Comprehensive analysis of methods used for the evaluation of compounds against **Mycobacterium tuberculosis**

Tuberculosis 92 (2012) 453–488

Contents lists available at SciVerse ScienceDirect

Tuberculosis

journal homepage: http://intl.elsevierhealth.com/journals/tube

MODEL SYSTEMS

Comprehensive analysis of methods used for the evaluation of compounds against *Mycobacterium tuberculosis*

Scott G. Franzblau [b], Mary Ann DeGroote [a], Sang Hyun Cho [b], Koen Andries [d,i], Eric Nuermberger [c,i], Ian M. Orme [a], Khisimuzi Mdluli [e,i], Iñigo Angulo-Barturen [f,i], Thomas Dick [g,i], Veronique Dartois [h,i], Anne J. Lenaerts [a,*]

[a] *Department of Microbiology, Immunology and Pathology, 200 West Lake Street, Colorado State University, Fort Collins, CO 80523-4629, USA*
[b] *Institute for Tuberculosis Research, College of Pharmacy, University of Illinois at Chicago, 833 S. Wood Street, Chicago, IL 60621-7231, USA*
[c] *The Johns Hopkins Medical School, 1550 Orleans Street, Baltimore, MD 21231, USA*
[d] *Tibotec, Pharmaceuticals Inc, Turnhoutseweg 30, Beerse, Belgium*
[e] *The Global Alliance for TB Drug Development, 40 Wall Street, New York City, New York 10005, USA*
[f] *GlaxoSmithKline, Tres Cantos Medicines Development Campus, Severo Ochoa 2, 28760 Tres Cantos, Madrid, Spain*
[g] *Department of Microbiology, Yong Loo Lin School of Medicine, National University of Singapore 5 Science Drive 2, Singapore 117597, Singapore*
[h] *Public Health Research Institute, 225 Warren St, Newark, NJ 07103, USA*

ARTICLE INFO

Article history:
Received 10 May 2011
Received in revised form
2 July 2012
Accepted 9 July 2012

Keywords:
Tuberculosis
In vivo
In vitro
Animal models
Screening methods
Drugs
Efficacy testing
Relapse
Historic review

SUMMARY

In drug development, there are typically a series of preclinical studies that must be completed with new compounds or regimens before use in humans. A sequence of *in vitro* assays followed by *in vivo* testing in validated animal models to assess the activity against *Mycobacterium tuberculosis*, pharmacology and toxicity is generally used for advancing compounds against tuberculosis in a preclinical stage. A plethora of different assay systems and conditions are used to study the effect of drug candidates on the growth of *M. tuberculosis*, making it difficult to compare data from one laboratory to another. The Bill and Melinda Gates Foundation recognized the scientific gap to delineate the spectrum of variables in experimental protocols, identify which of these are biologically significant, and converge towards a rationally derived standard set of optimized assays for evaluating compounds. The goals of this document are to recommend protocols and hence accelerate the process of TB drug discovery and testing.

Data gathered from preclinical *in vitro* and *in vivo* assays during personal visits to laboratories and an electronic survey of methodologies sent to investigators is reported. Comments, opinions, experiences as well as final recommendations from those currently engaged in such preclinical studies for TB drug testing are being presented. Certain *in vitro* assays and mouse efficacy models were re-evaluated in the laboratory as head-to-head experiments and a summary is provided on the results obtained. It is our hope that this information will be a valuable resource for investigators in the field to move forward in an efficient way and that key variables of assays are included to ensure accuracy of results which can then be used for designing human clinical trials. This document then concludes with remaining questions and critical gaps that are in need of further validation and experimentation.

1. Introduction

The primary goals of drug development for tuberculosis (TB) are to shorten and simplify treatment of active TB, provide safer and more efficacious treatments for drug-resistant TB, eliminate drug–drug interactions for TB/HIV co-infections, and improve treatment for latent TB infections. Successful development of new, safe and effective TB therapies faces a number of challenges, some unique to TB drug discovery and development, many with implications for other therapeutic indications. For the first time in 40 years there is a portfolio of novel anti-tuberculosis compounds which are in need of extensive testing as single agents and in combination regimens before moving on to human clinical trials.[1,2] Around ten drug candidates are now in clinical development and we currently have the largest pipeline of early-stage projects and compounds in history.[1–3] Although there is definitely a reason for some optimism, the current pipeline is likely not sufficient considering the high attrition rates of compounds in clinical studies.[4] Therefore, there is a definite need to increase the numbers of compounds at all levels of

* Corresponding author. Tel.: +1 970 4913079.
E-mail address: Lenaerts@colostate.edu (A.J. Lenaerts).
[i] These authors are contributing authors.

1472-9792/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.tube.2012.07.003

the TB drug discovery process, as well as to improve testing methodologies to identify new active compounds. This need is especially urgent in light of the recent efforts by the Critical Path to New TB Regimens (CPTR) program, which strives to develop a novel combination drug regimen for tuberculosis based on 3–4 new drugs with a novel mechanism of action.[5]

Typically there are a series of preclinical studies with new chemical entities that must be completed before use in humans. A sequence of *in vitro* assays followed by *in vivo* testing in validated animal models to assess the activity against *Mycobacterium tuberculosis*, pharmacology and toxicity is generally used for advancing compounds in a preclinical stage.[6] A plethora of different assay systems and conditions are used to study the effect of drug candidates on the growth of *M. tuberculosis*, making it difficult to compare data from one laboratory to another. The Bill and Melinda Gates Foundation recognized the scientific gap to delineate the spectrum of variables in experimental protocols, identify which of these are biologically significant, and converge towards a rationally derived standard set of optimized assays for evaluating compounds. The question that we are asking here is whether there are more efficient ways of testing new candidate agents and re-evaluating the current methodologies of compound testing. The goals of this document are to recommend protocols and hence accelerate the process of TB drug discovery and testing. This document is not meant to be a substitute for the regulatory requirements for drug approvals, but rather to provide insight into key variables of preclinical assays for the benefit of all investigators in the field. In 2001, a scientific blueprint was published to guide the process of preclinical studies, and this document detailed ways in which to close the gaps that exist in TB drug discovery and development.[5,7] The blueprint celebrates this year as its 10th anniversary. This document can be seen as an extension of this blueprint, however, the focus will mainly be on the *in vitro* and *in vivo* methodologies required for activity testing against *M. tuberculosis* and less on methodologies that are utilized in drug development in general (such as ADME, pharmacokinetics, toxicology). In addition, it includes a focused review of the current and historic methodologies. In the chapters that follow, an assessment of the current state of activities ongoing in laboratories engaged in TB drug testing methodologies as well as the lessons learned from experiences in both academic and industrial laboratories is presented. Data gathered from preclinical *in vitro* and *in vivo* assays during personal visits to laboratories and an electronic survey of methodologies sent to investigators is reported. Comments, opinions, experiences as well as final recommendations from those currently engaged in such preclinical studies for TB drug testing are being presented. Certain *in vitro* assays and mouse models were re-evaluated in our laboratories as head-to-head experiments and a summary is provided on the results obtained. It is our hope that this information will be a valuable resource for investigators in the field to move forward in an efficient way and that key variables of assays are included to ensure accuracy of results which can then be used for designing human clinical trials. This document then concludes with some remaining questions that are in need of further validation and experimentation.

2. *In vitro* evaluations of TB drugs

2.1. Historic review

Although the tubercle bacillus was first described by Koch in 1882, it would be six more years before media were devised which allowed for sufficient growth to enable studies of bacillary composition and antigenicity. The key was the introduction of glycerol as a carbon source, first in the presence of protein/peptone[8,9] and also in its absence,[10] the latter being preferred for immunologic studies. The minimal requirements for growth were defined as early as 1894 by Proskauer and Beck[11] and consisted of ammonium carbonate, mono-potassium phosphate, magnesium sulfate and glycerol although they also recognized that simple amino acids could substitute for ammonium salts.

In the first two decades of the 20th century several groups focused on further studies of the nutrition of the tubercle bacillus with much attention devoted to amino acids.[12] This followed the discovery that it was actually small amounts of (contaminating) free amino acids in the protein/peptone preparations available at that time that supported growth in the early culture media. A number of single amino acids could serve equally well as the sole source of nitrogen, most notably arginine, histidine, alanine, leucine and asparagine but not including those containing a benzene ring (phenyalanine, tryptophan and tyrosine).[13] During this time the H37 strain (originally isolated in 1905) was widely employed and the extent of growth was usually described as scant, moderate or luxuriant.

By the time the first useful anti-TB agents were discovered in the early 1940s there were already numerous culture media developed for the cultivation of *M. tuberculosis*. The effects of various media constituents on growth and inhibition were however still being described concurrently with the discovery of streptomycin (SM), para-aminosalicyclic acid (PAS) and isoniazid (INH).[14] Crimm and Martos[15] found growth of H37 to be optimal when using an enzymatic hydrolysate of casein ("Amigen") as an alternative to peptone, individual amino acids or combinations of amino acids. Of the latter, the combination of phenylalanine, tryptophan and threonine gave the highest growth yield.

Following the description of the stimulation of oxygen uptake by salicyclic and benzoic acids,[16] the Swedish scientist Jorgen Lehmann evaluated more than 50 analogs for the ability to inhibit the growth of *Mycobacterium bovis* bacillus Calmette-Guérin (BCG) using Sauton's medium[17] and in doing so identified para-aminosalicylic acid as the most potent. This was published together with animal toxicity data and clinical trials results. Compounds at 10 mM to 1 μM were evaluated by the extent of reduction in either pellicle area or dry weight at a time (12–20 days) when controls had formed a confluent pellicle. Pellicle area could be determined by comparison to paper circles of known radii.[18]

The effects of variables such as inoculum size and growth phase on static versus cidal anti-TB activity were also being elucidated at this time. *Mycobacterium smegmatis* (then referred to as the avirulent strain *M. tuberculosis* hominis 607) was often used in parallel with the $H_{37}Rv$ strain. It was during this period that the use of non-ionic detergents[19] was first employed for the purpose of dispersed growth,[14] and this allowed for the quantitative estimate of growth in liquid media by measuring turbidity.[20,21] Youmans[22] argued that testing of compounds be preferably conducted in simple, chemically defined media such as Proskauer–Beck (a basal medium containing glycerol, asparagine and salts) and concluded, based on the testing of over 3000 compounds by the mid-1940s, that *in vitro* testing was of value in eliminating inactive compounds from further consideration.

Dubos used the closely related Kirchner medium to identify chaulmoogra oil (used clinically for leprosy at the time) as being active against *Mycobacterium avium* via turbidometric measurements and then further modified the medium by adding citrate, Tween 80, casein hydrolysate and bovine serum albumin for dispersed growth of *M. tuberculosis*. The Dubos lab[23] described the protective effect of serum albumin (particularly fraction V) against the previously recognized toxicity of free fatty acids, although Youmans and Youmans later found that there was a very narrow concentration window within which fatty acids would support growth in the absence of glycerol or protein.[24] Dubos also

recognized that the toxicity associated with impure Tween 80 in the absence of serum or albumin was due to contamination with small amounts of fatty acids although this toxicity was not apparent in media containing glycerol. Youmans and Youmans[19] noted that Tween 80 increased or decreased the potency by >3-fold of compounds tested in a basal, protein-free, medium (containing only glycerol, asparagine, magnesium citrate, KH_2PO_4 and K_2SO_4, pH 7.0). These effects were abolished for some of the compounds when albumin was added to the medium. Regardless of this study and previous reports on both positive and negative effects of Tween 80 on MICs, this detergent was retained in media used by most groups[25] as turbidometric measurement was the only viable broth dilution-based quantitative substitute for colony forming unit-based (CFU) assays at that time and to use the former requires well dispersed growth. The potential effects, however, were evaluated whenever a new candidate lead was being profiled by testing in the presence and absence of Tween 80.[26]

The one exception to the use of visual or photometric determinations of growth described in the early literature was the use by the Youmans lab of total nitrogen measurement. This was done using a micro Kjeldahl method for comparison of growth in media with and without Tween 80 as they recognized the difficulty in visually comparing dispersed and clumped growth.[27]

While some groups chose to use a single culture medium[28] in their discovery efforts, by the late 1940s several groups began using multiple media, sometimes including egg-based media,[18,29] in their *in vitro* assessments of new TB drugs, a practice that continued thereafter.[18,30,31] Drain et al.[25] compared the activity of PAS analogues in 4 different media: Proskauer and Beck (P&B), P&B with 10% human serum, Kirchner synthetic medium and Tween-albumin liquid medium. Although there were differences in growth rate (Tween-albumin = Kirchner > P&B+10%serum > P&B), there was no significant difference in inhibitory activities. Steenken and Wolinsky[32] used Tween-albumin, P&B and P&B plus 10% (beef) serum to evaluate the *in vitro* activity of terramycin (oxytetracycline) against *M. tuberculosis* H37, noting a reduction in activity with the latter medium. They later extended this approach to a wide array of antimicrobial agents.[33] It apparently became common practice at this time to evaluate compounds in the presence and absence of 10% animal or human serum[18,34] if media were being used that were otherwise devoid of albumin.

Inocula at this time were commonly described in terms of dry weight, ranging from 1 μg/ml[25] to 6 μg/ml,[35] however with the advent of media supporting dispersed growth, it became feasible to standardize inocula photometrically.[36]

There is little mention of the use of organic solvents to evaluate water-insoluble compounds, however Drain et al.[25] noted that propylene glycol could be used for this purpose without effect on the growth of *M. tuberculosis* at concentrations up to 1% and this solvent was also employed by Klip[18] in the investigation of calciferol.

While most screening was conducted in liquid media, the activity of viomycin was first observed at Parke–Davis in 1951 via cross-streaks of *Streptomyces floridae* on solid medium with both sensitive and SM-resistant *M. tuberculosis* H_{37}Rv as well as *M. smegmatis*.[36] Unfortunately, the medium used was not described for the cross-streak nor for the subsequent MIC tests (although it was noted that the latter contained serum) that showed the peptide to be more active against mycobacteria than other bacteria. More detailed studies at Pfizer, where viomycin had been independently discovered, determined MICs in Dubos–Davis liquid medium with 0.02% Tween 80 and 0.5% albumin using a visual readout after 12 days of growth. Inoculum consisted of a 4–6 day liquid culture, adjusted photometrically.

The various liquid media and methods described above were used throughout the 1950s in studies of analogs of anti-TB active small synthetic molecules,[30–34,37,38] nutrients[18,39,40] or natural substances found in serum,[41,42] urine,[43] tissue[44–48] or from microorganisms[26,32,49–53] including mycobacteria.[54,55]

There were few major changes with respect to *in vitro* methodology in the 1960s. The stimulus for developing newer variations of existing culture media came largely from efforts to improve primary culture with respect to time and for isolation of drug-resistant and auxotrophic clinical isolates (e.g. adding catalase and biotin). These were incorporated into the commercialized Dubos Broth and Middlebrook 7H9-based formulations.

Readouts were still largely limited to visual observation or spectrophotometric measurement of turbidity in tube cultures through the mid-1970s until the description of an automatable radiometric assay for determination of the growth of *M. tuberculosis* – the precursor of the BACTEC 460 system.[56] Several years later the BACTEC 460 was demonstrated to be usable for clinical drug susceptibility testing.[57,58] Although expensive, radiometric and low throughput by today's standards, it was by far the most rapid assay available, returning susceptibility data in 5–7 days. As such it represented the state-of-the-art for TB drug discovery until the dawn of the non-radiometric, metabolism-based assays in the 1990s. These new automated clinical drug susceptibility tests retained the large volume, individual tube/bottle format but now monitored growth/inhibition via fluorometric detection of oxygen consumption (MGIT), colorimetric detection of carbon dioxide (MB/BacT) or detection of a pressure change due to oxygen consumption (Versa TREK, formerly ESP Culture System II).[59] Most of those involved with drug discovery began to develop and utilize microplate-based assays[60] that employed redox dyes such as resazurin/Alamar Blue,[61] luminescent reporter proteins[62] and fluorescent reporter proteins.[63]

2.2. Contemporary survey of current practices

2.2.1. Survey results: overview

A ninety five question survey of protocols and opinions was conducted of which about half of the questions were devoted to *in vitro* and the other half to *in vivo* methodologies. Responses were received from 30 investigators from 25 institutions including government laboratories (18.5%), academic laboratories (56%) and from the pharmaceutical industry (11%). Of these, 79% of respondents were willing to share their protocols. The survey can be viewed in its entirety at the following website: www.mrl.colostate.edu/acceleration.

Those laboratories known to be involved in tuberculosis drug discovery at the level of *in vitro* evaluation were surveyed with respect to the use of bacterial strains, culture media, testing format, assays types (e.g. primary screening, minimum inhibitory concentration (MIC), etc.), data interpretation and scale of testing. The groups surveyed were quite diverse with respect to their TB drug discovery efforts. Some engage in high throughput screening (HTS) with hit to lead and lead optimization follow-up while others are focused on one or a few classes of compounds and therefore have no need for HTS-compatible assays. The diversity of approaches was reflected in the survey results.

Most (70%) of the respondents performed whole cell screening while 41% performed isolated molecular target assays. This may have reflected a bias in selection of labs thought to be conducting cellular assays. A smaller percentage (7.4%) performed multicomponent screening (Figure 1).

2.2.1.1. Choice of bacterial strains. As expected the majority of laboratories used the H_{37}Rv strain from ATCC (#27294). Some did however report the use of clinical isolates (Figure 2). Only three groups reported differences in compound activity between strains

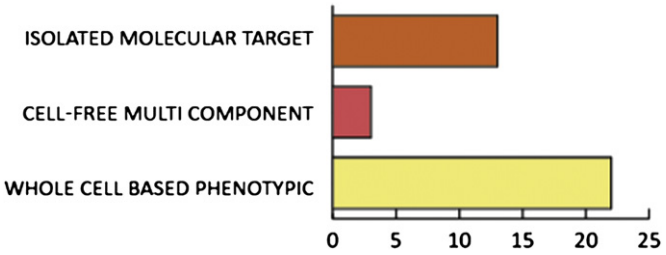

Figure 1. Survey responses regarding the type of *in vitro* assay methodology used in TB drug testing (presented in percentages). Survey question: Which type of *in vitro* assay(s) have been used in your laboratory?

of *M. tuberculosis*. One found differences between $H_{37}Ra$ versus $H_{37}Rv$ susceptibility to new compounds. Another reported at least 2 compounds to which $H_{37}Ra$ is more sensitive than $H_{37}Rv$, one being rifampin (RIF). One laboratory observed that BCG was more susceptible to nitroimidazoles versus their standard $H_{37}Rv$ strain; however they did not speculate on the mechanism. Another group found that >90% of hits on BCG also had activity for *M. tuberculosis*. There was no mention of the percentage of TB-active compounds that were inactive against non-tuberculous surrogates, (i.e. would be missed in a surrogate-based screen), however, a very recent publication indicated these figures to be 50% and 21% for *M. smegmatis* and BCG, respectively.[64]

2.2.1.2. Choice of media. Most of the laboratories utilize a Middlebrook 7H9-based medium with specific supplements (Table 1). A smaller number utilize glycine–alanine-salts (GAS) or Sauton's medium, both of which are devoid of protein, and investigators assume higher sensitivity using poorer media. Both filtration and autoclaving were mentioned as means of sterilization. While filtration is routinely used for medium supplements with well recognized heat lability, this method of sterilization was also used by Becton Dickinson (Salman Siddiqi, personal communication) for the preparation of the complete commercial BACTEC 12B medium (the non-radiometric equivalent of Middlebrook 7H12 medium), which contains the heat-stable 7H9 broth base and supplements, some of which are heat-labile. In a previous head-to-head comparison of autoclaved 7H9 base plus filter-sterilized supplement versus filter-sterilized complete medium, we noted superior growth of *M. tuberculosis* in the latter (S. Franzblau, unpublished observation). The use of filter-sterilized complete medium also obviates the need for making 10× concentrated supplement stocks where solubilizing bovine serum albumin fraction V can be problematic.

Table 1
Composition of the Middlebrook media series and supplements used*.

Common components		Variable components	
Na_2HPO_4	2.5 mg	Glycerol	2 mg/ml (0.2%)
KH_2PO_4	1.0 mg	Oleic acid	50 μg/ml
NH_4SO_4	0.5 mg	Palmitic acid	5.6 μg/ml
Na glutamate	0.5 mg	Albumin, Bovine	5 mg/ml (0.5%)
Na citrate	0.5 mg	Dextrose	2 mg/ml (0.2%)
Ferric NH_4 citrate	40 μg	Catalase	3 μg/ml
$MgSO_4$	50 μg	Casitone	0.1 mg/ml
$CuSO_4$	1.0 μg	Tween 80	0.5 mg/ml (0.05%)
$ZnSO_4$	1.0 μg	WR1339	0.25 mg/ml
$CaCl_2$	0.5 μg	NaCl	0.85 mg/ml
Pyridoxine	1 μg	pH 6.6 ± 0.2	
Biotin	0.5 μg		

*Handbook of Microbiological Media, 2nd Edition, Atlas and Parks (Eds.), CRC Press.[257]

2.2.1.3. Incubation and readout. In contrast to the overnight incubations used for testing for growth inhibition of the majority of gram-positive and gram-negative bacteria, mycobacteria require longer incubations due to their longer generation times. All reported incubation temperatures were at 35–37 °C. Of 21 respondents performing phenotypic assays with *M. tuberculosis*, 12 used microplates with incubation times ranging from 5 to 14 days but most often between 7 and 10 days. Those reporting use of a tube format mostly used a visual observation readout with somewhat longer incubation times of up to 21 days.

Overall approximately one third of respondents used visual readouts in liquid culture (Figure 3). Alamar Blue or resazurin reduction was the most frequently used endpoint for those performing microplate assays, with about half using a fluorometric readout and half using an absorbance or visual readout. At least 5 groups reported use of intracellular ATP measurement and two of these groups also used luciferase reporter strains. At least thirty percent of respondents used colony forming units (CFU) readouts. It is possible that this is an underestimate since some groups may only have conveyed what they use for primary or secondary testing and CFU readouts may be used only as a confirmatory readout. At least 8 groups reported using more than one endpoint.

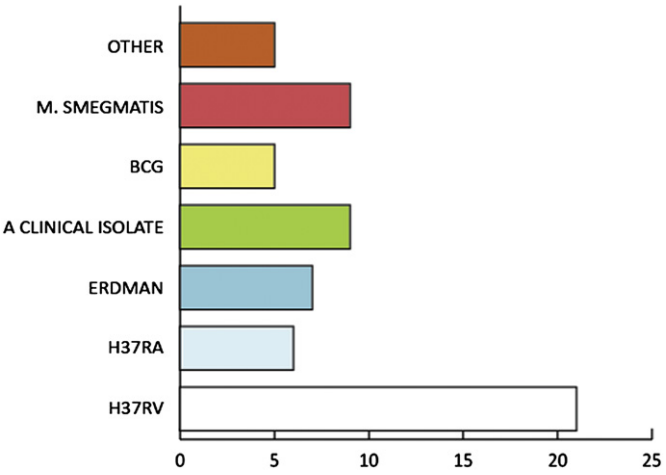

Figure 2. Survey responses regarding the mycobacterial strain used in anti-TB drug screening assay (presented in percentages). Survey question: Which *M. tuberculosis* strain is being used in your MIC assays?

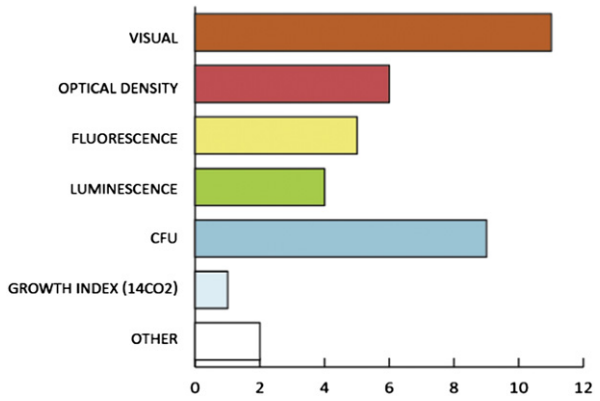

Figure 3. Survey responses regarding the readout of the *in vitro* assay for evaluation of TB compounds (presented in percentages). Survey question: what readout/endpoint is being used in your MIC assays?

The incubation time between reagent addition and reading ranged from minutes for luciferase assays to at least 16 h, and more typically 24 h for the Alamar Blue assay method. Many felt that the signal to noise ratio was affected by the time of incubation. The inoculum, temperature, evaporation and compound precipitation were cited as factors potentially influencing results.

2.2.1.4. Defining a "hit". For those groups that conduct primary screening prior to determining MICs, maximum concentrations ranged from 2 to 10 μg/ml or in one case 30 μM. One group mentioned to screen at multiple concentrations from 0.12 to 2 μg/ml. A 10-fold reduction or 90% inhibition in viability/metabolism compared to DMSO (drug-free) controls was commonly used to define a hit.

2.2.1.5. Control compounds. MICs obtained with clinical anti-TB agents are used to quality control assay performance. The most commonly used positive control drugs were isoniazid (INH) and RIF but SM, ethambutol (EMB), moxifloxacin (MXF), ofloxacin (OFL), levofloxacin (LVF), amikacin (AK) and linezolid (LZD) were also cited. DMSO alone, erythromycin (EMY) and vancomycin (VAN) were reported as negative controls. For hypoxic or anaerobic assays, most laboratories use metronidazole and isoniazid as positive and negative controls, respectively. When testing is performed in microplates, these agents are either included in the same plate with test samples (40% of respondents) or in separate plates (60% of respondents).

2.2.1.6. Minimum inhibitory concentration (MIC). Some groups conduct primary screens at fixed concentrations but many determine MICs for every sample. The former is often conducted for large-scale screens followed by MICs for all hits. MICs were determined by at least one group on 2-fold serial dilutions of compounds using visual observation of the reduction of a redox dye such as resazurin or Alamar Blue. MICs were variably defined, in some cases the visual observation of a total inhibition of growth while many calculate/extrapolate the lowest concentration yielding 50 or 90 percent inhibition with the aid of various programs including PRISM, WHONET etc. Other examples include the lowest concentration for which luminescence is lower than a 1:100 dilution of test culture with no drug. Agar dilution MIC was also reported, based on the absence of growth on drug-containing 7H11 plates. Incubation time, pH, inoculum, compound solubility, batch of compound, use of $H_{37}Ra$ (versus $H_{37}Rv$), media age, compounds used after prolonged storage, growth phase of the cells, initial OD and inoculum size are variables found to affect the MIC for standard anti-TB drugs or experimental compounds at fixed screening concentrations.

2.2.1.7. Source of compound libraries for HTS. Most of the compound libraries screened have been commercially available synthetic (e.g. Chembridge, Tripos, Nanosyn), or natural product (e.g. Analyticon) collections or proprietary libraries (e.g. Novartis archive). Biological extracts, usually from academic organizations, have also been screened (Figure 4).

2.2.1.8. Other assays. Macrophage assays to assess intracellular activity against *M. tuberculosis* are performed with a wide variety of cell sources including rat cell lines, bone-marrow derived macrophages, J774 and THP-1 cells. About half of respondents determine MBCs by sub-culture from the MIC test while others conducted it as a separate test.

2.2.1.9. Non-replicating or slowly replicating bacterial assays. Low oxygen, starvation and nitric oxide (NO)-release assays are used to measure the activity of new compounds against non-replicating (NR) or slowly replicating (SR) TB bacilli. *M. tuberculosis*, *M. smegmatis* and *M. bovis* BCG are usually used for these studies. One group uses the $H_{37}Rv$ strain of *M. tuberculosis* for NR and SR assays as well as BCG and all were chosen because they are standard laboratory strains and the genome sequences are available. The positive and negative controls for the NR and SR assays included such agents as metronidazole, isoniazid (INH), MXF and RIF.

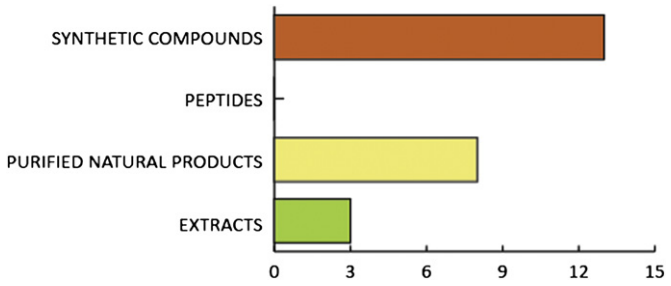

Figure 4. Survey responses regarding the type of compound libraries used for the evaluation of TB compounds for activity in *in vitro* assays (presented in percentages). Survey question: what type of compounds are being used in your high throughput screening assays?

Comments regarding throughput indicated these assays to be relatively low throughput. Laboratories reported conducting experiments once every three weeks with only 12–20 compounds per run, another did 5–10 compounds/week once per week and another reported doing 2–3 experiments per week.

The major variables that affect the outcome of the NR or SR assay include the head space ratio, abrupt stirring, the rotation per minute (rpm) of stirring cultures and sealing of vials. Variation in results is seen with INH (showing slight activity), or with experimental compounds; the latter is usually due to solubility issues. Other assay variables included the starting inoculum, degree of anaerobiosis and clumping (for nutrient starvation model). Weaknesses of the NR or SR assay system include the fact that the *in vivo* relevance of these assays is not known and variable outcomes between assays. Also mentioned is the lack of suitability for HTS and lower reproducibility than the regular MIC assay. Assay optimization was mentioned to be critical for the Wayne model to obtain meaningful results. The strength is its strong ability to differentiate between sterilizing compounds (such as RIF) and non-sterilizing compounds (such as INH) against NR/SR *M. tuberculosis*. A perceived strength is that it is a good indicator of efficacy against non-replicating, persistent bacilli which may identify compounds that could potentially shorten the treatment of TB.

2.3. Issues related to in vitro testing methodology

The issues discussed below were derived from the post-Gordon Research Conference Workshop held on August 21, 2009 in Oxford, UK as well as visits to GSK in Spain, NITD in Singapore, AstraZeneca in India and Tibotec/J&J in Belgium.

2.3.1. Media and supplements

It appears that most of the pharmaceutical companies engaged in TB drug discovery employ some version of the Middlebrook 7H media series (Table 1), all of which contain bovine serum albumin but these then vary with respect to the supplements utilized. Perhaps the most important observation and illustrative example of the effect of carbon source was the experience of NITD with glycerol-containing media. A number of compounds identified by NITD as hits in glycerol-containing media were subsequently found to be both glycerol-dependent *in vitro* and inactive *in vivo* despite

adequate exposure. These compounds were later shown to inhibit an enzyme involved in glycerol metabolism, resulting in accumulation of a toxic intermediate.[65] (GSK also observed an excessive number of glycerol-dependent hits using BCG with possibly some hits missed as well.) As glycerol is perhaps the most commonly used carbon source for culturing *M. tuberculosis* in a variety of protein-containing and protein-free media, in order to avoid pursuing glycerol-dependent series it then behooves us to consider using alternative carbon sources and/or to confirm early on that all hits obtained in glycerol-containing media are not in fact glycerol-dependent.

