with *M. tuberculosis.* Responses were generally low (SI<10), inconsistent over time, and did not differentiate between vaccine groups, or predict progression with disease.

DISCUSSION

There is a growing need for improved preclinical models for vaccine assessment with sensitive and discriminatory readouts that would act as correlates of protection as work continues to develop a new more effective TB vaccine. Ideally, models should be highly reproducible, include the same endpoints as those used in human clinical trials, and be validated to predict efficacy in humans. This study developed and characterised an aerosol challenge model of *M. tuberculosis* in rhesus macaques and assessed immunological and clinical read-outs for utility in vaccine evaluation studies. Aerosol challenge models have been used previously to evaluate the efficacy of TB vaccines in NHP (2, 9, 14, 27, 30) but no standardised or validated protocol has emerged. However, it is clear that the disease which develops following low-dose aerosol challenge shows marked animal to animal variation, particularly when the read-out of disease severity is survival, and much larger group sizes will be required if survival is used as an end-point in future vaccine evaluation studies. Even more pronounced heterogeneity in progression to disease was observed when very low doses were administered by the intra-tracheal route suggesting that the lower the initial inoculum, the less predictable is the outcome of the host-pathogen interaction from one animal to the next (8).

Our study is the first example of TB vaccine evaluation in rhesus macaques where long term survival was one of the primary endpoints. The study was taken to a point where disease progressed in all unvaccinated control animals to allow outcome in vaccinated animals to be compared using a variety of pathological and clinical readouts. Previous studies in which TB vaccine candidates have been evaluated for efficacy in the rhesus macaque have

been ended at a fixed point, 8-17 weeks after either aerosol (1, 2, 3, 14, 27) or intratracheal (15, 32) challenge. Longer term studies of vaccine efficacy with post-challenge investigation periods in excess of 16 weeks have been conducted in the cynomolgus macaque (23, 24, 26). In these latter studies, survival after challenge has provided a read-out of vaccine efficacy, although only one of the four efficacy studies reported (22) was sufficiently long to allow the entire unvaccinated group to progress to end-stage disease i.e. not survive the challenge. Our study has shown that the heterogeneity in long-term disease progression following relatively low-dose aerosol challenge is not compatible with survival as an end-point of vaccine efficacy in rhesus macaques.

In NHP vaccine efficacy studies with group sizes of about six, as used here, survival appears not to be a reliable endpoint. In this study, the limitation by cost of group size meant that a survival difference between the BCG vaccinated group and the naïve control group was of borderline statistical significance and dependent on the statistical analysis used. Comparison of the survival time after challenge of BCG-MVA85A vaccinated animals with that of BCG vaccinated animals and unvaccinated animals using the Cox regression, or the Kaplan Meier Log Rank statistical tests, did not show a statistically significant difference between survival in any of the three experimental groups. The conclusion to be drawn from the current study is that survival does not discriminate between different groups when small numbers of animals are used and this is supported by a retrospective analysis of numerous guinea pig studies, even though larger group sizes can be more easily achieved for guinea-pigs (36). Furthermore, survival is a complex

723 endpoint with many contributing factors implicated. In human efficacy trials,
724 death would not be an endpoint; rather, disease defined by positive culture or
725 presentation of clinical symptoms, signs or abnormal investigations such as a
726 chest radiograph would be used. Whilst animal models provide excellent test
727 systems for the evaluation of new vaccines, there are constraints which
728 prevent these models simulating every aspect of the natural disease in
729 humans. For example, only a small percentage of humans who are exposed
730 to tuberculosis will get disease. In animal models, the uniform development of
731 disease within a reasonable time frame after challenge in unvaccinated
732 control groups is essential as the process of vaccine evaluation relies on the
733 definition of differences in disease burden between vaccinated and
734 unvaccinated animals. These requirements necessitate the use of challenge
735 doses sufficient to ensure uniform disease induction as well as the use of
736 complex readouts.
737 The evaluation of vaccine efficacy was also shown to be complex and not
738 clear cut. Other clinical-, pathological- and bacteriology-based measures were
739 confounded by groups having a mixture of survivor and non-survivor animals.
740 The impact of this is that the animals that reach their humane endpoint will
741 have CFU counts and pathology scores at the maximum for the readouts due
742 to the late stage accelerated bacterial growth and inflammation whereas
743 survivors will have lower levels. Some read-outs (total pathology score, chest
744 radiograph and bacterial load) showed a reduction in surviving animals, but
745 lacked sufficient power to differentiate between study groups. A significantly
746 higher number of lesions were counted in the BCG vaccinated animals
747 compared to the unvaccinated. This finding likely reflects difficulties

associated with manual counting of both discrete and coalesced lesions as a readout of disease burden, particularly in animals with more severe disease, where the individual lesions became difficult to distinguish, and eventually too numerous to count. The disease in the unvaccinated animals was more severe than in BCG vaccinated animals and a greater proportion of coalescing lesions and consolidated lobes led to an apparently low total lesion count. In contrast, the lesions in the BCG vaccinated animals tended to be smaller and remained discrete, and so were more easily countable.

