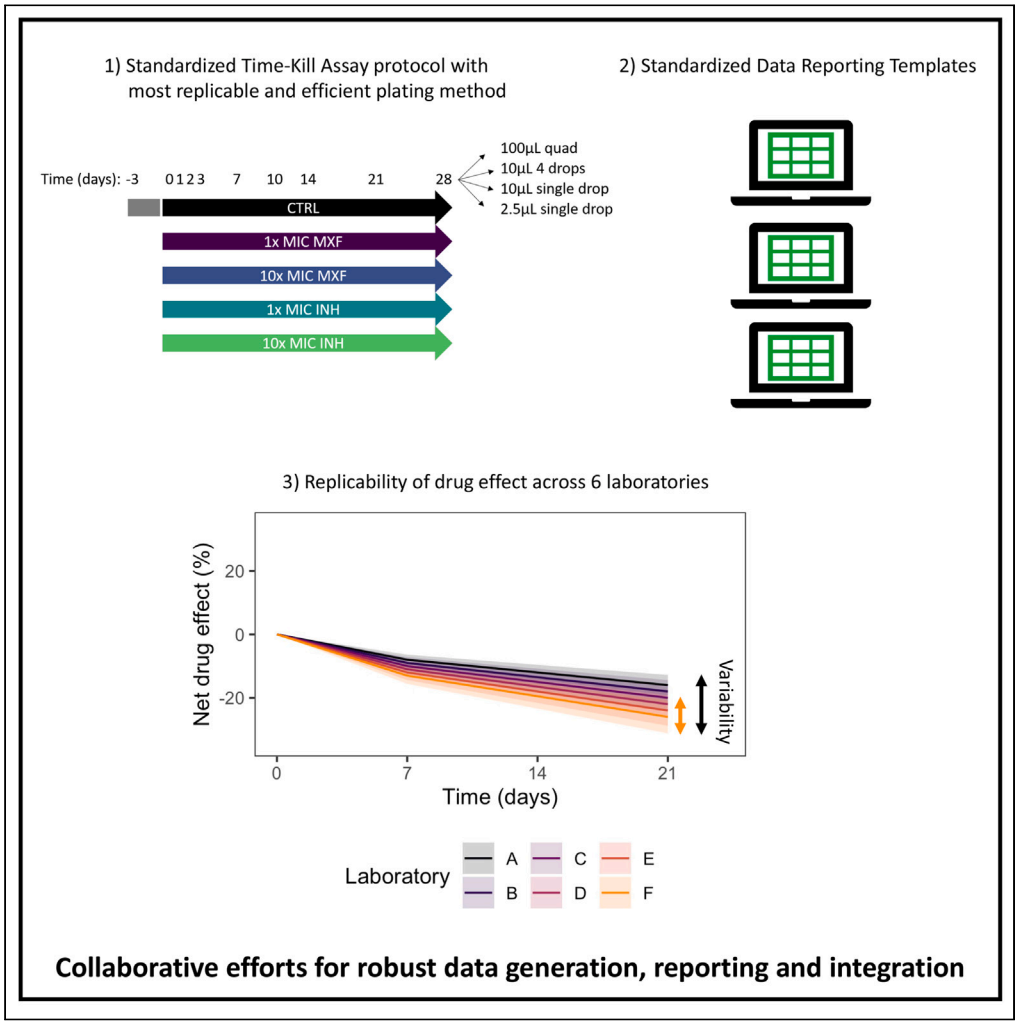
iScience

CellPress
OPEN ACCESS

Article

Implementing best practices on data generation and reporting of *Mycobacterium tuberculosis in vitro* assays within the ERA4TB consortium



Rob C. van Wijk, Ainhoa Lucía, Pavan Kumar Sudhakar, ..., Ulrika S.H. Simonsson, Santiago Ramón-García, ERA4TB consortium

santiramon@unizar.es

Highlights

Best collaboration practices are shared on a large consortium initiative (ERA4TB)

Robust standardized time kill assay protocol for *Mycobacterium tuberculosis*

Variability—inherent to the assay—was successfully quantified and minimized

An innovative 2.5 µL plating technique performed similar to the 100 µL standard

van Wijk et al., iScience *26*, 106411
April 21, 2023
https://doi.org/10.1016/j.isci.2023.106411

iScience

CellPress
OPEN ACCESS

Article

Implementing best practices on data generation and reporting of *Mycobacterium tuberculosis* *in vitro* assays within the ERA4TB consortium

Rob C. van Wijk,[1,13] Ainhoa Lucía,[2,3,13] Pavan Kumar Sudhakar,[4] Lindsay Sonnenkalb,[5] Cyril Gaudin,[6] Eik Hoffmann,[6] Bérénice Dremierre,[6] Diana Angélica Aguilar-Ayala,[2] Michael Dal Molin,[7,8] Jan Rybniker,[7,8,9] Stefano de Giorgi,[10] Laura Cioetto-Mazzabò,[10] Greta Segafreddo,[10] Riccardo Manganelli,[10] Giulia Degiacomi,[11] Deborah Recchia,[11] Maria Rosalia Pasca,[11] Ulrika S.H. Simonsson,[1] Santiago Ramón-García,[2,3,12,14,*] and on behalf of the ERA4TB consortium

SUMMARY

Tuberculosis (TB) is the historical leading cause of death by a single infectious agent. The European Regimen Accelerator for Tuberculosis (ERA4TB) is a public-private partnership of 30+ institutions with the objective to progress new anti-TB regimens into the clinic. Thus, robust and replicable results across independent laboratories are essential for reliable interpretation of treatment efficacy. A standardization workgroup unified *in vitro* protocols and data reporting templates. Time-kill assays provide essential input data for pharmacometric model-informed translation of single agents and regimens activity from *in vitro* to *in vivo* and the clinic. Five conditions were assessed by time-kill assays in six independent laboratories using four bacterial plating methods. Baseline bacterial burden varied between laboratories but variability was limited in net drug effect, confirming 2.5 μL equally robust as 100 μL plating. This exercise establishes the foundations of collaborative data generation, reporting, and integration within the overarching Antimicrobial Resistance Accelerator program.

INTRODUCTION

Tuberculosis (TB) caused by *M. tuberculosis* (*Mtb*) is the historical leading cause of death by a single infectious agent and, nowadays, second after COVID-19. In 2021, TB was responsible for 1.6 million casualties.[1] TB mortality has increased because of reduced access to care due to COVID-19 devastating effects on TB services.[2]

Combination therapy is required for TB treatment in order to prevent resistance emergence and reduce treatment duration.[3] To accelerate the development of new treatments against TB, the Innovative Medicines Initiative (IMI) (www.imi.europa.eu)[4] funded the public-private partnership European Regimen Accelerator for Tuberculosis (ERA4TB) consortium (https://era4tb.org/).