There was little discussion of other experiences with media component effects other than the results presented from the head-to-head experiments conducted at University of Illinois (discussed under head-to-head comparisons). One such result was that MICs of INH in Middlebrook media containing commercial ADC supplement were 1.8–2.1 μM, which was 6–7× higher than in all other media compositions not using this particular supplement. Similar MICs for INH in ADC-containing media were observed at Vertex Pharmaceuticals.

The overall impression gained from the visits and the workshop was that while there may have been an initial lack of appreciation at individual institutions for the effects of media components, especially carbon source, that this has changed based on direct experience. In contrast, today some groups are determining MICs in parallel with different carbon sources (such as glucose vs. acetate) as supportive evidence that growth inhibition may be due to inhibition of a particular target/pathway. There are efforts now in designing assays more closely resembling *in vivo* situations (looking into changing media with components that are part of pathways of reported genes essential for *in vivo* growth). An example is using both glucose and acetate as carbon sources in the media.

Tween 80 and occasionally other non-ionic detergents are commonly used to minimize clumping of mycobacteria during growth. This is especially important for being able to aliquot uniformly and for determining growth/inhibition by turbidity. Most groups use these supplements in media used to grow the inoculum and some retain it in the culture medium during exposure to test compounds. Other groups, however, are concerned about potentiating the activity of test compounds presumably by "artificially" increasing the permeability of compounds and some instances were cited of MIC differences of several orders of magnitude. There was, however, little indication of specific compounds for which this effect had been noted. As discussed below, the UIC comparative studies identified the potentiating effect of Tween 80 only with RIF while an antagonistic effect was noted with ethionamide and ketoconazole.

There is a need for clarity regarding terminology used for Middlebrook-based media. It is very common for investigators to report the use of "Middlebrook 7H9" medium without indicating which supplements are added or omitted (Table 1). It is recommended that each supplement and its concentration be specified when using the Middlebrook 7H9 broth base.

2.3.2. Readouts

The most common readouts appear to be a) fluorometric or spectrophotometric measurement of resazurin reduction, using either the commercial Alamar Blue formulation (microplate Alamar Blue or MABA) or purified resazurin (often referred to as REMA), b) measurement of intracellular ATP using a commercial kit containing luciferin, luciferase and a lysing reagent, or c) measuring turbidity at a visible wavelength between 570 and 600 nm. There was some dissatisfaction with the signal:background when using resazurin/Alamar Blue although it was not clear if this was related to spectrophotometric measurement where sensitivity is reduced due to the spectral overlap of the oxidized resazurin with the resorfin product. In a comparison of different Middlebrook-based media (see section 2.6) using the MABA, signal:background ratios approaching 20 were observed with media containing the OADC supplement. At least two major companies have used measurement of intracellular ATP in BCG for high throughput screening. This appears to be a sensitive and reliable viability readout that would likely be more widely employed if reliance on the relatively expensive kits could be circumvented. Modern microplate spectrophotometers have the sensitivity to detect very low levels of absorbance in the visible light spectrum, therefore some labs choose to use this as measure of growth/inhibition. No mention was made of a problem here with signal:background although in comparative experiments (see Section 2.6) the absorbance readout gave an unacceptable Z' value.

2.3.3. Protein binding

There appears to be a growing consensus that determining functional protein binding via MIC shift is more important than just determining percentage bound by equilibrium dialysis or similar methods. Only one group, however, made mention of this during our visits. Most groups use Middlebrook-based media containing 0.5% bovine serum albumin in their routine susceptibility assays whereas the albumin content of blood is approximately 4% w/v. Although albumin is the serum protein most commonly involved in significant binding of drugs, others such as alpha 1-acid glycoprotein, lipopeptides and globulins may also be involved. The most common method of determining protein MIC shift is through the use of serum supplements. In our head-to-head experiments (see Section 2.6), we are evaluating both serum/plasma and pure albumin.

2.3.4. Minimum bactericidal concentration (MBC)

In general most groups determine MBC by subculture onto drug-free solid media from the same incubations used in MIC determination and determine the lowest concentration effecting either a 1 (MBC_{90}) and/or 2 $\log_{10}$ (MBC_{99}) decrease in CFU relative to that existing prior to compound addition. No mention was made of the use of metabolic or reporter gene CFU surrogates for MBC determination, likely because a) this is normally not a high throughput assay, b) lack of sufficient sensitivity to see a 1–2 log decrease from the starting inoculum and c) excessive stability of some reporter proteins.

2.3.5. NRP assays

In general, assays to detect activity against slow or non-replicating cultures of *M. tuberculosis* use oxygen or nutrient limitation or low pH with or without nitric oxide. Very recently, models have been described that use biofilms[66] or a streptomycin-requiring strain,[67] however, the newer assays were beyond the scope of discussion during the visits or the workshop. Several of the companies we visited have worked with more than one NRP model, typically at least a Wayne-type model and a starvation or stationary phase model. All use CFU readouts but at least one group includes intracellular ATP as an alternate viability indicator in their Wayne model. There was no mention of conducting high throughput screening with an NRP assay although this would not be feasible with CFU readouts. It is not clear how many groups examine compound activity at low pH as this was only mentioned by one group and in that case the liquid broth method used was the BACTEC with the polyoxyethylene stearate (POES) supplement. RIF, TMC-207 and pyrazinamide (PZA) are considered to be relevant controls. Overall there is no consensus regarding the relative value of the different NRP models.

2.3.6. Strain panel

There is general consensus regarding the need for a panel of globally representative strains of *M. tuberculosis* to confirm that new drug classes would be useful against all major genetic clades found throughout the world. Currently some institutions have clinical isolate collections but to our knowledge none have all of the desired attributes:

- ○ Originate from single colonies (which has the advantage of eliminating mixed infection, however, there is a risk of limiting diversity by isolating only a subclone).
- ○ Contain representatives from all major genetic clades as determined by at least 2 typing systems
- ○ Have full genome sequences
- ○ Have multiple aliquots of freeze-dried stocks for long term storage
- ○ Complete chain of custody
- ○ Early passage number from original clinical isolate

2.3.7. Macrophage assay

The macrophage assay is considered critical by some groups and of little to no importance/usefulness by others. Part of this disagreement stems from differences observed between the location of *M. tuberculosis* in mouse models and in humans. However, although the bacilli appear to be mostly intracellular in most mouse models, there is skepticism by some about using the macrophage assay as a filter before testing in mice. The groups that place the most emphasis on this model point out the importance of variables such as passage number or cell type or (in the case of PZA) activation status and whether conditions support bacterial growth or persistence. One group mentioned that in their laboratory *in vivo*-active compounds did not always show TB activity in macrophages. In fact, three of their most active compounds in the mouse model did not show any activity in the macrophage model. The reason for this was not clear, but might (at least in part) have to do with the solubility of the compound in an *in vitro* system. Readout is typically CFU, although Institute Pasteur Korea uses an automated confocal microscopy imaging system,[68] and therefore can conduct medium to high throughput screening. In general, activity in macrophage cultures is considered a profiling assay that often is not included among those on the critical path.

2.3.8. Whole blood assay

Although not discussed at the post-conference workshop described above, the whole blood assay is essentially an intracellular infection model that can be used to determine the activity of anti-TB agents.[69] This involves incubation of *M. tuberculosis* in heparinized whole blood collected from drug treated individuals (or alternatively the blood samples are ex-vivo exposed to anti-TB agents) and subsequent determination of the viability of *M. tuberculosis* by time to positivity in the BACTEC 460 or MGIT 960 systems. Interestingly the aminoglycosides, which are widely considered inactive in macrophage monolayers, demonstrated potent activity in this assay.[70] A recent publication describes the use of this assay to assess combinations of experimental and established anti-TB drugs for activity against XDR-TB.[71]

2.4. Overview of the in vitro methodology and key definitions in the field

The major variations in the *in vitro* testing of experimental compounds against replicating cultures include culture media composition and readout. Culture media are usually based on a) a defined minimal media with glycerol as the carbon source and devoid of protein (Sautons, Proskauer-Beck, Glycerol–alanine-salts), or b) albumin containing media that use fatty acids and/or glycerol as the primary carbon source (Middlebrook or Dubos). Either type of media may be used with or without a non-ionic surfactant such as Tween 80 to promote dispersed growth.

Because non-CFU methods of growth determination are significantly faster and less labor intensive, they are widely used in TB drug discovery efforts. In addition fluorescent and luminescent readouts from dye reduction or reporter gene expression offer increased sensitivity for growth detection and in the latter case, fewer problems with background signal from the test compounds. There are however, few direct comparisons of different readouts when testing compounds against *M. tuberculosis*.

Although most non-replicating methods use a CFU readout, they vary markedly with respect to the means of restricting bacterial growth due to the culture and environmental conditions used. Because the published methods also differ markedly with respect to culture medium and incubation conditions, no attempt was made to determine the impact of each variable. Instead methods were reproduced from the existing publications with the primary intent of determining if test compounds would be ranked similarly in the different methods or if particular compounds would only be active in specific assays.

Text box 1.

Lessons learned about *in vitro* assays from the workshop and visits to laboratories

- Absorbance, redox and intracellular ATP assays most commonly used for MIC determination
- BCG considered a useful surrogate for HTS
- Glycerol-dependent hits need to be recognized early
- Tween 80 can profoundly influence activity
- Relative value of individual NRP models uncertain
- Macrophage assay considered highly informative by some, of no value by others and a profiling assay but not on critical path by most

2.5. Head to head comparisons of the in vitro methodology

During visits to labs involved in TB drug discovery, we discussed numerous variables impacting primary screening, MICs, MBCs macrophage and non-replicating assays. These included media composition, formats, readouts, species/strains and means of achieving a non-replicating (NRP) state. We chose to investigate all of these except for macrophage assays and species/strains; the latter was omitted since there was a study in progress that was evaluating the major surrogates, *M. bovis* BCG and *M. smegmatis*.[64] In addition to other media components, we also examined the effect of serum and increased protein concentration since existing culture media do not use protein concentrations that reflect physiological concentrations. There is also a lack of information in the literature regarding high protein concentrations.

All evaluations of the culture media composition, readouts and protein/serum on MIC were performed using an 8 μm filtered, frozen, $H_{37}Rv$ ATCC 27294 inoculum at a final density of $1–3 \times 10^5$ CFU/ml in 96-well format. Microplate cultures were incubated without sealing at 37 °C for 1 week and signals acquired in a Victor-3 multilabel reader. MIC was defined as the lowest concentration effecting a reduction in signal of ≥90% relative to untreated controls as determined by interpolation from the dose–response curve.

Text box 2.

Key Definitions on *in vitro* assays

- Minimum inhibitory concentration (MIC): Minimum concentration effecting a reduction in growth at the end of the incubation period relative to untreated controls of 50, 90 or 99% and often indicated as an MIC_{50}, MIC_{90} or MIC_{99}, respectively. Readouts can be cfu on solid media or absorbance, fluorescence or luminescence of liquid cultures. The terminology of MIC_{50} and MIC_{90} is also routinely used in clinical microbiology to indicate the concentration of a compound effecting a defined percent inhibition of a percentage (50 or 90) of clinical isolates. Therefore in order to avoid confusion it is strongly recommended that the criteria (percent inhibition) for determining the MIC of samples against a particular strain be defined separately in a manuscript or presentation and then indicated as MIC (only) and not MIC_{50} or MIC_{90}.
- Minimum bactericidal concentration (MBC): The minimum concentration to reduce viability (as determined by cfu) by 90 (MBC_{90}) or 99% (MBC_{99}) relative to the starting inoculum.
- Protein or serum/plasma shift: The fold increase in MIC in the presence of a protein or serum/plasma concentration that exceeds that used in the basal testing medium. Since media commonly used for susceptibility testing of *M. tuberculosis* can contain from 0 to 0.5% w/v protein, descriptions of protein shift should indicate the protein concentrations and protein source of both the basal and supplemented media.
- Non-replicating *M. tuberculosis* (NR-TB) or non-replicating persistent *M. tuberculosis* (NRP-TB): With respect to *in vitro* models, this refers to culture conditions that preclude a net increase in cfu during incubation. Existing models utilize oxygen or nutrient deprivation, low pH, NO or a combination of low pH/NO to restrict growth.

Compounds were selected based on either clinical use in treatment of tuberculosis or experimental evidence of *in vitro* or *in vivo* activity. With a couple of exceptions, only one compound from a chemical class was included so as to maximize chemical (and presumably target) diversity. The compounds included isoniazid (INH), rifampin (RIF), ethambutol (EMB), moxifloxacin (MXF), streptomycin (SM), capreomycin (CAP), linezolid (LIZ), ethionamide (ETH), cycloserine (CS), PA-824, TMC-207, clofazimine (CLF), clarithromycin (CLA), nitrofurantoin (NFT), thiacetazone (TAC), pyrazinamide (PZA), ketoconazole (KETO) or econazole (ECO).

2.5.1. Media composition and readouts

The Middlebrook 7H series is widely used in clinical drug susceptibility testing (DST) and in drug discovery efforts. It is based on a common basal medium (7H9 broth base) with a variety of possibilities for supplementation as indicated in Table 1. The major variations in the Middlebook media series concern the major carbon source [typically either glycerol or oleic acid (or palmitic acid in the case of 7H12 medium)], and the presence or absence of non-ionic surfactants [such as Tween 80 (polysorbate 80), WR1339 or tyloxypol]. In addition, some variations include Casitone, a casein digest. Therefore, these variations were assessed for the effect on both signal:background and on compound MICs. Media were broadly classified as either 7H12-based (7H9 broth common components + 5.6 μg/ml palmitic acid as carbon source, Table 2: media 1–5) or 7H9-based (7H9 broth common components + glycerol and/or 50 μg/ml oleic acid as carbon source, Table 2, media 6–13). MICs were determined using the microplate Alamar Blue assay (MABA).

2.5.1.1. Effect of Middlebrook media components on signal:background (S:B) ratio in MABA. After one week incubation, fluorescence measurements of cultures containing only bacteria (without any added drugs) indicated that signal:background in general was similar in 7H12-based media and 7H9 media containing the commercial OADC supplement.[18–20] There was a modest reduction in media containing the ADC supplement (Table 2, media 7 and 12) and a very significant reduction in when Tween 80 was added to either 7H12 or 7H9-based media (Table 2, media 3 and 13, respectively).

2.5.1.2. Effect of media components on MICs in MABA (Table 3). Tween 80 was found to significantly potentiate the activity of RIF, confirming previous observations dating back to 1969.[72] This was observed in both Middlebrook-based media (Table 3) as well in protein-free media (Table 4). It was also the only example of a medium component decreasing the MIC of any drug. The same supplement was found to significantly increase the MIC of ETA and KETO, again in both Middlebrook-based and protein-free media. Less antagonism was (only) noted in the former media with TMC-207 and CLF. Middlebrook media incorporating the commercial

Table 2
Effect of Middlebrook media formulations on signal:background (S:B) in the microplate alamar blue assay (MABA).

Media	Middlebrook 7H9 broth + bovine serum albumin supplemented with							Signal:background[§]	
	Glycerol	Palmitic acid	Oleic acid	Dextrose	Catalase	Casitone	Tween 80	Mean	SD
1 (7H12)		x			x	X		18.06	3.15
2		x			x			16.20	3.50
3		x			x	x	x	3.72	0.58
4	x	x			x	x		18.69	2.66
5		x				x		19.20	2.34
6	x				x			16.06	4.80
7*	x			x	x			13.35	1.51
8†	x		x	x	x			19.38	3.26
9†	x		x	x	x	x		20.61	1.82
10†			x	x	x	x		19.01	3.20
11†			x	x	x			21.11	2.70
12*	x			x	x	x		14.07	1.59
13	x			x	x		x	6.61	1.65

* Commercial ADC supplement.
† Commercial OADC supplement.
§ From 2 to 3 independent experiments.

Table 3
MABA MIC of anti-TB agents in various formulations of Middlebrook media.

Media*	MIC (μM)															
	RIF	INH	ETH	PA	SM	MXF	TMC	CLF	CAP	NFT	EMB	CS	CLA	LIZ	ECO	KETO
1	0.06	0.24	2.35	0.12	0.37	0.22	0.06	0.13	0.93	43.75	6.81	23.67	3.41	0.96	11.39	7.60
2	0.05	0.25	0.86	0.23	0.22	0.23	0.06	0.13	0.95	25.90	6.84	16.62	*0.36*	0.83	10.78	7.70
3	0.02	0.48	**33.62**	0.39	0.28	0.22	0.19	0.46	0.97	35.80	3.92	24.02	1.59	1.19	22.65	**43.47**
4	0.06	0.24	1.97	0.06	0.32	0.21	0.12	0.23	1.64	42.02	3.62	22.96	2.13	0.90	11.35	10.13
5	0.04	0.24	2.32	0.12	0.43	0.23	0.09	0.23	0.78	37.37	6.12	23.86	1.30	0.87	11.14	7.59
6	0.07	0.29	1.73	0.14	0.24	0.23	0.12	0.28	0.54	26.60	4.08	22.10	*0.73*	0.89	11.71	10.87
7	0.10	**3.43**	5.04	0.24	0.67	0.23	0.11	0.16	1.79	27.27	9.40	45.62	*0.48*	0.81	9.81	18.14
8	0.06	0.36	4.71	0.16	0.44	0.24	0.11	0.23	1.38	36.05	8.19	22.88	*0.30*	0.88	13.16	24.14
9	0.06	0.27	5.06	0.12	0.48	0.23	0.07	0.12	1.76	44.07	13.40	24.26	*0.26*	0.78	10.83	24.06
10	0.06	0.44	4.69	0.21	0.50	0.24	0.08	0.18	0.96	48.88	14.96	24.76	*0.37*	0.81	11.73	23.25
11	0.04	0.38	4.73	0.21	0.34	0.23	0.08	0.19	0.96	34.59	7.42	22.29	*0.30*	0.78	10.06	23.35
12	0.07	**1.98**	3.75	0.12	0.49	0.22	0.09	0.17	1.01	33.24	7.53	46.56	12.41	0.65	7.88	**33.06**
13	*0.01*	**3.97**	**25.05**	0.30	0.91	0.25	**0.37**	**0.67**	1.97	73.86	15.70	74.96	9.56	1.65	23.69	**51.16**

In bold: ≥4-fold vs. 7H12 medium.
In italic: ≤4-fold vs. 7H12 medium.
* Medium compositions in Table 2.

ADC supplement had higher MICs for INH and CS than did 7H9 versions using either the commercial OADC supplement or 7H12 media (which contains albumin, catalase and palmitic acid). Dextrose appears to modestly increase MICs of ETH and SM in Middlebrook-based media. MICs of EMB were significantly increased in all protein-free media (Table 4) compared to the protein-containing Middlebrook-based media. Finally alanine, present in glycerol–alanine-salts medium, antagonized the activity of CS in the protein-free media, an effect described more than 40 years ago[73] that is consistent with the mechanism of action of this drug.

2.5.1.3. Readouts. Three readouts were evaluated for MIC testing in both 7H12 and 7H9 + OADC media: absorbance at 570 nm, fluorometric measurement of Alamar Blue reduction and luminescent measurement of intracellular ATP (Table 5). In 7H12 medium, MICs for CLA, INH, MXF and CS were >3–4 fold higher in the ATP assay vs. either MABA or A_{570} but the former assay returned a lower MIC for CLF. MICs for both RIF and CLA in 7H9+OADC were >4-fold higher when assessed by A_{570} vs. MABA or ATP. Overall the results suggest that the measurement of intracellular ATP is a reliable method for determining MICs across compound classes especially when using a medium such as 7H9+OADC. Recent results (data not shown) conducted in parallel with CFU analysis, confirmed excellent correlation between the two readouts.

Table 4
MIC of anti-TB agents in protein-free vs. protein-containing media.

Drug	MIC (μM)				
	GAST Fe*	GAST†	Sauton + Tw 80	Proskauer Beck 2	7H9 w glycerol + ADC
RIF	0.01	0.01	0.01	0.10	0.13
INH	0.12	0.24	0.19	> 8	3.63
ETH	18.50	17.75	69.45	1.84	4.84
PA-824	0.24	0.26	0.24	0.26	0.36
SM	1.61	1.10	0.30	1.99	0.78
MXF	0.24	0.35	0.33	0.97	0.33
TMC-207	0.12	0.12	0.14	0.11	0.16
CLF	0.10	0.15	0.24	0.18	0.22
CAP	3.18	3.31	0.72	3.49	1.00
NFT	37.39	37.64	69.39	68.07	23.22
EMB	27.44	29.91	31.07	> 32	14.50
CS	91.50	88.93	23.37	47.66	65.49
CLA	1.26	3.74	1.61	2.98	1.73
LIZ	1.55	1.91	2.91	2.44	0.97
ECO	10.34	12.11	11.21	14.77	10.96
KETO	27.54	35.50	36.69	12.20	12.72
TAC	14.27	94.69	28.42	4.08	191.38

* Glycerol-alanine-salts with Tween 80 and iron.
† Glycerol-alanine-salts with Tween 80.

2.5.2. Serum shift assays

Of those established drugs known to be significantly bound to serum proteins, the vast majority are bound by albumin with a smaller proportion bound by other proteins. It is, however, the on/off rates of protein binding that determine whether such binding will diminish drug availability and thus activity. Experiments in Middlebrook media confirmed the historical observations that some protein is important for growth of *M. tuberculosis* in media utilizing fatty acids as the carbon source. Human serum up to 50% and bovine and human serum albumin concentrations up to 3 and 4%, respectively, had no deleterious effect on growth (versus the 0.5% present in Middlebook-based medium) as determined by CFU (Figure 5). However, all did show a dose-dependent attenuation of fluorescent or luminescent signal in the MABA and ATP assays, respectively. At the highest concentrations (4% albumin or 50% serum), after 7 days of incubation this resulted in ATP signal:-background of approximately 150 and 51 with albumin (bovine or human) and serum, respectively. Using the MABA, these values were 3–4 and 2 for albumin and serum, respectively. Preliminary drug studies with 50% human serum and 5% human albumin in 7H12 and 7H9 + OADC using MABA showed an MIC shift of +3-fold

Table 5
Effect of readout on MIC of anti-microbial agents in Middlebrook media.

Drug	MIC (μM)					
	7H12			7H9 + OADC		
	A_{570}	MABA	ATP	A_{570}	MABA	ATP
RIF	0.10	0.10	0.09	0.25	0.03	0.05
INH	0.24	0.24	0.90	0.24	0.24	0.45
PA-824	0.04	0.07	0.09	0.31	0.09	0.21
SM	0.56	0.44	0.86	0.78	0.39	0.56
MXF	0.25	0.26	0.79	0.23	0.22	0.36
TMC-207	0.07	0.07	0.05	0.14	0.06	0.20
CLF	0.27	0.28	0.08	0.41	0.18	0.32
CAP	1.02	1.02	0.89	1.57	0.59	1.39
NFT	53	42	86	79.46	33.42	51.81
EMB	2.53	2.84	7.89	8.92	7.62	15.08
CS	24	24.17	85.25	44.17	21.61	23.71
CLA	8.36	6.41	44.87	19.82	0.27	3.48
LIZ	0.90	0.73	0.93	0.96	0.64	0.97
KETO	6.20	5.97	9.00	26.10	12.00	22.29

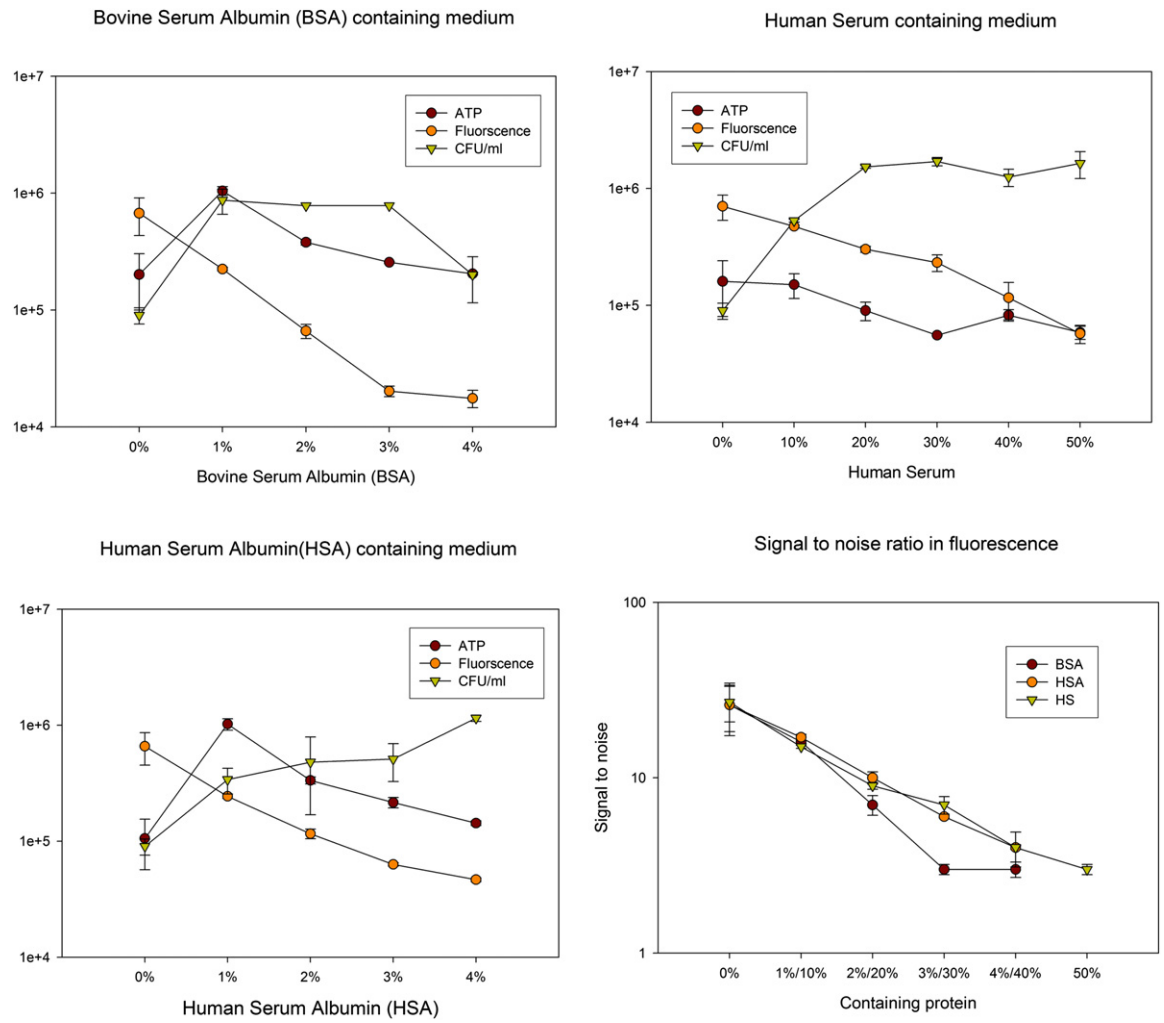

Figure 5. Effect of serum and albumin concentration on *M. tuberculosis* viability (CFU), Alamar Blue-mediated fluorescence and intracellular ATP-dependent luminescence.

for RIF with both protein sources (Table 6). In contrast, PA-824 had a +7-fold shift with serum and +3-fold shift with albumin. SM and EMB had +2–3-fold shifts in serum but no shift with albumin. Isoniazid and MXF showed no MIC shifts. There were no differences between the two Middlebrook media versions. In general the results are consistent with higher protein binding of hydrophobic compounds. Although serum effects a larger signal attenuation then albumin when using metabolic readouts, it contains proteins other than albumin such as alpha 1-acid glycoprotein, lipopeptides and globulins, which may be important in protein binding or metabolism of test compounds. The ATP readout is sufficiently strong to minimize the impact on signal:background and the concurrent evaluation of both albumin and serum can provide additional information regarding the nature of any observed reduction in activity.

2.5.3. Non-replicating persistence assays

We compared head-to-head, selected assays utilizing hypoxia (in both tube[74] and microplate[75] versions), starvation,[76] stationary phase[77] and low pH/NO release.[78] The methods used were taken

Table 6
Effect of serum and albumin on MABA MIC of anti-TB agents in Middlebrook media.

Drug	MIC (μM) using the following media:							
	7H12 media containing				7H9 media + OADC containing			
	Human serum		Human albumin		Human serum		Human albumin	
	0.5%	50%	0.5%	4%	0.5%	50%	0.5%	4%
RIF	0.11	0.35	0.05	0.20	0.06	0.21	0.03	0.13
INH	0.47	0.28	0.24	0.42	0.42	0.24	0.45	0.43
SM	0.50	1.53	0.38	0.50	0.52	1.50	0.69	0.71
MXF	0.41	0.24	0.39	0.46	0.37	0.24	0.37	0.42
PA-824	0.12	0.85	0.20	0.75	0.21	1.39	0.25	0.77
EMB	7.79	14.80	5.05	6.13	7.65	21.36	12.79	7.85
None, RFU*	144,882	76,441	382,737	22,263	513,767	88,521	224,325	19,167

* Net relative fluorescent units.

directly from the published reports (and in the case of the low pH/NO method, additional personal communications) and used to compare several assays for the *in vitro* evaluation of 12 antimicrobial agents against slowly or non-replicating *M. tuberculosis*. With the exception of the use of *M. tuberculosis* $H_{37}Rv$, there was no attempt to control any single variable. All assays were evaluated (only) by CFU and all drugs were tested at 1, 10 and 50 μg/ml (Table 7).

In nearly all cases, dose responses were observed in all assays. Of the few exceptions, two were observed with ethambutol (low pH/NO and UIC low oxygen). Greater than 90% inhibition at the highest concentration of 50 μg/ml was observed for 8–9 of the drugs by the low oxygen assays, for 5 drugs by the low pH/NO and stationary phase assays and for 4 drugs by the starvation assay. The two low oxygen assays returned very similar results. At lower concentrations the CSU assay appeared slightly more sensitive to RIF and MXF while the UIC assay appeared more sensitive to TMC-207, CLF and CAP and at the highest concentration, MET as well. PZA failed to effect a 90% reduction in viability in any of the assays with strongest activity being 72% inhibition at 50 μg/ml in the low pH/NO assay. Considering the pH-dependence of this drug and the relative low concentrations used in this study, perhaps this is not surprising. Likewise EMB, as expected, failed to effect a 90% reduction in any assay at any concentration. INH, also considered to be an inactive "control" only demonstrated >90% inhibition in the low oxygen assays and then only at 50 μg/ml. RIF, PA-824 and MXF were active (>89–90% reduction in viability) in all assays at 10 μg/ml. CLF demonstrated activity in all assays at some concentration with the exception of the starvation assay. Assays that appeared to be uniquely sensitive to specific drugs included the low oxygen assays (especially the UIC version) for capreomycin and the low pH/NO (and to a lesser extent, the starvation assay) for NTZ.