The level of PPD-specific proliferative responses induced by vaccination did not correlate with protection. For example, K44 from the BCG-vaccinated group made the largest and most consistent proliferative responses after vaccination and yet was the first animal in this group to reach the humane endpoint criteria following challenge. Similarly, PPD specific proliferative responses of a similar magnitude were detected in two animals (K65 and K69) in the BCG -MVA85A vaccinated group two weeks after vaccination with MVA85A, but the outcome of *M. tuberculosis* infection was very different. The disease in K69 progressed rapidly and met humane end-point criteria within ten weeks of challenge, whereas K65 successfully controlled disease for the duration of the study.

IFN-γ is considered to be an essential cytokine for the control of TB. Three assays were used to assess the responses made to vaccination and challenge. These were the ELISpot assay used in the clinical trials with MVA85A (20) to measure the frequency of IFN-γ secreting cells, an ELISA to quantify of IFN-γ secreted in response to the antigenic stimulation of whole blood (5), and intracellular cytokine staining followed by flow cytometry. The

response profiles defined by the three approaches in each animal showed similar trends but none of the IFN-γ profiles following vaccination correlated with protection. Interestingly, the animal (K65) that showed the highest IFN-γ response as measured both by ELISpot and ELISA, and a high level of Ag-85A specific polyfunctional $CD4^+$T cells, one week after vaccination with MVA85A, controlled infection most effectively after challenge, showed fewer clinical symptoms during the post-challenge phase of the study, and had the lowest disease burden when euthanized at the end of study. During the period after MVA85A boost and prior to challenge, the PPD-specific responses in the group vaccinated with BCG-MVA85A were not significantly different to those seen in the group vaccinated with BCG alone. This is in contrast with BCG vaccinated mice that were boosted with MVA85A) and this vaccine regimen was shown to protect mice against aerosol challenge with *M. tuberculosis* (13). In our study the lack of a post-MVA85A boost response in the other five animals in the BCG-MVA85A group perhaps explains why survival in this group was not improved as reported by Verreck *et al.*, (32). However, ELISpot responses induced immediately after the MVA85A boost in the animals in the Verreck study were not measured, and so this cannot be confirmed. Further work is required to understand why the immune response to MVA85A in five of the animals in this study are considerably lower than those seen in humans, despite the NHPs in this experiment having received a higher dose of vaccine (20). Whilst a trend for a larger IFN-γ early response in the BCG-MVA85A group was suggested, the validity of continued comparison of responses later than eight weeks after challenge was confounded by the loss of some animals from the study. The only readout to reliably distinguish vaccinated and

798 unvaccinated NHPs was MR stereology at the end of the study. The even
799 distribution of pulmonary disease following aerosol challenge reduced
800 sampling error associated with readouts reliant on sampling less than the
801 entire lung, and allowed the successful application of the MR stereology
802 technique without compromising the use of other histopathology-based read-
803 outs.
804 Any long term study that relies on end-of-study readouts may suffer when
805 combining data from survivors and non-survivors. Therefore, longitudinal 'in
806 life' measurements that can include all study subjects are required. 'In life'
807 longitudinal measures of the immune response were used in this study;
808 however, the currently proposed key markers were not shown to correlate with
809 protection. We conclude that, given the lack of statistical power, it is generally
810 not warranted to conduct long-term survival studies on NHPs for ethical and
811 financial reasons, and shorter experiments with defined endpoints are
812 recommended. Radiological imaging offered a reliable correlate even at the
813 end of study and work is in progress to investigate the potential of 'in life'
814 imaging as a readout of vaccine efficacy.

815 **ACKNOWLEDGEMENTS**

816 This work was supported by Department of Health, UK. The views expressed
817 in this publication are those of the authors and not necessarily those of the
818 Department of Health. We thank the staff of the Biological Investigations
819 Group at HPA for assistance in conducting studies; Graham Hatch for
820 aerobiology assistance, Nicola Alder for statistical support, Arturo Reyes-
821 Sandoval for the staining panel used for polyfunctional T cell analysis and
822 Susan Gray for histology support.

823

824 **REFERENCES**

825 **1. Anacker R.L., W. Brehmer, W.R. Barclay, W.R. Leif, E. Ribi, J.H.**
826 **Simmons, A.W. Smith.** 1972. Superiority of intravenously administered BCG
827 and BCG cell walls in protecting rhesus monkeys (*Macaca mulatta*) against
828 airborne tuberculosis. Z Immunitatsforsch Exp Klin Immunol. May;**143**:363-
829 76.

830 **2. Barclay W.R., R.L., Anacker, W. Brehmer, W. Leif, E. Ribi**. 1970.
831 Aerosol-induced tuberculosis in sub-human primates and the course of the
832 disease after intravenous BCG vaccination. Infect Immun. **2** 574-582.