Activities within the consortium are integrated following the typical progression sequence in pharmaceutical development. As such, the development and quality control of standardized practices and protocols are a high priority to facilitate collaboration within the profiling activities of the consortium and ensure data robustness among the different laboratories. In this context, a Standardization Workgroup (StWG) was established to curate the process of gathering input from respective areas of expertise and sharing best practices on different *in vitro* protocols, data templates, and reporting processes to determine compounds/combinations *in vitro* activities against *Mtb*. Each laboratory assigned a representative, and technique-specific sub-groups were formed to develop consensus protocols. The StWG met monthly and generated a white document with common standardized *in vitro* protocols which is now used internally as a reference manual. In line with the IMI Antimicrobial Resistance (AMR) Accelerator program, which overarches the ERA4TB and other IMI consortia, the StWG together with Data Management and Modeling and Simulation teams of the ERA4TB consortium developed data templates for data aggregation, reporting, sharing, and analysis following FAIR principles.[5]

[1]Department of Pharmaceutical Biosciences, Uppsala University, 75124 Uppsala, Sweden

[2]Department of Microbiology, Faculty of Medicine, University of Zaragoza, 50009 Zaragoza, Spain

[3]CIBER Enfermedades Respiratorias (CIBERES), Instituto de Salud Carlos III, 28029 Madrid, Spain

[4]Data collaboration center, Critical Path Institute Ltd, D11 YNR2 Dublin, Ireland

[5]Molecular and Experimental Mycobacteriology, Research Center Borstel Leibniz Lung Center, 23845 Borstel, Germany

[6]Center for Infection and Immunity of Lille, Institut Pasteur de Lille, 59019 Lille, France

[7]Department I of Internal Medicine, Division of Infectious Diseases, University of Cologne, 50937 Cologne, Germany

[8]Faculty of Medicine, Center for Molecular Medicine Cologne (CMMC), University of Cologne, 50931 Cologne, Germany

[9]German Center for Infection Research (DZIF), Partner Site Bonn-Cologne, 50931 Cologne, Germany

[10]Department of Molecular Medicine, University of Padova, 35121 Padova, Italy

[11]Department of Biology and Biotechnology Lazzaro Spallanzani, University of Pavia, 27100 Pavia, Italy

[12]Research & Development Agency of Aragon Foundation (Fundación

Continued

In TB, and other infectious diseases where the pathogen and its response to treatment can be studied outside the patient, *in vitro* experiments are convenient to elucidate drug effects. Among these, time kill assays (TKA) are employed to quantify the antibiotic concentration-effect relationship in a time-dependent manner. In a TKA, bacteria are exposed to different drug concentrations including a negative control or natural growth condition, and the bacterial burden is determined longitudinally at multiple timepoints. In contrast to the static single timepoint metric of a minimum inhibitory concentration (MIC), TKA data reflects the dynamics of the antibiotic effect over time on cell viability.[6,7] Therefore, TKA data is informative for quantifying a concentration-effect relationship that can be translated from the *in vitro* context to *in vivo*, including prediction of human efficacy. TKA observations are commonly performed by quantifying colony-forming units (CFU), a cumbersome methodology in TB research due to the slow growth of the bacteria (that takes 2-4 weeks to form colonies) and the need of working in Biosafety Level 3 (BSL3) facilities. Inherent to its design with multiple timepoints, a TKA requires more experimental observations than MIC determinations; therefore, more resources including incubator space and consumables, but also physical workload, add limitations when it comes to the design of TKA with a large number of conditions. In addition, CFU measurements can be variable due to experimental conditions, number of dilutions plated, methods of colony counting, and the tendency of mycobacterial cells to form aggregates.[8,9]

In order to meet ERA4TB's demand on *in vitro* analyses, assay capacity needs to be built across partner laboratories. In the case of TKA, these will have to be performed at different locations in parallel, stressing the need for consistency between partners and standardization of a robust and replicable TKA protocol in order to build confidence in data generation for pharmacometric modeling. To address this challenge, the StWG optimized four different CFU plating methodologies to limit variability and increase throughput and capacity. Our objective here was thus to (i) quantify the replicability of a common standardized TKA protocol between six laboratories by determining variability in CFU performed using four plating methods, and (ii) share lessons learnt on how to collaborate between partners in an IMI consortium to harmonize, implement and standardize experimental protocols and data management for consistent reporting and sharing within the AMR Accelerator program.

RESULTS

Four different plating methods yielded robust and reproducible outcomes

Four CFU plating methods with decreasing culture volume and incubator plate space requirements were included in the standardized TKA protocol and evaluated for replicability and consistency across laboratories: i) 100 μL of each dilution plated on quad plates (100 μL/quad section) (standard plating method); ii) 10 μL of each dilution spread over 4-6 drops using a multichannel pipette (10 μL 4 drops); iii) 10 μL of each dilution in a single spot drop (10 μL single drop) and; iv) 2.5 μL of each dilution in spot drop (2.5 μL single drop) (Figures 1 and S1).

To assess the replicability and consistency of the four methods and the standardized TKA protocol, the antimycobacterial activity of moxifloxacin (MXF) and isoniazid (INH) was evaluated against *Mtb* at two concentrations (1x MIC and 10x MIC) by six different ERA4TB laboratories. The untreated group showed the normal growth of *Mtb* in the assay media over three to four weeks with an average of 2.5 log10-fold increase in bacterial burden. Treated groups showed an inhibited growth for the lower concentrations (1x MIC) and bacterial killing of an average of 1.4 log10-fold decrease in bacterial burden, consistently at 10x MIC of MXF. Regrowth after an initial decline in bacterial burden was observed at both INH exposure concentrations. Graphically, curve profiles overlap regardless of the plating method, indicating similar performance in techniques (Figure 2). When executed following the standardized protocol, all four plating methods resulted in robust and reproducible outcomes (Table 1). The coefficient of variation as measure for method variability was smaller than 35% for all methods. The 10 μL 4 drops method showed the smallest variability of 29.1% and the 2.5 μL single drop the largest with 34.7%, a difference of only 5.6 percent point. None of the methods was statistically significantly different from the 100 μL quad standard. The innovative 2.5 μL single drop plating method was not inferior to the standard 100 μL quad plating method, while requiring nine times less plate and incubator space.

Time kill assay variability was limited among laboratories

The quantification of bacterial burden through time-kill CFU-based assays is inherently variable, and replicability between laboratories is of utmost importance for robust conclusions in antibiotic drug development from these pivotal experiments. In our trial, variability between laboratories was aimed to be limited