When comparing the results obtained here in a single lab vs. published reports, in general we observed higher levels of activity for some drugs in some assays. For example, although our data were for the most part consistent with those previously reported by Gergenbacher et al. for the starvation assay,[76] we did see higher activity with TMC-207. Likewise we observed activity at lower concentrations than those reported by Lenaerts et al. for RIF, INH, MXF and PA-824[74] although in both cases MET appeared to be devoid of potent activity even at 50 μg/ml. Conversely we observed less activity for RIF and INH in the stationary phase assay reported by Sulochana et al.[77]

Table 7
Activity of antimicrobial agents against slowly or non-replicating *M. tuberculosis*.

>90% inhibition* @	Starvation[76]	Stationary[77]	Low pH/NO[78]	Low O_2 UIC[75]	Low O_2 CSU[160]
1 μg/ml		MXF		PA-824 CLF RIF SM TMC-207 CAP	RIF MXF
10 μg/ml	PA-824 RIF	RIF PA-824	RIF PA-824 NTZ CLF MXF	MXF	TMC-207 SM CLF PA-824
50 μg/ml	TMC-207 NTZ	SM CLF		MET INH	CAP INH
>50 μg/ml	MXF† CLF SM CAP EMB INH PZA MET	TMC-207 CAP NTZ INH EMB MET PZA	INH TMC-207 PZA CAP SM MET EMB	EMB NTZ PZA	NTZ EMB MET PZA

* All readouts are CFU.
† 87–89% inhibition at 50 and 10 μg/ml respectively.

2.6. Future directions

The comparisons done here are restricted by the limited chemical diversity and number of corresponding molecular targets. As additional compounds with confirmed mechanisms become known, these should be evaluated in a similar manner. In some cases, HTS with diverse chemical libraries would be very valuable although the number of variables evaluated would need to be limited. Newer reporters such as the mCherry fluorescent protein and *luxABCDE* construct may be superior to those constructs used in *M. tuberculosis* reporter strains to date. The *luxABCDE* codes for both luciferase and the n-decanal substrate and so, like the fluorescent reporters, does not require the addition of a substrate and enables the observation of the kinetics of growth/inhibition.[79] Preliminary data obtained in our lab with the *luxABCDE* construct (provided by Siouxsie Wiles, Imperial College, London) corroborates the utility in determining MICs and we have recently used this construct to screen a 100,000 compound library. The utility of this reporter however extends beyond MIC determination to include MBCs and activity in macrophage culture[79] and we are currently evaluating its use in the low oxygen recovery assay.[75] Although the MIC protein-shift models described above can be used with data normalization, additional rapid readouts (e.g. fluorescent protein reporters, turbidity, *luxABCDE* reporter) or means of neutralizing the protein effects should be assessed to lessen or eliminate signal attenuation. At least two additional NRP models, one utilizing a streptomycin-dependent strain[67] and another resulting in formation of a biofilm[66] have been recently reported and it would be of value to evaluate these in head-to-head comparisons as well. The achievement of an NRP state certainly is not unique to the $H_{37}Rv$ strain, nor to the *M. tuberculosis* species. The recent demonstration of *M. bovis* BCG having reasonably good concordance with *M. tuberculosis* with respect to the sensitivity of replicating cultures to diverse compounds,[64] suggests that a direct comparison of the sensitivity of NRP cultures of these two species should be conducted.

Determination of synergy or antagonism, although of obvious interest for a disease that requires treatment with multiple drugs, was not assessed in this study largely because there is little evidence that existing methods are predictive of these effects *in vivo*. Such determinations are most commonly performed via checkerboards and calculation of a fractional inhibitory concentration (FIC) index. The detection of synergy using this method can only be done as sub-MIC concentrations. Alternatively these interactions can be determined using time-kill curves where synergy can be detected at concentrations greater than the MIC which may be more physiologically relevant. The disadvantage is the requirement of a CFU readout. The sensitivity and dynamic range afforded by intracellular ATP measurements or use of the *luxABCDE* reporter could make this a viable alternative readout for kill curves.

2.7. Final recommendations

During initial screening, especially HTS, it is important to maximize sensitivity as it is unlikely that these compounds will ever be re-tested. From this perspective it is preferable to

Text box 3.

Lessons learned from the *in vitro* laboratory experiments

- Tween 80 can both potentiate and attenuate compound activity
- Use of the OADC supplement in Middlebrook media vs. the ADC supplement yields higher signal:background ratios in MABA and lower MICs for some drugs.
- MICs of particular compounds can vary with different media/readout combinations.
- 50% serum or 3–4% serum albumin can be used in MIC protein shift evaluations without compromising viability/growth but with a dose-dependent inhibition of fluorescent or luminescent readouts; an ATP readout maintains a better signal:background ratio.
- The low oxygen assays demonstrate sensitivity to a greater number of existing/experimental TB drugs than do other non-replicating/slowly replicating assays. While RIF, PA-824 and MXF appear highly active in all tested assays, the relative activity of a few drugs appears to be assay-specific.

encounter false positives versus false negatives as long as the former can be rapidly and efficiently identified in a counter-screen. Overall, it is recommended to use a medium devoid of glycerol and Tween 80 and employing a fatty acid as the carbon source. Simply using Middlebrook 7H9 + OADC without any additional supplements will give a good signal:background in growth assays. It is likely (but not conclusively determined here) that the dextrose and catalase components are inconsequential and could be eliminated if the supplement is being prepared from individual components (vs. purchasing commercial products). Either intracellular ATP or redox-based assays are recommended. Cost concerns aside, the intracellular ATP assay might be optimal for HTS due to the low probability of background luminescence with subsequent confirmation by a redox dye method such as MABA or REMA. Our experimental results suggest that using absorbance measurements alone should be done with caution due to a very low signal:background, however most MICs obtained with this method did correspond with those obtained via MABA.

Protein MIC shift experiments can be conducted with confidence at 50% serum or 3–4% albumin, however data need to be normalized if using fluorescent or luminescent readouts. Intracellular ATP maintains a good signal:background while that for the MABA (and possibly other redox-based readouts) is significantly reduced. Both serum and albumin can be used concurrently to differentiate albumin-specific effects from those possibly caused by other serum constituents. It should be noted that larger protein shifts could likely be noted if beginning with protein-free media (versus Middlebrook media with 0.5% BSA). Therefore, such protein MIC shifts should be interpreted in a relative manner with respect to a series of control compounds.

Based on the limited information available at the time of this writing, it is difficult to make specific recommendations regarding NRP assays. We can only encourage research groups to run multiple assays on the largest number of compounds, especially those with known mechanisms and to report this data so as to build up a cumulative body of knowledge. As more of these compounds reach animal models and clinical trials and are assessed for the ability to reduce duration of treatment, we will slowly gain an understanding of the relevance of such *in vitro* models.

3. *In vivo* evaluations of TB drugs

Methodologies for preclinical testing of drugs against *M. tuberculosis* in animal models do vary significantly from laboratory to laboratory; however, it is unknown whether the heterogeneity in the experimental parameters of these methods affects the results of these studies, or can be considered an advantage. Would two different laboratories studying the same agent or drug regimen arrive at different conclusions regarding efficacy? For instance, can the mouse strain or strain of TB determine the outcome of drug efficacy trials? These questions are of great importance to rate the ability of the different TB infection models to rank order new regimens and compare them to the standard regimen for drug susceptible TB. To begin the process of understanding the state of affairs in a variety of academic and industry laboratories, we began a multi-pronged approach to understand the various protocols currently used in the field. This included personal visits to laboratories, an electronic survey, a multi-national workshop and a series of head-to-head animal model experiments in one laboratory.

The historical and contemporary methodology was identified from the literature. Protocols and philosophies from top universities, companies and institutions engaged in TB drug discovery and development were collected via an electronic survey and personal visits. The survey was designed and conducted to elicit annotated, detailed information on protocols which is generally not present in publications. We included 95 questions with one-half appropriated to *in vitro* screening and the other half to *in vivo* assays. Personal visits were conducted with many TB drug development representatives to discuss screening strategies, issues and recurring problems as well as receiving recent updates on *in vivo* projects. The visits, survey and literature searches then allowed us to prioritize the questions raised regarding animal models in need of re-evaluation in head-to-head experiments in one laboratory. In the head-to-head animal model experiments, we compared carefully varying key assay variables (such as strain of mouse, route of infection, timing and formulation of drugs) in order to ascertain the relative contribution to outcome of the animal trial. For this Drug Accelerator grant effort by the Gates Foundation, we were instructed to only focus on mouse models currently used in the field. The three most commonly used mouse models by experts in the field were at the start of the project; the intravenous injection model, a low dose and a high dose aerosol infection model. We compared results in two of the most commonly used strains of mice. This is the first report on side-by-side comparisons of assays and animal models used in the field.

3.1. Historic overview of animal models

3.1.1. Early animal models

In the development of new drugs and vaccines for *M. tuberculosis*, animal models have always been a crucial component.[6,80,81] In many of the earlier studies, larger animal species such as guinea pigs and rabbits were primarily chosen, and often studies with high animal numbers were employed. Since the early 1900s, animal studies for drug evaluations provided information on histopathology, bacterial pathogenesis as well as penetration and effect of compounds such as thiocyanate and gold.[82,83] While insight was gained mostly on the biology of the disease, most drug therapies being evaluated at the time were usually disappointingly ineffective against the infection.[84] Many historic studies made remarkable contributions by thoroughly detailing histopathology of *M. tuberculosis* infected animals.[85,86] The progression of pathology after a TB infection in mice was beautifully described and illustrated by Youmans in 1948.[87–89] The authors mostly presented this data, consisting of the gross lung lesions (showing

photographs) and survival time, as a metric to assess drug efficacy. The 1930s through the 1950s brought more promise of new agents for the treatment of tuberculosis. It was only in later studies beginning in the late 1940s and 1960s that the emphasis shifted from the larger animal models for TB drug evaluations to using the mouse model for tuberculosis.[87,90–94]

At times the instrumentation for performing *in vivo* inoculations with *M. tuberculosis* was homemade, but the infection of the animal could be observed (by gross pathology and changes in body weight) and showed to be reproducible. Often, the infection procedure was conducted under a laminar flow cabinet with a strong updraft, which represents an early environmental control system for biosafety purposes. The optimization of culture and propagation conditions of *M. tuberculosis* on solid agar media[95,96] allowed for the identification of colony morphology, bacillary enumeration and drug resistance evaluation, and therefore made it possible to use well-defined inocula for infection of the animal models. Animals were infected via various routes of inoculation such as subcutaneous, intracutaneous, intravenous injections (IV), orally, intraperitoneally, eye inoculation, intracerebrally and aerogenically (by aerosol, intranasal or intratracheal routes)[81,90–93,97–105] The IV route proved practically difficult in guinea pigs given the paucity of superficial vessels. This method was easier in the rabbit using the marginal ear vein, and was often performed in mice via the lateral tail vein.[92,93] The IV route allowed for a uniform deposition of the TB bacilli in the lung via the hematogenous route.

An important era ensued when a careful evaluation and cataloging of *M. tuberculosis* strains was started around 1945 after a meeting of the American Thoracic Society and the National Institutes of Health (NIAID).[106,107] For instance, a collection of mycobacterial strains was established at the Trudeau sanatorium in Saranac Lake NY in 1946, under the guidance of Dr. William Steenken.[108] For each bacterial isolate in the collection, the species name and strain was recorded; its history and source; date of deposition at Trudeau; its virulence patterns in mice, guinea pigs, and rabbits if available. In addition, a series of *in vitro* tests were conducted to determine growth characteristics, biochemical profile and drug susceptibility, and the data were then compared to the properties of a 'hypothetical median strain of TB'.[107–109] Another important advance in the 1950s was the development of *in vitro* culture methods that allowed enumeration of mycobacteria from organ homogenates as a measure for drug efficacy from animal infection studies.[92,93] However, it was the discovery of streptomycin (SM) and its subsequent successful testing in Guinea pigs that represented a major landmark in the animal modeling field.[110]

3.1.2. The guinea pig model

The first use of Guinea pigs in tuberculosis research was for TB diagnostic studies that involved the inoculation of animals with human sputum. Guinea pigs were the main animal studied by Koch, when he first saw and described the "Koch phenomenon" [caused by a delayed-type hypersensitivity response] similar to that observed in humans.[111–113] The descriptive pathology was often illustrated with highly detailed and annotated drawings of microscopic images in the early days of the field, including by Koch himself.[114] Due to its high susceptibility to intravenous and aerosol routes of infection with the tubercle bacillus, the Guinea pig was the favored animal model for many years especially by the Mayo Clinic investigators.[115–118] One of the first reports of an *in vivo* active compound was by Rich and Follis in the late 1930s, who tested sulphanilamide in Guinea pigs.[100] Other early studies included testing of SM as monotherapy and in combinations with drugs such as para-amino salycilic acid (PAS).[119] In these early studies, measures for drug efficacy other than enumeration of bacterial load in lungs were adopted, such as weight or radiography.[20,21,120,121] In addition, survival, changes in animal weight and gross necropsy observations were used as crude measures of drug efficacy.[122,123] Another quantitative measure was developed by Mitchison, and called the "Root Mean Virulence assay for Guinea pigs" [calculated as the square root of the pathology score, divided by the survival time in days] which related lung pathology to animal survival times.[124–127] Guinea pigs were often inoculated subcutaneously and at times developed lesions at the site of the inoculation. Infection of the Guinea pigs was occasionally checked by the tuberculin sensitivity skin test to assess adequacy of infection and this was also used as a marker of drug efficacy.[128]

In the 1940s, more antimicrobial agents started to become available that deterred the disease in guinea pigs (e.g. the sulphanilamides). Due to this expansion of available compounds, a careful assessment of the methodology was put forward around that time. Variables included the animal species used, short term versus long term treatment experiments, recording of necropsy outcomes, the number of animals needed based on the desired study, and one started to put forth criteria for judging an effective agent. Feldman was the first to establish a document in 1945 detailing a suggested outline of laboratory procedures for testing antituberculous substances in experimentally infected animals.[98] In 1944 Salman Waksman (who together with Albert Schatz discovered SM[122,129,130]), sent SM to Feldman, who then used the drug to treat guinea pigs after they had been infected with the TB strain H_{37}Rv.[122] The drug worked spectacularly well and reduced mortality and the number of lesions. An important outcome of those studies was the development of the first systematic scoring system based on gross pathology in guinea pig organs (Figure 6).[115–117,119] Feldman established some basic guidelines for the evaluation of chemotherapeutic agents, which included: ensuring the virulence of the TB strain used, starting treatment when the infection is well established, recording survival time and showing evidence of histological improvement with therapy, continuing to observe the animals post therapy and monitoring for the emergence of drug resistance.[86] Also in 1944, a Frenchman, Chorine, who had studied leprosy in rats, injected the *M. bovis* strain Ravenel into rats and treated the infected rats with nicotinamide (Vitamin B).[131,132] Based on his positive results in rats, the same approach was tried in guinea pigs infected via a subcutaneous injection with *M. tuberculosis*, and noted a reduction in dissemination of the disease with derivatives of Vitamin B and this was seen by others as well.[25,133–135]

Many compounds tested, such as PAS, streptovaricin (SV) and capreomycin (CM), demonstrated activity with a clear dose-dependent response in Guinea pigs (as well as in mice), which was reflected in improvement of gross lesion involvement of lungs, livers, spleens and at the site of inoculation.[118,119,136] The readout method which combined gross lung lesion observation and survival time proved useful to compare relative efficacy among single drugs and drug combinations. Such gross necropsy endpoints were replaced with a quantitative evaluation of the reduction in bacterial load and the use of a microbial enumeration technique only became a more common practice in the 1950s–1960s in mouse models.[90,91]

Although the investigator Gardner Middlebrook is best remembered for assembling the composition of 7H9 and 7H10 media for mycobacterial growth,[23,101,137] he also performed several classical guinea pig studies optimizing several delivery methods of infection in order to investigate TB pathogenesis, drug evaluations, as well as skin testing. In addition, he made an important contribution to the animal model field by designing the first, of what ultimately became a widely used, aerosol infection apparatus.[138,139] In the 1950s and 1960s, the mouse model for tuberculosis became more common likely due to practical and economic considerations.

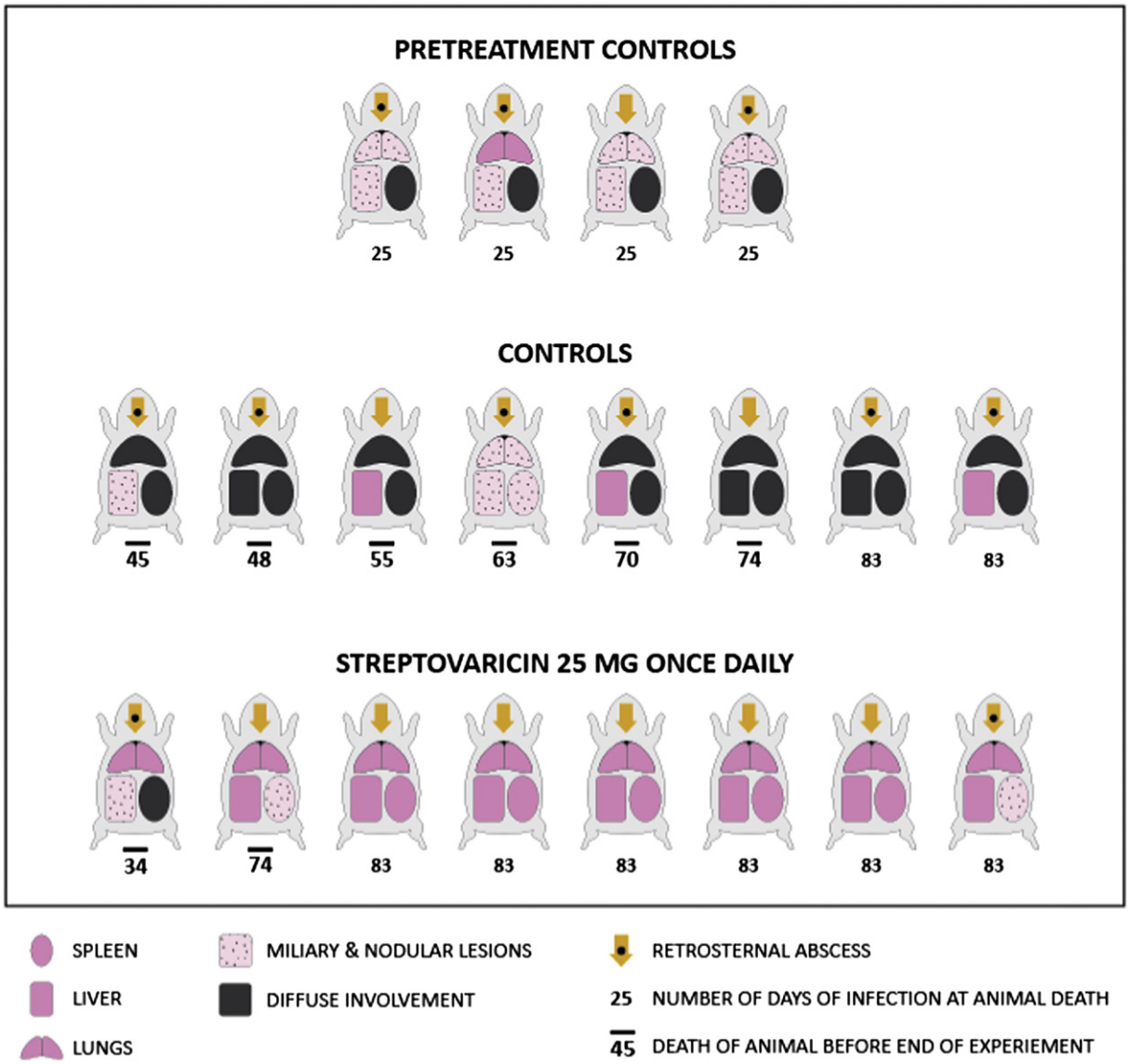

Figure 6. Schematic of disease involvement in lungs, site of injection, spleen liver of guinea pigs infected by injection of *M. tuberculosis* in suprasternal notch and treated with streptovaricin (SV) (lower panel) or untreated (upper panel). Diagram of gross pathologic findings and date of death observed in guinea pigs treated with anti-tuberculosis test agent SV. The black dots indicate lesions visible on necropsy and the number under each animal was the number of days since infection at animal death. The solid bar under the number indicates the number of days the animal survived. In liver, spleen, lungs the degree of involvement is indicated by the number and intensity of lesions. Death of the animal is indicated by a black horizontal bar and the day below (illustration reproduced from the original reference[119]).

3.1.3. The mouse model

Mice were first used in the TB field mainly to study the host immune defense of a tuberculosis infection and to understand the possible role of the reticulo-endothelial system in resistance to a TB (or *M. bovis*) infection. In the 1940s, a more thorough analysis was reported including the details of the mouse infection protocols, on the susceptibility of mice to infection[87–89] and important differences in pathology between mice, guinea pigs and humans were described.[140–142] Investigators like Dickinson, Mitchison and Grosset and their colleagues did much of the pioneering work in TB drug testing using the mouse model.[143–145] Mouse studies in the 1950s and 1960s were mostly employed to test the drug effectiveness by means of both lengthening of survival time and improvement of histopathology, which then resulted in the ability to rank active compounds.[145] In most studies, the drugs were administered to mice by administration in the diet or by injection at a fixed, high dosage. Often little pharmacokinetic information was available to allow for analysis of proper drug exposure. A major contribution in the mouse model field was made by Mitchison and his colleagues, who demonstrated in the 1970s that mouse data evaluating drug regimens based on isoniazid (H, INH), rifampin (R, RIF) and pyrazinamide (Z, PZA) were very consistent in predicting the human clinical trial results of the British Council [using the outbred ICR mouse IV model[146]]. In the 1950–60s, Canetti, Grumbach and Grosset carried out numerous large scale mouse experiments (often containing over 100 animals infected intravenously with high inocula of *M. tuberculosis* $H_{37}Rv$) which helped understand the role of anti-tuberculosis agents in the course of successful therapy and molded modern day chemotherapy.[143,147,148] Most early studies were performed with outbred mice such as ICR, albino or white mice and the use of inbred strains of mice in TB drug development is relatively new (during the last decades).

Another era started with addressing the issue of relapse of infection after termination of drug treatment with the important

studies by McCune, McDermott and Thompsett at Cornell University in the 1950s. McCune et al. published a key series of papers detailing the microbial enumeration technique, and a well-documented analysis of microbial persistence after drug treatment.[92,93,149,150] McCune performed many long term drug treatment trials using large numbers [hundreds] of mice as well as guinea pigs which were pivotal to form the understanding of the 'non-culturable' TB bacillary population after drug treatment that can reactivate after termination of drug treatment. It became an appreciated fact that despite persisting in the face of drug therapy, these bacilli were susceptible to antimicrobials when tested *in vitro*.[92] Relapse of infection in the animals was either spontaneous after completion of the drug regimen (usually 90 days) or by immunosuppression of the host. If no re-growth of bacteria was observed in lungs and spleens upon immunosuppression, it was concluded that the drug regimen was truly "sterilizing". This mouse model is mostly referred to as the 'Cornell model', named after the institute where McCune made his discoveries. Many of the methods described in his comprehensive publications are still in use today, such as the intravenous challenge, macroscopic lesion analysis scoring, histopathologic assessments, and propagation of *M. tuberculosis* through mouse passage to maintain virulence.

The predictive value of mouse models for clinical trials has been studied but also questioned throughout history. Most experimental infection models of tuberculosis in mice have been of great help in the development of chemotherapy for human infection. The long-term mouse model, developed first at Cornell University, then at the Pasteur Institute in Paris, led directly in 1970 to the testing of the drug combination of RIF and PZA in one of the most important of the British Medical Research Council randomized clinical trials.[151] In addition, the mouse model was also highly predictive in showing the poor sterilizing activity of INH and SM.[143,147]

Based on an editorial by Dr. D. Mitchison,[152] there also have been occasions where mouse studies might have misdirected human clinical trials, such as in the case of (1) substituting moxifloxacin for isoniazid in the standard drug regimen and (2) the use of PZA and RIF for treatment for a latent TB infection (LTBI). The mouse studies that may have misdirected human clinical trials arose from experiments comparing the substitution of moxifloxacin for the main drugs (INH, EMB) in the standard TB regimen.[153,154] The moxifloxacin-containing regimens were found to be superior in mice when substituting MXF for INH. In the particular mouse model used, the removal of INH from INH-RIF-PZA resulted in a more effective two drug combination of RIF-PZA suggesting an antagonism between INH and the combination of RIF-PZA in the mouse model used.[155] Therefore the superior efficacy seen in the moxifloxacin-containing regimens in the mouse model used was mainly caused by removal of INH and not by adding moxifloxacin as was also later shown in mice and published.[156] There was also suggestive evidence at that time that showed the finding of antagonism to be due to incomplete absorption of drugs in the mouse in combination treatment. These studies in mice indicated that RIF serum concentrations were lower when RIF was given with INH and PZA than with PZA alone.[155] We will show in a later section that the antagonism between INH, RIF and PZA is in fact not a general phenomenon and only seen with one specific *M. tuberculosis* strain. Antagonism in the standard TB drug regimen was never observed in any of our mouse model experiments (see Section 2.7: Head to head comparisons of methodologies).

Although the editorial by Mitchison[152] mentions that the mouse model findings have led perhaps to a premature use of the RIF and PZA combination in trials of preventive therapy, this statement is only partially correct and requires some explanation. RZ did not fail as an LTBI regimen due to a lack of efficacy. The combination is highly efficacious in patients but failed because of toxicity issues.[157,158] The mouse models have yielded good results in ranking LTBI regimens on the basis of efficacy and relative duration of treatment needed to cure LTBI,[159] although clinical successes for LTBI are difficult to determine. The only way the mouse was poor in predicting the clinical outcome for RIF-PZA in LTBI is that it did not predict in early mouse models the severe hepatotoxicity observed in latently infected people. These reports of fatal and severe liver injury associated with RIF-PZA treatment regimens for LTBI prompted a later study of liver enzyme levels in mouse models after high drug exposures.[160] After high doses of RIF and PZA, a highly significant increase of three liver enzymes was also observed in uninfected and *M. tuberculosis* infected mice.[160] The mouse model might have potential to assess hepatoxicity following extensive exposure to tuberculosis drugs, however, this still would need further evaluation using more comparative TB drug regimens.

A few important papers were recently published addressing the importance of including relapse studies in TB drug development[161–164] and on genetic variability in the *M. tuberculosis* $H_{37}Rv$ strain,[165] urging investigators to be careful in preclinical studies, and in translating animal model data to humans. A recent paper showed that bactericidal activity of TMC-207-containing drug regimens did not necessarily predict sterilizing activity[166] (Table 8), indicating that such relapse studies are critical to assess the sterilizing properties of drug regimens prior to going to clinical trials. Both of these topics are highly relevant and will therefore be addressed in more detail later (Section 2.6: Lessons learned and future directions). One important caveat that cannot be modeled in the classical mouse is the advanced pathology features seen in TB patients.

The recent observation that C3HeB/FeJ mice develop necrotic granulomas in response to *M. tuberculosis* infection has raised hopes that this strain may offer more human-like pathology.[167–169] The development of necrotic granulomas is determined by a susceptibility allele in the *sst1* locus which leads infected macrophages down a necrotic, as opposed to apoptotic, cell death pathway.[170] The lesions exhibit central necrosis and hypoxia characteristic of other larger species, including humans.[168] Although there has been little work characterizing the response to antibiotics in this species, several significant efforts are currently underway.[168,169,171]

In summary, while keeping in mind that they are only "models", mice have been highly instrumental in guiding TB drug development and represent the only animal model that has been validated to date by testing regimens and treatment durations known to work in patients.

Table 8
Ranking of TB drug regimens based on relative activity during the 'bactericidal phase of infection' versus the relative activity during the 'relapse of infection phase'. Rank order of various TB treatment regimens determined by $\log_{10}$ CFU in lungs of mice. Relapse is defined as the number of mice with positive organ cultures three months after discontinuation of therapy. R = rifampin, J = TMC-207, P = rifapentine, M = MXF, Z = pyrazinamide, H = isoniazid.

Regimen	Rank order bactericidal activity	Rank order relapse
JZP	1	1
JZPM	1	1
J1/2ZP	1	2
PZM	2	3
JZM	1	4
HRZ	3	5

Data comparison from Andries et al.[166]

3.1.4. Other animal models

Other animal species studied for tuberculosis infections in earlier studies by different investigators often included hamsters, rabbits, rats, voles, chick embryos, and more.[100,172] Despite the track record of the mouse model, the lack of pathological hallmarks of human TB such as caseating granulomas, cavitary disease, and fibrosis is a known limitation of this model.[6,81] Unlike the mouse, other larger animal models such as guinea pigs, rats, rabbits and non-human primates develop necrosis, fibrosis, mineralization and, in the case of rabbits and non-human primates, cavitary disease.[173] Bacterial persistence has been hypothesized to result, at least in part, from environmental conditions such as hypoxia and nutrient deprivation, as well as reduced penetration or diffusion of drugs into lesions. The mouse model may not entirely represent these specific environments which might be required to assess the sterilizing activity of a new TB drug or combination.[6,81] In addition, the most widely used mouse strains only develop inflammatory, but non-necrotic, lesion types in the lungs where the bacilli reside primarily intracellularly.[174] This is not what is observed in progressive disease in other animal models,[174] nor in advanced human TB in which the vast majority of bacilli are extracellular inside necrotic lesions and the liquefactive core of cavities.[175]

These limitations of the mouse model are largely theoretical at this point, as bacterial persistence is in fact a phenomenon also observed in mice. Drug treatment in mice is highly effective in the first weeks of treatment but is far less effective to eradicate the last 1% of bacilli. This results in a biphasic killing curve, and ultimately it takes just as long to cure a mouse infected with TB as it takes to cure a TB patient.[176] No other animal model has been validated as thoroughly as the mouse up to date, or has demonstrated to give similar or a better prediction of drug regimens being evaluated for clinical use. More comparative studies using various animal models are urgently needed to evaluate whether they show sufficient superior predictive value to justify the excess cost of using the larger animal models. The availability of several promising new drugs and regimens currently in clinical trials is the ideal time for these comparative studies. In the section below, a selection of different animal models currently being studied by investigators are listed with their strengths and weaknesses, as well as their potential promise in future trials evaluating new regimens. More recently, zebra fish,[177] minipigs,[178] the Kramnik mouse model[169,170] non-human primates and macaques[173,179] have entered the field of animal models for tuberculosis.