833 **3. Barclay W.R., W.M. Busey, D.W. Dalgard, R.C. Good, B.W. Janicki, J.E.**
834 **Kasik, E. Ribi, C.E. Ulrich, E. Wolinsky**. 1973 Protection of monkeys against
835 airborne tuberculosis by aerosol vaccination with bacillus Calmette-Guerin.
836 Am Rev Respir Dis. **107**:351-8.

837 **4. Basaraba R.J**. 2008. Experimental tuberculosis: the role of comparative
838 pathology in the discovery of improved tuberculosis treatment strategies.
839 Tuberculosis (Edinb). **88** Suppl 1:S35-47.

840 **5. Black GF, P.E.M. Fine, D.K. Warndorff, S. Floyd, R.E. Weir, J.M.**
841 **Blackwell, L. Bliss, L. Sichali, L. Mwaungulu, S. Chaguluka, E. Jarman, B.**
842 **Ngwira, H.M. Dockrell**. 2001. Relationship between IFN-gamma and skin
843 test responsiveness to *Mycobacterium tuberculosis* PPD in healthy, non-BCG-
844 vaccinated young adults in Northern Malawi. Int J Tuberc Lung Dis **5**:664-
845 672.

846 **6. Beveridge N.E., D.A. Price, J.P. Casazza, A.A. Pathan, C.R. Sander,**
847 **T.E. Asher, D.R. Ambrozak, M.L. Precopio, P. Scheinberg, N.C. Alder, M.**

Roederer, R.A. Koup, D.C. Douek, A.V. Hill, H. McShane. 2007. Immunisation with BCG and recombinant MVA85A induces long-lasting, polyfunctional *Mycobacterium tuberculosis*-specific CD4+ memory T lymphocyte populations. Eur J Immunol **37**: 3089-100

7. Blanchard TJ, A. Alcami, P. Andrea, G.L. Smith. 1998. Modified vaccinia virus Ankara undergoes limited replication in human cells and lacks several immunomodulatory proteins: implications for use as a human vaccine. J Gen Virol. **79**: 1159-67

8. Capuano S.V., D.A. Croix, S. Pawar, A. Zinovik, A. Myers, P.L. Lin, S. Bissel, C. Fuhrman, E. Klein, J.A. Flynn. 2003. Experimental *Mycobacterium tuberculosis* infection of cynomolgus macaques closely resembles the various manifestations of human *M. tuberculosis* infection. Infect Immun. **71**: 5831-5844

9. Chaparas S.D., R.C. Good, B.W. Janicki. 1975. Tuberculin-induced lymphocyte transformation and skin reactivity monkeys vaccinated or not vaccinated with bacille Calmette-Guerin, then challenged with virulent *Mycobacterium tuberculosis*. Am. Rev. Respir. Dis. **112**: 43-47.

10. Colditz G.A., T.F. Brewer, C.S. Berkey, M.E. Wilson, E. Burdick, H.V. Fineberg , F. Mosteller. 1994. Efficacy of BCG vaccine in the prevention of tuberculosis. Meta-analysis of the published literature. Jama. **271**: 698-702.

11. Druett H.A. 1969 A mobile form of the Henderson apparatus. *J Hyg(Lond)*. **67**:437-48.

12. Good R.C. 1968. Simian tuberculosis: Immunologic aspects. Annals NY Acad Sci. **154**:200-213

872 **13. Goonetilleke N.P., McShane H., Hannan C.M., Anderson R.J.,**
873 **Brookes R.H., Hill A.V.** 2003 Enhanced immunogenicity and protective
874 efficacy against Mycobacterium tuberculosis of bacille Calmette-Guérin
875 vaccine using mucosal administration and boosting with a recombinant
876 modified vaccinia virus Ankara. J Immunol. **171**(3):1602-9.
877
878 **14. Janicki B.W., R.C. Good, P. Minden, L.F. Affronti, W.F. Hymes.** 1973.
879 Immune responses in rhesus monkeys after bacille Calmette-Guerin
880 vaccination and aerosol challenge with *Mycobacterium tuberculosis*. *Am. Rev.*
881 Respir Dis. **107**:359-66
882 **15. Langermans J.A.M., P. Anderson, D. van Soolingen, R.A.W. Verenne,**
883 **P.A. Frost, T. van der Laan, S. Kroon, I. Peekel, S. Florquin, A.W.**
884 **Thomas**. 2000. Divergent effect of bacillus Calmette-Guerin (BCG)
885 vaccination on *Mycobacterium tuberculosis* infection in highly related
886 macaque species: implications for primate models in tuberculosis vaccine
887 research. PNAS. **98**: 11497-11502.
888 **16. Lin P.L., S. Pawar, A. Myers, A. Pergu, C. Fuhrman, T.A. Reinhart,**
889 **S.V. Capuano, E. Klein and J.L. Flynn**. 2006. Early events in
890 *Mycobacterium tuberculosis* infection in cynomolgus macaques. Infect Immun.
891 **74**: 3790-3803.
892 **17. McMurray D.N., F.M. Collins, A.M. Dannenberg, Jr. and D.W. Smith**.
893 1996. Pathogenesis of experimental tuberculosis in animal models. *Curr Top*
894 Microbiol Immunol. **215:**157-79
895 **18. McMurray D.N**. 2000. A non-human primate model for the preclinical
896 testing of new tubeculosis vaccines. *Clin Infect Dis.* **30**:(Suppl 3): S210-212