ARAID), 50009 Zaragoza, Spain

[13]These authors contributed equally

[14]Lead contact

*Correspondence: santiramon@unizar.es

https://doi.org/10.1016/j.isci.2023.106411

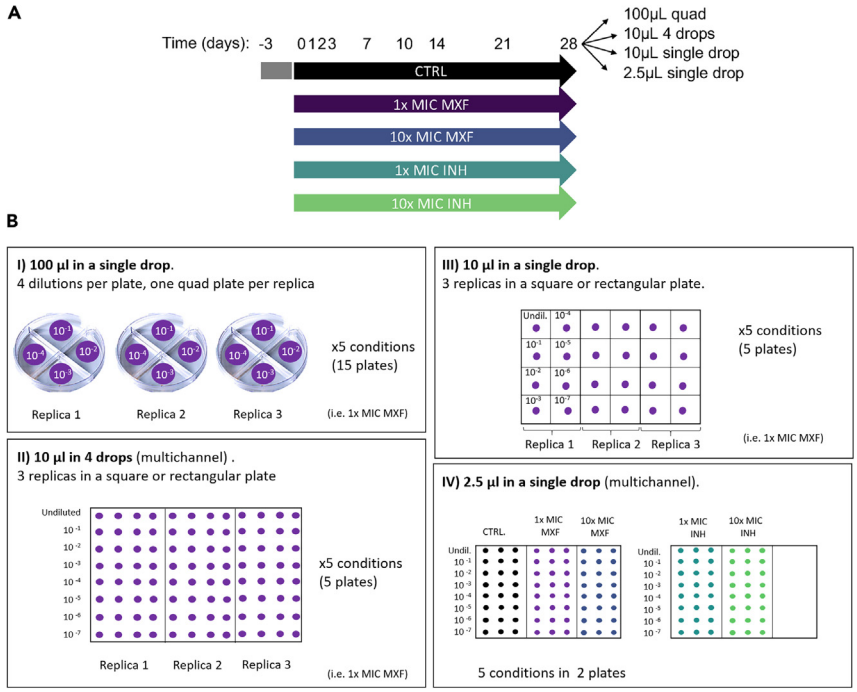


Figure 1. Overview of the standardized time kill assay protocol

(A) Schematic of the experimental design including treatment groups (colored arrows), timepoints, and plating methods. (B) Detailed comparison of the four plating methods shown for three replicates, five conditions, and 4 to 8 dilutions for 100 µL quad (I), and 10 µL 4 drops (II) or 10 µL single drop (III) or 2.5 µL single drop (IV), respectively. Representative images of actual dilution plates are shown in Figure S1.

by following the standardized protocol as well as by having the same original *Mtb* strain (cultivated at one of the laboratories and then distributed to the rest) and consumables from the same providers. Figure 3 shows the bacterial burden over time for the six laboratories stratified by treatment group. Across treatment groups, bacterial burden changes over time are consistent between laboratories, as the within-laboratory variability largely overlaps with the between-laboratory variability. Within-laboratory variability was limited and ranged between 23.0% and 40.5%, with a median of 31.05% (Table 2). Laboratories A, B, and C had more BQL data in the 10x MIC MXF and the 1x and 10x MIC INH groups; nevertheless, their profiles were able to capture the bacterial kill (10x MIC MXF) and regrowth for 1x MIC INH although not for 10x MIC INH (Figure 3).

Baseline bacterial burden per laboratory ranged from 3.29 to 5.97 log10 CFU/mL (Figure 4). Laboratories A, B, and C had the lower baseline bacterial burden, which correlates with their higher BQL frequency. Variability in baseline bacterial burden (Table 3) was slightly lower than for all samples (Table 2). Four out of six laboratories took an additional sample at the pre-inoculum timepoint ($t = -3$ days), which enabled observing *Mtb* growth dynamics between pre-inoculum and drug addition ($t = 0$ days). On average, this growth ranged from −0.58 log10 CFU/mL to 1.6 log10 CFU/mL.

Drug effects were consistent across laboratories

TKA data generated under a common protocol should support the same conclusion on the effect of the drug(s) tested, independently of where it was generated. Net drug effects of INH and MXF against *Mtb* at the two concentrations tested were consistent with expected activity (Figure 5). The protocol was able to capture the variability in different profiles including limited drug effect as seen in 1x MIC MXF, bacterial killing (10x MIC MXF), and initial decline and regrowth (1x and 10x MIC INH). In summary, variability between laboratories and between plating methods was acceptable with coefficients of variation less than 50%.

These findings supported the implementation of the standardized protocol within ERA4TB. The standardized data template was optimized and consistently used by all laboratories to capture their data and

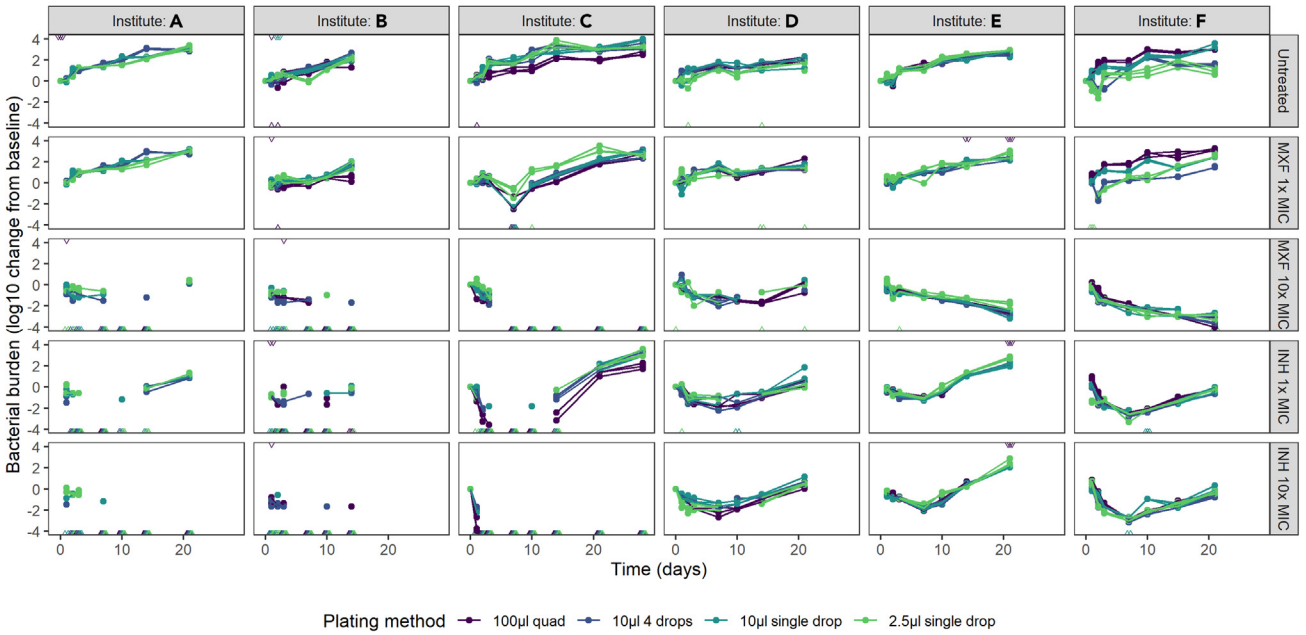


Figure 2. Bacterial burden in log10 change from baseline over time quantified by four plating methods (colors) for different laboratories (columns), untreated (top row), and different treatment conditions (bottom rows)

INH, isoniazid; MIC, minimum inhibitory concentration; MXF, moxifloxacin; quad, quad plate.

transfer it to the other partners. The standardized output allowed for R-script-based data analysis for fast and reproducible graphical and numerical reporting. Transfer of the relevant information reported in the data template to the AMR accelerator repository was designed based on the data template and loaded and maintained in the DDIM-grit42 platform[10] in the AMR Pre-Clinical results template structure.