Rabbits have been used for testing drugs, but the inoculation methods used in the past (e.g. including injection into the eye) would likely not be permitted today. Lurie extensively studied the rabbit model and used inbred strains of New Zealand White (NZW) rabbits with different susceptibilities to virulent *M. tuberculosis* ($H_{37}Rv$ or the bovine Ravenel strain).[180] These particular NZW rabbit strains became extinct, but inbred rabbit strains which are more susceptible compared to outbred animals have been raised again and show lung pathology closely resembling that seen in humans.[181] Rabbits are highly susceptible to virulent strains of *M. bovis* (e.g., the Ravenel strain), rapidly developing caseation necrosis and pulmonary cavities and dying within 5–6 weeks of infection. In contrast, infection with laboratory strains of *M. tuberculosis*, such as $H_{37}Rv$ and Erdman, does not result in a productive infection, giving rise to only 1 granuloma per 50-200 bacilli, depending on the strain used.[182–184] More recently, however, the more virulent *M. tuberculosis* HN878 strain has been shown to cause a productive, robust infection reaching more than 6 $\log_{10}$ CFU in the lungs after a nose-only inoculation with 1000 CFU.[185] Bronchoscopic instillation or transthoracic injection of *M. tuberculosis* into previously sensitized rabbits may promote more rapid and consistent cavitation.[186,187] Despite the characteristic pathology, the high expense of maintaining rabbits is an obstacle to their use. Additional drawbacks include pathogen containment concerns, since rabbits shed *M. tuberculosis* in the urine, and intercurrent infections such as with *Pasteurella* and *Bordetella bronchiseptica* [the latter also seen in guinea pigs]. However, rabbits show to be well suited for evaluating drug distribution, penetration and cellular accumulation in different lesion types, using high sensitivity MALDI-MRM-MS technology as was recently described.[188–190]

Rats are primarily used in drug development for pharmacokinetic testing of new compounds but have also been used as well in the past for therapeutic trials for drug testing against *M. tuberculosis*. Studies in the 1920s with “albino rats” found that far higher inocula of TB were needed to kill rats compared to guinea pigs. As a result, rats were considered resistant to a tuberculosis infection,[98] and by implication, not suitable for research. Most recently, NITD has developed a Wistar rat model in which rats are infected via the endotracheal route. Bacterial multiplication follows for the next 2–3 weeks, but the onset of adaptive immunity is associated with progressive clearance of *M. tuberculosis* $H_{37}Rv$.[191] This spontaneous clearance of infection makes evaluation of regimens for active TB challenging but offers prospects for a latent TB infection model. The cotton rat has resurfaced as a model for TB drug testing,[192] as well as other rat strains.[193–195]

Non-human primates are much more closely aligned to humans and have obvious advantages in mimicking the pathology of a human infection.[196] At present, the non-human primate model can most closely represent a naturally occurring latent TB infection. Macaques are highly susceptible to *M. tuberculosis* and can also be infected with SIV, which makes them an attractive model for TB-HIV co-infection studies.[197] Although there are multiple difficulties associated with the use of non-human primates [low availability, cost, ethical concerns], the model has been used for TB drug testing and has renewed interest given the similar pathology to human infection.[105,179,198] Cynomolgus macaques infected with *M. tuberculosis* via bronchoscopic instillation of few bacilli can be followed by tuberculin skin testing and serial clinical, laboratory and radiographic examinations to identify those developing active TB or latent infection (which occur in roughly equal proportions). To date, however, these endpoints have not been shown to be useful in discriminating between highly active drug treatment regimens with differing potency. An evaluation of the activity of metronidazole, alone and in combination with first-line drugs is ongoing. Marmosets develop similar pathology as macaques in response to *M. tuberculosis* infection, but are smaller in size and therefore more economical. As of late, the interest for the marmoset model has increased for their use in transmission experiments, as they have the advantage of frequent twinning which provides ideal control groups. Although NHPs remain very expensive and are limited in supply, they may be best employed to answer defined scientific questions related to microbial persistence, prevention of transmission or sterilizing activity of selected new agents as well as biomarker research. The use of animals in tuberculosis research changed dramatically in the past hundred years. Due to extensive research and analysis a greater understanding of the immunology, pharmacology, treatment responses and pathology has been elucidated. This effort focuses primarily on the mouse model which will likely remain the main animal model in TB Drug development for now due to its low requirement for drug material, economic advantages and ease of use.

3.2. Survey of current practices

3.2.1. Survey results: overview

A ninety five question survey of protocols and opinions was conducted of which about half of the questions were devoted to *in vitro* and the other half to *in vivo* methodologies. The survey can be viewed in its entirety at the following website: www.mrl.colostate.edu/acceleration.

3.2.2. Toxicology and pharmacology

About half of the responders mentioned the use of *in vitro* cytotoxicity assays. The pharmaceutical industry uses a wider array of cell lines compared to academia. For instance one laboratory evaluates compounds typically against a panel of 4 different cell lines: THP-1 (human monocyte cell line, suspension), BHK21 (Syrian Golden hamster kidney cells, adherent), HepG2 (Human Liver Carcinoma, adherent) and C6 Glioma (Rat brain glioma, adherent), all acquired from ATCC. In addition, the industry has easy access to additional toxicity assays and will therefore also implement these sooner in the screening process, such as genotoxicity (Ames test) and cardiotoxicity (hERG binding and patch clamp) assays. Extensive *in vivo* toxicology is usually not performed at the stage of early compound optimization, but most investigators perform an acute *in vivo* toxicity study with any new compound to screen for any adverse effects (behavior, weight, food intake) and gross necropsy findings (liver, spleen, kidney pathology analysis) after single or multiple dose administration. A few laboratories administer new compounds for toxicity testing for one week to evaluate weight loss, at higher than anticipated concentrations than those used in efficacy studies. In other laboratories, acute and chronic Maximum Tolerated Dose studies are performed before any *in vivo* studies are initiated.

Before testing the efficacy of an experimental compound *in vivo*, most investigators evaluate its oral bioavailability to a lesser or greater extent depending on the laboratories involved. Usually the pharmaceutical industry, who has access to a formulation and pharmacology department, performs extensive absorption, distribution, metabolism and elimination (ADME) and PK analysis prior to any efficacy testing of novel compounds. For the industry ADME and PK analyses of experimental compounds are the key elements of the *in vivo* approach, and only important leads will advance to *in vivo* efficacy testing. Other institutions and academic groups include a more basic evaluation of oral bioavailability where the analysis of drug levels is performed by bioassay or by HPLC. In academic settings, compounds generally tend to advance more rapidly to *in vivo* efficacy testing models due to the lack of access to ADME and PK capabilities. Before any long term efficacy mouse study is initiated a more formal pharmacokinetics analysis is performed by most investigators to establish the appropriate dose and dosing schedule, aiming to achieve an equivalent dose compared to humans for known TB drugs or drug classes. One laboratory reported routine pharmacological assessment of 6–8 time points following single doses or after steady-state dosing, obtaining serum or plasma by cardiac puncture using 3 mice per time point, and this laboratory also mentioned obtaining data on bioequivalence for first-line drugs and some second-line drugs. Based on the *in vivo* PK data, the investigators reported using non-compartmental analysis such as WinNonLin software to calculate the PK parameters and provide a simulation of a range of drug doses and dosing schedule. PK indices (AUC/MIC, Cmax/MIC, T $>$ MIC) are then determined at each dose in order to select doses expected to be efficacious in the infection model. Some investigators described extensive pharmacokinetic analyses using combination drug studies to rule out significant PK drug interactions. In another laboratory, a maximum tolerated dose (MTD) and PK/PD assays of all compounds are performed before initiating any efficacy animal studies.

Most respondents do not routinely test the ADME profile of experimental compounds in early stages of lead development. In most laboratories, the following assays are used at some point in the lead optimization process: the solubility of the compounds between pH 1.0-6.8, the parallel artificial membrane permeability (PAMPA) assay, $CaCO_2$ assay (to assess the permeability and efflux), liver microsomal assays using mouse and human microsomes (to evaluate metabolic stability), and CYP450 inhibition. At a late preclinical stage, more extensive PK studies under good laboratory practices (GLP) conditions are conducted in rats and one non-rodent species, which can include tissue distribution studies with radio-labeled compounds. Drug formulation studies are also reported to be initiated only later in preclinical development, after a lead compound has been identified.

As mentioned before, this document will primarily focus on the *in vitro* and *in vivo* methodologies specific for the testing of activity against *M. tuberculosis* and less on methodologies that are utilized in drug development in general (such as ADME, pharmacokinetics, toxicology, formulation optimization).

3.2.3. Choice of the mouse infection model for M. tuberculosis for drug testing

Most investigators performing *in vivo* studies reported using the mouse model. About equal numbers of investigators use either Balb/c or C57BL/6 mice (15–18%) and 11% used the outbred Swiss mice for *in vivo* efficacy trials. Only 4 laboratories reported the use of guinea pigs for drug evaluations (Figure 7). The mouse model for efficacy testing of experimental compounds was favored by most investigators because it is economical, widely available and mice

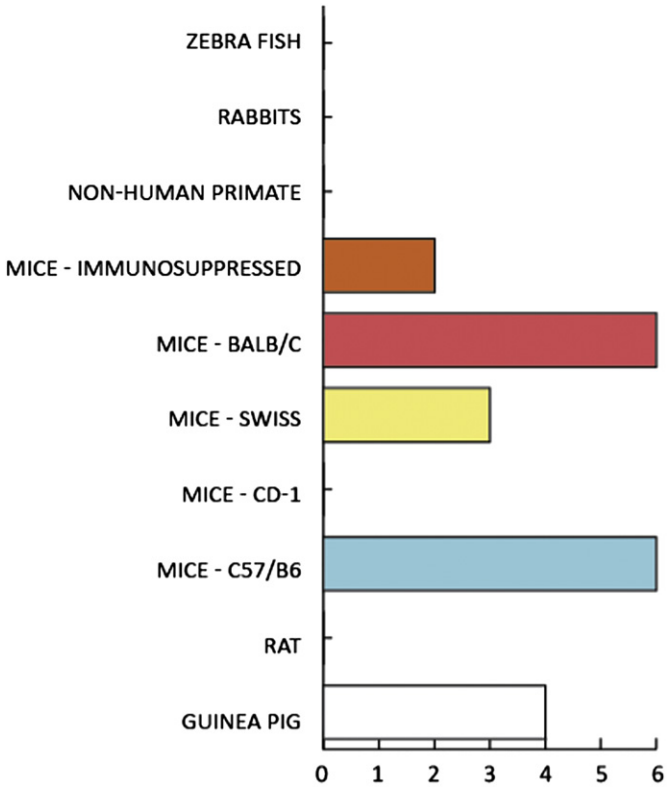

Figure 7. Survey responses regarding the animal species used for the evaluation of TB compounds *in vivo* (presented in percentages). Survey question: Which animal species is being used in your laboratory for TB drug evaluations?

are easy to handle. Mice were also reported to be commonly used for *in vivo* toxicity studies on lead compounds. One group of investigators mentioned some disadvantages of outbred mice (such as Swiss mice): the higher variability in CFU counts, requiring a higher number of animals to reach sufficient statistical power. In addition, the Swiss mice grow larger and therefore the model requires larger amounts of compound. The outbred nature of the Swiss mice, however, was seen by a few investigators as an advantage by reflecting better the diversity also seen in the TB patient population, as well as the wealth of data available from past studies. The advantage of using inbred mice is that results show small variations within animal groups and will strengthen the statistical power of the study. We have to mention here that a head-to-head comparison was in fact never performed to study the variance of infection between inbred and outbred mice. However, when comparing results from published studies on Swiss mice infected via the IV route versus our results on Balb/c mice infected via the IV route, we see a slight trend of higher variances in the Swiss mouse results (A. Lenaerts, unpublished observations). The survey reported mouse strain differences, for instance the Balb/c mice do not control the infection to the extent seen in C57BL/6 mice. Balb/c mice develop a productive infection with stable plateau in bacterial loads in lungs and spleens as opposed to declining bacterial numbers after initiation of the adaptive immune response as is seen in other inbred, more resistant mouse strains (such as C57BL/6 mice). The characteristics of the Balb/c were felt to be similar to the Swiss mice, which are less expensive and are outbred. It was commented that C57BL/6 mice are more sensitive to adverse effects seen by some drugs than other mouse strains. The main disadvantage of the mouse model was reported to be the lack of lung pathology seen in typical human TB pathology, as reported by several investigators. In addition, pharmacokinetics and drug metabolism in rodents are different from humans and cross species bridging studies/calculations have to be performed. The currently used mouse strains may not allow observation of the full range of alternative physiological states of bacilli present in human TB, such as in latency. In most cases, mice harbor mainly intracellular organisms. One investigator in the survey reported that the aerosol infection of a susceptible animal species (such as the guinea pig) mimics human granuloma formation, which is useful for studying extrapulmonary dissemination and hematogenous reseeding of the lung, and allows the evaluation of drug effects under conditions relevant to man.

When conducting animal models for efficacy testing of compounds, the route of infection as well as the start inoculum of *M. tuberculosis* differed significantly across the different research groups. Regarding the route of infection, the aerosol infection was mentioned by several groups as the most realistic mode of infection when compared to natural infection, and the guinea pig was found by several of the responding investigators to be most reflective of the human pathology and disease. At least one group uses the intranasal route of infection for mice which is felt to at least partially mimic the natural infection. Others pointed to the lymphatic and hematogeneous dissemination of TB bacilli from the lesion in favor of infecting via the intravenous route of inoculation in the IV infection model.

The infection inocula of *M. tuberculosis* reported by the different research groups ranged from 10 to 20 CFU per animal (guinea pigs) to more than 7 logs (intravenous injection in mice). Most investigators use either an aerosol or intranasal infection with 100–3500 CFU per animal, and only a few investigators reported use of the intravenous infection with a high bacterial number. One group using the intranasal inoculation defined their model as a low dose infection model, because the chosen inoculum had been shown to be the minimal dose capable of establishing a lung infection. The investigators pointed out that when using a very low inoculum that the bacteria are cleared before reaching the lungs, as was indicated by an increase in the standard deviation when less than 500 bacteria were used. Some laboratories aim at inoculating with 10 CFU during an aerosol challenge, others prefer 200–300 CFU/mouse for drug testing purposes because the bacteria grow faster to higher titers in the lung and spleen. The use of an even higher inoculum (3.5–4 logs) implanted by an aerosol infection was felt to be important by one group in order to reach a bacterial burden of 7–8 logs (which is hypothesized to be similar to a human cavity) at the start of the combination therapy 14 days later. In this latter model, it takes approximately 6 months of treatment with the standard HRZ-based regimen to sterilize most mice, which is similar to the length of treatment used in TB patients. The intravenous route where the early bacterial number is high also allows for the extrapolation of the antibiotic treatment results to the treatment schedule in TB patients. Investigators stated that the IV route is a more difficult technique to perform compared to an aerosol infection. For safety considerations, the intravenous route was aborted by at least one laboratory. It was pointed out that often the particular method of the mouse infection model depends whether a compound is an early hit or in the lead optimization stage of TB drug development. For instance, a low-dose aerosol is highly relevant for use in monotherapy trials as it will not give rise to emergence of resistance, and it will better represent a chronic, established TB infection. Two thirds of laboratories include a placebo gavage group and most have not observed any adverse outcomes of the placebo. Almost all laboratories include INH or RIF as control compounds in every study, and all mention having good reproducibility of the results in their validated mouse infection model.

3.2.4. Choice and propagation of the TB strain for use in vivo

In the survey, a high variability was reported for the origin and propagation of the *M. tuberculosis* strains used for *in vivo* experiments. The origin of *M. tuberculosis* strains in many of the laboratories is not entirely clear (many report the strain being a gift from colleagues in the past) and in 40% of the laboratories there was no record kept regarding the passage number for their *M. tuberculosis* strain. Only one laboratory has compared the infectivity of their *M. tuberculosis* strains and found that *M. tuberculosis* Erdman was more virulent than *M. tuberculosis* H_{37}Rv in guinea pigs and mice. One reported propagation method was as follows; the *M. tuberculosis* H_{37}Rv strain was obtained from ATCC, then grown to mid-log phase in liquid culture, homogenized, sonicated briefly, and filtered to obtain a single-cell suspension before freezing at −80 °C and enumerating the bacterial number from a thawed aliquot. A second laboratory also mentioned filtering their bacterial suspensions through a 5 micron syringe filter. One laboratory uses an early passage number of *M. tuberculosis* Erdman from ATCC grown as a pellicle and frozen as seed stocks. The seed stock is then expanded over three passages to 100–200 ml cultures in Proskauer Beck, up to an optical density (OD) of 0.5–0.7, and subsequently aliquoted followed by a virulence test in mice. Others mentioned to passage their strain in mice, then sub-culture and freeze. In another laboratory, *M. tuberculosis* is grown to an OD_{600} of 0.3–0.5 in 7H9-ADS medium, subsequently the cells are harvested by centrifugation (10 min, 3300 rpm), washed once in 7H9 medium at 37 °C, resuspended to obtain an OD_{600} of 1.0 in 7H9-ADS medium supplemented with 15% glycerol, and frozen at −80 °C. The viable bacterial number of each batch is determined by plating, and before infection, a vial is thawed, diluted and sonicated (25 s in a sonicating water bath). The following media were used to propagate the bacterial culture: 1) Dubos broth with ADCC enrichment, 2) 7H9, without Tween80, 3) 7H9-ADC media with glycerol and 0.05%

Tween80, 4) Difco 7H9 with OADC and Tween80, 5) 7H9-ADS medium containing 0.05% Tween80, or 6) 7H9 broth with 0.02% Tween80.

In 80% of the laboratories, frozen bacterial stocks are prepared for the inoculation of the mice (typically in 7H9) and stored in small aliquots at −80 °C for several years. This method allows a precise duplication of the infection level of mice every time an experiment is performed, and there appears to be no detectable change in infectivity or virulence over many months to a couple of years. Some investigators report to have noticed a drop in the viability after prolonged storage at −80 °C (e.g. 3–4 years), while others found that freezing the inoculum does not adversely affect the viability (with remarkable reproducibility of this method over decades). The titers of frozen stocks are by most laboratories confirmed by plate counts and the virulence is evaluated in mouse models. Others favor infecting with an actively growing *M. tuberculosis* culture, and one group standardizes the inoculum according to McFarland suspensions. It was reported that when a fresh culture is used the optical density determinations are not always reliable due to clumping of bacteria.

A wide range of survey responses were obtained on questions relating to control measures on the infectivity of bacterial stocks. One of the investigators mentioned not to enumerate bacterial loads in the lung the day after infection in most routine experiments unless the challenge strain changes, and another group agreed with this notion as long as clumps are removed by filtration. Still, most investigators perform all appropriate controls to track the bacterial loads in the mice at the day of infection or one day after and at the start of treatment, and multiple responders stated that they use a minimum of 5 mice at each sacrifice time point including for control groups.

3.2.5. Readouts for drug efficacy

The most widely used method for determining drug efficacy in mice remains enumeration of the bacterial load in lungs and spleens, by counting the colony forming units (CFU) of the organ homogenates on agar plates. For the testing of single compounds as well as short term mouse experiments, some investigators mentioned plating whole lungs while others plate the homogenate of a single lung lobe from each animal. After long treatment regimens when low bacterial numbers are expected, generally the whole lung is homogenized and a sizeable fraction of up to one-half of the homogenate is plated. One group homogenizes whole lungs in a total of 2.5 ml PBS and then plates the entire homogenate on five 7H11 plates (0.5 ml per plate). Another group uses whole lungs and total spleen homogenized in 4 ml PBS supplemented with 0.05% Triton X-100. Only a minority of investigators saves organs for histopathology, and of these, most used formalin fixation from 2 to 10 days to decontaminate the histology samples. Other readout methods besides CFU enumeration are only occasionally used in efficacy trials (pathology in 15% of the institutions, gross necropsy in 22%, Ziehl-Neelsen staining in 15% and only one laboratory uses auramine rhodamine staining) (Figure 8).

Relapse experiments have currently been performed by only a few of the respondents, and the time periods allowed for relapse vary from 12 to 40 weeks following cessation of the drug regimen. The relapse studies are at times performed with immunosuppression (using hydrocortisone, dexamethasone or cyclophosphamide), but are mostly performed without. Readouts for relapse of infection include observation of clinical signs, CFU enumeration, histopathology, spleen weights, and gross necropsy observation of lung lesions.

All respondents use statistical methods in conjunction with their animal trials and most do not consult with a statistician. Treatment effects are mostly analyzed by a one-way ANOVA, followed by a *T*-test, Tukey or Dunnett's test. For relapse studies a Fisher Exact or Chi-square test is applied to compare relapse proportions between groups.

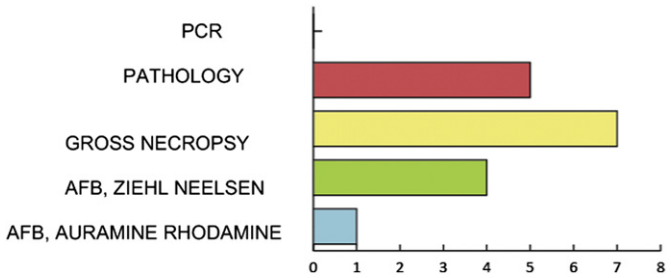

Figure 8. Survey responses regarding the type of additional or supportive readouts used for the evaluation of TB compounds in animal models (presented in percentages). Survey question: Are other readouts used to confirm or supplement the CFU data?

There were some unexpected data from animal studies mentioned in the survey. One investigator mentioned that while metronidazole and PZA demonstrated activity in some laboratories in mice, these drugs were not found effective in other laboratories. Some groups observed antagonism between the three standard drugs INH, RIF and PZA when administered in combination, while this was not seen by other researchers.

3.3. Issues related to in vivo testing methodology

The current views on *in vivo* testing methodologies as well as issues in the field are the result of a series of visits to both academic and industry laboratories, as well from a TB Drug Accelerator Workshop on "Assay Standardization", which was held after the Gordon Research Conference on Tuberculosis Drug Development titled 'Targets, Technologies and Trials', in the summer of 2009 in Oxford, UK. Since the time of the visits, however, we learned about the fact that more groups became involved in TB Drug development which unfortunately due to constraints of time and resources could not be included in this document. These current views on *in vivo* testing methodologies together with its issues were then used as a starting point for the head-to-head studies described in Section 2.5.

3.3.1. The relevance of the mouse model

The discussion during the Workshop focused on the use and value of the mouse model in TB Drug development despite its lack

Text box 4.

Lessons learned from the survey on *in vivo* models

- Numerous laboratories are willing to collaborate and share ideas, challenges and protocols
- There is often uncertainty regarding the origin, history and propagation records of the TB strain used *in vivo*
- Infection methods of animals vary from laboratory to laboratory, with the inhalation method currently being the most widely used
- Only in rare occasions were there unexpected results reported for the mouse efficacy trials
- Mice are the species most commonly used for TB drug testing (although our survey may have underrepresented other animal models due to the focus of this effort on mouse models)

of advanced pathology resembling that of TB patients. The idea that was put forward during the Workshop is that we do *not necessarily have to mimic human disease but try to mimic in animal models the response of M. tuberculosis to drugs*. Certain research groups aim to achieve in their mouse model similar relapse rates for the HRZ regimen as those seen in TB patients, or they aim to have a similar number of bacilli at the start of treatment in mice as is seen in cavitary lesions in TB patients, or to have a similar means of exposure to *M. tuberculosis* as in humans. The mouse model has its limitations (only one lesion type is present, and most bacilli are intracellular) and most investigators felt there is a need to investigate additional animal models with more progressive disease to address whether these shortcomings in pathology are real issues. The acute mouse model used by several investigators has severe limitations in reproducing human pathology and disease, as the mice have primarily an actively replicating bacillary load, show a lack of certain pathological characteristics and metabolic stages of bacilli. The use of this acute model was discussed to be appropriate to obtain an *in vivo* proof of concept (POC), evaluate PK features, tissue distribution and *in vivo* activity. Despite its many flaws, most attendees expressed that the mouse is a reasonable compromise to qualitatively and rapidly assess the efficacy of new compounds in a medium throughput manner. In addition, it is at the current time the only animal model that has been demonstrated to be predictive of clinical trial data (see also Section 2). A significant controversy still exists among key opinion leaders in the field over the most relevant animal model. It was felt that the state (intra- or extracellular), the number of bacilli and the host's immune response can influence the outcome of treatment studies. The non-human primate has received much attention in the last years because it recapitulates human pathology. It is still the mouse, however, which was considered to be the most practical and appropriate species to use for early TB drug development.

How relevant is the mouse model? It was generally agreed upon that the mouse is very useful but that one should not over-extrapolate mouse model data. The main purpose of the mouse model is providing information of the efficacy of a novel drug regimen for tuberculosis over another. This was agreed on by most, the mouse model being primarily useful in "*determining a rank order of novel drug regimens to increase confidence for a selected drug regimen to be superior to the standard drug regimen in clinical studies*". To that effect it will be highly important to confirm that rank orders of bactericidal and sterilizing potencies obtained in one mouse model are predictive for humans.[164,166] In addition, the mouse model can provide information on the contribution of a single drug within the regimen in studies of compounds with similar background regimens. However whether such results are predictive of human outcomes is currently not well known as this clinical data is not always available. It was emphasized that historically the mouse model has performed well (a few examples were also described in Section 2). For instance, the mouse model has shown that RIF and PZA are active against slowly replicating bacilli, which correlates well with the use of these drugs in the treatment of human TB. Conversely, INH and SM do not clear these slowly replicating bacteria in humans and this was predicted in the mouse.

3.3.2. The various mouse infection models

The two most discussed variables in infection methodologies were the route of infection with *M. tuberculosis* and the size of the inoculum. Some laboratories use the "realistic" low dose aerosol model; the human infectious inoculum is more closely recapitulated by a low dose aerosol rather than the high dose aerosol infection model. Although even an inoculum of a so-called "low dose aerosol" is usually 1 or 2 orders of magnitude higher than the infectious bacillary dose thought to produce natural infection in humans. But in a sense, realism is not actually a necessity if one merely wishes to see if a drug is active or not. Aerosol infection models are more expensive due to the purchase and maintenance of a specialized aerosol apparatus. The low dose aerosol model never achieves a bacterial load greater than 6–6.5 $\log_{10}$ CFU in immune-competent animals. On the other hand, the high dose mouse infection model aims to achieve a bacterial burden of 7–8 $\log_{10}$ CFU, thereby aiming to simulate the bacterial burden in a human cavitary lesion. In humans though there are usually only one or a few cavitary lesions, while in a mouse with an equivalent high burden there will be hundreds of individual inflammatory lesions.

It was also discussed that an aerosol with very few bacteria is possible, but this is constrained by the inherent variation in delivery. Ideally the ability to deliver just one bacillus into one mouse lung would be desirable, but the technology just does not exist. Higher inocula rapidly induce the host immunity, but often overwhelm the immune system and the mice succumb to disease. The lymphocytic-predominant granuloma seen in the mouse intravenous (IV) model bears little resemblance to the necrotic granuloma seen in the guinea pig aerosol model. Clearly drug tolerance is observed in mouse models as around 1% of the bacterial population is less responsive to drugs and requires long periods of treatment in order to eradicate. Whether drug tolerance (as seen in the classical mouse models) or persistence due to the environment (as seen in animal models with progressive disease) are one and the same remains to be investigated. Some believe that the emergence of 'persistent' bacilli in cavities is the basis of the long therapy period required in TB patients. Animal model data shows though that a small inoculum of *M. tuberculosis* administered by aerosol to guinea pigs produces a persistent bacillary subpopulation located in a specific microenvironment within primary lesions (in the necrotic caseous core region and rim of the granuloma).[199]

Another point of discussion, though rarely discussed, was the age of the mouse. It should be noted that mice are used in immunology studies when 6–8 weeks of age, and one could argue that at this age mice are still not fully immunocompetent. A related point of discussion concerned body weight and blood volume. Some investigators used 4 week-old mice to administer intravenous infections in volumes of 0.5 ml, which is large for mice of this young age.

3.3.3. Readouts for drug efficacy

Several issues regarding the endpoints and the mouse models themselves were discussed. Regarding accurate endpoints in treatment trials, the method used for enumerating bacteria from organ homogenates should be carefully (re)considered given the possibility that some non-culturable bacteria in samples of animal tissues may not form colonies on solid agar. Liquid media, such as BACTEC or MGIT, was suggested as a method to enhance sensitivity for finding low numbers of CFU and to gain some insight into the state of the bacillus at the time point examined. Liquid testing using the MGIT system is also now the method of choice for sputum evaluation to assess culture conversion for clinical trials. Several automated liquid culture systems have shown greater sensitivity than the traditional solid-media cultures the acknowledged increased in the mycobacterial recovery rate of liquid media, which is likely due to a more mycobacterial populations being able to recover in liquid culture than on solid media. This is supported by *in vivo* and *in vitro* observations that subpopulations of *M. tuberculosis* with different states of metabolic activity co-exist in old liquid cultures,[200] as well as in liquid cultures with bacterial growth from chronic infected mice that do not grow on solid media.[201]

Another topic of discussion in at least two laboratories was the issue of drug carry-over at the time of enumeration of bacterial colonies in the organs. CLF and TMC-207 both have long half-lives, high tissue distribution and tissue binding, and therefore drug might still be present at the time of sacrificing the animals. The first indication that drug carry-over is an issue is observed when dilutions of organ homogenates do not have the expected reduction in bacterial number. Another indication might also be the lack of correlation between the bacterial number and the gross pathology observation. Several methods have been described to reduce the carry-over of drug in agar plates, including using LJ medium or 7H11 with 5% bovine serum albumin (BSA) for TMC-207,[202] and using 0.4% activated charcoal for CLF.[203] Relapse studies will then show the true sterilizing potential of these drugs and drug combinations.

The guinea pig model shows primarily extracellular bacteria in the central caseum and in a rim of fibrosis or non-mineralized necrosis. Some bacteria occur singly, but most are in small clusters. Methods for obtaining bacterial load data from particular guinea pig tissues that allow cultivation of single organisms may be needed in order to assess the true effect of chemotherapy. It was pointed out that colony forming units (CFU) measures by plating organ homogenates on solid agar are often inaccurate since they do not necessarily include just single bacilli but rather small clumps or conglomerations. There was discussion around the hypothesis that these clusters represent a type of biofilm. Hatfull et al. found these *M. tuberculosis* biofilms *in vitro* to consist of bacteria surrounded by a layer of free mycolic acids.[66] Although the biofilm hypothesis was not accepted by everyone, it was agreed that the extracellular bacteria in these micro-environments are unique and should be targeted by drug treatment. Another topic of discussion was about the drug refractory nature of these persisting bacilli being a drug penetration issue, by drugs not getting through this fibrous rim and extracellular lipid matrix. The ability to determine the penetration of TB drugs into a granuloma was seen as an important gap in the current knowledge about TB drugs.

3.3.4. Long term treatment studies and relapse

The main goal of long term models is to assess the activities of new regimens in terms of shortening the length of treatment versus the standard regimen. This might be best modeled in the high dose aerosol model because it resembles a similar treatment length as those used in the clinic (the intensive and continuation phase), whereas in the low dose aerosol infection model shorter treatment times would have to be investigated and correlated to the clinical conditions. Relapse studies were discussed as being a critical component of drug evaluations to assess the sterilizing potential of new drugs and regimens, however, some laboratories are reluctant to consider long term relapse models due to the lack of clinical data that proves relapse data in mice are required. A poster presented at the GRC meeting (which is now published[166]), however, showed that bactericidal activity is not always predictive of long term sterilizing properties and thereby highlights the importance of relapse studies. Our head-to-head studies showed a similar result for the moxifloxacin containing regimens (see also Section 3.5.3).