897 **19. McShane H., R. Brookes, S.C. Gilbert, A.V. Hill.** 2001. Enhanced
898 immunogenicity of CD4(+) t-cell responses and protective efficacy of a DNA-
899 modified vaccinia virus Ankara prime-boost vaccination regimen for murine
900 tuberculosis. Infect Immun. **69**: 681-6
901 **20. McShane H., A.A. Pathan, C.R. Sander, S.M. Keating, S.C. Gilbert, K.**
902 **Huygen, H.A. Fletcher, A.V. Hill**. 2004. Recombinant modified vaccinia virus
903 Ankara expressing antigen 85A boosts BCG-primed and naturally acquired
904 antimycobacterial immunity in humans. Nat Med. **10**: 1240-4
905 **21. A Study of MVA85A in Healthy Infants.** 2009
906 http://clinicaltrials.gov/ct2/show/NCT00953927 August 18, 2009.
907 **22. Pathan A.A., C.R. Sander, H.A. Fletcher, I. Poulton, N.C. Alder, A.V.**
908 **Hill, H. McShane.** 2007. Safety and immunogenicity of boosting BCG
909 vaccinated subjects with BCG: comparison with boosting with a new TB
910 vaccine, MVA85A. PLoS ONE . **2**: e1052
911 **23. Okada M., Y. Kita, T. Nakajima, N. Kanamaru, S. Hashimoto, T.**
912 **Nagasawa, Y. Kaneda, S. Yoshida, Y. Nishida, H. Nakatani, K. Takao, C.**
913 **Kishigami, Y. Inoue, M. Matsumoto, D.N. McMurray, E.C. Dela Cruz, E.V.**
914 **Tan, R.M. Abalos, J.A. Burgos, P. Saunderson, M. Sakatani**. 2009**.** Novel
915 prophylactic and therapeutic vaccine against tuberculosis. Vaccine. **26**:3267-
916 70. Epub 2009 Feb 5.
917 **24. Okada M., Y. Kita, T. Nakajima, N. Kanamaru, S. Hashimoto, T.**
918 **Nagasawa, Y. Kaneda, S. Yoshida, Y. Nishida, R. Fukamizu, Y. Tsunai, R.**
919 **Inoue, H. Nakatani, Y. Namie, J. Yamada, K. Takao, R. Asai, R. Asaki, M.**
920 **Matsumoto, D.N. McMurray, E.C. Dela Cruz, E.V. Tan, R.M. Abalos, J.A.**
921 **Burgos, R. Gelber, M. Sakatani**. 2007. Evaluation of a novel vaccine (HVJ-

922 liposome/HSP65 DNA+IL-12 DNA) against tuberculosis using the cynomolgus
923 monkey model of TB. Vaccine. **20**:2990-3.
924 **25**. **Orme I.M., D.N. McMurray and J.T. Belisle.** 2001. Tuberculosis vaccine
925 development: recent progress. Trends Microbiol **9:**115-8
926 **26. Reed S.G., R.N. Coler, W. Dalemans, E.V. Tan, E.C. Dela Cruz, R.J.**
927 **Basaraba, I.M. Orme, Y.A.W. Skeiky, M.R. Alderson, K.D. Cowgill, J-P.**
928 **Prieels, R.M. Abalos, M-C. Dubois, J. Cohen, P. Mettens, and Y. Lobet**.
929 2009. Defined tuberculosis vaccine, Mtb72F/AS02A, evidence of protection in
930 cynomolgus monkeys. *PNAS.* **106**: 2301-2306.
931 **27. Ribi E., R.L. Anacker, W.R. Barclay, W. Brehmer, S.C. Harris, W.R.**
932 **Leif, J. Simmons.** 1971. Efficacy of mycobacterial cell walls as a vaccine
933 against airborne tuberculosis in the rhesus monkey. J infect Dis. **123:** 527-
934 538.
935 **28. Rodrigues L.C., V.K. Diwan, J.G. Wheeler**. 1993. Protective effect of
936 BCG against tuberculous meningitis and miliary tuberculosis: a meta-analysis.
937 Int. J Epidemiol. **22**: 1154-8

938 **29. Sharpe S.A, E. Eschelbach, R.J. Basaraba, F. Gleeson, G.A. Hall, A.**
939 **McIntyre, A. Williams, S.L. Kraft, S. Clark, K. Gooch, G. Hatch, I.M Orme,**
940 **P.D. Marsh, M.J. Dennis**. 2009. Determination of lesion volume by MRI and
941 stereology in a macaque model of tuberculosis. Tuberculosis. Nov;**89**:405-16.
942 Epub 2009 Oct 30.