DISCUSSION

TB drug development is a collaborative endeavor, with consortia consisting of different partners with various areas of expertise working together to find new combination therapies to treat all forms of TB. Replicability of results is crucial in drug development in general, and specifically between consortium members working toward the same objectives.[11–13] Here, we share our lessons from a collaborative working group created amongst research partners at the IMI ERA4TB consortium to harmonize and standardize experimental protocols, data management, and reporting. For this, the StWG developed standardized protocols for TKA (Supplemental information - TKA protocol, Figure S2) and quantified variability between six laboratories to assess the robustness of the protocol implementation.

TKA are a key *in vitro* methodology in early antimicrobial drug development; they provide longitudinal information on the pharmacodynamic properties of a compound against different bacterial pathogens based on standard CFU plating in agar petri dishes. In the case of the TB field (the scope of this work), this was traditionally performed by distributing 0.1-1 mL culture dilutions on the agar surface, a time and resource-demanding procedure due to the need to work in biosafety containment level 3 facilities.[14,15] In addition, due to the hydrophobic nature of the mycobacterial cell wall and the tendency to form cell aggregates, it was believed that large volumes (typically 100 µL) are needed when plating to ensure proper enumeration of the CFUs.

In this work, we aimed to challenge this perception; we proposed four alternatives CFU plating methods (Figures 1 and S1) already implemented at any of the six laboratories participating in this trial. Each plating method used less resources (culture volume, media, plate, incubator space), ranging from 3 plates in the 100 µL quad method to 1/3rd plate in the 2.5 µL single drop per condition assuming triplicates. Methodological and reporting procedures were agreed within the StWG and the activity of two anti-tuberculosis drugs (MFX and INH) tested at two concentrations using the four plating methods (Figure 2). No statistically significant differences were found among the plating methods when compared to the 100 µL quad, our internal gold standard (Table 1). These results highlighted that just plating 2.5 µL (instead of 100 µL) was

Table 1. Variability per plating method in bacterial burden over time in CFU/mL reported as coefficient of variation (CV) across laboratory, treatment group, timepoints and replicates

Plating method	Variability (CV)
100 µL quad	31.3%
10 µL 4 drops	29.1%
10 µL single drop	31.8%
2.5 µL single drop	34.7%

A statistically significant difference was only found between the 10 µL 4 drops and the 2.5 µL single drop methods (p value = 0.014, Dunn test with Bonferroni correction).

equally robust thus largely increasing the throughput capacity of TKA, a remarkable finding when managing limited resources under high-capacity demand.

This is the first study to show that variability between CFU quantification techniques is negligible, proving that larger volumes are not always necessary. The innovative 2.5 µL single drop technique requires 40-fold less bacterial culture, 15-fold less consumables, and less BSL-3 incubator space than the 100 µL quad, an extraordinary increase in assay capacity; it is thus recommended for further experiments within the ERA4TB consortium as it guarantees the highest experimental throughput, while remaining robust and reproducible.

In addition to the comparison of variability per plating method, variability was quantified at different levels, for example per laboratory or per treatment group. Variability of measured bacterial burden within laboratories showed a large overlap with the variability between laboratories for the same treatment group. Laboratories with smaller within-laboratory variability (i.e., A, E, and F) seem to show more distinct profiles in the time-kill assay. However, when quantifying the drug efficacy, which remains the overall purpose of performing these experiments, the resulting drug effect profiles were largely replicable across laboratories; this is remarkable, especially given the nature of bacterial growth-based assays and its inherent variability. In general, the quantified variability was acceptable with coefficients of variation around 30%. Part of this variability is the random noise for which the number of technical replicates accounts. Therefore, the number of technical replicates of at least three is appropriate to quantify the within experiment variability at different levels in addition to the noise.

One source of variability of interest here when comparing *in vitro* efficacy data across laboratories is influenced by the bacterial strain background used for the study and how the pre-inoculum is handled, i.e., cell numbers and physiological state of the bacteria. The StWG attempted to limit it by sharing the same strain

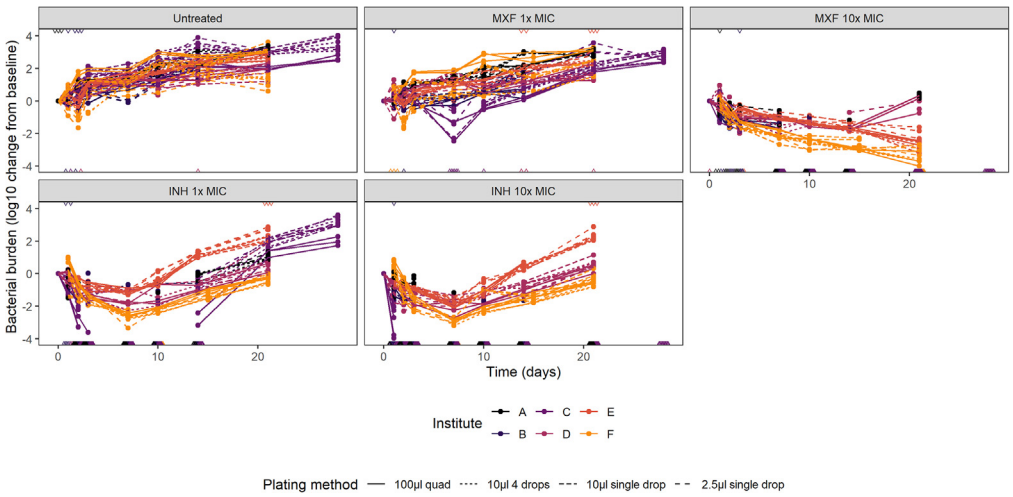


Figure 3. Bacterial burden in log10 change from baseline over time for different treatment conditions (panels)
Colors depict different laboratories; line type depicts different plating methods. INH, isoniazid; MIC, minimum inhibitory concentration; MXF, moxifloxacin; quad, quad plate.

Table 2. Variability per laboratory in bacterial burden over time in CFU/mL reported as coefficient of variation (CV) across plating methods, treatment group, timepoints and replicates

Laboratory	Variability (CV)
A	23.0%
B	37.4%
C	34.7%
D	40.5%
E	27.4%
F	27.1%

under a common protocol addressing handling and pre-inoculum set up. Despite such efforts, baseline bacterial burden was variable between laboratories. Most laboratories took observations at both timepoints t = −3 and t = 0 (when drugs are added), which enabled the recommended quantification of the initial bacterial growth prior to treatment. The size of the baseline bacterial burden, or pre-inoculum, can impact the drug effect with higher bacterial burdens attenuating the drug effect.[16] Therefore, the true concentration-effect relationship may not be apparent, which warns caution when interpreting supposed-to-be key decision-making data.

Development of a pharmacometric non-linear mixed effects model can be leveraged to quantify the concentration effect relationship taking into account the inoculum.[16–18] Pharmacometric models can also account for observations that are either below (empty plate) or above (uncountable plate) a quantification level. Well-established pharmacometric modeling methods developed to account for concentration data outside of the quantification range determined by analytical chemistry methods can be applied to bacterial burden data as well.[19,20] These methods combine continuous (CFU/mL) and categorical (BQL, AQL) data by characterizing the probability of an observation being outside of the quantification range, given the underlying functions for bacterial growth and/or drug effect. It is therefore important to record an observation that is BQL or AQL including the corresponding quantification limit, as these observations still contain information that can support the quantification of a drug effect.