It was discussed that relapse studies using mice in fact cannot be realistic as in the clinic relapse rates of 5% with standard therapy are observed. If in animal experiments relapse rates reached merely 5%, the experiment will not achieve sufficient statistical power to find any statistical beneficial effect of a new drug regimen over the standard regimen. Investigators commented on this by explaining the methods used in their laboratories to obtain higher relapse rates; this was achieved by either infecting animals via the IV route which gives higher relapse rates than infection by inhalation, or by including different lengths of drug treatment. Using one of both strategies will ensure a sufficiently powered experiment which can then be used to rank order drug regimens versus the standard drug regimen. An example was discussed, where data from relapse studies was used to show an intermittent treatment regimen was as effective as daily dosing (such as the Denver regimen).[204]

3.3.5. The bacterial strain

The overwhelming variety of *M. tuberculosis* strains used by different laboratories was the first striking realization when visiting the different laboratories involved in TB drug evaluations in animal models. In addition, a difference in virulence of these strains was also reported which could be caused by each investigator propagating their strain using in-house protocols instead of using a single reference method. Issues regarding cultivation methods have been the topic of papers for close to a century, mouse passage and pellicle growth are usually the methods used to generate virulent stocks.[107]

Relying on one bacterial strain for efficacy studies was felt to be a less desirable practice due to possible unique properties of a particular strain; rather, several diverse strains including a current clinical strain should be included in animal studies. To have well characterized *M. tuberculosis* strains analyzed by whole genome sequencing and metabolomics, proteomics and drug susceptibility testing was considered to be a necessity. The consensus was to validate and standardize the bacterial strains

Text box 5.

Lessons learned from the workshop: on *in vivo* efficacy models

- The mouse model is still the animal model of choice as it represents the only animal model that has been validated to date by testing regimens and treatment durations known to work in patients
- Regarding the route and inoculum of TB infection in mice, realism is not actually a necessity if one merely wishes to see if a drug is active or not
- The main role of the mouse model is to establish a rank order among novel drug regimens in order to increase the confidence for a selected drug regimen to be superior to the standard drug regimen in clinical studies
- Relapse studies are critical as the bactericidal activity in mouse model is not always predictive of sterilizing activity
- Liquid media may be a more sensitive growth indicator
- Diversity of animal models is desirable, standardization of protocols and TB strains within a laboratory is a requirement
- Publishing all methodological details is encouraged
- Confirmation of results of *in vivo* activity in another infection model and against another TB strain from a characterized strain collection is strongly recommended
- Include all controls (including the two drug combination when there is a substitution of one agent in a triple drug regimen)
- Always test and report MIC data on the *M. tuberculosis* strain used in the *in vivo* studies
- Confirm findings in a second animal model to address evaluation of efficacy against persisting bacilli in advanced lesions. [Currently of theoretical concern only: good comparative studies between mouse and other models are missing]

used for *in vivo* testing and to keep the diversity of mouse infection models, since there is great advantage to be gained from confirming data of certain drugs and drug regimens in different models. Plus there needs to be more discussion on what criteria a bacterial strain needs to have to be defined as "well characterized". Of importance, the *in vitro* drug susceptibility of the strains should be reported for the reference drugs, to ascertain that results from one laboratory to another can be compared.

3.3.6. Statistical analysis

Statistical methods are of key importance and different methods are used depending on the question asked. Methods used were found to be very similar across laboratories, and often assistance of a statistician was provided to answer new questions. Power analysis prior to the experiments is required to determine the number of animals used in the experiment. The primary data analysis for mouse models is usually a one-way analysis of variance (ANOVA) of the $\log_{10}$ CFU bacterial loads, with *t*-distribution based contrasts comparing individual treatments. A Dunnett's test is an appropriate method for comparing all new compounds to untreated controls, while controlling the overall error rate in that set of comparisons at the usual 0.05 level. Also Bonferroni and Tukey statistical tests are appropriate tests for a pairwise comparison between treatment groups. Evidence of differential relapse based on detection (yes/no) of TB bacteria would be established using a Fisher's exact test comparing rates of relapse between two experimental groups. In experiments with $n = 5–10$ per group, power to identify significantly different relapse rates is very low. Therefore in order to increase the statistical power, more treatment groups are implemented with various lengths of treatment, or a higher relapse rate is aimed for in order to see significant differences between different treatment regimens. Relapse information obtained from mouse models should always be interpreted with great care and seen as trends in relapse differences between treatments.

3.3.7. Screening sequence

Recommendations were made regarding the evaluation of single drugs for which short term models can be utilized, and for the evaluation of the sterilizing activity of single drugs to use a minimum of a 4-week treatment regimen. To assess drug combinations against tuberculosis, a 3–6 month long study using a chronic infection mouse model was recommended, and for relapse studies a 3-month drug-free interval should be included. Internal controls should always be included, such as an HRZ combination arm for drug regimens and an INH arm for single drug studies. It was discussed during the workshop that studies should be repeated in a different laboratory to confirm the results and to rule out effects of specific differences in methodology on outcomes before an agent is taken into the clinic. In light of our later findings, we can refine that statement. To confirm *in vivo* results we now strongly recommend that the repeat experiment is performed with a different TB strain, in another infection model (changing the route of infection, time of treatment) which is preferably confirmed in a second laboratory (see also Section 2.7: Recommendations).

The choice of a drug dose *in vivo* was discussed with two different research groups, and it was mentioned that for initial *in vivo* evaluations a high drug dose is appropriate to give the compound the best chance of success. For many studies in the mouse model, a dose is chosen based on having a similar AUC in human patients if this information is available. For preclinical work with unknown lead compounds, the PK characteristics and toxicology will initially determine the drug dose. Subsequently, dose fractionation studies with TB drug candidates are essential to assess the predictive parameters that determine *in vivo* activity (such as AUC/MIC, or Time over MIC). In addition, the hollow fiber method might give additional information at a later stage of TB drug development regarding dose selection and optimal dosing regimen of single drugs.[205] Although dose selection is a critical component of TB drug development, during our visits not many groups had lead compounds in that stage of TB drug development. The CPTR (Critical Path for a New TB Regimen) effort will however release a white paper focusing on the gaps and issues in advancing drug regimens to the clinic and this will address specifics on the hollow fiber model, dose selection and more.

In several recent reviews and during visits of the various laboratories in TB drug development, different strategies in TB Drug R&D became apparent.[6,81] Novartis spoke about following a strict "Critical Path" with an emphasis of MIC, ADME and PK followed only then by testing *in vivo* efficacy in mice. At J&J the emphasis was clearly to increase the selectivity index of their initial hits [the ratio of cell cytoxicity over inhibitory concentration against *M. tuberculosis* (CC50/IC50)] of the diarylquinoline compound series until no improvement of the *in vitro* activity could be reached. This was followed by extensive ADME and PK profiling of a few compounds (resulting in 'compound polygraphs'), and *in vivo* testing in the mouse model. Data of the selected compounds were then summarized in a 'compound monograph'. The J&J strategy by Koen Andries was rapidly known as the "3M" approach focusing on "MIC, mice and man". Although the MBC and NRP activity were considered by K. Andries as useful additive *in vitro* data, these would only come as secondary screens. At GlaxoSmithKline the emphasis lies heavily on PK/PD *in vivo* analysis and only good leads will be considered to go forward for *in vivo* efficacy testing. Most laboratories use some variety of a short term mouse model initially (starting treatment a few days after infection, or against actively replicating bacteria), followed by an established infection model to look at the sterilizing activity of the compound.

3.3.8. Scientific questions

An interesting question was raised whether an immunocompromised mouse such as the GKO mouse model could completely eradicate the disease upon drug treatment, or whether an intact immune response is required for a durable cure. The idea behind the published GM-CSF mouse model to initially rapidly kill more than 90% of the bacilli so that only drug tolerant bacilli remain, and then to evaluate the activity of single compounds against the drug tolerant bacilli.[206] The use of immunocompromised mice for this purpose will shorten the treatment periods significantly and therefore the model could be used early in TB drug development to see if a compound has any sterilizing activity against slowly replicating or drug tolerant bacteria.

There are multiple issues regarding assessment of new compounds, and not all of these are currently adequately addressed. Pharmacokinetic and pharmacodynamic issues are paramount, such as determining the extent of drug penetration into TB lesions [not many results yet], understanding how the local environment [pH, hypoxia, etc.] influences activity, and identifying which drug of a drug regimen participates at what time during treatment? There are also many issues surrounding the study of drug combinations that need to be addressed (proper PK, antagonism, toxicity issues, formulations, timing of drug delivery etc.).

Another question that came up several times during visits was the usefulness of the MBC value of a new compound in regards to its *in vivo* killing kinetics. Most investigators see the MBC as additional but not critical data. Whether the MBC is reached at the site of infection *in vivo* will largely depend on the PK features of the molecule. For instance, RIF has a higher MBC (2–4 μg/ml) compared to its MIC (0.03 μg/ml) and therefore the ratio of the MBC/MIC is relatively high. *In vivo* however these drug levels are in fact reached due to the favorable PK and protein binding features of

RIF. The same can be said for the results obtained in NRP assays as a good predictor for sterilizing activity *in vivo*. Since a good sterilizing agent should display good activity against actively replicating and other subpopulations, it is better to have a compound with a low MIC and reasonably good activity against NRP, as well as displaying good tissue distribution. It is important to combine and interpret all data on a compound in context to each other.

Another topic of discussion was the problem with dogmas in the tuberculosis field, and one of these is about the location of bacilli in a TB infection. A lot of speculation is ongoing about the possible low pH in the necrotic lesions, and most of the bacteria being extracellular while there is little proof for this. All recognize that these would be important studies to conduct. In addition, most investigators realize that most animal models do not reflect these features.

3.3.9. Standardization of infection models?

The issue of the need to standardize assay conditions is a difficult one and in the end the diversity was seen as being positive. There are many issues regarding assay standardization and there will likely be no harmonization amongst laboratories performing studies in a certain manner. It is in fact likely that this diversity will allow discovery of a greater number of compounds. It was acknowledged that there is definitely added value in having investigators use only standardized and reproducible protocols within their laboratories,and that the infection model be highly reproducible and be able to recapitulate the already available clinical data. Investigators should calculate statistical power prior to a study in order to have high enough numbers of animals to make meaningful conclusions, and they should thoroughly record all variables relevant to the assessment of the animal. Regarding methods, it was recommended to report every detail in publications such that all possible variables could be reproducible. There will be a continuous effort to find alternatives to animal models such as computer modeling, the hollow fiber or artificial granuloma models, but these would have to be critically assessed and validated thoroughly with existing drugs and compared to the mouse which has been fairly predictive for many years. There are many gaps and questions in TB drug development where these models might actually provide an answer (establishing an appropriate drug dosing regimen, selection of drugs based on penetration in the granuloma, etc.), but all of these new methods will need more validation. Important is the continuing need to analyze the current state of affairs and understand the limitations and strengths of each method.

3.4. Overview of the methods and key definitions

There are two main mouse models of infection for TB drug development which are primarily addressed in this strategic plan. Both methods differ in the route of infection, and these were the intravenous and the aerosol infection model. The two methods were the most reported ones by the participants at the time of the survey. We realize that more recently several groups are also utilizing the intratracheal and intranasal infection approaches. Investigators have reported reproducible results to us for the intranasal infection when the inoculum is higher than 200 CFU. Unfortunately, we were constrained under this effort to focus on the most widely used approaches which are the intravenous and aerosol models.

Both of the studied infection models differ significantly in their approach, on the site of infection and the activation of the immune response. After intravenous infection, assuming the bacterial inoculum is correctly prepared as a single cell suspension, there is a characteristic pattern of 90% of the bacilli being deposited in the liver, 10% in spleen and only 1–2% will initially reach the lung. In terms of the spleen, the bacilli are directly delivered into a lymphoid organ, resulting in rapid induction of both innate and later acquired immunity in this organ. For the aerosol infection all bacteria are delivered in the lung and only after the initial days will there be dissemination via the bloodstream to other organs. The usual inoculum used in mice is ~100–5000 bacilli by aerosol infection, but a range of inoculums have been reported. The immune response is initially activated in the lung.

3.5. Head to head comparisons of the in vivo methodology

Methodologies for preclinical animal models testing of drugs against *M. tuberculosis* vary from laboratory to laboratory; however it is unknown whether this heterogeneity in experimental parameters results in different outcomes regarding new TB agents used alone or in combination. The objectives in identifying the most important variables in the mouse infection models are two-fold. The first goal is to identify the most important variables that can influence results in animal models, thereby informing new investigators. The second goal is to improve the TB drug development process by implementing the most effective and accurate methods for drug evaluations to eventually identify highly effective drugs with demonstrated sterilizing activity.

In order to determine whether the results obtained in the most commonly used mouse models are comparable to one another, a series of head-to-head comparisons of drug regimens in three of the most commonly used mouse models was studied under this project [the intravenous injection, a low dose and a high dose aerosol infection model of *M. tuberculosis*] in two commonly used strains of mice. The main goal was to determine if outcomes in mouse efficacy models can vary based on the methodology used. A number of variables were identified from the start of the project, others were included during the duration of the project, and the following variables were selected for the head-to-head evaluations: the influence of the mouse strain on drug efficacy, the route of infection (aerosol, intravenous infections evaluated on drug efficacy and relapse), inoculum conditions (fresh versus frozen cultures, water versus broth as diluent, the size of inoculum used for infection), the preparation and formulation of drugs, the timing of drugs when administered in combinations and the *M. tuberculosis* strain used for infection of the mice.

3.5.1. Inoculum of M. tuberculosis

In a few initial mouse experiments, a couple of variables regarding the *M. tuberculosis* inoculum used for infecting the animals were studied. A fresh bacterial culture versus a frozen aliquot of *M. tuberculosis* was compared for infecting the mice. The results showed that both inoculums gave the same *in vivo* bacterial growth kinetics in lungs and spleens. There was also no difference seen in the *in vivo* bacterial growth depending on the diluents used to prepare for the inoculum. The same results were achieved with bacteria either in H_2O or 7H9 media, or when a fresh bacterial culture was frozen and thawed, and diluted with H_2O. When incubating the organ dilutions on 7H11 plates for enumeration of bacteria in organ homogenates, incubation in presence or absence of CO_2 had no effect on the bacterial number (CFU). The only difference was the faster growth of the bacteria in presence of CO_2 (colonies were readable in 2.5 weeks instead of 3.5 weeks without CO_2 incubation).

3.5.2. Choice of mouse strain

A head-to-head experiment was aimed at evaluating drug efficacy of the standard drug regimen INH, RIF and PZA in two of the most commonly used mouse strains. The mouse strains were

Text box 6.

Key definitions on *in vivo* efficacy models

- Acute infection: defines an *M. tuberculosis* infection in animals with actively replicating bacteria. In standard laboratory mice, this is an infection established in less than two weeks (which is before the adaptive immune response is activated).
- Chronic infection: defines an *M. tuberculosis* infection in animals where the bacterial load progresses to a steady-state in CFU due to the adaptive immune response (and is considered the opposite of an acute infection). The bacilli are considered to be (non- or) slowly replicating.
- Low dose infection: This infection model will progress slowly into a chronic infection in standard laboratory mice, and the animals are able to control the infection. Typically this is obtained after delivery by aerosol of 50–100 CFU of *M. tuberculosis*.
- High dose infection: This infection model will progress rapidly and develop a high bacterial burden, and the animals will succumb to disease 2–4 weeks after infection. Typically this is obtained after delivery of a high inoculum by the intravenous route, or after 3.5 $\log_{10}$ CFU by the aerosol route.
- Drug exposure: Defined as the overall drug levels that the animal is subjected to treat an infection which is typically based on the serum or plasma levels of the drug.
- Initial/intensive phase of treatment: the initiation of therapy which consists typically of 3–4 drugs associated with a large fall in bacterial numbers. This phase typically lasts for 2 months.
- Continuation phase of treatment: the prolonged phase of therapy usually with two drugs only, aiming to eradicate a small persisting bacterial population. This phase typically lasts for 4 months.
- Latent infection: defines a clinical condition in which the person is PPD skin test positive and is thought to have been infected with TB. The TB patient has been able to control the infection, and does not show any signs or symptoms of active disease.
- Persisters: defined as bacteria that can survive long term exposure to drugs to which they are fully susceptible on the basis of MIC testing, and which represents itself as phenotypic tolerance. Persisters can occur in response to drugs, the immune response and hostile micro-environments (such as hypoxia, nutrient depletion, etc.).
- Dormancy: a term often used in the past, which might not be appropriate for *M. tuberculosis* as metabolic processes have been shown to continue in persistent organisms *in vitro*, although to a lesser extent.
- Cure or sterilizing activity: is synonymous with observing no relapse of infection in mouse models after termination of treatment. Relapse in mice should be tested at least 3 months after therapy is completed. [These terms are preferred over the misused term 'durable cure'. 'Sterilizing activity' is preferred over 'sterilization' as sterilization cannot be confirmed experimentally, whereas sterilizing activity is a measurable activity against persisting bacteria that can result in cure].
- Bactericidal activity: This term refers to the activity of a compound or drug regimen as assessed during or immediately after termination of drug treatment. [This term is preferred over the often misused term 'sterilizing activity' to describe the activity of a compound during treatment].

chosen based on the survey results. Mice were infected by low dose aerosol with *M. tuberculosis* and the three-drug combination was administered for 4 months using either Balb/c or C57BL/6. The results showed that the bactericidal efficacy of the standard drug regimen was not significantly different in the two mouse strains infected by aerosol.[164] These results show that for the bactericidal activity of the tested drug regimen no difference was observed.

The question whether relapse rates after drug treatment would be similar in various mouse strains was not investigated, and this should still be addressed in future as the role of the immune response will be more prominent at this stage. It has been well established that mouse strains have varying resistance towards a *M. tuberculosis* infection,[207] and therefore relapse of infection might be delayed in a more resistant mouse strain. However, the more important question is to what extent the ranking of drug efficacies for different drug regimens in relapse studies might change for the various mouse strains, if at all, and this remains to be evaluated. A study comparing relapse results of different regimens using inbred mice versus Swiss mice [the mouse strain used in many historical experiments validating the IV mouse model] would answer these questions.

3.5.3. Route of *M. tuberculosis* infection

In a next series of head-to-head mouse model studies, a comparison was performed between three mouse models based on the route of *M. tuberculosis* infection and inoculum size. The mouse infection models included the HDA, LDA and IV models (Figure 9). Three treatment regimens were evaluated for *in vivo* efficacy during the initial treatment stage and during relapse of infection. The standard drug regimen (INH, RIF and PZA), and the 2-drug combination (RIF and PZA) were evaluated, as well as the combination regimen where moxifloxacin (MXF) was included as a replacement for INH in the standard regimen (MXF, RIF and PZA). Drugs were administered for a duration of 4–6 months, and relapse was assessed 3 months after termination of treatment. The reason for inclusion MXF was that the fluoroquinolone is currently in phase III clinical trials for TB treatment.[208] This study also tried to address the observation seen by others using mouse models of the antagonism between the three standard drugs and thereby seeing a beneficial effect of replacing INH with moxifloxacin, which has been a controversial finding in the field.[152]

The results of the studies were published and therefore are just summarized here.[164] When comparing the efficacy of the drug regimens tested in the LDA mouse model versus the IV model, it was clear that there was a delay in bactericidal activity in mice infected by the IV route versus LDA by the standard drug regimen. The efficacy results during drug treatment showed statistically similar activity for both three-drug combinations (MRZ and HRZ) for all time points tested, and this in all three mouse infection models. Therefore, the ranking of the tested drug regimens was found to be similar during treatment for all three infection models. Relapse results showed statistically similar relapse rates for the following drug regimens and lengths of treatment: 2HRZ/4HR = 2MRZ/2 MR for both the HDA and IV infection models, and 2HRZ/2HR = 2MRZ/1 MR for the LDA. Of importance, 2MRZ/2 MR showed improved sterilizing activity by showing significantly less relapse when compared to 2HRZ/3HR (only statistically significant for the IV infection model). For the HDA group, a similar trend was observed but this was not statistically significant due to a low number of mice relapsing infection. In the LDA model virtually no relapse was observed after 2MRZ/1MR treatment, while there was 10% relapse seen after 2HRZ/2HR treatment. These results showed that the bactericidal activity does not necessarily reflect the relapse results of a drug regimen (as is discussed in De Groote et al.[164]).

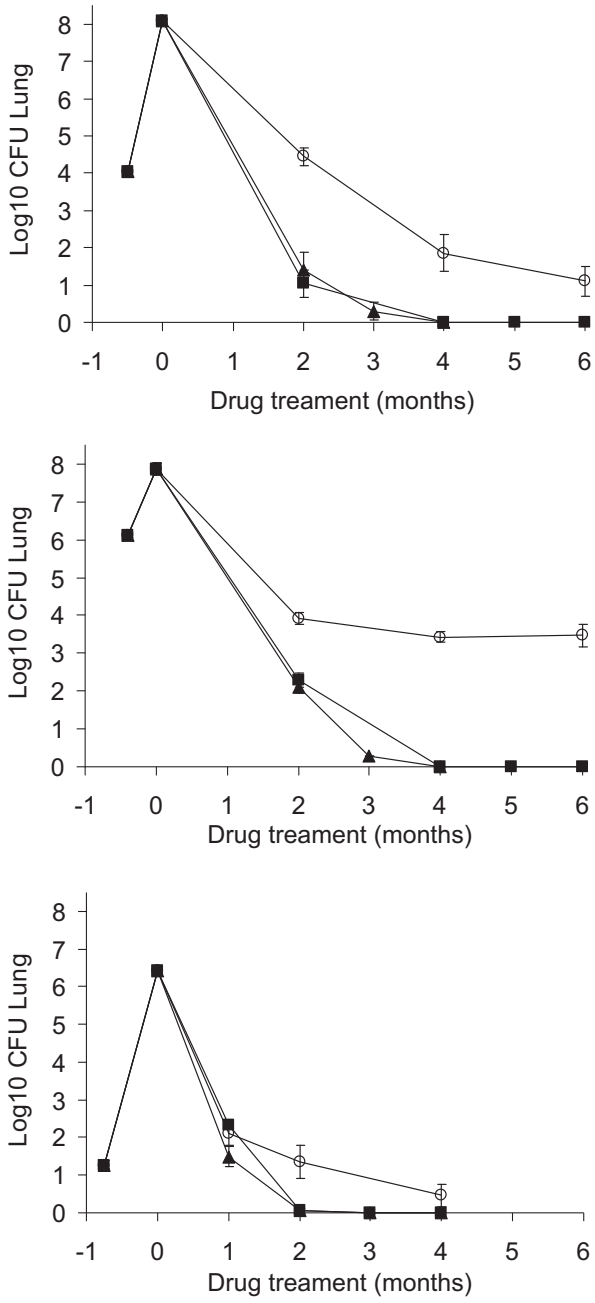

Figure 9. Bacterial numbers in lungs of Balb/c mice after HDA (top), IV (middle), or LDA (bottom) infection with *M. tuberculosis* and 1, 2, 3, 4, 5 or 6 months of drug treatment. Drug treatment included either two months INH, PZA and RIF, followed by INH and RIF (■); or two months MXF, PZA and RIF, followed by MXF and RIF (▲); or 6 months of PZA and RIF (○). Taken from reference [164].

When the relapse rates were compared between the different mouse infection models, the IV groups showed significantly more relapse than for the groups from both aerosol models. This observation has been made by others as well. It is possible that the difference in immune response at the initial site of infection is responsible for this. After an IV infection, a high number of bacilli is introduced in liver and spleen (only 1% is found early in the lung) which may have an effect on the immune response thereby influencing the frequency of relapse. A delayed bactericidal activity in the IV model might also contribute to a faster relapse of infection. The increase in relapse rate was not caused by a higher number of resistant colonies in the IV infected groups.

In conclusion, the results comparing the *in vivo* efficacy of drug regimens in three mouse infection models gave the same end conclusion and therefore showed us that the route of infection was not a variable that changed the outcome of the *in vivo* efficacy study. These studies showed similar results in bactericidal efficacy for the MXF-containing regimen versus the standard drug regimen, regardless of the mouse infection model used (either low or high dose aerosol, or the IV infection model). In contrast to observing similar bactericidal activity of MRZ and HRZ over the initial months of treatment, the relapse rates for the MXF-containing regimen were significantly lower than these for the standard drug regimen in the IV infection model. These results confirm the importance of evaluating both the bactericidal and sterilizing efficacy of drug regimens in mouse models in order to select the drug regimen that can lead to cure. In this case the IV infection model was the only model showing statistical significant superiority of the sterilizing activity of the MXF-regimen over the standard regimen, which was not observed for both aerosol infection models due to lower relapse rates.

Clinical trials have been conducted to assess the benefit of MXF in standard TB drug regimens. Tuberculosis Trials Consortium Study 28 was conducted based on the mouse model data of substituting MXF for one of the TB drugs in the standard regimen (HRZ).[154,209] A trend in improvement with MXF (99/164 were culture negative vs. 90/164 for INH) was seen at 8 weeks, but it did not reach statistical significance ($p = 0.37$) when substituting INH in the standard regimen for MXF. The results of our mouse studies support the findings of human TBTC study 28, where no significant difference was seen in bactericidal activity when MXF was substituted for INH in a daily regimen for 8 weeks. The mouse studies discussed here predict that MXF might only show significant benefit in later stages of the clinical trial. However, there are also other reported IV infection studies using Swiss mice which do not show improved relapse rates of the MXF-containing regimen over the standard regimen.[210] In that report, the relapse results obtained with the MXF-containing regimens in the high dose aerosol infection model[209,211] could not be confirmed in the IV infection mouse model.[166,210] Ultimately, the long term follow-up data of the ReMoxTB clinical trial will be highly informative regarding the predictive value of the mouse infection models and relapse studies, and is currently underway (described at www.tballiance.org).

Since the focus of this project was to concentrate on drug combination studies, it is certainly possible, and perhaps likely, that treatment with single drugs would show a difference in drug effect between the different animal models. It would not be surprising that the actively replicating bacterial population in the IV and HDA infection models behave differently for compounds with a certain mode of action, when compared to bacteria in a chronic state after an LDA infection. Additional experiments with a panel of TB drugs targeting certain bacterial populations would be required in order to address this question.

3.5.4. Drug preparation, formulation and dosing schedule

The experiment aimed at determining the effect of formulation and dosing schedules of drugs in combination treatment regimens on the efficacy in mouse infection models. Balb/c mice were infected by HDA and drug treated with the standard regimen of INH, RIF and PZA, with and without MXF, and compared to the MXF, RIF and PZA combination. Two different RIF formulations with alternate schedules of drug delivery were compared in this study. In the first method RIF was dissolved in water and dosed at least 1 h prior to other drugs, and in a second method RIF was administered in 5% DMSO 1 h after combinations of INH or MXF and PZA. The results of the studies were published and therefore are just

summarized here.[164] In conclusion, there was no significant difference seen in drug efficacy of the tested drug regimen depending on the RIF formulation or the dosing schedules used here. Analysis of drug concentrations was performed on plasma samples of mice to assess the pharmacokinetics of INH, RIF and PZA after a single drug administration of RIF (using the same RIF formulation methods as described above in the efficacy study). There were some slight differences seen in the kinetics of the concentration–time curves of RIF depending on whether the drug was ground in a mortar in water or prepared as a DMSO solution, but overall drug exposures of RIF were not significantly different. Overall, no significant differences were seen in the AUC of the drugs depending on the RIF formulation or the dosing schedules used. The results here show that at least for the drug combination tested, the formulation was not the critical variable that changed the efficacy results in a significant manner. Grinding RIF in water showed at least as good, and possibly even better *in vivo* activity than when formulated in DMSO. Therefore, we recommend using water and careful grinding of RIF to smaller particle size as the optimal formulation for RIF.

3.5.5. Strain of *M. tuberculosis*

After testing many variables, it became clear that the main variable that was not yet addressed was (the source and propagation of) the *M. tuberculosis* strain used for *in vivo* infections, and that this sole variable might in fact explain some of the discrepancies between experiments performed under this project and by others. In the head-to-head studies described in this section, the combination of HRZ was in all three infection mouse models significantly more effective than the RZ dual therapy. In none of the three infection models was an antagonism of H in the HRZ combination observed. Antagonism between the three standard drugs has been shown by others,[212] where it was demonstrated that the 2-drug regimen of RZ after removal of INH performed better than the standard 3-drug regimen HRZ. In order to prove that the bacterial strain might be the most important variable to explain the discrepancy between laboratories (in this case the antagonism between HRZ), an experiment was initiated where all other variables (such as inoculum, route of infection, mouse strain, drug formulation, dosing schedule, and others) were kept identical except for the *M. tuberculosis* strain used for infecting the mice. The drug regimens HRZ, RZ, MRZ and HMRZ were evaluated in Balb/c mice infected by HDA with either *M. tuberculosis* $H_{37}Rv$ (Origin: Frank Collins, mouse passaged[176,209]) or *M. tuberculosis* Erdman strains (Origin: Trudeau, pellicle grown[164]). The results of the studies were recently published[213] and are therefore just summarized here. The results showed that there was a significant antagonism observed in HRZ in the mice infected with the *M. tuberculosis* $H_{37}Rv$ used, while this was not seen in mice infected with *M. tuberculosis* Erdman. At the time of this finding, the result could not just be explained by the strain difference of using either $H_{37}Rv$ versus Erdman as other investigators using *M. tuberculosis* $H_{37}Rv$ for infecting mice have never reported seeing antagonism in the standard drug regimen [such as using the Swiss IV mouse model with *M. tuberculosis* $H_{37}Rv$ of uncertain derivation (obtained from Laboratoire de Bacteriologie, Faculte de Medecine Pitie-Salpetriere, Paris, France)[166,214,215]].

Therefore, a final head-to-head study was performed to confirm that the antagonism is a strain specific phenomenon, and included clinical strains. Ultimately, the most important question is to know which drug regimens will perform well or better than the standard drug regimen against clinical TB strains. In a recent trial, a comparative study was conducted which evaluated 5 bacterial strains: 2 *M. tuberculosis* $H_{37}Rv$ [of two different origins[164,176,209]], *M. tuberculosis* Erdman, HN878 and CDC1551. The recently published data[213] indicates that 4 out of 5 *M. tuberculosis* strains never showed any antagonism in HRZ (with HRZ being significantly better than RZ), whereas one of the $H_{37}Rv$ strains (Origin: Frank Collins, mouse passaged[216]) showed an entirely different drug profile with RZ being as effective as HRZ.