943 **30. Shen Y., D. Zhou, L. Qiu, X. Lai, M. Simon, L. Shen, Z. Kou, Q. Wang,**
944 **L. Jiang, J. Estep, R. Hunt, M. Clagett, P.K. Sehgal, Y. Li, X. Zeng, C.T.**
945 **Morita, M.B. Brenner, N.L. Letvin, Z.W. Chen.** 2002. Adaptive immune

responses of Vγ2Vδ2+ T cells during mycobacterial infections. Science. **295:**2255-2258.

31. Trunz B.B., P. Fine, C. Dye. 2006. Effect of BCG vaccination on childhood tuberculous meningitis and miliary tuberculosis worldwide: a meta-analysis and assessment of cost-effectiveness Lancet **367**:1173-80.

32. Verreck F.A.W., R.A.W. Vervenne, I. Kondova, K.W. van Kralingen, E.J. Remarque, G. Braskamp, N.M. van der Werff, A. Kersbergen, T.H.M. Ottenhoff, P.J. Heidt, S.C. Gilbert, B. Gicquel, A.V.S Hill, C. Martin, H. McShane and A.W. Thomas. 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: 1-11.

33. Vordermeier H.M., S.G. Rhodes, G. Dean, N. Goonetilleke, K. Huygen, A.V. Hill, R.G. Hewinson, S.C. Gilbert. 2004. Cellular immune responses induced in cattle by heterologous prime-boost vaccination using recombinant viruses and bacille Calmette-Guérin. Immunology **112**: 461-70

34. Vordermeier HM, Villarreal-Ramos B, Cockle PJ, McAulay M, Rhodes SG, Thacker T, Gilbert SC, McShane H, Hill AV, Xing Z, Hewinson RG. 2009. Viral booster vaccines improve Mycobacterium bovis BCG-induced protection against bovine tuberculosis. Infect Immun. **77**:3364-73. Epub 2009 Jun 1.

35.Williams A., N.P. Goonetilleke, H. McShane, S.O. Clark, G. Hatch, S.C Gilbert, A.V. Hill. 2005. Boosting with poxviruses enhances *Mycobacterium*

970 *bovis* BCG efficacy against tuberculosis in guinea pigs. Infect Immun. **73**:
971 3814-6.
972 **36**. **Williams A., Y.Hall, I.M. Orme.** 2009. Evaluation of new vaccines for
973 tuberculosis in the guinea pig model. Tuberculosis (Edinb). **89**:389-97. Epub
974 2009 Oct 7.
975 **37. WHO**. 2008. Global tuberculosis control: surveillance, planning, financing:
976 WHO Report. WHO/htm/TB/2008.393

977 Table 1. Aerosol challenge doses of *M. tuberculosis* delivered to rhesus

978 macaques.

979

Vaccine Group	Animal Identification number	Presented Dose (cfu)	Estimated Retained Dose (cfu)
Group A BCG + MVA85A	K20	1293	60
	K61	898	40
	K65	1271	60
	K69	858	40
	K79	1299	60
	K86	812	40
Group B BCG	K43	1263	60
	K44	1368	65
	K50	932	45
	K52	1275	60
	K59	884	40
	K80	817	40
Group C No vaccine	K32	1286	60
	K47	1237	60
	K54	821	40
	K62	821	40

980

981 Table 2

982 Mann Whitney test results: statistical analysis of the differences between test

983 groups using different vaccine efficacy readouts

Readout	P value		
	Unvaccinated v BCG	Unvaccinated v BCG / MVA85A	BCG v BCG / MVA85A
Lesion volume to Lung volume ratio	0.011	0.011	0.631
Lesion number	0.024	0.199	0.337
Total pathology score	0.017	0.198	0.33
Lung pathology score	0.079	0.083	0.466
Chest X-ray	0.174	0.285	0.371
Bacterial load (CFU/g)	0.136	0.522	0.297
% weight loss	0.199	0.055	0.747
ESR	0.087	0.157	0.515
Vaccinated significantly lower than unvaccinated. **P< 0.05**		Vaccinated significantly higher than unvaccinated **P< 0.05**	

984

985

Figure 1 Experimental study design.

Figure 2 Immune response to vaccination:
Panel A shows the frequency of PPD-specific IFN-γ secreting cells measured by ELISpot. Panel B shows the frequency of Ag85A-specific IFN- γ secreting cells measured by ELISpot; panel C shows the quantity of IFN- γ secreted by PBMC during 6 day culture of PBMC with PPD measured by ELISA; panel D shows the frequency of $CD3^+$ T cells that secrete IFN-γ following stimulation with PPD. Vaccination with BCG at week 0 is indicated by the dotted line and vaccination with MVA85A indicated by the dashed line at week 12.

Figure 3 Polyfunctional T cell analysis after vaccination. Panels A, B, C show the responding $CD4^+$ T cell populations following stimulation with Ag85A peptide pools Panels D, E, F, G, H show the responding $CD4^+$ T cell populations following stimulation with PPD and panels I and J show the responding $CD8^+$ T cells following stimulation with PPD.

Figure 4 Kaplan Meier plot of survival of vaccinated and unvaccinated NHPs after challenge with *M. tuberculosis*.