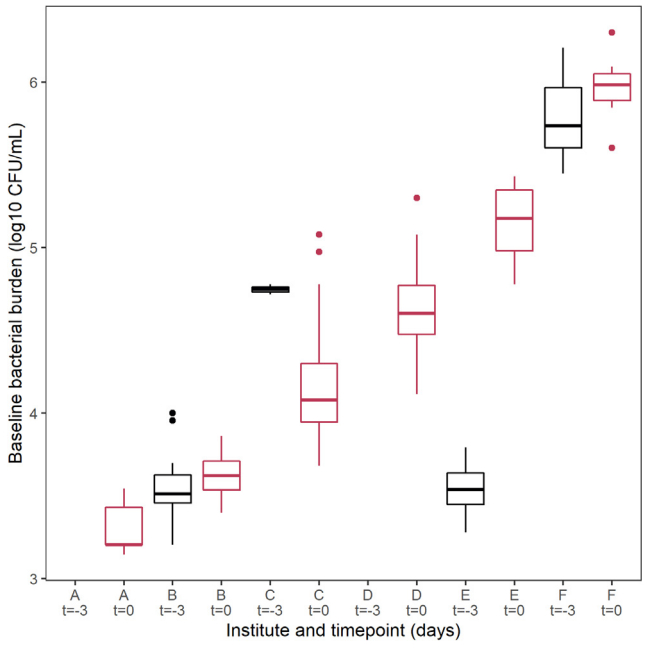


Figure 4. Baseline bacterial burden in log10 CFU/mL for different laboratories and timepoints (colors)
Boxplot shows median (thick line) and first and third quartile (hinges), with the whisker connecting to the largest value not further than 1.5x interquartile range, data beyond which are shown as symbols.

Table 3. Baseline bacterial burden in CFU/mL and corresponding variability reported as coefficient of variation (CV) per laboratory, across plating methods, treatment group, and replicates

Laboratory	Mean baseline t = −3 (log10 CFU/mL)	Variability (CV)	Mean baseline t = 0 (log10 CFU/mL)	Variability (CV)
A	NA	NA	3.29	8.24%
B	3.56	46.8%	3.63	31.1%
C	4.75	10.1%	4.17	23.6%
D	NA	NA	4.65	32.5%
E	3.55	17.5%	5.15	25.5%
F	5.78	21.5%	5.97	36.8%

Characterization of the normal growth of the pathogen is essential to distinguish the drug effect from the untreated condition.[21,22] The standardized protocol, therefore, includes an untreated control at each time-point, in addition to the observations from pre-inoculum to treatment start. The untreated control was here utilized to calculate the net drug effect of MXF and INH, as it was not the focus of this work to characterize the concentration-effect relationship of these well-studied drugs. With a larger range of drug concentrations, pharmacometric modeling can be leveraged to characterize the concentration-effect relationship *in vitro*, which is fundamental to translating the drug efficacy to the *in vivo* and clinical context.[23–25] Part of these methods is the quantification of different levels of variability, which can be applied to quantify the differences in variability between laboratories in more detail.

Developing new combination therapies against TB is a collaborative effort between many players in the antimicrobial discovery and development field, including experimentalists and modelers in pharmacology. The ERA4TB consortium falls within the broader scope of the larger IMI AMR Accelerator program, which comprises three pillars: Capability Building Network, Tuberculosis Drug Development Network; and Portfolio Building Networks. Under one structure, the AMR Accelerator addresses many of the scientific challenges of AMR, contributing to the European action plan on AMR[26] and supporting the development of new ways to prevent and treat AMR. Projects in the AMR Accelerator are bound to the development of new pre-clinical tools and methods to assist in the translation of novel anti-infective agents' pre-clinical data to the clinic. To this end, data management guidelines and an IT infrastructure enable the collection, aggregation, storage, sharing, and analysis of datasets generated by AMR Accelerator projects. The StWG was able to integrate discrepancies that are typically found between academic experimentalist data generators and end users, i.e., industry, modelers, and regulators.

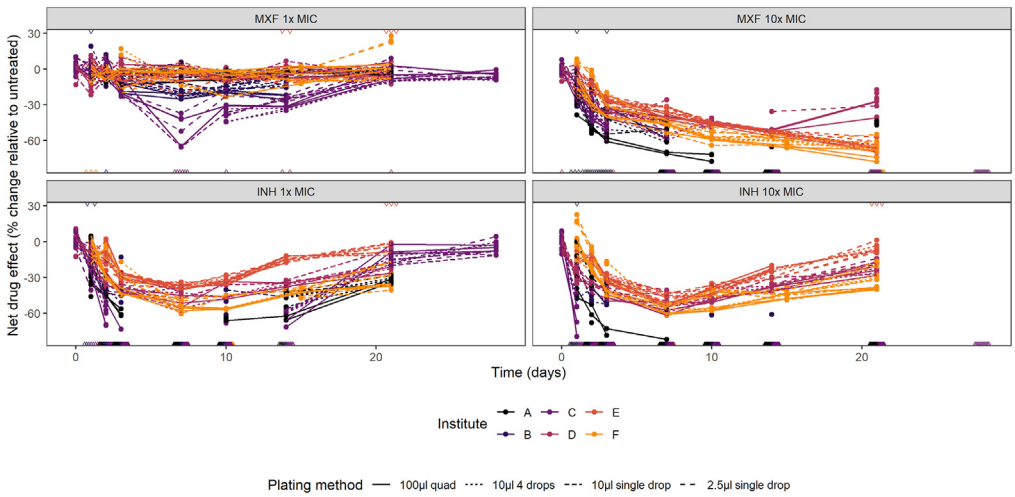


Figure 5. Net drug effect reported as bacterial burden in the treatment group of interest relative to the untreated group, for different treatment conditions (panels)

Colors depict different laboratories; line type depicts different plating methods. INH, isoniazid; MIC, minimum inhibitory concentration; MXF, moxifloxacin; quad, quad plate.

A key endeavor of the StWG was to provide a safe forum for data generators and end users to establish a constructive dialogue to meet raised expectations by both parties. Partner commitment and fluid communication through e-mail exchanges and frequent videoconferences were critical to achieve a safe working place. Taking TKA as an exemplary methodology, *ad hoc* data templates were generated to ease experimentalist data reporting load while facilitating the automatization of data transfer into tabular formats required for integration into the AMR Accelerator program. The objective of the template development was thus to create a transparent and traceable dataset containing all information for data analysis, but with limited change in operations by experimentalists. Staying close to the experimentalist' previous data-capturing operations, by basing the standardized template on a typical example of a data acquisition sheet, was expected to result in highest acceptance and adherence to the new standardized template. Data managers and data analysist worked on function- and macro-based exports of the relevant data for data storage and analysis. Data templates included essential information up to regulatory submission standards (Supplemental information – Raw data template.xlm & Guidelines), allowing for data provenance including audit trails; for this, an electronic Laboratory Notebook was implemented in the different laboratories.

The standardized data template is subject to data transfer to the AMR repository, with data availability to all AMR Accelerator consortia. Data availability will improve collaboration within and across consortia. The standardized data template was developed for TKA data, but is currently being expanded to include all experimental procedures performed in ERA4TB, including assays such as artificial caseum, efflux pump inhibition, resistance evolution assays, frequency of resistance, granuloma-like structure, hERG safety, hollow fiber system for TB, growth inhibition assays, or *in vitro* pharmacokinetics.

In summary, here we share lessons learnt on how to build a collaborative space between partners in an IMI consortium to harmonize, implement and standardize experimental protocols and data management for consistent reporting and sharing within the AMR Accelerator program, and described how different areas of expertise come together to rise to the challenge. Standardization to achieve robust and replicable results is the first step to improve and streamline the elements in the developmental process to bring new combination therapies into the clinic for the treatment of all TB forms.

Limitations of the study

Despite protocol standardization and shared *Mtb* strain, baseline bacterial burden was variable between laboratories. Baseline bacterial burden can impact TKA data, and consequently the interpretations of drug effects. The slow growth rate of *Mtb* implies that, in TKA, CFU data of the first timepoint is counted almost after the last timepoint is taken (typically 21 days), preventing pre- or mid-experiment adaptations.

Only two dose levels per drug were utilized in these experiments, limiting the possibilities to extract a reliable exposure-response relationship of the drug tested; although this was not the scope of this exercise since the activity of both drugs is well characterized against *M. tuberculosis*. As such, drugs and dose levels tested here were selected to reflect bactericidal and sterilizing effects in the context of a TKA. In TKA experiments performed within the consortium larger dose ranges are tested to quantify the full exposure-response relationships using pharmacometric modeling.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Standardized bacterial strain & growth conditions
  - Standardized drugs & plasticware
- METHODS DETAILS
  - Standardization workgroup
  - Standardized time kill assay

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.106411.

ACKNOWLEDGMENTS

The project leading to this publication has received funding from the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement No 853989. The JU receives support from the European Union's Horizon 2020 Research and Innovation Program and EFPIA and Global Alliance for TB Drug Development Non-Profit Organisation, Bill & Melinda Gates Foundation, University of Dundee. This work reflects only the author's views, and the JU is not responsible for any use that may be made of the information it contains.

We would especially like to thank the technical support of Larissa Mohr, Silvia Maass, Tanja Niemann, and Vanessa Mohr from the Research Center Borstel Leibniz Lung Center, Begoña Gracia Díaz and Ana Picó Marco from the University of Zaragoza and Jason Chhen from the University of Cologne.

AUTHOR CONTRIBUTIONS

RCvW, AL, USHS, and SRG conceptualized the study. LS, CG, EH, BD, DAAA, MDM, SG, LCM, GS, GD, and DR performed the experiments. RCvW, ALQ, LS, and PK designed the data template and completion guidelines. RCvW and USHS conducted data analysis. RCvW, AL, PKS, USHS, and SRG draft the article. LS, CG, EH, DAAA, MDM, and GD contributed to the final version of the article. JR, RM and MRP contributed to the standardization process and critically reviewed the article. All authors read and approved the final version of the article.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: July 4, 2022
Revised: November 14, 2022
Accepted: March 10, 2023
Published: March 15, 2023

REFERENCES

1. Geneva: World Health Organization (2022). Global Tuberculosis Report 2022.

2. Pai, M., Kasaeva, T., and Swaminathan, S. (2022). Covid-19's devastating effect on tuberculosis care — a path to recovery. N. Engl. J. Med. *386*, 1490–1493. https://doi.org/10.1056/nejmp2118145.

3. Kerantzas, C.A., and Jacobs, W.R. (2017). Origins of combination therapy for tuberculosis: lessons for future antimicrobial development and application. mBio *8*, e01586-16. https://doi.org/10.1128/mBio.01586-16.

4. Laverty, H., and Meulien, P. (2019). The innovative Medicines initiative —10 Years of public-private collaboration. Front. Med. *6*, 275. https://doi.org/10.3389/fmed.2019.00275.

5. Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., et al. (2016). Comment: the FAIR Guiding Principles for scientific data management and stewardship. Sci. Data *3*, 160018. https://doi.org/10.1038/sdata.2016.18.

6. Mueller, M., de la Peña, A., and Derendorf, H. (2004). Issues in pharmacokinetics and pharmacodynamics of anti-infective agents: kill curves versus MIC. Antimicrob. Agents Chemother. *48*, 369–377. https://doi.org/10.1128/AAC.48.2.369.

7. Rathi, C., Lee, R.E., and Meibohm, B. (2016). Translational PK/PD of anti-infective therapeutics. Drug Discov. Today Technol. *21–22*, 41–49. https://doi.org/10.1016/j.ddtec.2016.08.004.

8. Pathak, S., Awuh, J.A., Leversen, N.A., Flo, T.H., and Asjø, B. (2012). Counting mycobacteria in infected human cells and mouse tissue: a comparison between qPCR and CFU. PLoS One *7*, e34931. https://doi.org/10.1371/journal.pone.0034931.

9. Gillespie, S.H., Gosling, R.D., and Charalambous, B.M. (2002). A reitrerative method for calculating the early bactericidal activity of antituberculosis drugs. Am. J. Respir. Crit. Care Med. *166*, 31–35. https://doi.org/10.1164/rccm.2112077.

10. Pérez, J.C., Sosa, V.J., and Vlas, J.G. (2020). Instantiation of EU-Based Drug Development Information Management (DDIM) System - ERA4TB Initial Report.

11. Rotelli, M.D. (2015). Ethical considerations for increased transparency and reproducibility in

the retrospective analysis of health care data. Ther. Innov. Regul. Sci. *49*, 342–347. https://doi.org/10.1177/2168479015578155.

12. Wang, S.V., Verpillat, P., Rassen, J.A., Patrick, A., Garry, E.M., and Bartels, D.B. (2016). Transparency and reproducibility of observational cohort studies using large healthcare databases. Clin. Pharmacol. Ther. *99*, 325–332. https://doi.org/10.1002/cpt.329.

13. Freedman, L.P., and Gibson, M.C. (2015). The impact of preclinical irreproducibility on drug development. Clin. Pharmacol. Ther. *97*, 16–18. https://doi.org/10.1002/cpt.9.

14. Hazan, R., Que, Y.A., Maura, D., and Rahme, L.G. (2012). A method for high throughput determination of viable bacteria cell counts in 96-well plates. BMC Microbiol. *12*, 259. https://doi.org/10.1186/1471-2180-12-259.

15. Jett, B.D., Hatter, K.L., Huycke, M.M., and Gilmore, M.S. (1997). Simplified agar plate method for quantifying viable bacteria. Biotechniques *23*, 648–650. https://doi.org/10.2144/97234bm22.

16. Nielsen, E.I., Khan, D.D., Cao, S., Lustig, U., Hughes, D., Andersson, D.I., and Friberg, L.E. (2017). Can a pharmacokinetic/pharmacodynamic (PKPD) model be predictive across bacterial densities and strains? External evaluation of a PKPD model describing longitudinal in vitro data. J. Antimicrob. Chemother. *72*, 3108–3116. https://doi.org/10.1093/jac/dkx269.

17. Salas, J.R., Jaberi-Douraki, M., Wen, X., and Volkova, V.V. (2020). Mathematical modeling of the "inoculum effect": six applicable models and the MIC advancement point concept. FEMS Microbiol. Lett. *367*, fnaa012. https://doi.org/10.1093/femsle/fnaa012.

18. Bhagunde, P., Chang, K.T., Singh, R., Singh, V., Garey, K.W., Nikolaou, M., and Tam, V.H. (2010). Mathematical modeling to characterize the inoculum effect. Antimicrob. Agents Chemother. *54*, 4739–4743. https://doi.org/10.1128/AAC.01831-09.

19. Beal, S.L. (2001). Ways to fit a PK model with some data below the quantification limit. J. Pharmacokinet. Pharmacodyn. *28*, 481–504. https://doi.org/10.1023/A:1012299115260.

20. Bergstrand, M., and Karlsson, M.O. (2009). Handling data below the limit of quantification in mixed effect models. AAPS J. *11*, 371–380. https://doi.org/10.1208/s12248-009-9112-5.

21. Clewe, O., Aulin, L., Hu, Y., Coates, A.R.M., and Simonsson, U.S.H. (2016). A multistate tuberculosis pharmacometric model: a framework for studying anti-tubercular drug effects in vitro. J. Antimicrob. Chemother. *71*, 964–974. https://doi.org/10.1093/jac/dkv416.

22. van Wijk, R.C., van der Sar, A.M., Krekels, E.H.J., Verboom, T., Spaink, H.P., Simonsson, U.S.H., and van der Graaf, P.H. (2020). Quantification of natural growth of two strains of Mycobacterium marinum for translational anti-tuberculosis drug development. Clin. Transl. Sci. *13*, 1060–1064. https://doi.org/10.1111/cts.12793.

23. Nielsen, E.I., Cars, O., and Friberg, L.E. (2011). Pharmacokinetic/pharmacodynamic (PK/PD) indices of antibiotics predicted by a semimechanistic PKPD model: a step toward model-based dose optimization. Antimicrob. Agents Chemother. *55*, 4619–4630. https://doi.org/10.1128/AAC.00182-11.

24. Susanto, B.O., Wicha, S.G., Hu, Y., Coates, A.R.M., Simonsson, U.S.H., Biosciences, P., Pharmacy, C., Kingdom, U., Simonsson, U.S.H., Biosciences, P., et al. (2020). Translational model-informed approach for selection of tuberculosis drug combination regimens in early clinical development. Clin. Pharmacol. Ther. *108*, 274–286. https://doi.org/10.1002/cpt.1814.

25. van Wijk, R.C., Hu, W., Dijkema, S.M., van den Berg, D.-J., Liu, J., Bahi, R., Verbeek, F.J., Simonsson, U.S.H., Spaink, H.P., van der Graaf, P.H., and Krekels, E.H.J. (2020). Anti-tuberculosis effect of isoniazid scales accurately from zebrafish to humans. Br. J. Pharmacol. *177*, 5518–5533. https://doi.org/10.1111/bph.15247.

26. European Commission (2017). A European one health action plan against antimicrobial resistance (AMR) (Brussels), pp. 1–24.

27. Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. v, Eiglmeier, K., Gas, S., Barry, C.E., III, et al. (1998). Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature *393*, 537–544. https://doi.org/10.1038/31159.

28. O'Meara, P., Faizan, A., Sudhakar, P.K., and Liwski, R. (2021). Standardized Templates For Collection And Reporting Of Clinical And Preclinical Data Available To Consortium Members – ERA4TB Interim Report.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Mycobacterium tuberculosis strain H37Rv	GenBank	GenBank ID: NC_000962.3
Chemicals, peptides, and recombinant proteins		
Isoniazid	European Pharmacopeia	cat no. I0500000
Moxifloxacin	European Pharmacopeia	cat no. Y0000703
Glycerol	Fisher	158920010
Glycerol (Glycerin)	PanReac AppliChem	A2926,1000
Glycerol	Euromedex	EU3550
Tyloxapol	MERCK-SIGMA	T8761-50G
DMSO	MERCK-SIGMA	D5879-1L
Difco™ Middlebrook 7H10 Agar 500g	Becton Dickinson	262710
Difco™ Middlebrook 7H9 Broth 500g	Becton Dickinson	271310
OADC Middlebrook Enrichment 20mlx10	Becton Dickinson	211886
OADC Middlebrook Enrichment 500mL	Becton Dickinson	212351
Deposited data		
Raw data	This paper	https://doi.org/10.6084/m9.figshare.19766083
Software and algorithms		
R v.4.0.4	CRAN	https://www.r-project.org
RStudio v. 1.4.1106	POSIT	https://posit.co/products/open-source/rstudio/
Other		
Polypropylene 96-well plates	TPP; Nunc	92096; 167008
Tissue culture flask (25 cm^2)	TPP	90025
Polystyrene rectangular plates	ThermoScientific; Corning	264728; BP124-05

RESOURCE AVAILABILITY

Lead contact

Further information and request should be directed to the corresponding author Santiago Ramón-García at santiramon@unizar.es.

Materials availability

This study did not result in new microbiological or chemical material.

Data and code availability

- Bacterial burden over time data have been deposited at Figshare and are publicly available as of the date of publication. Accession numbers are listed in the key resources table.
- This paper does not report original code.
- Any additional information required to reanalysed the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Standardized bacterial strain & growth conditions

Experiments were carried out with the *Mtb* strain H37Rv (GenBank ID: NC_000962.3) from standardized stocks distributed by FZB to all six partner laboratories involved in the TKA.[27] The *Mtb* strain has no sex assigned. Bacterial stocks were further sub-cultured following a common agreed sub-cultivation procedure and stored at -80°C. After 14 days at -80°C, CFU titers were determined by selecting three stock tubes per box (n= 81). CFU enumeration is performed every six months to ensure that freeze stocks do not lose cell viability. Difco™ Middlebrook 7H9 broth supplemented with 0.5% glycerol, 10% BBL™ Middlebrook OADC Enrichment and 0.025% tyloxapol was used for bacterial expansion and stock generation. 7H9 broth supplemented with 10% OADC and 0.5% glycerol, without tyloxapol, (7H9-OADC-Gly, from now on) was used for TKA. 7H10 agar supplemented with 10% OADC and 0.1% tyloxapol (7H10-OADC-Tx, from now on) was normally used for growth on solid medium.

Standardized drugs & plasticware

All partners purchased TKA drugs from the European Pharmacopeia: moxifloxacin (MXF, cat no. Y0000703) and isoniazid (INH, cat no. I0500000). Stock solutions of 1-20 mg/ml were prepared for both compounds in distilled water and sterilized through 0.22 μm filter membrane, or in DMSO (D2650-5x5mL, MERCK-SIGMA), and stored at -20° for up to one month.

The 96 well plates used to perform the dilutions were made of polypropylene (ref. 92096, TPP; ref. 167008, Nunc). The 25cm^2 culture flasks used for bacterial growth were made of polystyrene (ref. 90025, TPP). Quad petri dishes were made of polystyrene (ref. 200210, Deltalab; ref. 29069, FL Medical), and rectangular plates were made of polystyrene (ref 264728, ThermoScientific; ref. BP124-05, Corning).

METHODS DETAILS

Standardization workgroup

The StWG was established with representatives of six experimental laboratories (University of Zaragoza [UNIZAR], Research Center Borstel Leibniz Lung Institute [FZB], Institut Pasteur de Lille [IPL], University of Cologne [UKÖ], University of Padova [UNIPD], and University of Pavia [UPV]) supported by representatives of the computational (Uppsala University [UU]) and data management (Critical Path Institute [C-PATH]) partners in the consortium. Experimental procedures were standardized by consensus among laboratories, leading to a standardized TKA protocol (**Supp Info – TKA protocol**). The StWG met monthly to discuss experimental procedures. StWG sub-groups were in charge of developing consensus protocols and data templates on specific *in vitro* assays (**Supp Info – Raw data templates & Guidelines**). Data templates were established to facilitate consistent reporting and automated data transfer to script-based computational analysis, while minimizing deviations from the laboratories' respective established reporting processes.

Standardized time kill assay

TKA included a control group to quantify natural growth, and four experimental groups to measure exposure to 1x MIC and 10x MIC of either MFX and INH (Figures 2A and S1). Set reference concentrations for MXF and INH were 0.06 μg/mL and 0.5 μg/mL, respectively. First, a characterized bacterial stock was thawed and distributed in the required volume to culture flasks at a bacterial suspension of 10^4 CFU/ml in 7H9-OADC-Glyc. This pre-inoculum (t = -3 days of the TKA) was incubated at 37°C for 3 days in static conditions, thus reaching a concentration of ca. 10^5 CFU/ml (t = 0 days of the TKA). Then, after this three-day recovery time, 10 mL of the pre-inoculum were distributed in each 25 cm^2 tissue culture flask, drugs from stocks added at the desired concentrations and flasks incubated at 37°C in static conditions without CO_2. At t(days)=-3, 0, 1, 2, 3, 7, 10, 14, and 21, cultures were thoroughly mixed, a 100 μL aliquot from each condition withdrawn and 10-fold serial dilution performed in PBS + tyloxapol 0.1%. Dilutions were plated on 7H10-OADC-Tx, in triplicates for each plating condition and CFU enumerated after 14 and 21 days.

Standardized data reporting

Data templates are designed to support all *in vitro* experiments conducted within the ERA4TB consortium. Each experiment requires two mandatory worksheets, including "General_Info" and "Compounds" sheets to be filled, and experiment specific worksheets to report the data in its original structure/format. The

"General_Info" sheet collects the administrative information as well as the (meta)data which will remain constant throughout the experiment. Specifically, for TKA experiments the following sheets are expected to be reported in addition to (i) General_Info and (ii) Compounds: (iii) CFU_Calculation and; (iv) CFU_Summary. Data from the "CFU_Calculation" sheet is summarized and transcribed automatically into the "CFU_Summary" sheet using excel formulas. In general, the worksheets or columns which are necessary to be filled within the template are colour coded (highlighted in red), these colour coding also indicates that this information will be extracted for downstream data transformation activities. For CFU calculation, excel formulas were composed and prepopulated in the template itself to ease the reporting process and minimise potential erroneous entries and miscalculations by taking parameters like plate concentration or the dilution factor used. Results were captured at a more granular level by reporting CFU measurements of all individual replicates along with the timepoints. Statistical parameters, like mean or standard deviation, were calculated and recorded within the template to aid data analysis. When the results observed were outside a reliable observation range, the term "AQL" (Above Quantification Limit) or "BQL" (Below Quantification Limit) was reported; this was based on the experimental setup as an agreed and observed result value.

The template is also designed to preserve higher degree of consistency by maintaining a dictionary of predefined values for all the categorical values to be reported such as, "Result Type", "Experiment Type", "Medium", "Strain", etc., to minimise the chance of human errors like wrong spellings, incorrect cases, etc. The dictionary option also allows the reporting of only agreed terminology within the consortium. For traceability, an identification number is assigned for all individual replicates; similarly, the compound(s) manufacturing batch number is also reported for all compounds used within the experiment.

Within ERA4TB, TKA experiments are conducted under different experimental conditions by changing the strain, medium or compound concentration, however, each distinct experimental condition is uniquely identified by a Group identifier. The GroupID is defined in the "Compounds" worksheet and serves as a reference across the template to identify a particular experimental condition. This minimises the redundant entries of experimental conditions in multiple worksheets across the template and it also helps merging the worksheets in downstream data transformation activities.

A more detailed explanation about the template structure can be found in the ERA4TB deliverable D1.8.[28]

QUANTIFICATION AND STATISTICAL ANALYSIS

Quantification of variability

Bacterial burden for each timepoint (t=i) was reported on absolute scale as CFU/mL and on relative scale as log10 change from baseline with the baseline set at t=0.

$$log_{10}\, CFU_{relative,t=i} = log_{10}\, CFU_{t=i} - log_{10}\, CFU_{t=0}$$

Variability was reported as coefficient of variation (CV), or relative standard deviation.

$$CV_j = \frac{sd(CFU_j)}{mean(CFU_j)}$$

Variability was reported for different levels j. The lowest level is over the technical replicates (n=3), in which the standard deviation and the mean in the equation above are calculated based on the CFU data from the replicates. This will result in one CV value per all other covariates (here: laboratory, plating method, treatment, timepoint). The CV can also be calculated at higher levels, for example to quantify the variability per plating method. In that case, the standard deviation and the mean in the equation above are calculated based on all CFU data per plating method (laboratory, treatment, timepoint, technical replicate). This will result in one CV value per plating method. For quantification of variability in the drug effect, the treatment group data was compared to the control data at the corresponding timepoints. The net drug effect approximates drug efficacy.

$$Net\ drug\ effect = \frac{log_{10} CFU_{treated} - log_{10}\, CFU_{untreated}}{log_{10}\, CFU_{untreated}} \cdot 100\%$$

Statistics

Variability was reported graphically and numerically. Data was blinded for laboratory. Three technical replicates were taken for each condition (i.e. laboratory, treatment, plating method, timepoint). Data above the quantification limit (AQL, uncountable) or below the quantification limit (BQL, zero colonies) were excluded from numerical analysis but retained in graphical analysis by placing them on the top or bottom x-axis, respectively. Parametric statistical tests were performed for data approximating a normal distribution, and non-parametric tests otherwise with significance level α=0.05, details of which are reported in Table 2 legend. Variability less than 50% CV was considered acceptable. Graphical and numerical analysis was performed in R (v. 4.0.4) through user interface RStudio (v. 1.4.1106).