The most surprising result was that both $H_{37}Rv$ strains gave a different result in this study. Although many laboratories use $H_{37}Rv$, there are many variations of this lab-adapted strain in circulation. $H_{37}Rv$ was initially derived from a clinical isolate, named H37, which was obtained in 1905 from a patient E.R. Baldwin with pulmonary tuberculosis (Figure 10). The $H_{37}Rv$ (ATCC 25618) strain is maintained at the Trudeau Institute in New York, although unfortunately, the original H37 clinical isolate has been lost. Strain ATCC 27294 (Trudeau Mycobacterial Collection; TMC102) is also frequently used as a representative of $H_{37}Rv$ in studies, and treated equivalently in the literature. ATCC 25618 and ATCC 27294 were both isolated from the same patient in different years, and both are fully drug susceptible. The TMC102 strain was derived from the 1905 human-lung isolate H37 in 1934 by W. Steenken, and deposited by GP Kubica. ATCC 25618 was identically derived, but deposited by AG Karlson. Despite these differences within $H_{37}Rv$ both maintain virulence.[165,217] Besides these different sublineages of $H_{37}RV$, the propagation of the bacterial strain in a laboratory can also lead to genetic alterations. A recent paper details the elaborate study of a group of investigators to assess the consistency of the genome sequences among $H_{37}Rv$ strains in use by six different laboratories and the extent to which the strains have diverged from the original sequenced $H_{37}Rv$ strain by whole-genome sequencing.[165] It was clear from the results of the six strain varieties that besides different polymorphisms, nucleotide substitutions, insertion/deletions, and multiple IS*6110* transposition events were observed (Table 9 and Figure 10). Two of the ATCC 25618 strains demonstrated frameshift mutations + GC in the *mas* gene (encoding mycocerosic acid synthase). The authors here concluded that the genome-wide catalog of genetic differences can help explain any phenotypic differences that might be found, including the observed frameshift mutation in the mycocerosic acid synthase gene which caused these two of the tested strains to be deficient in biosynthesis of the surface glycolipid phthiocerol dimycocerosate (PDIM), however, both are known to retain virulence.

The results described here show the importance of the *M. tuberculosis* strain in mouse efficacy studies, and therefore we highly recommend for investigators to use only highly characterized strains with known background and confirm any drug regimen results in mouse infection models against a second *M. tuberculosis* strain, preferably a currently relevant, characterized clinical isolate.

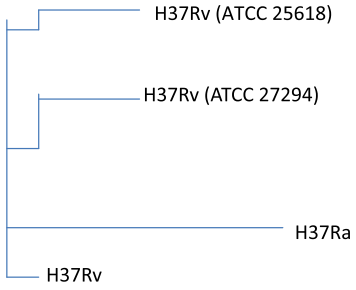

Figure 10. Phylogenetic tree of $H_{37}Rv$ and $H_{37}Ra$, adopted and redrawn from reference [165]. H37 was isolated from a human patient in 1905 and propagated until at the Trudeau Laboratories. When the avirulent $H_{37}Ra$ was separated, the two $H_{37}Rv$ and $H_{37}Ra$ were separately maintained and the parent was no longer available. Within the $H_{37}Rv$ from various laboratories there has been divergence based on single nucleotide polymorphisms (SNPs).[165]

Table 9
Descriptions of the two different *M. tuberculosis* $H_{37}Rv$ strains deposited in the ATCC. (See description in Ref. [165].)

	ATCC 27294	ATCC 25618
Origin	Steenken, 1934	Steenken, 1934
Depositor	Kubica	Karlson
IS6110	777777**757**60771 777777**777**60771	777777477760771
PDIM	No mutations in mas gene	Frameshift in mas gene

3.6. Lessons learned and future directions

Under this effort, certain parameters regarding mouse infection models for tuberculosis were re-evaluated in head-to-head experiments in the laboratory. However, during the course of these *in vivo* experiments as well as through discussion at the laboratory visits and during the workshop, additional questions or remaining topics of discussion have arisen on animal models for TB drug testing. Some of these are listed below together with the lessons learned.

3.6.1. Is standardization needed?

In preclinical testing of experimental compounds against tuberculosis, there are various mouse models used by different investigators in the field as described above. By surveying investigators, visiting laboratories and performing in-house experiments it has become clear that the standardization of the methods within a laboratory is definitely recommended. Certain crucial parameters of mouse infection models have been identified through our efforts that can change the outcome of drug efficacy trials in TB drug development. These results have shown that within the same laboratory it is important to standardize the mouse infection models used so that results can be compared from experiment to experiment. Standardization in terms of using a certain inoculum of bacteria, start of infection, appropriate qualitative controls to track the infectivity and growth characteristics of the strain in every experiment, inclusion of drug controls, etc. In addition, every laboratory should work with a well characterized *M. tuberculosis* strain with known origin and have an SOP for culturing and propagating their bacterial strain. A log is highly recommended for the TB strain used *in vivo* to track the virulence performance after repeated culturing, the passage number should be noted (and should be limited to 3–6 passages), and results of positive controls tracked.

As for standardization of mouse models across laboratories, several discussions were held on this topic. The consensus was that there was far more to be learned from multiple mouse infection models used across different laboratories instead of relying on one single infection model used by every laboratory. Confirmation of a superior activity of one drug combination in multiple animal models in different laboratories was only seen as increasing the confidence to move with that drug regimen further into clinical trials.

3.6.2. The in vivo testing sequence

Among mouse infection models, there are clear differences in bacterial responses seen to drugs depending on the stage of infection, the effect of the immune response, the size of infection inoculum used and on when the start of treatment occurs. Bacteria are actively replicating in acute infection models, when the infection is performed with a high inoculum, when treatment is started before the adaptive immunity is activated, or when immunocompromised mice are used.[174] In a chronicly infected animal, studies have indicated that bacteria are replicating more slowly.[218,219] Different research groups have shown that INH has the best activity in the acute infection models,[174] whereas PZA and mefloquine (MFQ) (unpublished results from CSU as well as NITD) have better activity in chronic infection models. Overall compounds showed far greater activity in an acute mouse model [such as the GKO mouse model] when compared to a chronic mouse model [using immunocompetent mice], such as described in Hoff et al.[174] Another study evaluated the acute (3 day incubation period) and chronic (28 day incubation period) mouse models after 4 weeks of treatment and showed that drug activity of the single drugs tested ranked differently in both models.[220] The largest discrepancy was observed with PZA, which showed it to be highly bactericidal in the chronic model but permitted bacterial multiplication in the acute model. Differences with RIF were similar, though not as dramatic. On the other hand, INH and MXF ranked at the top in the acute model but near bottom in the chronic model. AstraZeneca showed very similar data on a poster at ICAAC 2010 on this topic, and confirmed the effect of PZA. It is our recommendation to use acute mouse models for early drug discovery projects because the models show a greater sensitivity in identifying active compounds. In a lead identification program, a chronic infection model is more appropriate in order to select a lead compound with activity against slower replicating bacteria.

One needs to recognize that most mouse models are evaluating activity of single compounds or drug regimens against intracellular bacteria, as standard laboratory mice primarily develop inflammatory lesion types. It is important for TB drug development to realize that the majority of the bacilli in an advanced disease state are in fact extracellular organisms located in necrotic lesions,[174] as is seen in human cavitary disease.[221] On the premise that bacterial persistence is determined, at least in part, by microenvironmental conditions (e.g. hypoxia, nutrient depletion, low pH) found in caseating and necrotic lung lesions, it has been suggested that the mouse model may not faithfully represent sterilizing activity of new tuberculosis drugs. Up to date, however, there is no data available yet that supports this hypothesis. Therefore, the evaluation of the predictive value of mouse and other animals models for clinical trials definitely warrants further investigation which will be possible in future with more new drugs being currently evaluated in clinical trials.

In conclusion, the differences in host immune response among the different animal models infected with *M. tuberculosis* result in a wide variety of granulomatous lesion types which will influence the location, metabolism and state of the bacilli and which eventually will determine the drug responsiveness. Increasing our understanding of lesion morphologies as well as the location of bacilli among the different animal models is vital in designing different levels of stringency for testing new drugs.

3.6.3. Choice of the M. tuberculosis strain

M. tuberculosis $H_{37}Rv$ was the TB strain most commonly referenced in our survey for both *in vitro* and *in vivo* models for TB drug testing. The notion of $H_{37}Rv$ as a 'standard reference' strain should be used with some caution though. After our conversations with investigators, it was often not entirely clear what the exact origin and background was of the $H_{37}Rv$ strain used (usually it was reported to us as being 'received from a colleague'), and at times no special precautions were taken into account for strain culturing and propagation to retain virulence or to limit the passage number. Although a recent publication showed that animal passaging of *M. tuberculosis* strains prior to quantitative virulence testing in mouse or guinea pig models does not enhance or restore potency to strains [that may have lost virulence due to *in vitro* passaging], the authors urge investigators to verify virulence of parental strains before genetic manipulations are undertaken or other comparisons

Table 10
Characterized *M. tuberculosis* clinical isolates available to researchers worldwide (strains kindly provided by Drs. M. Kato-Maeda and P. Small, Francis J Curry National Tuberculosis Center, University of California, San Francisco, CA).

Resources for researchers			
Identifier	Lineage	Available through ATCC/BEI	Publications
X001354	East African Indian	Yes	Gagneux et al. PNAS 2006, Hershberg et al. PLoS Biology 2008, Comas et al. PLoS ONE 2009, Comas et al. Nature Genetics 2010
X004439	East Asian	Yes	Gagneux et al. PNAS 2006, Hershberg et al. PLoS Biology 2008, Comas et al. PLoS ONE 2009, Comas et al. Nature Genetics 2010
X003899	IndoOceanic	Yes	Gagneux et al. PNAS 2006, Hershberg et al. PLoS Biology 2008, Comas et al. PLoS ONE 2009

are made.[222] Another recent paper details the elaborate study of a group of investigators which assessed the divergence of the genome sequences among $H_{37}Rv$ strains in use by six different laboratories.[165] The idea of having well-established reference *M. tuberculosis* strains with a full chain of custody, good growth characteristics and low passage number is definitely a necessity.

This leads, however, to the more important question of how confident are we that the newer drug combinations will be effective against current clinical strains. *M. tuberculosis* is a member of the *M. tuberculosis* complex, and genetic diversity has recently been linked to clinical, pathogenic, and immunologic heterogeneity in disease progression and outcomes.[223–229] Strains are currently being classified into six major phylogenic clades: East African, East Asian, Euro-American, Indo-Oceanic, and two West African (*Mycobacterium africanum*) types.[230,231] Several studies have reported different biological, clinical, or epidemiological behaviors of the various strains, defined by their geographical location or strain type.[165,225,232,233] In addition, strains have also been shown to vary significantly in virulence in animal infection models as measured by survival, bacterial loads in target organs, histopathology,[234,235] and the severity of pulmonary and extrapulmonary lesions, as well as the immune response engendered after infection. All these recent reports show that *M. tuberculosis* strains are more genetically diverse than was previously recognized.[223,226,233,236] Results from this project thus far, strongly urge investigators to repeat and confirm findings against another *M. tuberculosis* laboratory or current clinical isolate. For this purpose, the *M. tuberculosis* clinical isolates do not necessarily have to be drug resistant in order to rank order drug regimens for bactericidal activity or relapse of infection, as there is no evidence of a correlation between virulence and drug resistance in present clinical isolates.[234] For relapse studies, the presence of regulatory T-cells which suppress T-cell activation[237–239] might be of importance and this immune cell population is not expressed after infection by all TB strains.[237,240] In Table 10, we list a few resources that can provide characterized *M. tuberculosis* strains (clinical strains) to be used by investigators worldwide. Table 11 lists websites and resources available to researchers and drug developers available from the NIH. More studies using various clinical strains are required to increase our basic knowledge on disease pathogenesis in relation to drug treatment.

3.6.4. Enumeration methods of bacteria

The activity of an investigational drug or regimen has been in recent years mostly determined by the reduction of colony forming units (CFU) of *M. tuberculosis* by dilution of organ homogenates on solid agar 7H11 plates. Plating of the organ homogenates has always been the gold standard for quantifying drug efficacy *in vivo*. However, drug discovery efforts are often times held up by this time consuming step requiring an incubation period of the bacterial plates of 3–4 weeks. For early drug discovery efforts, more efforts are being investigated lately for indirect, but more rapid, methods to 'measure' the bacterial load: such as the luciferase readout[241,242] or fluorescence.[243] These novel detection methods will undoubtedly accelerate the TB drug discovery process by delivering an immediate readout on the efficacy of an experimental compound at either time of sacrifice (for luciferase readout) and even in live animals in real time (fluorescence). Thorough validation, however, is not available as of yet and would be required before these indirect methods can replace the enumeration by CFU for all drug classes.

Another discussion took place at several occasions during this project, pointing out that perhaps not all bacteria can be identified, cultured or visualized by current methods. The discussion started with the recent introduction of liquid culture media for the diagnostic evaluation of clinical specimens for suspected tuberculosis.[244] The question then became whether drugs or regimens which preferentially kill certain bacillary populations would give rise to a 'flawed' readout by only culturing certain subsets of bacteria on solid agar. If liquid cultures of organ homogenates allow the growth of TB subpopulations that will not grow on solid media, the parallel evaluation in liquid and in solid cultures might offer an opportunity to study drug effects on subpopulations in different metabolic states. This idea was proposed for EBA clinical trials,[244]

Table 11
NIH/NIAID Services and contracts available to researchers.

Services for researchers available research programs at the NIH/NIAID		
Activity	NIH Services Resource/URL	NIH Contact
Investigator-initiated grant applications in all areas of TB	NIAID's Microbiology and Infectious Diseases Biological Resources Repository; Partnership initiatives for the development of product candidates	T. Pickett
HIV/TB research. Response to currently active initiatives for fundamental and translational science in MDR/XDR TB	Screening, microbiological activity, animal testing and preclinical development of new TB chemotherapeutics: /LabsAndResources/ resources/dmid/ pretheraagents/, and/Labsandresources/resources/dmid/invitro/ Pages/invitro.aspx, and http://nihroadmap.nih.gov/raid/	T. Parker B. Spinelli M. Kurilla M. Ussery M. Nasr J. Boyce
Evaluate the pharmacology of existing and new drug candidates in animal models of TB disease.	*In vitro* and *in vivo* PK/PD models, testing of individual drugs and drug regimens	T. Parker
Characterize the genome and drug resistance markers of drug resistant *Mtb* strains	Sequencing, molecular genetics, bioinformatics: niaid.nih.gov/LabsAndResources/resources/ brc/ niaid.nih.gov/LabsAndResources/resources/gsc/	M. Giovanni

but holds true for animal model testing of drugs or regimens as well and therefore should be tested.

This question might even be more relevant in animal models with a greater variety in granulomatous lesion types where bacilli are located in different environments (such as in necrotic, closed or cavitary lesions) and hence might have different metabolic stages. Classical studies have reported on those non-culturable bacilli from closed lesions.[245] In those times it was difficult to grow bacteria from such lung lesions, often as not the calcified lesions from both lungs and lymph nodes were found not to be infectious, but certain long term culturing exercises could regrow bacteria in closed lesions. In a way this is returning to the "viable but non-culturable" debate of old, but in a new light of current drug development.

3.6.5. PK/PD issues

As we mentioned before, this document mainly focuses on the *in vitro* and *in vivo* methodologies required for activity testing against *M. tuberculosis* and less on methodologies that are utilized in drug development in general such as compound preparation and storage, ADME, pharmacokinetics and pharmacodynamics and impact of drug formulation on *in vivo* efficacy. However, there are just a few points of discussion relating to TB specifically that are addressed here.

Generally, initial *in vivo* efficacy experiments will use monotherapy with high drug doses in order to give the new chemical entity the best chance of success and provide *in vivo* proof of concept. For combination therapy trials, the drug dose is established after pharmacokinetic analysis based on a dose with equivalent AUC in humans if available. Dose fractionation studies are also feasible in mice and can determine the PK features that drive bacterial killing (pharmacokinetic effect, PD). Recent such studies have improved our understanding of the PK-PD of INH, RIF and the fluoroquinolones.[246,247] These experiments may inform clinical development with regard to optimal dosing strategies or might be required to understand clinical data. In the case of PA-824 where in a clinical EBA trial no dose response was observed,[248] dose-fractionation studies in mice by E. Nuermberger *et al.* at JHU have shown that the major PK parameter driving bactericidal activity is time above MIC,[249] which explains the clinical EBA data entirely. The same was the case for TMC-207 which did poorly in the initial days of the EBA clinical trial,[250] which was repeated and showed the same results in a mouse model.[250] These experiments are very labor-intensive however, and not very amenable to long-term models including relapse. TB drugs will always be used in combination, and therefore PK analysis should also address drug–drug interaction and evaluating PK parameters driving bactericidal killing when drugs are administered in combination. Of significance, these PK parameters might be very well different for the various stages during TB treatment, either in the earlier bactericidal stage or later when killing the last 1% persisting bacilli. The question is how far to take these PK/PD studies in animal models prior to going to clinical trials, or do we only utilize these labor-intensive studies when certain questions arise? The CPTR (Critical Path for a New TB Regimen) effort will release shortly a white paper focusing on gaps and issues in advancing drug regimens to the clinic and will also address specifics on the dose selection of TB drugs in regimens in mice and man.

3.6.6. Relapse studies

During our visits with laboratories involved in TB Drug development, a recurring point of discussion was whether relapse studies in mice are required for the preclinical testing of a new lead compound in a drug regimen. Relapse studies entail labor-intensive mouse studies of up to 9 months with large mouse numbers per group to have sufficient statistical power. They usually also require treatment regimens of various lengths, making this a time-consuming and expensive phase of preclinical testing of a new drug entity in a regimen. Until more recently it was mostly perceived that the initial bactericidal activity of a TB drug regimen was generally speaking correlated with relapse of infection after cessation of treatment. No convincing published data was available that showed relapse studies are indeed adding information to the bactericidal activity of a new regimen. Recently, an important publication by Andries et al showed that relapse studies give additional information regarding the sterilizing properties of a regimen.[166] In the described study, a rank order was determined of the bactericidal and sterilizing potencies of several combination regimens, and results showed that certain drug regimens with very good bactericidal properties did not necessarily have good sterilizing properties (Table 8). Another example from Johns Hopkins University using drug combinations of standard drugs and PA-824 demonstrated that bactericidal activity and sterilizing activity, while being close, did not exactly correlate.[251] Our head-to-head experiments (described in Section 2.6) also showed that bactericidal activity and sterilizing activity for the drug combinations tested did not always correlate.[164] We observed similar efficacy of MRZ and HRZ over the initial months of treatment, whereas the relapse rates for the MXF-containing regimen were significantly lower than these for the standard drug regimen for IV infected mice. Clinical trials to evaluate the superiority of the MXF-containing regimens to the standard regimens will provide the definite answer to the question on the predictive value of the mouse model and are currently underway.

Of significance, the lack of correlation between bactericidal and sterilizing potencies seen in the mouse model also questions the usefulness of the current biomarkers in clinical trials for relapse-free cure (sterilization). Current biomarkers are based on bactericidal potency measures such as 2-month sputum conversion of serial sputum colony counts in TB patients.[252] A discrepancy between bactericidal and sterilizing potencies of regimens has previously been observed in clinical trials, in which the relative proportions of culture conversion after 8 weeks of treatment were not predictive of relapse rates.[253] It was suggested that there is a link between the 2-month sputum culture conversion and relapse, but for various reasons that correlation is not straightforward.[254]

A recent initiative by the Gates Foundation (Critical Path to New TB Regimens, or CPTR) will stimulate development of new combination regimens (of new investigational drugs along with existing TB drugs) to avoid developing each drug sequentially over decades to come.[255] Although there is no doubt that the CPTR approach will lead to more effective drug regimens, it will become crucially important to select the most promising combinations from preclinical studies for both bactericidal and sterilizing activity to advance and be tested clinically. Since the regimens will mainly consist of new drug entities, relapse studies in different mouse models will be required to confirm the improved efficacy over the standard regimen. Different drugs have varying activities against the various metabolic states of TB subpopulations observed in the lung (in sputum, necrotic or cavitary lesions, closed lesions), and therefore a careful dissection of the bactericidal and sterilizing contributions of a single drug of the novel regimen during the length of treatment would be preferred to be able to rationally design the most potent regimen.

Our head-to-head experiments focused primarily on the comparison of three mouse infection models using Balb/c mice.[164] In most of the early studies though, investigators were primarily using the outbred Swiss mice which were inoculated generally via the IV route.[147,166,210] From our studies and by others, it is clear that it takes longer treatment times to achieve cure in the IV infected

Balb/c[164] or Swiss mouse model,[166] when compared to the Balb/c mouse model infected by aerosol.[164,209,211] For instance, the drug regimen [RPT, MXF and PZA] required only 10 weeks of treatment to obtain a relapse-free cure in the aerosol mouse model,[211] versus > 4 months in the IV mouse model.[166] It is also apparent in the IV mouse models that the relapse rates change less dramatically compared to the Balb/c mice when increasing the treatment times, which might limit the ability to document statistically significant differences depending on the time and number of relapse time points taken. A typical result with RHZ for instance in reported studies with Swiss mice may show 80% relapse at 4 months, 50% at 5 months, and 25% at 6 months of treatment. These results are barely showing sufficient differences in activity over the timepoints used to produce statistically significant differences in relapse rates. With Balb/c mice after a high dose aerosol infection, typical relapse rates are 75% after 4 months of HRZ, 17% at 5 months, and 0% after 6 months. These data illustrate that estimates of minimal treatment durations are different in the various mouse models, and therefore should be carefully assessed. The most important question, however, is to know whether *the rank order of the sterilizing activity of drug regimens* would be the same regardless of the mouse infection model used. A head-to-head 2 × 2 factorial comparison of drug regimens in relapse studies for instance, using IV infected Swiss mice (outbred) versus aerosol infected Balb/c mice (inbred) has not been done and would be of great interest. In the meantime, we recommend the evaluation of the rank order of sterilizing efficacies of new treatment regimens preferably in two different mouse models which might give the ultimate confirmation prior to going to clinical trials.

3.6.7. Reporting of animal model data

Investigators will often not publish negative results and optimism bias permeates the field. A recommendation for publication of negative results and adoption of uniform reporting requirements should be put forth. There should be a rigorous systematic review or meta-analysis before embarking on human trials as has been recommended in the past. Hackam *et al.* in a recent editorial about the issues in translating animal research into clinical benefit,[256] points out the presence of methodological biases in animal experimentation as well as inadequacy of animal models resembling human pathophysiology as the major culprits. The suggestions that are put forward can be applied to TB research. A better uniform reporting system is needed where all animal data is collected, positive and negative, and available for investigators, regulators and funding agencies. If multiple data sets are available this would allow for more systemic reviews of a drug regimen in order to do meta-analysis studies as are performed in clinical trials. Another suggestion from Hackam et al. is to register animal experiments which would reduce publication bias and therefore provide a more comprehensive, informed view before proceeding to clinical trials. In the meantime, it is prudent to be skeptical about the applicability of animal model data to the clinical domain, which is something investigators in TB drug evaluations are well aware of.

3.7. Summary and final recommendations

Since new drugs need to be tested in animals before going to man it is crucial that the important parameters of animal models be identified and variables of experimental methodology be understood. This is especially important now that more compounds than ever will advance to clinical trials and in light of the CPTR program. We have re-evaluated protocols and collected information that hopefully can assist laboratories, both new and experienced, to accelerate the process of TB drug discovery and development. We have identified historical and contemporary methodologies and collected protocols and publications from universities, companies and institutions engaged in this work via an electronic survey and personal visits. In order to make progress in a field of diverse methodologies, different parameters in assays and mouse models were identified and tested in head-to-head experiments in the laboratory. To that effect, we compared three of the most widely used mouse infection models, which are the low dose aerosol infection model, a high dose aerosol infection model and the intravenous infection model for tuberculosis. Results of *in vivo* efficacy testing of the standard drug combination were shown to be similar regardless the strain of mice used. The results of both the bactericidal and sterilizing activity for the drug combinations tested were found to be independent of the route of infection, inoculum size, the formulation of RIF, or the timing of administration for RIF in combination (either 1 h prior or after) for the drug regimens tested. The results showed that the most important variable in our experiments was the *M. tuberculosis* strain used for infection. Whether this variability is related to the origin of the strain, the method of propagation, the passage number of the strain still needs to be investigated further. The following suggestions and recommendations resulted from this work:

- ○ According to most of those who participated in this effort, the mouse model is still the main animal model of choice for TB drug development as it represents the only animal model that has been validated to date by reproducing regimens and treatment durations known to work in patients.
- ○ Different laboratories should embrace the variety in mouse models used for TB drug development. The consensus on the topic of assay standardization throughout this effort was that there was far more to be learned from multiple mouse infection models used across different laboratories instead of relying on one single infection model used by every laboratory.
- ○ Standardization of the protocols and methods used within a laboratory *is definitely* recommended. Standardization in terms of using a certain inoculum of bacteria, start of infection,

Text box 7.

Remaining issues and questions on *in vivo* efficacy models

- Standardization is required within every laboratory; but the variety of different mouse models across laboratories is recognized to be valuable
- The origin of *M. tuberculosis* strains with limited passage number and the propagation conditions to retain virulence are critical parameters. More studies are required to study the relationship of clinical strains and drug treatment
- Bactericidal efficacy is not always predictive for sterilizing activity of a drug regimen in mice; therefore relapse studies are required
- Further studies are required to see whether ranking of newer drug regimens for sterilizing activity are similar regardless of IV and aerosol infection
- Are other animal models more predictive of clinical outcome compared to the mouse model?
- Further studies are required to re-evaluate the CFU readout as a determinant of drug efficacy; liquid media may be a more sensitive growth indicator
- Should there be a standard way of reporting animal model data, and collection all *in vivo* data in a central database?

start of treatment, appropriate qualitative controls to track the infectivity and growth characteristics of the *M. tuberculosis* strain in every experiment and inclusion of appropriate drug controls.

- In the *in vivo* head-to-head experiments described here, the *M. tuberculosis* strain was found as one of the most important variables that can change the outcome of the *in vivo* efficacy trial in mice. Therefore, we highly recommend the confirmation of *in vivo* efficacy results against a different *M. tuberculosis* strain than the one used in an initial experiment. If possible, the animal study should be repeated with a current clinical *M. tuberculosis* strain [belonging to one or more different clades[230]]. Of importance, it is recommended to only use *M. tuberculosis* strains that are fully characterized (genetically), with good growth characteristics, limited passage number, propagated appropriately to retain virulence, and which have a complete chain of custody.
- There is not one single strategic plan or testing sequence that can be followed for every new TB compound, and this is especially true for late stage of preclinical testing (some examples include PA-824, MXF, TMC-207 as mentioned above). Early in drug discovery, the testing for *in vivo* activity of a single compound in mice is straightforward. In the stage of drug regimen optimization, however, a more tailored approach is required for every new investigational entity. All features including PK/PD, protein binding, activity against NRP bacilli, Minimal Effective Dose, Minimal Bactericidal Dose, *in vivo* tolerability, drug–drug interactions, etc. will determine the strategy for advancing the compound into combination studies in mice.
- For all drug combination mouse studies, the recommendation is not to repeat the evaluation of drug regimens in the exact same mouse infection model, but to perform the confirmatory studies in a mouse model sufficiently different from the initial one (using a different mouse strain, other route of infection, other TB strain). This is to increase confidence and ensure the new drug combination is more potent compared to the standard drug combination prior to advancing to clinical trials. Final confirmation of results in a second laboratory would therefore be highly encouraged.
- Of utmost importance, any unexpected antagonism between new lead entities should be repeated before removal of a potentially potent combination (e.g. antagonism in the standard regimen, and between TMC-207 and PA-824 (PA). The same applies for any unexplained synergistic activities between drugs in the mouse model.
- The confirmation of results in a second animal model [such as the guinea pig, the rabbit, non-human primates, the Kramnik mouse model, etc.] to address evaluation of efficacy against persisting bacilli in advanced lesions is currently of theoretical concern only, as comparative studies between the mouse and other animal models are largely missing and their predictive value to clinical data has not been shown yet.
- Relapse studies are an essential part of evaluating the potency of a drug regimen given that the ranking of the bactericidal and sterilizing potencies of the regimens can be different.
- In any combination treatment study where a substitution is made with a new drug in a regimen (e.g. removal or an addition of another agent), it is important to include all appropriate control arms.
- It is critically important to continue to repeat and confirm results by others by duplicating every detail of the original work.
- Early discussions and interactions with representatives of regulatory agencies are encouraged and can provide critical guidance on the types and design of studies (both preclinical and clinical) likely to facilitate TB drug development.

In summary, given the wide heterogeneity of human clinical TB it may not necessarily be a good practice to rely on a uniform set of harmonized experimental methods. Protocols used by each laboratory that have developed and standardized their methods, typically generate reproducible results in that laboratory. We therefore recommend preclinical animal studies for TB drug development not be standardized as such but that critical studies are confirmed in a second laboratory using a different mouse, different strain(s) of TB, with or without infecting via another route and include the appropriate treatment control groups, prior to moving to expensive human trials.

Acknowledgements

The authors would like to thank Dr. Phillip Chapman (Statistics Department, CSU) for statistical assistance. We are grateful to Jena Valdez, Courtney Hastings, Elizabeth Donegan, Amanda Marsh, Elizabeth Brooks, Laura Nold, Janet Gilliland, Lisa Woolhiser, Colby Wells, Jordan Clark for technical assistance and the Laboratory Animal Resources staff for assistance with care and analysis of the animals and Mia Huang for graphic design. We gratefully acknowledge Dr. Scott Irwin for proofreading the manuscript. This project was supported by the Bill and Melinda Gates Foundation under Drug Accelerator grant ID number 42589 titled: “Assay Standardization”.

Funding: None.

Competing interests: K. A. is a full time employee at Janssen (Johnson and Johnson).

Ethical approval: Not required.

Appendix A. Supplementary material

Supplementary data related to this article can be found online at doi:10.1016/j.tube.2012.07.003.

References

1. Ginsberg AM. Drugs in development for tuberculosis. *Drugs* 2010;**70**:2201–14. http://dx.doi.org/10.2165/11538170-000000000-00000.
2. Ginsberg AM. Tuberculosis drug development: progress, challenges, and the road ahead. *Tuberculosis* 2010;**90**:162–7. http://dx.doi.org/10.1016/j.tube.2010.03.003.
3. Nuermberger EL, Spigelman MK, Yew WW. Current development and future prospects in chemotherapy of tuberculosis. *Respirology* 2010;**15**:764–78. doi:RES1775.
4. Payne DJ, Gwynn MN, Holmes DJ, Pompliano DL. Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat Rev Drug Discov* 2007;**6**:29–40. doi:nrd2201.
5. Global Alliance for TB drug development. http://www.tballiance.org/cptr/.
6. Nuermberger E. Using animal models to develop new treatments for tuberculosis. *Semin Respir Crit Care Med* 2008;**29**:542–51. http://dx.doi.org/10.1055/s-0028-1085705.
7. Tuberculosis. Scientific blueprint for tuberculosis drug development. *Tuberculosis* 2001;**81**(Suppl. 1):1–52.
8. Nocard M, Roux E. Sur la culture du bacille de la tuberculose. *Ann Inst Pasteur* 1888;**1**:19–29.
9. Hammerschlag A. Bakteriologisch-chemische Untersuchungen uber Tuberkelbacillen. *Zentr klin Med* 1891:9–18.
10. Kune W. *Zeitschr F Biol* 1892;**30**:221.
11. Proskauer B, Beck M. Beitrage zur Ernahrungs-physiologie des Tuberkelbacillus. *Z Hyg Infectionskrankh* 1894;**18**:128–52.
12. Armand-Delille P, Mater A, Schaeffer G, Terroine E. Culture du bacille de Koch en milieu chimiquement defini. *Compt Rend Soc Biol* 1913;**74**:272–4.
13. Long ER. The nutrition of acid-fast bacteria. *Am Rev Tuberc* 1922;**5**:857–69.
14. Dubos RJ, Davis BD. Factors affecting the growth of tubercle bacilli in liquid media. *J Exp Med* 1946;**83**:409–23.

15. Crimm PD, Martos VF. Effects of Amigen and amino acids on the growth of tubercle bacilli. *Am Rev Tuberc* 1944;**49**:94–104.
16. Bernheim F. The effect of salicylate on oxygen uptake of the tubercle bacillus. 1940;**92**:204.
17. Lehman J. Para-aminosalicylic acid in the treatment of tuberculosis. *Lancet* 1946;**250**:15–6.
18. Klip W. The tuberculostatic action of vitamin D2. *Antonie Van Leeuwenhoek* 1952;**18**:217–26.
19. Youmans AS, Youmans GP. The effect of "Tween 80" in vitro on the bacteriostatic activity of twenty compounds for Mycobacterium tuberculosis. *J Bacteriol* 1948;**56**:245–52.
20. Smith MI. Chemotherapeutic testing in experimental tuberculosis. *Am Rev Tuberc* 1947;**56**:377–81.
21. Smith MI. Quantitative measurement of Mycobacterium tuberculosis. *Proc Soc Exp Biol Med* 1947;**64**:36–40.
22. Youmans GP. Test tube evaluation of tuberculostatic agents. *Am Rev Tuberc* 1947;**56**:376.
23. Davis BD, Dubos RJ. The binding of fatty acids by serum albumin, a protective growth factor in bacteriological media. *J Exp Med* 1947;**86**:215–28.
24. Youmans AS, Youmans GP. Studies on the metabolism of Mycobacterium tuberculosis. IV. The effect of fatty acids on the growth of M. tuberculosis var. hominis. *J Bacteriol* 1954;**67**:731–3.
25. Bavin EM, Drain DJ, Seiler M, Seymour DE. Some further studies on tuberculostatic compounds. *J Pharm Pharmacol* 1952;**4**:844–55.
26. Hobby GL, Lenert TF, Donikian M, Pikula D. The tuberculostatic activity of terramycin. *Am Rev Tuberc* 1951;**63**:434–40.
27. Sattler TH, Youmans GP. The effect of tween 80, bovine albumin, glycerol, and glucose on the growth of Mycobacterium tuberculosis var. hominis (H37Rv). *J Bacteriol* 1948;**56**:235–43.
28. Lehman J. On the effect of isomers of para-aminosalicylic acid and related substances on the tuberculostatic effect of PAS. *Experientia* 1949;**5**:365–7.
29. Beeuwkes H. On the use of an egg agar medium in the determination of the resistance of Mycobacterium tuberculosis to streptomycin. *Antonie Van Leeuwenhoek* 1950;**16**:359–65.
30. Van Der Stelt C, Voorspuij AJ, Nauta WT. Investigation into the tuberculostatic activity of some derivatives of p. amino-benzoic acid. *Antonie Van Leeuwenhoek* 1954;**20**:285–98.
31. Gootjes J, Zwart Voorspuij AJ, Nauta WT. Investigation into the tuberculostatic activity of some aromatic hydroxy compounds. *Antonie Van Leeuwenhoek* 1953;**19**:237–44.
32. Steenken Jr W, Wolinsky E. The tuberculostatic action of terramycin in vitro and in the experimental animal. *Ann N Y Acad Sci* 1950;**53**:309–18.
33. Steenken Jr W, Wolinsky E. Effects of antimicrobial agents on the tubercle bacillus and on experimental tuberculosis. *Am J Med* 1950;**9**:633–53. doi:0002-9343(50)90213-0.
34. Bavin EM, Rees RJ, Robson JM, Seiler M, Seymour DE, Suddaby D. The tuberculostatic activity of some thiosemicarbazones. *J Pharm Pharmacol* 1950;**2**:764–72.
35. Wolinsky E, Wetzel V, Steenken Jr W. Tuberculostatic effect of furacin in vitro and in vivo. *Proc Soc Exp Biol Med* 1949;**70**:483–7.
36. Bartz QR, Ehrlich J, Mold JD, Penner MA, Smith RM. Viomycin, a new tuberculostatic antibiotic. *Am Rev Tuberc* 1951;**63**:4–6.
37. Donovick R, Rake G, Titus EO. An approach to a correlation between in vivo and in vitro activities of antitubercular compounds. *Ann NY Acad Sci* 1950;**53**:50–8.
38. Muic N, Sudic D, Fles D. Tuberculostatic activity of some pyrrole-azo-derivatives in vitro. *Acta Med Iugosl* 1952;**6**:141–53.
39. Klip W, Emmerie A, Engel C. The tuberculostatic properties of the unsaponifiable fraction of cod liver oil. *Antonie Van Leeuwenhoek* 1952;**18**:227–37.
40. Abdel Kader MM, Zaki O. Tryptophane as a tuberculostatic agent. *Experientia* 1958;**14**:455–6.
41. Myrvik Q, Weiser RS. A tuberculostatic serum substance possessing lysozyme-like properties. *Am Rev Tuberc* 1951;**64**:669–74.
42. Fujita Y. Studies on antituberculous factors in the low molecular fractions of the sera of various animal species. I. Crude materials and their tuberculostatic effects. *Acta Tuberc Jpn* 1960;**10**:67–74.
43. Bjornesjo KB. Tuberculostatic factor in normal human urine. *Am Rev Tuberc* 1956;**73**:967.
44. Dubos RJ. A tuberculostatic agent present in animal tissues. *Am Rev Tuberc* 1951;**63**:119.
45. Bjornesjo KB. On the occurrence of tuberculostatic substances in bovine organ estracts. *Acta Tuberc Scand* 1952;**27**:134–9.
46. Bloom WL, Hudgins PC, Cummings MM. Tuberculostatic substances extracted from normal rat spleen. *J Infect Dis* 1953;**92**:70–3.
47. Patnode RA, Hudgins PC. Tuberculostatic effect of calf lung fatty acids on human and bovine tubercle bacilli. *Am Rev Tuberc* 1957;**75**:630–7.
48. Myrvik QN, Soto-Figueroa E. A new tuberculostatic protein isolated from bovine spleen. *Am Rev Tuberc* 1958;**78**:93–100.
49. Ballon HC, Guernon A. Tuberculostatic activity of certain products of growth of penicillin notatum. *Can Med Assoc J* 1950;**62**:277.
50. Jenkins DE. Mycomycin: a new antibiotic with tuberculostatic properties. *J Lab Clin Med* 1950;**36**:841–2.
51. Hobby GL, Lenert TF, Donikian M, Pikula D. The activity of viomycin against Mycobacterium tuberculosis and other microorganisms in vitro and in vivo. *Am Rev Tuberc* 1951;**63**:17–24.
52. Lee HF, Sewell E, Mervine D. Study of culture filtrates of pathogenic fungi and yeasts for tuberculostatic properties. *Am Rev Tuberc* 1952;**66**:623–5.
53. Yegian D, Budd V. Novobiocin: activity in vitro and in experimental tuberculosis. *Am Rev Tuberc* 1957;**76**:272–8.
54. Uesaka I, Yoshioka H, Wakuda E. On the antibiotic activity of Mycobacteria and Nocardia; preliminary report on a tuberculostatic substance contained in bacteria cell. *Acta Tuberc Jpn* 1954;**4**:1–6.
55. Uesaka I, Yoshioka H, Wakuda E. On the antibiotic activity of mycobacteria. I. The existence of a tuberculostatic substance in the bacterial cell body. *J Antibiot (Tokyo)* 1955;**8**:1–4.
56. Middlebrook G, Reggiardo Z, Tigertt WD. Automatable radiometric detection of growth of Mycobacterium tuberculosis in selective media. *Am Rev Respir Dis* 1977;**115**:1066–9.
57. Siddiqi SH, Libonati JP, Middlebrook G. Evaluation of rapid radiometric method for drug susceptibility testing of Mycobacterium tuberculosis. *J Clin Microbiol* 1981;**13**:908–12.
58. Snider Jr DE, Good RC, Kilburn JO, Laskowski Jr LF, Lusk RH, Marr JJ, Reggiardo Z, Middlebrook G. Rapid drug-susceptibility testing of Mycobacterium tuberculosis. *Am Rev Respir Dis* 1981;**123**:402–6.
59. Piersimoni C, Olivieri A, Benacchio L, Scarparo C. Current perspectives on drug susceptibility testing of Mycobacterium tuberculosis complex: the automated nonradiometric systems. *J Clin Microbiol* 2006;**44**:20–8. doi:44/1/20.
60. Primm T, Franzblau SG. Recent advances in methodologies for the discovery of antimycobacterial drugs. *Curr Bioact Compd* 2007;**3**:201–8.
61. Collins L, Franzblau SG. Microplate alamar blue assay versus BACTEC 460 system for high-throughput screening of compounds against Mycobacterium tuberculosis and Mycobacterium avium. *Antimicrob Agents Chemother* 1997;**41**:1004–9.
62. Arain TM, Resconi AE, Hickey MJ, Stover CK. Bioluminescence screening in vitro (Bio-Siv) assays for high-volume antimycobacterial drug discovery. *Antimicrob Agents Chemother* 1996;**40**:1536–41.
63. Collins LA, Torrero MN, Franzblau SG. Green fluorescent protein reporter microplate assay for high-throughput screening of compounds against Mycobacterium tuberculosis. *Antimicrob Agents Chemother* 1998;**42**:344–7.
64. Altaf M, Miller CH, Bellows DS, O'Toole R. Evaluation of the Mycobacterium smegmatis and BCG models for the discovery of Mycobacterium tuberculosis inhibitors. *Tuberculosis (Edinb)* 2010;**90**:333–7. doi:S1472-9792(10)00103-4.
65. Pethe K, Sequeira PC, Agarwalla S, Rhee K, Kuhen K, Phong WY, Patel V, Beer D, Walker JR, Duraiswamy J, Jiricek J, Keller TH, Chatterjee A, Tan MP, Ujjini M, Rao SP, Camacho L, Bifani P, Mak PA, Ma I, Barnes SW, Chen Z, Plouffe D, Thayalan P, Ng SH, Au M, Lee BH, Tan BH, Ravindran S, Nanjundappa M, Lin X, Goh A, Lakshminarayana SB, Shoen C, Cynamon M, Kreiswirth B, Dartois V, Peters EC, Glynne R, Brenner S, Dick T. A chemical genetic screen in Mycobacterium tuberculosis identifies carbon-source-dependent growth inhibitors devoid of in vivo efficacy. *Nat Commun* 2010;**1**:1–8. doi:ncomms1060.
66. Ojha AK, Baughn AD, Sambandan D, Hsu T, Trivelli X, Guerardel Y, Alahari A, Kremer L, Jacobs Jr WR, Hatfull GF. Growth of Mycobacterium tuberculosis biofilms containing free mycolic acids and harbouring drug-tolerant bacteria. *Mol Microbiol* 2008;**69**:164–74. doi:MMI6274.
67. Sala C, Dhar N, Hartkoorn RC, Zhang M, Ha YH, Schneider P, Cole ST. Simple model for testing drugs against nonreplicating Mycobacterium tuberculosis. *Antimicrob Agents Chemother* 2010;**54**:4150–8. doi:AAC.00821-10.
68. Christophe T, Jackson M, Jeon HK, Fenistein D, Contreras-Dominguez M, Kim J, Genovesio A, Carralot JP, Ewann F, Kim EH, Lee SY, Kang S, Seo MJ, Park EJ, Skovierova H, Pham H, Riccardi G, Nam JY, Marsollier L, Kempf M, Joly-Guillou ML, Oh T, Shin WK, No Z, Nehrbass U, Brosch R, Cole ST, Brodin P. High content screening identifies decaprenyl-phosphoribose 2′ epimerase as a target for intracellular antimycobacterial inhibitors. *PLoS Pathog* 2009;**5**:e1000645. http://dx.doi.org/10.1371/journal.ppat.1000645.
69. Wallis RS, Palaci M, Vinhas S, Hise AG, Ribeiro FC, Landen K, Cheon SH, Song HY, Phillips M, Dietze R, Ellner JJ. A whole blood bactericidal assay for tuberculosis. *J Infect Dis* 2001;**183**:1300–3. doi:10.1086/319679.
70. Janulionis E, Sofer C, Song HY, Wallis RS. Lack of activity of orally administered clofazimine against intracellular Mycobacterium tuberculosis in whole-blood culture. *Antimicrob Agents Chemother* 2004;**48**:3133–5. http://dx.doi.org/10.1128/AAC.48.8.3133-3135.2004.
71. Wallis RS, Jakubiec W, Mitton-Fry M, Ladutko L, Campbell S, Paige D, Silvia A, Miller PF. Rapid evaluation in whole blood culture of regimens for XDR-TB containing PNU-100480 (sutezolid), TMC207, PA-824, SQ109, and-pyrazinamide. *PloS One* 2012;**7**:e30479. http://dx.doi.org/10.1371/journal.pone.0030479.
72. Lorian V, Finland M. In vitro effect of rifampin on mycobacteria. *Appl Microbiol* 1969;**17**:202–7.
73. Zygmunt WA. Antagonism of D-Cycloserine inhibition of mycobacterial growth by D-Alanine. *J Bacteriol* 1963;**85**:1217–20.
74. Lenaerts AJ, Gruppo V, Marietta KS, Johnson CM, Driscoll DK, Tompkins NM, Rose JD, Reynolds RC, Orme IM. Preclinical testing of the nitroimidazopyran PA-824 for activity against Mycobacterium tuberculosis in a series of in vitro and in vivo models. *Antimicrob Agents Chemother* 2005;**49**:2294–301. doi:49/6/2294.
75. Cho SH, Warit S, Wan B, Hwang CH, Pauli GF, Franzblau SG. Low-oxygen-recovery assay for high-throughput screening of compounds against non-replicating Mycobacterium tuberculosis. *Antimicrob Agents Chemother* 2007;**51**:1380–5. http://dx.doi.org/10.1128/AAC.00055-06.

76. Gengenbacher M, Rao SP, Pethe K, Dick T. Nutrient-starved, non-replicating Mycobacterium tuberculosis requires respiration, ATP synthase and isocitrate lyase for maintenance of ATP homeostasis and viability. *Microbiology* 2010;**156**:81–7. http://dx.doi.org/10.1099/mic.0.033084-0.
77. Sulochana S, Mitchison DA, Kubendiren G, Venkatesan P, Paramasivan CN. Bactericidal activity of moxifloxacin on exponential and stationary phase cultures of Mycobacterium tuberculosis. *J Chemother* 2009;**21**:127–34.
78. de Carvalho LP, Lin G, Jiang X, Nathan C. Nitazoxanide kills replicating and nonreplicating Mycobacterium tuberculosis and evades resistance. *J Med Chem* 2009;**52**:5789–92. http://dx.doi.org/10.1021/jm9010719.
79. Andreu N, Fletcher T, Krishnan N, Wiles S, Robertson BD. Rapid measurement of antituberculosis drug activity in vitro and in macrophages using bioluminescence. *J Antimicrob Chemother* 2012;**67**:404–14. http://dx.doi.org/10.1093/jac/dkr472.
80. Young D. Animal models of tuberculosis. *Eur J Immunol* 2009;**39**:2011–4. http://dx.doi.org/10.1002/eji.200939542.
81. Lenaerts AJ, Degroote MA, Orme IM. Preclinical testing of new drugs for tuberculosis: current challenges. *Trends Microbiol* 2008;**16**:48–54. doi:S0966-842X(07)00249-1.
82. Baldwin ER, Petroff SA, Gardner LS, editors. *Tuberculosis: bacteriology, pathology, and laboratory diagnosis*. Philadelphia: Lea and Febiger; 1927.
83. Corper HJ. Studies in the biochemistry and chemotherapy of tuberculosis. XII. The action of sodium thiocyanate in tuberculosis. *J Infect Dis* 1915;**16**:38–46.
84. Schroder G. The treatment of tuberculosis with silicylricinolic-acid ethyl ester (Silogran). 1936;**88**:248–53.
85. Soltys MA, Jennings AR. The dissemination of tubercle bacilli in experimental tuberculosis in the guinea pig. *Am Rev Tuberc* 1950;**61**:399–406.
86. Feldman W. Experimental evaluation of antibacterial agents in tuberculosis therapy. *Irish J Med Sci* 1951:505–27.
87. Youmans GP, Raleigh GW. The use of mice in experimental chemotherapy of tuberculosis; the histopathologic assay of chemotherapeutic action. *J Infect Dis* 1948;**82**:221–5.
88. Raleigh GW, Youmans GP. The use of mice in experimental chemotherapy of tuberculosis; pathology and pathogenesis. *J Infect Dis* 1948;**82**:205–20.
89. Raleigh GW, Youmans GP. The use of mice in experimental chemotherapy of tuberculosis; rationale and review of the literature. *J Infect Dis* 1948;**82**:197–204.
90. McCune RM, Feldmann FM, Lambert HP, McDermott W. Microbial persistence. I. The capacity of tubercle bacilli to survive sterilization in mouse tissues. *J Exp Med* 1966;**123**:445–68.
91. McCune RM, Feldmann FM, McDermott W. Microbial persistence. II. Characteristics of the sterile state of tubercle bacilli. *J Exp Med* 1966;**123**:469–86.
92. McCune Jr RM, McDermott W, Tompsett R. The fate of Mycobacterium tuberculosis in mouse tissues as determined by the microbial enumeration technique. II. The conversion of tuberculous infection to the latent state by the administration of pyrazinamide and a companion drug. *J Exp Med* 1956;**104**:763–802.
93. McCune Jr RM, Tompsett R. Fate of Mycobacterium tuberculosis in mouse tissues as determined by the microbial enumeration technique. I. The persistence of drug-susceptible tubercle bacilli in the tissues despite prolonged antimicrobial therapy. *J Exp Med* 1956;**104**:737–62.
94. Youmans GP, Williston EH. The effect of streptomycin on well-established experimental tuberculosis of mice. *Proc Soc Exp Biol Med* 1948;**68**:661–4.
95. Middlebrook G, Cohn ML. Bacteriology of tuberuclosis: laboratory methods. *Am J Public Health* 1958;**48**:844–53.
96. Middlebrook G, Cohn ML, Dye WE, Russell WF, Levy D. Microbiologic procedures of value in tuberculosis. *Acta Tuberc Scand* 1960;**38**:66–81.
97. Acred P, Brown DM, Wright D. The antitubercular activity of 3-acetylthiopropylene sulphide and 3-(2-furoylthio) propylene sulphide. *Br J Pharmacol Chemother* 1960;**15**:496–9.
98. Feldman WH, Hinshaw HC. Chemotherapeutic testing in experimental tuberculosis. Suggested outline of laboratory procedures for testing antituberculous substances in experimentally infected animals. *Am Rev Tuberc* 1945;**51**:582–91.
99. Karlson AG, Gainer JH, Feldman WH. The effect of neomycin on tuberculosis in guinea pigs infected with streptomycin-resistant tubercle bacilli. *Am Rev Tuberc* 1950;**62**:345–52.
100. Hunter P, editor. *Experimental models in antimicrobial chemotherapy*. Academic Press; 1986.
101. Pierce C, Dubos RJ, Middlebrook G. Infection of mice with Mammalian tubercle bacilli grown in Tween-albumin liquid medium. *J Exp Med* 1947;**86**:159–74.
102. Woods AC, Burky EL. Studies in experimental ocular tuberculosis. The effect of promin and promizole in experimental ocular tuberculosis in the immune-allergic rabbit. *Trans Am Ophthalmol Soc* 1947;**45**:170–83.
103. Woods AC, Burky EL. Effect of promin and promizole in experimental ocular tuberculosis in the normal rabbit. *Arch Ophthalmol* 1948;**40**:1–13.
104. Francis J, Spinks A, Stewart GT. The toxic and antituberculous effects of two thiosemicarbazones and streptomycin in dogs, monkeys, and guineapigs. *Br J Pharmacol Chemother* 1950;**5**:549–64.
105. Schmidt LH. Studies on the antituberculous activity of ethambutol in monkeys. *Ann NY Acad Sci* 1966;**135**:747–58.
106. Kim TH, Kubica GP. Long-term preservation and storage of mycobacteria. *Appl Microbiol* 1972;**24**:311–7.
107. Kubica GP, Kim TH, Dunbar FP. Designation of H37Rv as the neotype of Mycobacterium tuberculosis. *Int J Syst Bacteriol* 1972;**22**:99–106.
108. Steenken W, Oatway WH, Petroff SA. Biological studies of the tubercle bacillus: Iii. Dissociation and Pathogenicity of the R and S Variants of the human tubercle bacillus (H(37)). *J Exp Med* 1934;**60**:515–40.
109. Trudeau Institute. Catalog of the Trudeau mycobacterial culture collection.
110. Feldman WH. Tuberculosis chemotherapy: reminiscence of in vivo research. *Scand J Respir Dis* 1969;**50**:186–96.
111. Rook GA, Stanford JL. The Koch phenomenon and the immunopathology of tuberculosis. *Curr Top Microbiol Immunol* 1996;**215**:239–62.
112. Sakula A. Centenary of the discovery of the tubercle bacillus. *Lancet* 1982;**1**:750. doi:S0140-6736(82)92671-X.
113. Sakula A. Robert Koch: centenary of the discovery of the tubercle bacillus, 1882. *Thorax* 1982;**37**:246–51.
114. Brock T, editor. *Robert Koch: a life in medicine and bacteriology*. American Society of Microbiology; 1999.
115. Feldman WH, Hinshaw HC, Karlson AG. Frequency of administration of streptomycin; its influence on results of treatment of tuberculosis in guinea pigs. *Am Rev Tuberc* 1947;**55**:435–43.
116. Feldman WH, Hinshaw HC, Mann FC. Studies in streptomycin in tuberculosis; the action of streptomycin in experimental tuberculosis. *Trans Annu Meet Natl Tuberc Assoc* 1947;**42**:64.
117. Feldman WH, Karlson AG, Hinshaw HC. Streptomycin in experimental tuberculosis: the effects in guinea pigs following infection in intravenous inoculation. *Am Rev Tuberc* 1947;**56**:346–59.
118. Karlson AG, Pfuetze KH, Carr DT, Feldman WH, Hinshaw HC. The effect of combined therapy with streptomycin, para-aminosalicylic acid and promin on the emergence of streptomycin-resistant strains of tubercle bacilli; a preliminary report. *Proc Staff Meet Mayo Clin* 1949;**24**:85–8.
119. Karlson AG. Streptovaricin: therapeutic effect on guinea pigs infected with tubercle bacilli resistant to streptomycin and to para-aminosalicylic acid. *Proc Staff Meet Mayo Clin* 1958;**33**:193–6.
120. Feldman WH, Hinshaw HC. Sulfapyridine in experimental tuberculosis. *Am Rev Tuberc* 1940;**41**:732–50.
121. Smith DG, Waksman SA. Tuberculostatic and tuberculocidal properties of streptomycin. *J Bacteriol* 1947;**54**:253–61.
122. Feldman WH, Hinshaw HC. Streptomycin; a valuable anti-tuberculosis agent. *Br Med J* 1948;**1**:87–92.
123. Hart PD. Chemotherapy of tuberculosis; research during the past 100 years. *Br Med J* 1946;**2**:805 [passim].
124. Ramakrishnan CV, Bhatia AL, Fox W, Mitchison DA, Radhakrishna S, Selkon JB, Subbaiah TV, Velu S, Wallace JG. The virulence in the guinea-pig of tubercle bacilli isolated before treatment from South Indian patients with pulmonary tuberculosis. 3. Virulence related to pretreatment status of disease and to response to chemotherapy. *Bull World Health Organ* 1961;**25**:323–38.
125. Subbaiah TV, Bhatia AL, Geraghty E, Mitchison DA, Radhakrishna S, Selkon JB. The virulence in the guinea-pig of isoniazid-sensitive tubercle bacilli isolated from South Indian patients before treatment and after three months of chemotherapy. *Bull World Health Organ* 1961;**25**:759–64.
126. Dickinson JM, Mitchison DA. Experimental investigations of bacteriological mechanisms in short course chemotherapy. *Bull Int Union Against Tuberc* 1976;**51**:79–86.
127. Dickinson JM, Mitchison DA. Bactericidal activity in vitro and in the guinea-pig of isoniazid, rifampicin and ethambutol. *Tubercle* 1976;**57**:251–8.
128. Adams JM, Kalajan VA, Mork BO, Rosenblatt M, Rothrock WJ, O'Loughlin BJ. Reversal of tuberculin reaction in early tuberculosis. *Dis Chest* 1959;**35**:348–56.
129. Waksman SA. Tenth anniversary of the discovery of streptomycin, the first chemotherapeutic agent found to be effective against tuberculosis in humans. *Am Rev Tuberc* 1954;**70**:1–8.
130. Wainwright M. Streptomycin: discovery and resultant controversy. *Hist Philos Life Sci* 1991;**13**:97–124.
131. McDermott W. The story of INH. *J Infect Dis* 1969;**119**:678–83.
132. Chorine V. Action de l'amide nicotinique sur les bacilles du genre mycobacterium. *Compt Rend Hebdomadaires Séances Acad Sci* 1945;**220**:150–1.
133. Bonfiglioli H, Prado JM. Effect of isonicotinic acid hydrazide on experimental tuberculosis in the guinea pig. *Rev Asoc Med Argent* 1952;**66**:376–7.
134. Dessau FI, Yeager RL, Burger FJ, Williams JH. Pyrazinamide (aldinamide) in experimental tuberculosis of the guinea pig. *Am Rev Tuberc* 1952;**65**:519–22.
135. Karlson AG, Feldman WH. Isoniazid in experimental tuberculosis of guinea pigs infected with tubercle bacilli resistant to streptomycin and to para-aminosalicylic acid. *Am Rev Tuberc* 1952;**66**:477–85.
136. Nakamura RM, Kanai K, Murohashi T. Effectiveness of capreomycin against experimental tuberculosis of mice and guinea pigs. *Kekkaku* 1964;**39**:161–5.
137. Dubos RJ. The effect of lipids and serum albumin on bacterial growth. *J Exp Med* 1947;**85**:9–22.
138. Cohn ML, Davis CL, Middlebrook G. Airborne immunization against tuberculosis. *Science* 1958;**128**:1282–3.
139. Middlebrook G. An apparatus for airborne infection of mice. *Proc Soc Exp Biol Med* 1952;**80**:105–10.
140. Basaraba RJ. Experimental tuberculosis: the role of comparative pathology in the discovery of improved tuberculosis treatment strategies. *Tuberculosis (Edinb)* 2008;**88**(Suppl. 1):S35–47. doi:S1472-9792(08)70035-0.

141. Youmans GP, Youmans AS. The assessment of antituberculous chemotherapeutic activity in mice, using virulent human-type tubercle bacilli. *Am Rev Tuberc* 1951;**64**:541–50.
142. Youmans GP, Youmans AS. The relation between the size of the infecting dose of tubercle bacilli and the survival time of mice. *Am Rev Tuberc* 1951;**64**:534–40.
143. Grumbach F, Canetti G, Grosset J. Further experiments on long-term chemotherapy of advanced murine tuberculosis, with emphasis on intermittent regimes. *Tubercle* 1964;**45**:125–35.
144. Lecoeur HF, Lagrange PH, Truffot-Pernot C, Gheorghiu M, Grosset J. Relapses after stopping chemotherapy for experimental tuberculosis in genetically resistant and susceptible strains of mice. *Clin Exp Immunol* 1989;**76**:458–62.
145. Dickinson JM, Mitchison DA. In vitro and in vivo studies to assess the suitability of antituberculous drugs for use in intermittent chemotherapy regimens. *Tubercle* 1968;**49**(Suppl.):66–70.
146. Mitchison DA. Treatment of tuberculosis. The Mitchell lecture 1979. J Roy Coll Phys Lond;**14**:91–5.
147. Grumbach F. Experimental basis for the chemotherapy of tuberculosis. *Antimicrob Agents Chemother* 1965;**5**:1058–64.
148. Grumbach F, Canetti G, Grosset J, le Lirzin M. Late results of long-term intermittent chemotherapy of advanced, murine tuberculosis: limits of the murine model. *Tubercle* 1967;**48**:11–26.
149. McDermott W, McCune Jr RM, Tompsett R. Dynamics of antituberculous chemotherapy. *Am Rev Tuberc* 1956;**74**:100–8.
150. McCune RM, Deuschle K, Jordahl C, Des Prez R, Muschenheim C, McDermott W. The influence of streptovaricin used alone and with isoniazid in an experimental tuberculous infection in animals, and some clinical observations. *Am Rev Tuberc* 1957;**75**:659–66.
151. Heffernan JF. Proceedings: latest results of East African/British Medical Research Council short course chemotherapy studies. *Tubercle* 1975;**56**:165.
152. Mitchison DA, Chang KC. Experimental models of tuberculosis: can we trust the mouse? *Am J Respir Crit Care Med* 2009;**180**:201–2. doi:180/3/201.
153. Burman WJ, Goldberg S, Johnson JL, Muzanye G, Engle M, Mosher AW, Choudhri S, Daley CL, Munsiff SS, Zhao Z, Vernon A, Chaisson RE. Moxifloxacin versus ethambutol in the first 2 months of treatment for pulmonary tuberculosis. *Am J Respir Crit Care Med* 2006;**174**:331–8. doi:200603-360OC.
154. Dorman SE, Johnson JL, Goldberg S, Muzanye G, Padayatchi N, Bozeman L, Heilig CM, Bernardo J, Choudhri S, Grosset JH, Guy E, Guyadeen P, Leus MC, Maltas G, Menzies D, Nuermberger EL, Villarino M, Vernon A, Chaisson RE. Substitution of moxifloxacin for isoniazid during intensive phase treatment of pulmonary tuberculosis. *Am J Respir Crit Care Med* 2009;**180**:273–80. doi:200901-0078OC.
155. Grosset J, Truffot-Pernot C, Lacroix C, Ji B. Antagonism between isoniazid and the combination pyrazinamide-rifampin against tuberculosis infection in mice. *Antimicrob Agents Chemother* 1992;**36**:548–51.
156. Almeida D, Nuermberger E, Tasneen R, Rosenthal I, Tyagi S, Williams K, Peloquin C, Grosset J. Paradoxical effect of isoniazid on the activity of rifampin-pyrazinamide combination in a mouse model of tuberculosis. *Antimicrob Agents Chemother* 2009;**53**:4178–84. doi:10.1128/AAC.00830-09.
157. Lenaerts AM, Chase SE, Chmielewski AJ, Cynamon MH. Evaluation of rifapentine in long-term treatment regimens for tuberculosis in mice. *Antimicrob Agents Chemother* 1999;**43**:2356–60.
158. Burman WJ, Reves RR. Hepatotoxicity from rifampin plus pyrazinamide: lessons for policymakers and messages for care providers. *Am Journal Respir Crit Care Med* 2001;**164**:1112–3.
159. Zhang T, Zhang M, Rosenthal IM, Grosset JH, Nuermberger EL. Short-course therapy with daily rifapentine in a murine model of latent tuberculosis infection. *Am J Respir Crit Care Med* 2009;**180**:1151–7. doi:200905-0795OC.
160. Lenaerts AJ, Johnson CM, Marrieta KS, Gruppo V, Orme IM. Significant increases in the levels of liver enzymes in mice treated with anti-tuberculosis drugs. *Int Journal Antimicrob Agents* 2005;**26**:152–8. http://dx.doi.org/10.1016/j.ijantimicag.2005.04.011.
161. Zhang M, Li SY, Rosenthal IM, Almeida DV, Ahmad Z, Converse PJ, Peloquin CA, Nuermberger EL, Grosset JH. Treatment of tuberculosis with Rifamycin-containing regimens in immune deficient mice. *Am J Respir Crit Care Med* 2011. http://dx.doi.org/10.1164/rccm.201012–201949OC.
162. Veziris N, Ibrahim M, Lounis N, Andries K, Jarlier V. Sterilizing activity of second-line regimens containing TMC207 in a murine model of tuberculosis. *PLoS One* 2011;**6**:e17556. http://dx.doi.org/10.1371/journal.pone.0017556.
163. Shang S, Shanley CA, Caraway ML, Orme EA, Henao-Tamayo M, Hascall-Dove L, Ackart D, Lenaerts AJ, Basaraba RJ, Orme IM, Ordway DJ. Activities of TMC207, rifampin, and pyrazinamide against Mycobacterium tuberculosis infection in guinea pigs. *Antimicrob Agents Chemother* 2011;**55**:124–31. doi:10.1128/AAC.00978-10.
164. De Groote MA, Gilliland JC, Wells CL, Brooks EJ, Woolhiser LK, Gruppo V, Peloquin CA, Orme IM, Lenaerts AJ. Comparative studies evaluating mouse models used for efficacy testing of experimental drugs against Mycobacterium tuberculosis. *Antimicrob Agents Chemother* 2011;**55**:1237–47. doi:10.1128/AAC.00595-10.
165. Ioerger TR, Feng Y, Ganesula K, Chen X, Dobos KM, Fortune S, Jacobs WR Jr., Mizrahi V, Parish T, Rubin E, Sassetti C, Sacchettini JC. Variation among genome sequences of H37Rv strains of Mycobacterium tuberculosis from multiple laboratories. J Bacteriol;**192**:3645–53. doi:JB.00166-10.
166. Andries K, Gevers T, Lounis N. Bactericidal potencies of new regimens are not predictive of their sterilizing potencies in a murine model of tuberculosis. *Antimicrob Agents Chemother* 2010;**54**:4540–4. http://dx.doi.org/10.1128/AAC.00934-10.
167. Pan H, Yan BS, Rojas M, Shebzukhov YV, Zhou H, Kobzik L, Higgins DE, Daly MJ, Bloom BR, Kramnik I. Ipr1 gene mediates innate immunity to tuberculosis. *Nature* 2005;**434**:767–72. http://dx.doi.org/10.1038/nature03419.
168. Harper J, Skerry C, Davis SL, Tasneen R, Weir M, Kramnik I, Bishai WR, Pomper MG, Nuermberger EL, Jain SK. Mouse model of necrotic tuberculosis granulomas develops hypoxic lesions. *J Infect Dis* 2012;**205**:595–602. http://dx.doi.org/10.1093/infdis/jir786.
169. Driver ER, Ryan GR, Hoff DR, Irwin SM, Basaraba R, Kramnik I, Lenaerts AJ. Evaluation of mouse model forming necrotic granulomas using C3HeB/FeJ mice, for the testing of M. tuberculosis drugs. *Antimicrob Agents Chemother* 2012.
170. Sissons J, Yan BS, Pichugin AV, Kirby A, Daly MJ, Kramnik I. Multigenic control of tuberculosis resistance: analysis of a QTL on mouse chromosome 7 and its synergism with sst1. *Genes Immun* 2009;**10**:37–46. http://dx.doi.org/10.1038/gene.2008.68.
171. Davis SL, Nuermberger EL, Um PK, Vidal C, Jedynak B, Pomper MG, Bishai WR, Jain SK. Noninvasive pulmonary [18F]-2-fluoro-deoxy-D-glucose positron emission tomography correlates with bactericidal activity of tuberculosis drug treatment. *Antimicrob Agents Chemother* 2009;**53**:4879–84. doi:10.1128/AAC.00789-09.
172. Dubos RJ. Gordon Wilson lecture: the experimental analysis of tuberculous infections. *Trans Am Clin Climatol Assoc* 1946;**58**:117–36.
173. Capuano 3rd SV, Croix DA, Pawar S, Zinovik A, Myers A, Lin PL, Bissel S, Fuhrman C, Klein E, Flynn JL. Experimental Mycobacterium tuberculosis infection of cynomolgus macaques closely resembles the various manifestations of human M. tuberculosis infection. *Infect Immun* 2003;**71**:5831–44.
174. Hoff DR, Ryan GJ, Driver ER, Ssemakulu CC, De Groote MA, Basaraba RJ, Lenaerts AJ. Location of intra- and extracellular M. tuberculosis populations in lungs of mice and guinea pigs during disease progression and after drug treatment. *PLoS One* 2011;**6**:e17550. http://dx.doi.org/10.1371/journal.pone.0017550.
175. Grosset J. Mycobacterium tuberculosis in the extracellular compartment: an underestimated adversary. *Antimicrob Agents Chemother* 2003;**47**:833–6.
176. Nuermberger EL, Yoshimatsu T, Tyagi S, Williams K, Rosenthal I, O'Brien RJ, Vernon AA, Chaisson RE, Bishai WR, Grosset JH. Moxifloxacin-containing regimens of reduced duration produce a stable cure in murine tuberculosis. *Am J Respir Crit Care Med* 2004;**170**:1131–4. http://dx.doi.org/10.1164/rccm.200407-885OC.
177. Adams KN, Takaki K, Connolly LE, Wiedenhoft H, Winglee K, Humbert O, Edelstein PH, Cosma CL, Ramakrishnan L. Drug tolerance in replicating mycobacteria mediated by a macrophage-induced efflux mechanism. *Cell* 2011;**145**:39–53. http://dx.doi.org/10.1016/j.cell.2011.02.022.
178. Gil O, Diaz I, Vilaplana C, Tapia G, Diaz J, Fort M, Caceres N, Pinto S, Cayla J, Corner L, Domingo M, Cardona PJ. Granuloma encapsulation is a key factor for containing tuberculosis infection in minipigs. *PLoS One* 2010;**5**:e10030. http://dx.doi.org/10.1371/journal.pone.0010030.
179. Via LE, Lin PL, Ray SM, Carrillo J, Allen SS, Eum SY, Taylor K, Klein E, Manjunatha U, Gonzales J, Lee EG, Park SK, Raleigh JA, Cho SN, McMurray DN, Flynn JL, Barry 3rd CE. Tuberculous granulomas are hypoxic in guinea pigs, rabbits, and nonhuman primates. *Infect Immun* 2008;**76**:2333–40. doi:IAI.01515-07.
180. Lurie MB. The use of the rabbit in experimental chemotherapy of tuberculosis. *Ann NY Acad Sci* 1949;**52**:627–36. illust.
181. Dorman SE, Hatem CL, Tyagi S, Aird K, Lopez-Molina J, Pitt ML, Zook BC, Dannenberg Jr AM, Bishai WR, Manabe YC. Susceptibility to tuberculosis: clues from studies with inbred and outbred New Zealand White rabbits. *Infect Immun* 2004;**72**:1700–5.
182. Dannenberg Jr AM. Roles of cytotoxic delayed-type hypersensitivity and macrophage-activating cell-mediated immunity in the pathogenesis of tuberculosis. *Immunobiology* 1994;**191**:461–73.
183. Dannenberg Jr AM. Pathogenesis of pulmonary Mycobacterium bovis infection: basic principles established by the rabbit model. *Tuberculosis* 2001;**81**:87–96. doi:10.1054/tube.2000.0260.
184. Manabe YC, Kesavan AK, Lopez-Molina J, Hatem CL, Brooks M, Fujiwara R, Hochstein K, Pitt ML, Tufariello J, Chan J, McMurray DN, Bishai WR, Dannenberg Jr AM, Mendez S. The aerosol rabbit model of TB latency, reactivation and immune reconstitution inflammatory syndrome. *Tuberculosis (Edinb)* 2008;**88**:187–96. doi:S1472-9792(07)00131-X.
185. Tsenova L, Ellison E, Harbacheuski R, Moreira AL, Kurepina N, Reed MB, Mathema B, Barry 3rd CE, Kaplan G. Virulence of selected Mycobacterium tuberculosis clinical isolates in the rabbit model of meningitis is dependent on phenolic glycolipid produced by the bacilli. *J Infect Dis* 2005;**192**:98–106. doi:10.1086/430614.
186. Nedeltchev GG, Raghunand TR, Jassal MS, Lun S, Cheng QJ, Bishai WR. Extrapulmonary dissemination of Mycobacterium bovis but not Mycobacterium tuberculosis in a bronchoscopic rabbit model of cavitary tuberculosis. *Infect Immun* 2009;**77**:598–603. doi:10.1128/IAI.01132-08.
187. Helke KL, Mankowski JL, Manabe YC. Animal models of cavitation in pulmonary tuberculosis. *Tuberculosis* 2006;**86**:337–48. http://dx.doi.org/10.1016/j.tube.2005.09.001.
188. Prideaux B, Dartois V, Staab D, Weiner DM, Goh A, Via LE, Barry 3rd CE, Stoeckli M. High-sensitivity MALDI-MRM-MS imaging of moxifloxacin

distribution in tuberculosis-infected rabbit lungs and granulomatous lesions. *Anal Chem* 2011;**83**:2112–8. doi:10.1021/ac1029049.
189. Kjellsson MC, Via LE, Goh A, Weiner D, Low KM, Kern S, Pillai G, Barry 3rd CE, Dartois V. Pharmacokinetic evaluation of the penetration of antituberculosis agents in rabbit pulmonary lesions. *Antimicrob Agents Chemother* 2012;**56**:446–57. doi:10.1128/AAC.05208-11.
190. Manier M, Reyzer M, Goh A, Dartois V, Via L, Barry Cr, Caprioli R. Reagent precoated targets for rapid in-tissue derivatization of the anti-tuberculosis drug isoniazid followed by MALDI imaging mass spectrometry. *J Am Soc Mass Spectrom* 2011;**22**:1409–19.
191. Singhal A, Aliouat el M, Herve M, Mathys V, Kiass M, Creusy C, Delaire B, Tsenova L, Fleurisse L, Bertout J, Camacho L, Foo D, Tay HC, Siew JY, Boukhouchi W, Romano M, Mathema B, Dartois V, Kaplan G, Bifani P. Experimental tuberculosis in the Wistar rat: a model for protective immunity and control of infection. *PloS One* 2011;**6**:e18632. http://dx.doi.org/10.1371/journal.pone.0018632.
192. Elwood RL, Wilson S, Blanco JC, Yim K, Pletneva L, Nikonenko B, Samala R, Joshi S, Hemming VG, Trucksis M. The American cotton rat: a novel model for pulmonary tuberculosis. *Tuberculosis (Edinb)* 2007;**87**:145–54. doi:S1472-9792(06)00086-2.
193. Sugawara I, Yamada H, Mizuno S. Pathological and immunological profiles of rat tuberculosis. *Int J Exp Pathol* 2004;**85**:125–34. http://dx.doi.org/10.1111/j.0959-9673.2004.00379.x.
194. Sugawara I, Yamada H, Mizuno S. Pulmonary tuberculosis in spontaneously diabetic goto kakizaki rats. *Tohoku J Exp Med* 2004;**204**:135–45. doi:JST.JSTAGE/tjem/204.135.
195. Sugawara I, Yamada H, Mizuno S. Nude rat (F344/N-rnu) tuberculosis. *Cell Microbiol* 2006;**8**:661–7. doi:CMI658.
196. Gardner MB, Luciw PA. Macaque models of human infectious disease. *ILAR J* 2008;**49**:220–55.
197. Batten CJ, De Rose R, Wilson KM, Agy MB, Chea S, Stratov I, Montefiori DC, Kent SJ. Comparative evaluation of simian, simian-human, and human immunodeficiency virus infections in the pigtail macaque (Macaca nemestrina) model. *AIDS Res Hum Retroviruses* 2006;**22**:580–8. doi:10.1089/aid.2006.22.580.
198. Flynn JL. Lessons from experimental Mycobacterium tuberculosis infections. *Microbes Infect* 2006;**8**:1179–88. doi:S1286-4579(05)00457-0.
199. Lenaerts AJ, Hoff D, Aly S, Ehlers S, Andries K, Cantarero L, Orme IM, Basaraba RJ. Location of persisting mycobacteria in a Guinea pig model of tuberculosis revealed by r207910. *Antimicrob Agents Chemother* 2007;**51**:3338–45. doi:10.1128/AAC.00276-07.
200. Mitchison DA, Coates AR. Predictive in vitro models of the sterilizing activity of anti-tuberculosis drugs. *Curr Pharm Des* 2004;**10**:3285–95.
201. Dhillon J, Lowrie DB, Mitchison DA. Mycobacterium tuberculosis from chronic murine infections that grows in liquid but not on solid medium. *BMC Infect Dis* 2004;**4**:51. doi:1471-2334-4-51.
202. Lounis N, Veziris N, Chauffour A, Truffot-Pernot C, Andries K, Jarlier V. Combinations of R207910 with drugs used to treat multidrug-resistant tuberculosis have the potential to shorten treatment duration. *Antimicrob Agents Chemother* 2006;**50**:3543–7. doi:AAC.00766-06.
203. Tasneen R, Li SY, Peloquin CA, Taylor D, Williams KN, Andries K, Mdluli KE, Nuermberger EL. Sterilizing activity of novel TMC207- and PA-824-containing regimens in a murine model of tuberculosis. *Antimicrob Agents Chemother* 2011;**55**:5485–92. doi:10.1128/AAC.05293-11.
204. Cohn DL, Catlin BJ, Peterson KL, Judson FN, Sbarbaro JA. A 62-dose, 6-month therapy for pulmonary and extrapulmonary tuberculosis. A twice-weekly, directly observed, and cost-effective regimen. *Ann Int Med* 1990;**112**:407–15.
205. Srivastava S, Gumbo T. In vitro and in vivo modeling of tuberculosis drugs and its impact on optimization of doses and regimens. *Curr Pharmaceutical Design* 2011;**17**:2881–8.
206. Woolhiser LK, Hoff DR, Marietta KS, Orme IM, Lenaerts AJ. Testing of experimental compounds in a relapse model of tuberculosis using granulocyte-macrophage colony-stimulating factor gene-disrupted mice. *Antimicrob Agents Chemother* 2009;**53**:306–8. doi:10.1128/AAC.01346-07.
207. Flynn JL, Goldstein MM, Triebold KJ, Sypek J, Wolf S, Bloom BR. IL-12 increases resistance of BALB/c mice to Mycobacterium tuberculosis infection. *J Immunol* 1995;**155**:2515–24.
208. U.S. Food and Drug Administration, http://www.clinicaltrials.gov/ct2/show/NCT00864383?term=moxifloxacin&rank=1; 2010.
209. Nuermberger EL, Yoshimatsu T, Tyagi S, O'Brien RJ, Vernon AN, Chaisson RE, Bishai WR, Grosset JH. Moxifloxacin-containing regimen greatly reduces time to culture conversion in murine tuberculosis. *Am J Respir Crit Care Med* 2004;**169**:421–6. http://dx.doi.org/10.1164/rccm.200310-1380OC.
210. Ibrahim M, Truffot-Pernot C, Andries K, Jarlier V, Veziris N. Sterilizing activity of R207910 (TMC207)-containing regimens in the murine model of tuberculosis. *Am Journal Respir Crit Care Med* 2009;**180**:553–7. http://dx.doi.org/10.1164/rccm.200807-1152OC.
211. Rosenthal IM, Zhang M, Almeida D, Grosset JH, Nuermberger EL. Isoniazid or moxifloxacin in rifapentine-based regimens for experimental tuberculosis? *Am J Respir Crit Care Med* 2008;**178**:989–93. http://dx.doi.org/10.1164/rccm.200807-1029OC.
212. Dickinson JM, Aber VR, Mitchison DA. Bactericidal activity of streptomycin, isoniazid, rifampin, ethambutol, and pyrazinamide alone and in combination against Mycobacterium Tuberculosis. *Am Rev Respir Dis* 1977;**116**:627–35.
213. De Groote MA, Gruppo V, Woolhiser LK, Orme IM, Gilliland JC, Lenaerts AJ. Importance of confirming data on the in vivo efficacy of novel antibacterial drug regimens against various strains of Mycobacterium tuberculosis. *Antimicrob Agents Chemother* 2012;**56**:731–8. doi:10.1128/AAC.05701-11.
214. Ibrahim M, Andries K, Lounis N, Chauffour A, Truffot-Pernot C, Jarlier V, Veziris N. Synergistic activity of R207910 combined with pyrazinamide against murine tuberculosis. *Antimicrob Agents Chemother* 2007;**51**:1011–5. http://dx.doi.org/10.1128/AAC.00898-06.
215. Lounis N, Gevers T, Van Den Berg J, Andries K. Impact of the interaction of R207910 with rifampin on the treatment of tuberculosis studied in the mouse model. *Antimicrob Agents Chemother* 2008;**52**:3568–72. doi:AAC.00566-08.
216. Rosenthal IM, Williams K, Tyagi S, Peloquin CA, Vernon AA, Bishai WR, Grosset JH, Nuermberger EL. Potent twice-weekly rifapentine-containing regimens in murine tuberculosis. *Am J Respir Crit Care Med* 2006;**174**:94–101. doi:200602-280OC.
217. American type culture collection. www.atcc.org.
218. Munoz-Elias EJ, Timm J, Botha T, Chan WT, Gomez JE, McKinney JD. Replication dynamics of Mycobacterium tuberculosis in chronically infected mice. *Infect Immun* 2005;**73**:546–51. doi:73/1/546.
219. Gill WP, Harik NS, Whiddon MR, Liao RP, Mittler JE, Sherman DR. A replication clock for Mycobacterium tuberculosis. *Nat Med* 2009;**15**:211–4. doi:nm.1915.
220. Ahmad Z, Fraig MM, Bisson GP, Nuermberger EL, Grosset JH, Karakousis PC. Dose-dependent activity of pyrazinamide in animal models of intracellular and extracellular tuberculosis infections. *Antimicrob Agents Chemother* 2011;**55**:1527–32. http://dx.doi.org/10.1128/AAC.01524-10.
221. Eum SY, Kong JH, Jeon BY, Cho SN, Kim J, Via LE, Barry Iii CE, Koh WJ. Metaplastic ossification in the cartilage of the bronchus of a patient with chronic multi-drug resistant tuberculosis: a case report. *J Med Case Rep* 2010;**4**:156. http://dx.doi.org/10.1186/1752-1947-4-156.
222. Converse PJ, Eisenach KD, Theus SA, Nuermberger EL, Tyagi S, Ly LH, Geiman DE, Guo H, Nolan ST, Akar NC, Klinkenberg LG, Gupta R, Lun S, Karakousis PC, Lamichhane G, McMurray DN, Grosset JH, Bishai WR. The impact of mouse passaging of Mycobacterium tuberculosis strains prior to virulence testing in the mouse and guinea pig aerosol models. PLoS One;**5**:e10289. doi:10.1371/journal.pone.0010289.
223. Weiner B, Gomez J, Victor TC, Warren RM, Sloutsky A, Plikaytis BB, Posey JE, van Helden PD, Gey van Pittius NC, Koehrsen M, Sisk P, Stolte C, White J, Gagneux S, Birren B, Hung D, Murray M, Galagan J. Independent large scale duplications in multiple *M. tuberculosis* lineages overlapping the same genomic region. *PloS One* 2012;**7**:e26038. http://dx.doi.org/10.1371/journal.pone.0026038.
224. Portevin D, Gagneux S, Comas I, Young D. Human macrophage responses to clinical isolates from the Mycobacterium tuberculosis complex discriminate between ancient and modern lineages. *PLoS Pathog* 2011;**7**:e1001307. http://dx.doi.org/10.1371/journal.ppat.1001307.
225. Homolka S, Niemann S, Russell DG, Rohde KH. Functional genetic diversity among *Mycobacterium tuberculosis* complex clinical isolates: delineation of conserved core and lineage-specific transcriptomes during intracellular survival. *PLoS Pathog* 2010;**6**:e1000988. http://dx.doi.org/10.1371/journal.ppat.1000988.
226. Borrell S, Gagneux S. Strain diversity, epistasis and the evolution of drug resistance in *Mycobacterium tuberculosis*. *Clin Microbiol Infect Off Pub Eur Soc Clin Microbiol Infect Dis* 2011;**17**:815–20. http://dx.doi.org/10.1111/j.1469-0691.2011.03556.x.
227. Gagneux S. Host–pathogen coevolution in human tuberculosis. *Philos Trans Roy Soc Lond Ser B Biol Sci* 2012;**367**:850–9. http://dx.doi.org/10.1098/rstb.2011.0316.
228. Gagneux S, Burgos MV, DeRiemer K, Encisco A, Munoz S, Hopewell PC, Small PM, Pym AS. Impact of bacterial genetics on the transmission of isoniazid-resistant Mycobacterium tuberculosis. *PLoS Pathog* 2006;**2**:e61. http://dx.doi.org/10.1371/journal.ppat.0020061.
229. Lari N, Bimbi N, Rindi L, Tortoli E, Garzelli C. Genetic diversity of human isolates of Mycobacterium bovis assessed by spoligotyping and Variable Number Tandem Repeat genotyping. *Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis* 2011;**11**:175–80. http://dx.doi.org/10.1016/j.meegid.2010.09.004.
230. Gagneux S, Small PM. Global phylogeography of Mycobacterium tuberculosis and implications for tuberculosis product development. *Lancet Infect Dis* 2007;**7**:328–37. doi:S1473-3099(07)70108-1.
231. Gagneux S, DeRiemer K, Van T, Kato-Maeda M, de Jong BC, Narayanan S, Nicol M, Niemann S, Kremer K, Gutierrez MC, Hilty M, Hopewell PC, Small PM. Variable host-pathogen compatibility in Mycobacterium tuberculosis. *Proc Natl Acad Sci USA* 2006;**103**:2869–73. http://dx.doi.org/10.1073/pnas.0511240103.
232. Burman WJ, Bliven EE, Cowan L, Bozeman L, Nahid P, Diem L, Vernon A. Relapse associated with active disease caused by Beijing strain of Mycobacterium tuberculosis. *Emerg Infect Dis* 2009;**15**:1061–7. http://dx.doi.org/10.3201/eid1507.081253.
233. Nahid P, Bliven EE, Kim EY, Mac Kenzie WR, Stout JE, Diem L, Johnson JL, Gagneux S, Hopewell PC, Kato-Maeda M. Influence of M. tuberculosis lineage variability within a clinical trial for pulmonary tuberculosis. *PloS One* 2010;**5**:e10753. http://dx.doi.org/10.1371/journal.pone.0010753.
234. Palanisamy GS, Duteau N, Eisenach KD, Cave DM, Theus SA, Kreiswirth BN, Basaraba RJ, Orme IM. Clinical strains of Mycobacterium tuberculosis display

a wide range of virulence in guinea pigs. *Tuberculosis (Edinb)* 2009. doi:S1472-9792(09)00009-2.
235. Barczak AK, Domenech P, Boshoff HI, Reed MB, Manca C, Kaplan G, Barry 3rd CE. In vivo phenotypic dominance in mouse mixed infections with Mycobacterium tuberculosis clinical isolates. *J Infect Dis* 2005;**192**:600–6. doi:10.1086/432006.
236. Luo T, Yang C, Gagneux S, Gicquel B, Mei J, Gao Q. Combination of single nucleotide polymorphism and variable-number Tandem repeats for genotyping a Homogenous population of Mycobacterium tuberculosis Beijing strains in China. *J Clin Microbiol* 2012;**50**:633–9. http://dx.doi.org/10.1128/JCM.05539-11.
237. Ordway D, Henao-Tamayo M, Harton M, Palanisamy G, Troudt J, Shanley C, Basaraba RJ, Orme IM. The hypervirulent Mycobacterium tuberculosis strain HN878 induces a potent TH1 response followed by rapid down-regulation. *J Immunol* 2007;**179**:522–31. doi:179/1/522.
238. Ordway DJ, Shanley CA, Caraway ML, Orme EA, Bucy DS, Hascall-Dove L, Henao-Tamayo M, Harton MR, Shang S, Ackart D, Kraft SL, Lenaerts AJ, Basaraba RJ, Orme IM. Evaluation of standard chemotherapy in the guinea pig model of tuberculosis. *Antimicrob Agents Chemother* 2010;**54**:1820–33. http://dx.doi.org/10.1128/AAC.01521-09.
239. Cooper AM, Kaplan G, Derbyshire K. TB day summit in upstate New York: key issues to address. *Eur J Immunol* 2009;**39**:1443–7. http://dx.doi.org/10.1002/eji.200990031.
240. Ordway D, Palanisamy G, Henao-Tamayo M, Smith EE, Shanley C, Orme IM, Basaraba RJ. The cellular immune response to Mycobacterium tuberculosis infection in the guinea pig. *J Immunol* 2007;**179**:2532–41. doi:179/4/2532.
241. Stover CK, Warrener P, VanDevanter DR, Sherman DR, Arain TM, Langhorne MH, Anderson SW, Towell JA, Yuan Y, McMurray DN, Kreiswirth BN, Barry CE, Baker WR. A small-molecule nitroimidazopyran drug candidate for the treatment of tuberculosis. *Nature* 2000;**405**:962–6. doi:10.1038/35016103.
242. Kampmann B, Tena GN, Mzazi S, Eley B, Young DB, Levin M. Novel human in vitro system for evaluating antimycobacterial vaccines. *Infect Immun* 2004;**72**:6401–7. doi:72/11/6401.
243. Kong Y, Yao H, Ren H, Subbian S, Cirillo SL, Sacchettini JC, Rao J, Cirillo JD. Imaging tuberculosis with endogenous beta-lactamase reporter enzyme fluorescence in live mice. Proc Natl Acad Sci USA;**107**:12239–44. doi:1000643107.
244. Diacon AH, Maritz JS, Venter A, van Helden PD, Andries K, McNeeley DF, Donald PR. Time to detection of the growth of Mycobacterium tuberculosis in MGIT 960 for determining the early bactericidal activity of antituberculosis agents. Eur J Clin Microbiol Infect Dis. doi:10.1007/s10096-010-1043-7.
245. Pai SR, Actor JK, Sepulveda E, Hunter Jr RL, Jagannath C. Identification of viable and non-viable Mycobacterium tuberculosis in mouse organs by directed RT-PCR for antigen 85B mRNA. *Microb Pathog* 2000;**28**:335–42. http://dx.doi.org/10.1006/mpat.2000.0353.
246. Jayaram R, Gaonkar S, Kaur P, Suresh BL, Mahesh BN, Jayashree R, Nandi V, Bharat S, Shandil RK, Kantharaj E, Balasubramanian V. Pharmacokinetics-pharmacodynamics of rifampin in an aerosol infection model of tuberculosis. *Antimicrob Agents Chemother* 2003;**47**:2118–24.
247. Shandil RK, Jayaram R, Kaur P, Gaonkar S, Suresh BL, Mahesh BN, Jayashree R, Nandi V, Bharath S, Balasubramanian V. Moxifloxacin, ofloxacin, sparfloxacin, and ciprofloxacin against Mycobacterium tuberculosis: evaluation of in vitro and pharmacodynamic indices that best predict in vivo efficacy. *Antimicrob Agents Chemother* 2007;**51**:576–82. http://dx.doi.org/10.1128/AAC.00414-06.
248. Diacon AH, Dawson R, Hanekom M, Narunsky K, Maritz SJ, Venter A, Donald PR, van Niekerk C, Whitney K, Rouse DJ, Laurenzi MW, Ginsberg AM, Spigelman MK. Early bactericidal activity and pharmacokinetics of PA-824 in smear-positive tuberculosis patients. *Antimicrob Agents Chemother* 2010;**54**:3402–7. http://dx.doi.org/10.1128/AAC.01354-09.
249. Ahmad Z, Peloquin CA, Singh RP, Derendorf H, Tyagi S, Ginsberg A, Grosset JH, Nuermberger EL. PA-824 exhibits time-dependent activity in a murine model of tuberculosis. Antimicrob Agents Chemother. doi:AAC.00849-10.
250. Rustomjee R, Diacon AH, Allen J, Venter A, Reddy C, Patientia RF, Mthiyane TC, De Marez T, van Heeswijk R, Kerstens R, Koul A, De Beule K, Donald PR, McNeeley DF. Early bactericidal activity and pharmacokinetics of the diarylquinoline TMC207 in treatment of pulmonary tuberculosis. *Antimicrob Agents Chemother* 2008;**52**:2831–5. doi:AAC.01204-07.
251. Nuermberger E, Rosenthal I, Tyagi S, Williams KN, Almeida D, Peloquin CA, Bishai WR, Grosset JH. Combination chemotherapy with the nitroimidazopyran PA-824 and first-line drugs in a murine model of tuberculosis. *Antimicrob Agents Chemother* 2006;**50**:2621–5. http://dx.doi.org/10.1128/AAC.00451-06.
252. Horne DJ, Royce SE, Gooze L, Narita M, Hopewell PC, Nahid P, Steingart KR. Sputum monitoring during tuberculosis treatment for predicting outcome: systematic review and meta-analysis. *Lancet Infect Dis* 2010;**10**:387–94. http://dx.doi.org/10.1016/S1473-3099(10)70071-2.
253. Controlled clinical trial of four short-course regimens of chemotherapy for two durations in the treatment of pulmonary tuberculosis. Second report. Third East African/British Medical Research Council Study. *Tubercle* 1980;**61**:59–69.
254. Wallis RS, Doherty TM, Onyebujoh P, Vahedi M, Laang H, Olesen O, Parida S, Zumla A. Biomarkers for tuberculosis disease activity, cure, and relapse. *Lancet Infect Dis* 2009;**9**:162–72. http://dx.doi.org/10.1016/S1473-3099(09)70042-8.
255. Spigelman M, Woosley R, Gheuens J. New initiative speeds tuberculosis drug development: novel drug regimens become possible in years, not decades. Int J Tuberc Lung Dis;**14**:663–4.
256. Hackam DG. Translating animal research into clinical benefit. *BMJ* 2007;**334**:163–4. doi:334/7586/163.
257. Atlas R, editor. *Handbook of microbiological media*. CRC Press; 2004.