Figure 5 Measures of tuberculosis-induced pulmonary and clinical disease burden. Panel A: the lung to lesion volume ratio determined using MR stereology; Panel B: number of lesions in the lung enumerated by serial sectioning and manual counting; Panel C: the total pathology score determined using a qualitative scoring system; Panel D: the scores attributed to the pulmonary component as part of the total pathology score; Panel E: score attributed to the chest radiogram on the day of euthanasia; Panel F: % weight loss from peak post-challenge weight on the day of euthanasia; Panel G: Erythrocyte sedimentation rate (ESR) on the day of euthanasia; Panel H: bacterial load in the lung.

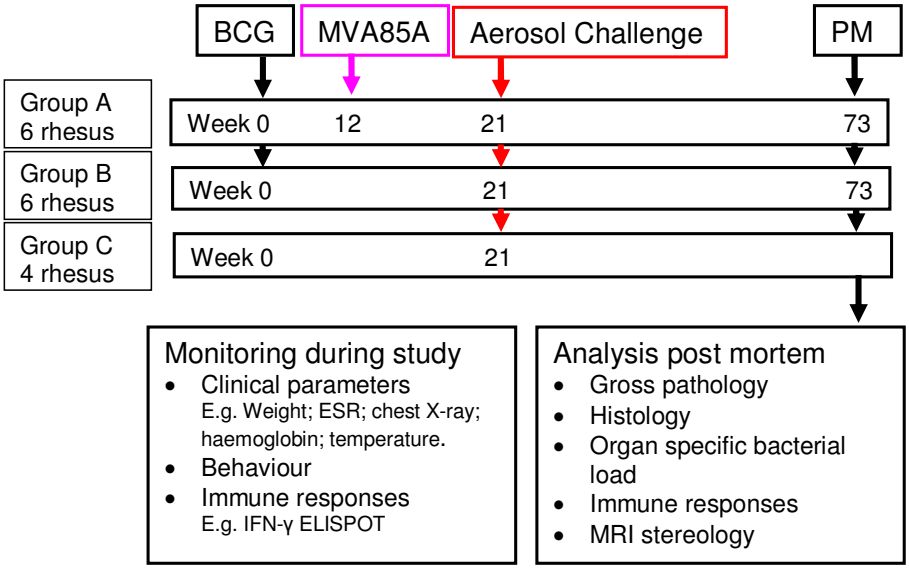
Figure 6 Immune responses after aerosol challenge with *M. tuberculosis*.
The frequency of mycobacterium-specific IFN-γ secreting cells measured after challenge measured by ELISpot is shown in panels A (ESAT6), B (Ag85A)

1020 and C (PPD). Panel D shows the profile of IFN-γ secretion by PPD-stimulated
1021 whole blood measured by ELISA.

1022 Figure 1:

1023

1024



1025

1026

1027

1028

1029

Figure 2:

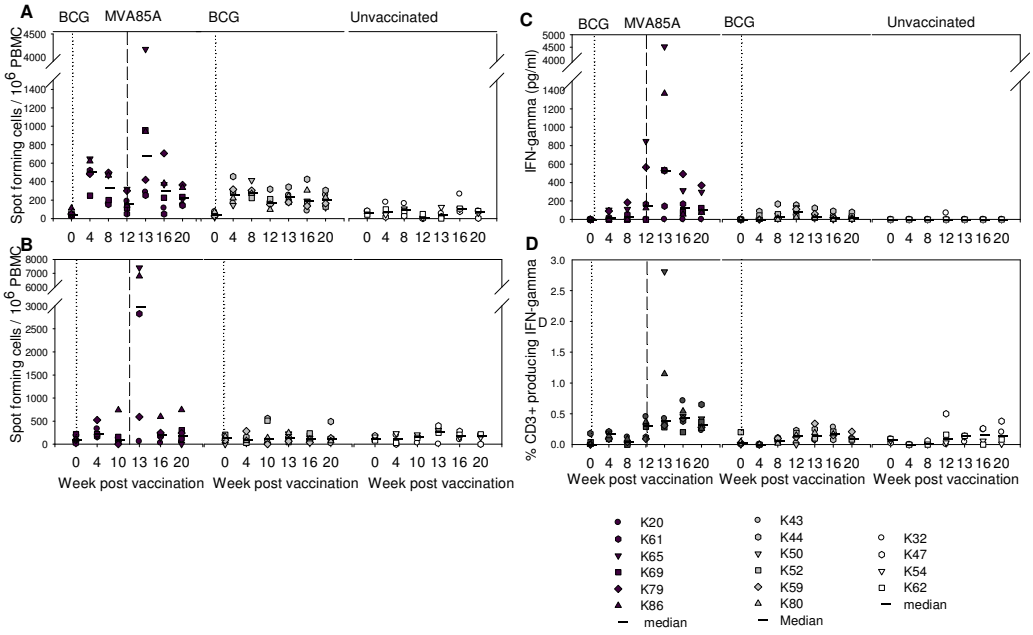


Figure 3:

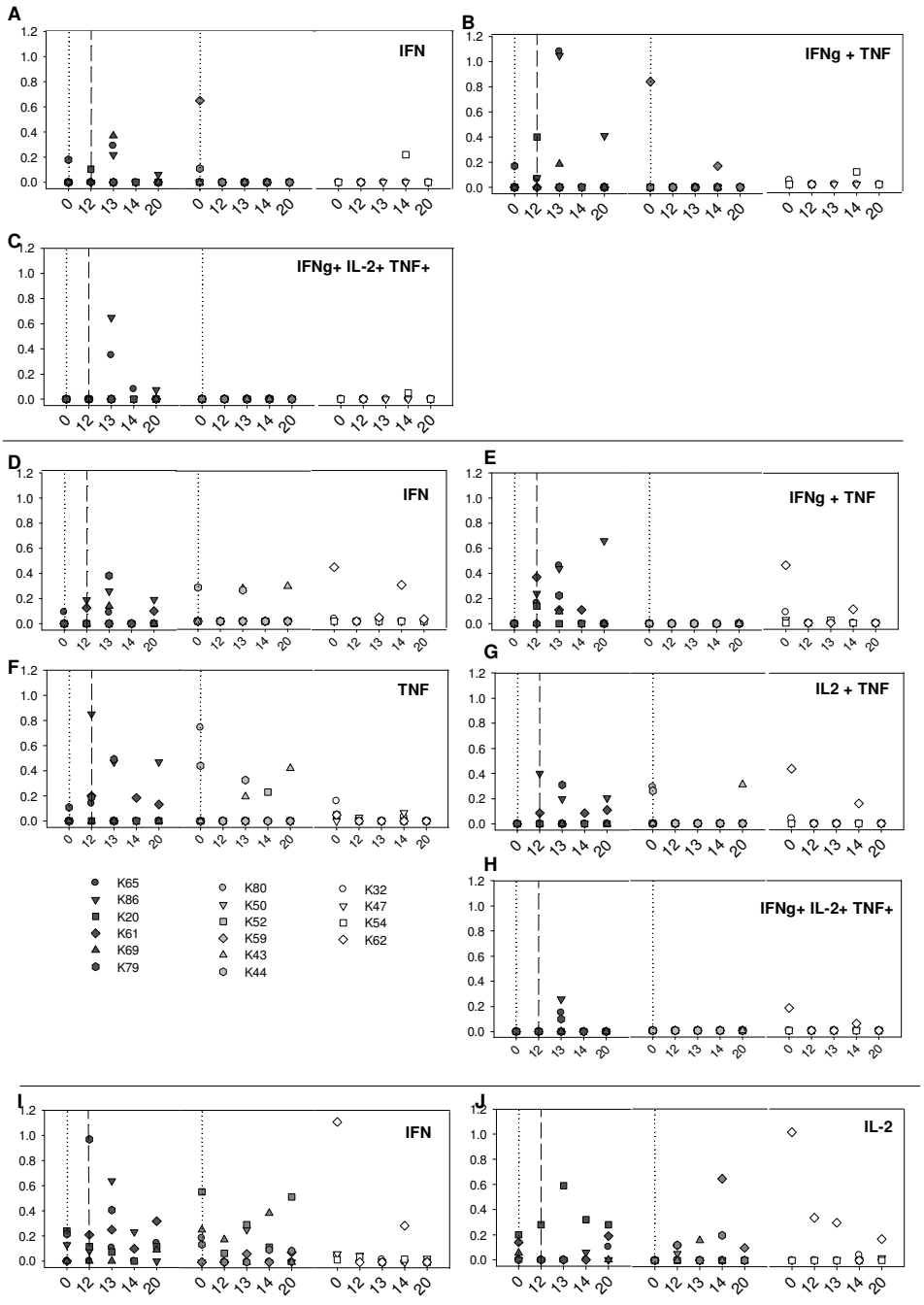


Figure 4

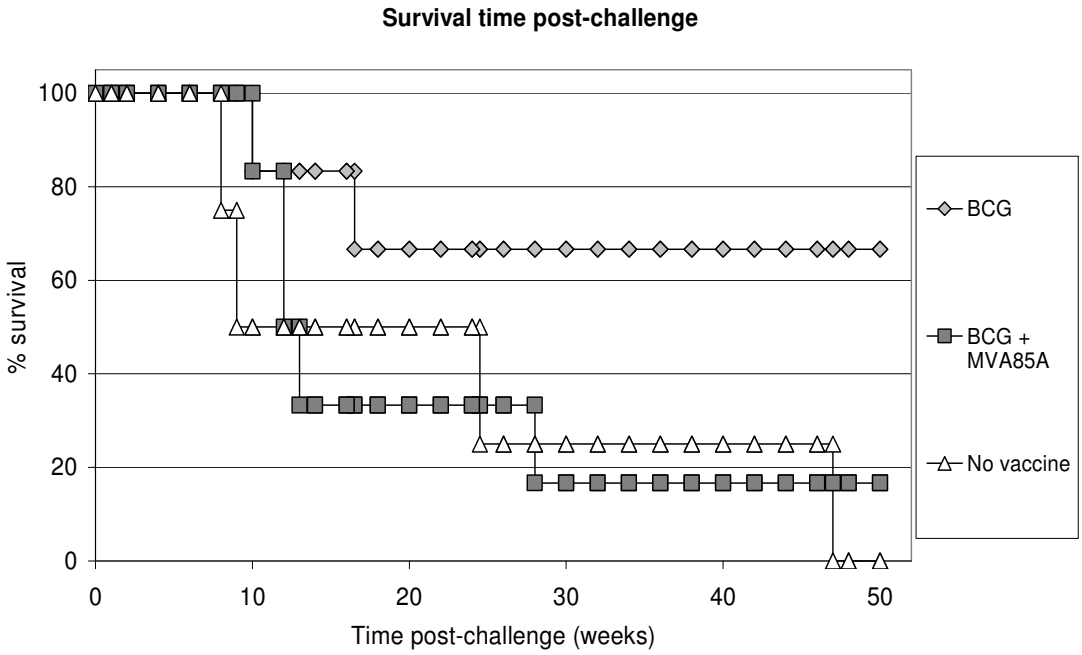
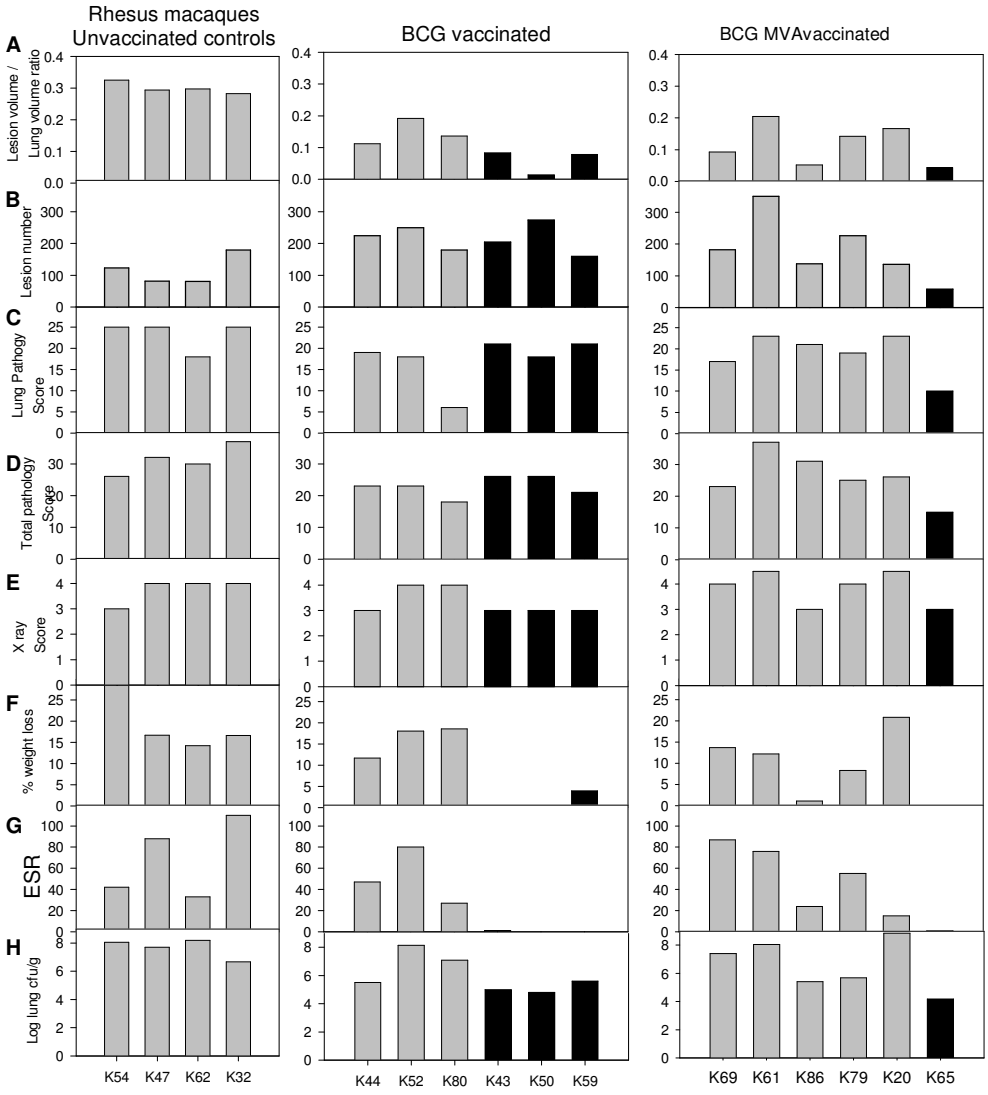


Figure 5



Black fills represent animals not showing adverse clinical or behavioural indicators at termination.

Figure 6: