Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A prospectively validated translational meta-model framework for in silico prediction of dose-ranging Phase 2B tuberculosis regimen outcomes

Eunsol Yang
University of California, San Francisco https://orcid.org/0000-0003-2581-349X
Rob C. van Wijk
University of California, San Francisco
Anu Patel
University of California, San Francisco
Sejung Hwang
University of California, San Francisco
Connie Silva
University of California, San Francisco
Niurys de Castro Suarez
University of California, San Francisco
Si-Yang Li
Johns Hopkins University School of Medicine
Eric L. Nuermberger
Johns Hopkins University School of Medicine
Rada M. Savic
rada.savic@ucsf.edu

University of California, San Francisco

Research Article

Keywords:

Posted Date: November 19th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11778751/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

A prospectively validated translational meta-model framework for *in silico* prediction of dose-ranging Phase 2B tuberculosis regimen outcomes

Eunsol Yang[1], Rob C. van Wijk[1], Anu Patel[1], Sejung Hwang[1], Connie Silva[1], Niurys de Castro Suarez[1], Si-Yang Li[2], Eric L. Nuermberger[2], Rada M. Savic[1]

[1] Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, California, United States of America

[2] Center for Tuberculosis Research, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America

Corresponding author: Rada M. Savic

University of California, San Francisco, 1700 4th Street, Rm 503C, UCSF Box 2552, San Francisco, CA 94158

E-mail: rada.savic@ucsf.edu

Abstract

Over two-thirds of drug candidates fail during mid-stage clinical development primarily due to insufficient efficacy. This challenge is particularly acute in tuberculosis, where new drugs must be evaluated in combination regimens at appropriate doses, and promising earlier findings have not consistently translated into conclusive Phase 2 results. Here we present a translational meta-model framework for quantitative prediction of dose-ranging Phase 2B tuberculosis regimen outcomes that integrates preclinical pharmacology with historically informed patient heterogeneity. We developed the methodology by back-translating four historical trials, followed by prospective validation through blinded predictions in a first-in-patient Phase 2 trial. Updating only drug-specific preclinical data, the validated framework predicts dose-response and performance versus control for novel regimens, with a quantitative efficacy benchmark further informing their treatment-shortening potential. Our platform enables rational regimen design and dose selection 3–4 years before clinical evaluation, de-risking mid-stage tuberculosis drug development, with potential for extension to other diseases.

Introduction

Fewer than one in ten drug candidates entering clinical development ultimately succeed in reaching regulatory approval.[1] The greatest attrition occurs during mid-stage clinical trials when efficacy is first formally evaluated in patients, a pivotal point for go/no-go decisions, dose selection, and progression to confirmatory trials.[1] By the completion of Phase 2, approximately 7 years of development and US$40 million have typically been invested.[2] Yet less than 30% of candidates progress from Phase 2 to Phase 3, primarily due to insufficient efficacy, resulting in a considerable loss of time and resources.[1,3] Given the persistent challenge of translating promising preclinical efficacy into mid-stage clinical success, anticipating trial outcomes at this critical stage could transform drug development decision-making by enabling earlier, evidence-based selection of candidates and doses to advance.

Tuberculosis (TB) exemplifies a particularly complex transition from preclinical to clinical development, given the inherent multidrug nature of its treatment.[4] As one of the world's deadliest infectious diseases, TB continues to present substantial unmet needs for shorter, safer, and more effective treatments, driving the rapid expansion of the drug candidate portfolio.[5,6] Novel TB agents in the pipeline therefore need to be assessed as part of combination regimens at appropriate doses for ultimate clinical approval, making Phase 2B trials the first critical decision point for regimen and dose selection. Yet outcomes at this stage have often been unfavorable or insufficiently informative for these decisions. Several recent major mid-stage trials, including ACTG A5362/Clo-Fast (NCT04311502) and PAN-TB XBQS (NCT05971602), were stopped early for futility despite encouraging earlier evidence.[7,8] Even in successful mid-stage trials, dose-ranging results have at times been inconclusive, for example, providing limited differentiation across doses or showing continued improvement with increasing doses, leaving optimal dose selection uncertain.[9,10] These challenges, particularly evident in TB drug development, raise a fundamental question: can mid-stage trial outcomes be quantitatively anticipated before clinical evaluation? Such prediction could prospectively characterize regimen performance and dose-response relationships in advance, enabling more efficient study design and subsequent clinical confirmation in patients.

Over the past decade, quantitative translational approaches have been increasingly applied across therapeutic areas to better inform clinical development decisions.[11,12] Yet their predictive performance is often validated retrospectively against known clinical data, leaving prospective evaluation of novel candidates insufficiently established.[13,14] At the same time, accumulated clinical evidence of patient heterogeneity in treatment response has driven a broader shift away from one-size-fits-all approaches and toward stratified medicine.[15,16] Translational models based primarily on preclinical data, however, have limited ability to capture such clinically observed variability, constraining the accuracy of quantitative predictions needed to guide regimen and dose selection.[17,18]

In this work, we present a translational meta-model framework for *in silico* prediction of dose-ranging Phase 2B TB regimen trials that integrates key determinants of clinical efficacy spanning preclinical pharmacology and historically informed patient heterogeneity. We show how the framework was developed through retrospective back-translation, prospective validation with blinded quantitative predictions in a first-in-patient Phase 2 trial under a prespecified analysis plan, and subsequent application to regimen ranking and Phase 2B efficacy benchmarking for treatment-shortening potential. Together, this work establishes a path for translating preclinical findings into prediction-informed regimen design and dose selection for mid-stage clinical development.

Results

Overview of a translational meta-model framework

Building on our established translational pharmacology platform,[19,20] we advanced the framework through back-translation of three regimens across four historical TB trials, incorporating new components informed by pharmacological insights and clinically observed heterogeneity to enable more reliable and accurate predictions. The resulting framework comprises six interconnected modules spanning disease immunology, drug pharmacokinetics (PK), pharmacodynamics (PD), biomarker conversion, patient phenotypes, and site-of-disease lesion coverage (Fig. 1).

Immune response, PK, and PD modules mechanistically translate treatment response from preclinical to clinical settings. First, our previously established immune response model characterizes host-mediated bacterial killing independent of treatment.[21] Next, extending the parent drug-based approach, the PK and PD modules disentangle the contributions of individual regimen components—the parent drug, active metabolite, and backbone combination—to overall treatment response. The PK model describes plasma exposure profiles for the drug of interest using preclinical plasma concentration-time data, with exposures scaled to humans allometrically or based on available clinical PK data. The PD model separately quantifies exposure-response relationships for the parent compound and its active metabolite using longitudinal lung colony-forming unit (CFU) counts from BALB/c mice treated with the combination regimen of interest while integrating PK and immune response dynamics. These relationships are then translated to humans using interspecies protein binding adjustments to reflect an effect driven by free drug concentrations.

Three newly introduced modules to this existing methodology—biomarker conversion, patient phenotype, and lesion coverage—account for clinical variability not captured by mechanistic translation. The biomarker conversion model addresses translational gaps arising from differences in bacterial burden measures across preclinical (CFU) and clinical (time to positivity [TTP]) settings as well as non-monotonic clinical TTP trajectories ("bouncers") that can prolong time to stable culture conversion (TSCC), the Phase 2B primary endpoint. With this, historical individual-level paired CFU-TTP data from 890 patients across eight TB trials, or a previously published algebraic equation,[22] are leveraged to convert participant-level baseline TTP into CFU equivalents as the initial bacterial burden for clinical simulation (Supplementary Fig. 1). Clinically observed bouncer behavior is further incorporated when deriving TSCC from simulated CFU trajectories. The patient phenotype model is grounded in clinically demonstrated heterogeneity in treatment response between easy-to-treat and hard-to-treat TB patients.[23–26] Clinical evidence on patient phenotypes from prior regimens is extrapolated to the novel regimen of interest to stratify simulated individuals and incorporate delays in culture conversion for hard-to-treat individuals. Finally, the lesion coverage model addresses uncertainty in extrapolating phenotype evidence from prior regimens by incorporating site-of-disease pharmacology, which has been demonstrated to predict treatment success in hard-to-treat patients.[27] Our previously established lung

lesion penetration tool predicts clinical lesion coverage using plasma and lesion concentration-time data from rabbits,[27] with ongoing efforts to extend the approach utilizing data from C3HeB/FeJ mice. Phenotype-specific delays in culture conversion are then adjusted based on dose-dependent lesion coverage of the novel regimen relative to prior regimens.

Outputs from all modules are integrated to quantitatively predict TSCC over 8 weeks. Trial design elements, including dosing regimens and sampling schedules, are implemented to mimic the target clinical trial *in silico*. A control comparator for predicted outcomes of novel regimens is further constructed by resampling baseline-matched patients from historical trial data. Drug-specific preclinical PK and PD data can be incorporated without modification of the overall framework, enabling forward application across the TB drug development pipeline. The framework was prospectively validated through blinded predictions of dose-ranging outcomes for a novel regimen, sorfequiline-pretomanid-linezolid (SPaL), in the first-in-patient Phase 2 NC-009 trial (NCT06058299) (Fig. 2).

Retrospective evaluation through back-translation of four historical trials

We retrospectively evaluated the framework by back-translating clinical observations from bedaquiline-pretomanid-linezolid (BPaL) in the Nix-TB (NCT02333799) and ZeNix (NCT03086486) trials,[28,29] pretomanid-moxifloxacin-pyrazinamide (PaMZ) in the STAND trial (NCT02342886),[30] and bedaquiline-pretomanid-moxifloxacin-pyrazinamide (BPaMZ) in the SimpliciTB trial (NCT03338621)[31] (Fig. 3a, Supplementary Table 1). We built on our previously published translational analyses of these three regimens[19,20] by incorporating the updated framework components—extended PK-PD, biomarker conversion, and patient phenotypes—into the predictions (Supplementary Fig. 2, Supplementary Tables 2-7). Trial predictions reflected participant-level baseline TTP and clinical and demographic characteristics and trial-specific dosing and sampling schedules. We also generated isoniazid-rifampin-pyrazinamide-ethambutol (HRZE) control comparators for the STAND and SimpliciTB trials by resampling baseline-matched participants from TBTC Studies 29 and 29X (NCT00694629).[32,33]

Across three regimens in four trials, our predictions showed overall agreement with the observed probability of stable culture conversion over 8 weeks of treatment (Fig. 3b, Supplementary Fig. 3). Quantitative

predictions closely aligned with the observed proportions achieving stable culture conversion at week 8, corresponding to the Phase 2B primary endpoint, with similar agreement observed at week 4. Dose-dependent profiles of culture conversion were reproduced in the ZeNix and STAND trials, both of which evaluated multiple dose levels.

Against HRZE control comparators, predicted hazard ratios for stable culture conversion by week 8 were well aligned with observed values in the STAND and SimpliciTB trials (Fig. 3c, Supplementary Fig. 4, Supplementary Table 8). The predictions captured relative regimen performance and whether the confidence intervals crossed the null value of 1, distinguishing regimens with and without evidence of improved performance relative to HRZE. Following back-translation, the framework was locked as Iteration 1.

Forward validation through blinded predictions in the Phase 2 NC-009 dose-ranging trial

We prospectively applied the validated Iteration 1 framework to the Phase 2 NC-009 trial (NCT06058299), which evaluated three dose levels of sorfequiline (25, 50, and 100 mg) in combination with PaL (SPaL) compared with two control regimens, BPaL and isoniazid-rifampin-pyrazinamide-ethambutol (HRZE)[34] (Fig. 3a, Supplementary Table 1). We incorporated sorfequiline-specific mouse PK and PD data and human PK data into the framework, while retaining all other components unchanged (Fig. 2, Supplementary Figs. 5 and 6, Supplementary Table 9). We generated blinded predictions of dose-ranging SPaL outcomes using participant-level baseline TTP values, clinical and demographic characteristics, and dosing and sampling schedules from NC-009, according to a prespecified analysis plan that was locked prior to the availability of trial results (Supplementary Material). The predefined primary hypotheses were that the framework could predict the dose-response relationship based on differentiation across doses and the primary trial endpoint, TSCC through 8 weeks of treatment.

The blinded predictions showed a clear dose-response relationship for the SPaL regimen, with response increasing across the three tested dose levels (Fig. 3d). Once trial outcomes became available, we conducted forward validation of these predictions against clinical observations. We confirmed the primary hypotheses, with the observed data matching the predicted dose-response relationship and generally falling within the 95% prediction intervals, with slight overprediction after week 6 for the 100 mg SPaL regimen.

To assess uncertainty in the predefined assumptions underlying the blinded predictions, we evaluated plausible alternatives to each assumption (Supplementary Tables 10-14). The predicted dose-response relationship for the SPaL regimen held across all alternative assumptions, with minimal to modest changes in quantitative predictions and no consistent improvement or worsening in agreement with clinical observations across dose levels (Supplementary Fig. 7). To evaluate the contribution of individual framework components, we performed sensitivity analyses by removing each component in turn. Predictions without individual components deviated further from clinical observations (Supplementary Fig. 8). Additional clinical evidence from NC-009, including phenotype effects and TTP bouncer behavior, was subsequently incorporated into the framework, resulting in Iteration 2.

We further evaluated the quantitative predictive performance of the Iteration 2 framework by predicting the Phase 2B primary endpoint—the proportion of patients with stable culture conversion at week 8—for the back-translated and forward-validated regimens, with predictions closely aligning with observed proportions across regimens and dose levels (Supplementary Fig. 9).

Prospective prediction of Phase 2B dose-ranging novel regimen outcomes

Extending the framework to TBAJ-587, an investigational drug that has not yet entered Phase 2B, we predicted dose-ranging outcomes of TBAJ-587 in combination with PaL (TBAJ-587-PaL) (Fig. 3a, Supplementary Table 1). TBAJ-587-specific mouse PK and PD data together with human PK data provided the drug-specific inputs for prospective prediction (Supplementary Figs. 10-11, Supplementary Table 15). The TBAJ-587-PaL regimen showed a clear dose-response relationship across the 25-200 mg dose range (Fig. 3e). Compared with the baseline-matched HRZE control, TBAJ-587-PaL at TBAJ-587 doses of 50 mg or higher was predicted to achieve more rapid stable culture conversion (Supplementary Fig. 12).

Regimen ranking and efficacy benchmarking to inform treatment-shortening potential

We next applied the Iteration 2 framework to rank Phase 2B efficacy across back-translated, forward-validated, and prospectively applied regimens, using the same participant-level baseline characteristics from NC-009 across regimens to enable direct comparison. We additionally included investigational regimens being evaluated in two Phase 2 trials—bedaquiline-moxifloxacin-pyrazinamide-delamanid (BMZD) and bedaquiline-

moxifloxacin-pyrazinamide-rifabutin (BMZRb) in TBTC Study 38/CRUSH-TB (NCT05766267)[35] and bedaquiline-pretomanid-sutezolid (BPaU) and bedaquiline-pretomanid-TBI-223 (BPaO) in ACTG A5409/RAD-TB (NCT06192160)[36]—by incorporating regimen-specific PK-PD inputs into the framework (Supplementary Tables 16-17). Predicted efficacy, defined by the proportion of patients with stable culture conversion at week 8, ranked regimens in line with their observed clinical performance (Fig. 4). Among regimens evaluated in late-phase trials, the 4-month BPaMZ regimen,[31] which met its primary efficacy endpoint, ranked highest, followed by BPaL,[28,29] whereas PaMZ,[30] which failed because of inadequate efficacy, ranked lowest. Among investigational regimens, BMZD and BMZRb showed the highest predicted efficacy, and BPaU and BPaO showed dose-dependent increases across the tested doses. TBAJ-587-PaL showed a predicted dose-response comparable to that of SPaL across matched doses.

Based on the clinically demonstrated efficacy of BPaMZ and isoniazid-rifapentine-pyrazinamide-moxifloxacin (HPZM) as 4-month shortened regimens in late-phase trials, we used the predicted BPaMZ outcome and baseline-matched observed HPZM outcome from TBTC Study 31/ACTG A5349 (NCT02410772)[37] to define a quantitative Phase 2B efficacy benchmark for identifying investigational regimens with treatment-shortening potential, corresponding to a proportion of 0.76-1.00 of patients achieving stable culture conversion at week 8 (Fig. 4). Among investigational regimens, BMZD, BMZRb, SPaL with sorfequiline doses of 100 mg or higher, and TBAJ-587-PaL with TBAJ-587 doses of 100 mg or higher fell within this benchmark.

Discussion

We established a translational meta-model framework for *in silico* Phase 2B TB trials, quantitatively predicting dose-ranging clinical outcomes of novel regimens using preclinical PK-PD data as drug-specific inputs. Through back-translation of four historical trials,[28–31] we identified and integrated elements informed by pharmacological rationale, clinically demonstrated patient heterogeneity,[23–26] and translational uncertainties as distinct modules within a unified framework, which recapitulated the observed trial outcomes. We then performed forward validation of this methodology in the Phase 2 NC-009 trial through blinded predictions that

captured the dose-response relationship and overall regimen performance. Finally, we successfully ranked regimens by their predicted outcomes in accordance with clinical performance and defined a quantitative Phase 2B efficacy benchmark anchored to clinically successful shortened regimens, which can inform the treatment-shortening potential of investigational regimens and support drug development decisions.

Integration of mechanistic and clinically informed modules was essential for accurate prediction of regimen outcomes (Supplementary Fig. 8). The disease immunology, PK, and PD modules provided the mechanistic basis for translating drug exposure–response relationships, as validated in our prior work.[19,20,38,39] Although available clinical PK was also used for translation in the present analysis, established approaches such as allometric scaling and physiologically based PK modeling can predict human drug exposure from animal PK data,[40,41] allowing the framework to be applied at the preclinical stage. We further extended the mechanistic modules to account for metabolite efficacy, particularly relevant for drugs with pharmacologically active metabolites. In our study, this applied to the diarylquinolines, a drug class central to current and investigational TB regimens.[42,43] This extension improved prediction compared with using the parent drug PK-PD alone (Supplementary Fig. 8b). However, because translation based solely on preclinical data cannot capture real-world clinical variability, we leveraged historical clinical data through back-translation to represent this heterogeneity within the framework. To account for disease phenotypes not represented by a single animal TB model,[44,45] we incorporated a patient phenotype module that extrapolates phenotype effects derived from clinically evaluated regimens to novel regimens. Risk factors defining hard-to-treat phenotypes, including cavitary disease, have been consistently identified,[23–26] supporting their potential generalizability across regimens. Because the magnitude of these effects may vary by regimen and dose, we complemented this clinical evidence with site-of-disease coverage, a lesion-centric pharmacology measure associated with treatment success in hard-to-treat patients.[27,46] This offered regimen-specific quantitative information for adapting historical phenotype effects accordingly. Beyond patient heterogeneity, clinical translation must also reconcile differences in how treatment response is measured and defined across preclinical and clinical settings. Bacteriological biomarkers such as CFU and TTP do not follow a single fixed relationship,[47] and initial culture negativity does not necessarily indicate stable culture conversion in patients because subsequent cultures may

revert to positive,[48] a pattern absent from preclinical models, where longitudinal bacteriological sampling of the same animal is challenging. Our biomarker conversion module addressed these complexities using participant-level data from previous trials, establishing rigorous quantitative links across bacteriological biomarkers and from these measures to the trial endpoint. Together, this modular architecture enables prospective application to novel regimens by combining clinically informed, drug-agnostic components with drug-specific preclinical inputs and supports iterative refinement and validation as new trial data emerge.

Most importantly, our meta-model framework was prospectively validated in the ongoing Phase 2 NC-009 trial under a prespecified analysis plan, representing the most stringent test of a translational predictive framework.[49] To our knowledge, this makes it the first-of-its-kind translational tool to quantitatively link preclinical data to mid-stage trial outcomes and demonstrate prospective predictive validity before trial readouts. The robustness of this validation was underpinned by methodological rigor, with key assumptions and uncertainties explicitly defined and plausible alternatives systematically evaluated through sensitivity analyses (Supplementary Material, Supplementary Table 11). Notably, blinded predictions confirmed the prespecified primary hypothesis that dose-response differentiation could be predicted, and this conclusion held across alternative assumptions (Supplementary Fig. 7). At the quantitative level, predictions at the 100 mg dose were somewhat optimistic after week 6. This discrepancy may reflect residual uncertainty in translating exposure–response relationships from mice to humans and in extrapolating clinically informed heterogeneity from historical regimens to novel regimens, as well as complexities in biomarker conversion. Integrating machine learning techniques with the current modules may leverage accumulated preclinical and clinical data to interrogate unresolved translational relationships and improve their quantitative characterization.[50,51] Continued back-translation and prospective validation will progressively reduce these uncertainties and iteratively advance the framework as forthcoming trial data become available.

The impact of our meta-model framework lies in providing prospective, quantitative predictions of dose-ranging Phase 2B TB regimen outcomes to guide regimen design and prioritize the most informative dose levels for clinical testing, de-risking mid-stage clinical development. Crucially, these predictions can be generated at the preclinical stage, providing actionable evidence 3 to 4 years ahead of Phase 2B evaluation. Shortening

treatment duration remains a major priority for TB drug development,[52] yet treatment-shortening trials have repeatedly failed for decades.[25,53] Using the proposed quantitative Phase 2B efficacy benchmark, our framework can assess the treatment-shortening potential of investigational regimens and doses, supporting rationale-driven decision-making earlier in development. Furthermore, novel regimen predictions can be compared with expected control outcomes derived from historical data, providing an estimate of relative treatment performance to support more informed trial design. Extending this capability, we aim to incorporate joint predictive models of longitudinal TTP and culture conversion in established control regimens, enabling more flexible and representative simulation of comparator outcomes and improved estimation of effect sizes and sample size requirements. Without modification of the framework, we prospectively ranked investigational regimens being evaluated in the Phase 2 Study 38/CRUSH-TB and A5409/RAD-TB trials by incorporating only regimen-specific PK-PD inputs. As trial-specific participant-level baseline characteristics become available, our validated framework will be applied to blinded prospective prediction of regimen outcomes in these trials, with subsequent extension to other ongoing and planned Phase 2 trials. With sufficient validation across multiple trials, we envision that the framework could support Phase 2B trials with fewer, more clinically relevant dose levels, shifting their role from dose finding toward dose confirmation and thereby reducing drug development resources.

Translational platforms have already begun to transform TB drug development.[20,39] Our established translational tool for predicting Phase 2A monotherapy early bactericidal activity (EBA) of anti-TB drugs[39] has enabled two novel drug candidates—sorfequiline and TBI-223—to bypass conventional EBA trials through *in silico* evaluation,[34,36] providing a precedent for broader application to future candidates. The present meta-model framework adds a complementary capability for mid-stage clinical trials. A further opportunity is to extend translational prediction to Phase 3. Beyond leveraging preclinical data, establishing reliable quantitative links between Phase 2B TSCC and Phase 3 outcomes would provide a critical bridge based on the most direct clinical assessment of regimen efficacy, particularly given that promising mid-stage bacteriological responses have not consistently translated into treatment success in late-phase trials.[54] Together, complementary frameworks spanning Phase 2A, Phase 2B, and Phase 3 could function as integrated *in silico* trial engines across

the TB clinical development pathway to enable evidence-based decision-making before major clinical investments, accelerating and de-risking drug development.[55]

Beyond TB, our meta-model framework represents a generalizable translational architecture with the potential to expand to other disease areas. Here, we demonstrate the integration of five core elements—disease biology, PK-PD, site-of-disease pharmacology, patient heterogeneity, and biomarker conversion—using TB as a compelling use case. These principles collectively shape how drug efficacy manifests in patients across therapeutic areas,[56–60] yet their integration within a single predictive framework remains limited. Our approach provides an example of how mechanistic and clinical data and models can be integrated to generate translational predictions with demonstrated prospective validity. Realizing such prediction-informed drug development across disease areas will demand data sharing and interdisciplinary collaboration throughout the development continuum.

In conclusion, we established a prospectively validated translational platform that enables quantitative prediction of dose-ranging Phase 2B outcomes for novel TB regimens from preclinical pharmacology data. By providing quantitative evidence before clinical evaluation, the platform can move regimen and dose decisions upstream to the preclinical stage, informing rational selection of combinations and doses for clinical advancement. This shift could make mid-stage clinical trials more informative and less risky, laying the foundation for an era of more predictable drug development.

Online Methods

Study design

Our previously established translational pharmacology platform[19,20] served as the basis for an iterative development and evaluation process comprising retrospective back-translation of four historical trials, forward validation in the Phase 2 NC-009 trial, and prospective application to a regimen not yet evaluated in Phase 2B. The framework was locked as Iteration 1 before blinded prediction of NC-009 outcomes and subsequently updated with additional NC-009 evidence to establish Iteration 2 (Supplementary Table 1). Finally, we defined a

quantitative measure of Phase 2B efficacy to inform the treatment-shortening potential of investigational regimens. Input data and models, framework components, and procedures for each stage are described below.

Data sources

Data were obtained from published and unpublished preclinical and clinical studies, with their sources and use by framework stage summarized in Supplementary Table 1. Drug-specific input data comprised preclinical PK and PD data and clinical PK information. Preclinical PK data consisted of dose-ranging plasma concentration-time measurements in BALB/c mice. Preclinical PD data comprised dose-ranging longitudinal lung CFU measurements in BALB/c mice infected with the H37Rv strain of *Mycobacterium tuberculosis* by aerosol and treated with the combination regimens of interest. Clinical PK profiles were simulated using previously published or unpublished population PK models. For sorfequiline and TBAJ-587, preclinical PK and PD data were obtained from experiments performed at Johns Hopkins University (Baltimore, MD, USA). For back-translated historical regimens, drug-specific inputs were sourced from our previous translational analyses.[19,20,39]

Drug-agnostic input data comprised historical clinical data informing biomarker conversion and patient phenotype effects. Paired baseline CFU and TTP measurements from 890 participants across eight TB trials (LIN-CL001 [NCT02279875], NC-002 [NCT01498419], NC-005 [NCT02193776], CL-010 [NCT00944021], TMC207-CL001 [NCT01215110], CL-007 [NCT00567840], NC-001 [NCT01215851], and NC-003 [NCT01691534]) were obtained from TB Alliance via personal communication (Supplementary Fig. 1). Participant-level baseline characteristics, longitudinal TTP measurements, and TSCC outcomes from four historical trials—Nix-TB (NCT02333799), ZeNix (NCT03086486), STAND (NCT02342886), and SimpliciTB (NCT03338621)—were sourced from our previous translational analyses,[19,20] and those from NC-009 (NCT06058299) were obtained from TB Alliance via personal communication.

Translational meta-model framework development

Immune response module. Bacterial dynamics in the absence of treatment were described using our previously published immune response model for *Mycobacterium tuberculosis* infection in BALB/c mice.[21] The model characterizes natural bacterial growth and death, along with time- and bacteria-dependent host immune control, reflecting its gradual development and more rapid activation at higher bacterial burdens. The model was first fit

to pretreatment bacterial dynamics data from each mouse PD experiment to quantify immune-mediated effects, as described in our previous translational analyses.[19,20,39]

Pharmacokinetics module. Mouse PK models were developed using plasma concentration-time data for the drug of interest. For drugs with an active metabolite (bedaquiline, sorfequiline, and TBAJ-587), parent and metabolite PK were jointly modeled, advancing the previous parent drug-based approach.[19,20,39] One-, two-, and three-compartment disposition models with first-order absorption with or without saturable absorption were tested. Clinical PK profiles were generated using human population PK models, previously available or developed as needed (Supplementary Table 1), at the doses evaluated in clinical trials or selected for prospective prediction.

Pharmacodynamics module. Drug effects were estimated using mouse CFU-time data after treatment initiation as previously described.[19,20,39] After accounting for host immunity, drug-mediated bacterial killing was linked to mouse drug exposure to quantify exposure-response relationships, with maximum efficacy (E_{max}) and sigmoidal E_{max} functions tested with and without a delayed effect. To better distinguish the contributions of individual regimen components, two extensions to this established method were introduced in Iteration 1. First, for drugs with an active metabolite, metabolite efficacy was incorporated alongside parent drug efficacy to account for interspecies differences in the metabolite-to-parent exposure ratio that could otherwise lead to over- or underprediction of clinical efficacy.[39] The relative *in vivo* potency (half-maximum effective concentration, EC_{50}) of the metabolite to the parent drug was assumed to equal their *in vitro* minimum inhibitory concentration (MIC) ratio (Supplementary Table 2). Second, the backbone effect was estimated from mouse CFU-time data for the backbone combination alone, where available, and incorporated when estimating exposure-response relationships for the drug of interest. This backbone effect component was identified through continued back-translation of historical trials after the NC-009 analysis plan was locked and added to the framework before blinded predictions were generated. The resulting mouse exposure-response relationships were translated to humans using clinical PK profiles based on unbound drug concentrations, accounting for interspecies differences in plasma protein binding.

Biomarker conversion module. Back-translation of historical trials revealed two gaps in translating bacterial dynamics to clinical trial outcomes: (1) the need to derive initial CFU values from clinical TTP measurements, as the PD module describes bacterial dynamics in CFU and (2) the need to account for bouncing TTP trajectories when deriving TSCC, which are not captured by the monotonic decline in simulated bacterial dynamics. First, baseline TTP values were converted to CFU equivalents by conditional sampling from historical paired baseline CFU and TTP data (n = 890). The paired data were divided into TTP quantile bins, each containing approximately 2% of observations (Supplementary Fig. 1). Each participant's baseline TTP in the trial of interest was assigned to the corresponding bin, and CFU values were sampled from its empirical distribution, with sampling repeated 200 times to capture variability. When the TTP distribution of the trial of interest was not adequately represented in the paired data, a previously published conversion equation used in our earlier translational analyses was applied instead.[19,20,22,39] The resulting CFU values served as initial bacterial burdens for simulating individual bacterial dynamics, with TSCC defined as the first time the simulated bacterial burden reached 1 bacterium. Second, trial data with available longitudinal TTP measurements were back-translated to quantify bouncer frequency and the associated delay in TSCC. In Iteration 1, bouncer behavior was informed by Nix-TB and ZeNix data (Supplementary Table 5), with NC-009 data subsequently incorporated in Iteration 2 (Supplementary Table 13). TTP bouncers were defined as participants whose TTP initially reached 42 days (the upper limit of quantification, indicating culture negativity) and subsequently fell below 42 days before achieving stable culture conversion. For each bouncer, the TSCC delay was calculated as the interval between the first negative culture result and the observed TSCC. Bouncer status was then probabilistically assigned to simulated participants based on the estimated frequency, and the median observed TSCC delay was added to predicted TSCC for bouncers.

Patient phenotypes module. Accumulated clinical evidence has demonstrated substantial heterogeneity in treatment response, with hard-to-treat patients disproportionately contributing to unfavorable outcomes.[23–26] However, the BALB/c mouse-based PD module represents an easy-to-treat phenotype, in part as evidenced by its lack of necrotizing lesions, limiting translation of treatment response to hard-to-treat individuals. In Iteration 1, historical trial data were analyzed using multivariable Cox regression to identify risk factors for delayed

culture conversion and quantify their associated effects on TSCC. STAND data were analyzed separately, while NC-005, Nix-TB, ZeNix, and SimpliciTB data were pooled to characterize phenotype effects across regimens containing the BPa backbone (Supplementary Table 6). These risk factors were used to define hard-to-treat individuals among simulated participants, with the corresponding delay added to predicted TSCC (Supplementary Table 7). The BPa-backbone phenotype effects were applied to SPaL predictions. In Iteration 2, NC-009 evidence was incorporated by adopting trial-defined phenotypes and observed between-phenotype differences in TSCC, which subsequently informed TBAJ-587-PaL predictions (Supplementary Table 14). Overall, phenotype-specific effects derived from prior regimens were extrapolated to novel regimens based on similarity in regimen composition (Supplementary Table 1).

Lesion coverage module. Lesion-centric pharmacology was incorporated to address uncertainty in extrapolating hard-to-treat phenotype effects from prior to novel regimens. Dose-dependent lesion coverage was simulated for the novel drug sorfequiline and the reference drug bedaquiline using published site-of-disease PK models[46] and compared based on the timing and depth of therapeutic target attainment. Compared with bedaquiline 200 mg, sorfequiline 25 mg showed shallower lesion coverage, 50 mg showed comparable coverage, and 100 mg achieved therapeutic targets more rapidly (Supplementary Fig. 13). Accordingly, the phenotype-specific TSCC delay was further increased by 2 weeks at 25 mg for patients with cavitation, maintained at 50 mg, and removed at 100 mg for patients whose only risk factor was bilateral cavitation (Supplementary Table 10). The impact of this adjustment was evaluated as a sensitivity analysis.

Clinical trial simulations

Each trial was simulated 200 times using participant-level baseline characteristics and trial-specific dosing and sampling schedules. Iteration 1 was used for retrospective evaluation of Nix-TB, ZeNix, STAND, and SimpliciTB and for blinded prediction of NC-009. Blinded NC-009 predictions were finalized and shared with TB Alliance before trial results were unblinded, according to the prespecified analysis plan (Supplementary Material). For framework validation, predictive performance was assessed graphically by comparing observed Kaplan-Meier profiles of the proportion without stable culture conversion over 8 weeks with the median and 95% prediction intervals of the simulated profiles, as well as by comparing the observed and predicted

proportions with stable culture conversion at weeks 4 and 8. Sensitivity analyses tested plausible alternatives to prespecified assumptions underlying the blinded NC-009 predictions (Supplementary Tables 10-14). Individual framework components were also removed in turn to assess their contribution relative to the full framework.

Iteration 2 was applied to prospective prediction of TBAJ-587-PaL outcomes, using NC-009 baseline characteristics as the reference population.

HRZE comparators were incorporated as a clinical reference to compare predicted efficacy of evaluated regimens with the control regimen. Participants from TBTC Studies 29 and 29X (NCT00694629)[32,33] were resampled 200 times by matching baseline characteristics to those of participants in the HRZE control arms of STAND and SimpliciTB for retrospective evaluation and of NC-009 for prospective prediction of TBAJ-587-PaL, based on (1) the combination of cavitation status and smear grade and (2) the distribution of each combination observed in each trial of interest.

Phase 2B efficacy benchmarking for treatment-shortening potential

To enable standardized comparison across regimens, Phase 2B outcomes were simulated using the Iteration 2 framework and NC-009 baseline characteristics as a reference population. All regimens evaluated using the framework (BPaL, PaMZ, BPaMZ, SPaL, and TBAJ-587-PaL), along with investigational regimens being evaluated in ongoing Phase 2 trials (BMZD, BMZRb, BPaU, and BPaO), were included in the comparison. Regimen efficacy was quantified as the proportion of patients with stable culture conversion at week 8, Phase 2B primary endpoint. The observed HPZM outcome from TBTC Study 31/ACTG A5349 (NCT02410772)[37] was baseline-matched to the reference population using the same approach as for the HRZE comparators and included as a clinical treatment-shortening reference. The Phase 2B efficacy benchmark, representing the reference range for investigational regimens and doses with treatment-shortening potential, was defined using the predicted BPaMZ outcome and baseline-matched observed HPZM outcome based on their clinically demonstrated treatment-shortening efficacy in late-phase trials.[31,37]

Software and statistical analysis

All data processing, statistical analyses, and visualization were conducted using R version 4.2.3 (R Foundation for Statistical Computing, Vienna, Austria). PK and PD model development and clinical trial simulations were

performed in NONMEM version 7.5.1 (ICON Development Solutions, Ellicott City, MD, USA) and Perl-speaks-NONMEM (PsN; version 5.3.1, Uppsala University, Uppsala, Sweden). Model parameters were estimated using first-order conditional estimation with interaction (FOCE-I), except when Beal's M3 method was used to handle data below the limit of quantification, for which the Laplacian estimation method was applied. Model selection was based on numerical and visual assessments, including the likelihood ratio test, precision of estimates, goodness-of-fit plots, and visual predictive checks.

Data availability

De-identified participant-level data from Nix-TB, ZeNix, STAND, and SimpliciTB trials, as well as from LIN-CL001, NC-002, NC-005, CL-010, TMC207-CL001, CL-007, NC-001, and NC-003 trials, are available to qualified researchers through the TB-Platform for Aggregation of Clinical TB Studies (TB-PACTS; https://c-path.org/programs/tb-pacts/). Data from the NC-009 trial are available from TB Alliance subject to its data-sharing policies. Preclinical PK and PD data for historical regimens and PD data and some PK data for sorfequiline and TBAJ-587 have been published previously (Supplementary Table 1). Unpublished preclinical PK data for sorfequiline and TBAJ-587 are available upon reasonable request, subject to applicable data-sharing agreements. Preclinical data underlying the regimen-specific PK-PD models for BMZD, BMZRb, BPaU, and BPaO are available upon reasonable request, subject to applicable data-sharing agreements.

Code availability

R scripts and NONMEM control streams supporting the findings of this study are available upon reasonable request. No custom software packages were developed for this study.

References

1. Thomas, D. et al. *Clinical Development Success Rates and Contributing Factors 2011–2020* (BIO, Informa Pharma Intelligence & QLS Advisors, 2021).
2. Sertkaya, A., Beleche, T., Jessup, A. & Sommers, B. D. Costs of Drug Development and Research and Development Intensity in the US, 2000-2018. *JAMA Netw Open* **7**, e2415445 (2024).
3. Morgan, P. *et al.* Can the flow of medicines be improved? Fundamental pharmacokinetic and pharmacological principles toward improving Phase II survival. *Drug Discov Today* **17**, 419–424 (2012).
4. Dartois, V. A. & Rubin, E. J. Anti-tuberculosis treatment strategies and drug development: challenges and priorities. *Nat Rev Microbiol* **20**, 685–701 (2022).
5. Pipeline | Working Group for New TB Drugs. https://www.newtbdrugs.org/pipeline/clinical.
6. Global tuberculosis report 2025. https://www.who.int/publications/i/item/9789240116924.
7. Sepulcri, C. *et al.* Recent advances and unmet needs in the treatment of drug-resistant tuberculosis. *Expert Rev Clin Pharmacol* **18**, 1037–1051 (2025).
8. Metcalfe, J. Z. *et al.* A 3-month clofazimine–rifapentine-containing regimen for drug-susceptible tuberculosis versus standard of care (Clo-Fast): a randomised, open-label, phase 2c clinical trial. *The Lancet Infectious Diseases* **26**, 46–54 (2026).
9. Olugbosi, M. *et al.* OA26-349-21 Safety and efficacy of TBAJ-876 In combination with pretomanid and linezolid: Week 8 analysis result of NC-009. Int J Tuberc Lung Dis 29 (11 Suppl 1), S1–S963 (2025).
10. Dawson, R. *et al.* Efficacy and safety of a 4-month quabodepistat, delamanid, and bedaquiline regimen for drug-susceptible pulmonary tuberculosis: a multicentre, open-label, randomised, proof-of-concept, non-inferiority, phase 2b/c trial. *Lancet Infect Dis* **26**, 1010–1021 (2026).
11. Marshall, S. *et al.* Model-Informed Drug Discovery and Development: Current Industry Good Practice and Regulatory Expectations and Future Perspectives. *CPT Pharmacometrics Syst Pharmacol* **8**, 87–96 (2019).
12. Lavé, T., Caruso, A., Parrott, N. & Walz, A. Translational PK/PD modeling to increase probability of success in drug discovery and early development. *Drug Discov Today Technol* **21–22**, 27–34 (2016).

13. Wendler, A. & Wehling, M. Translatability scoring in drug development: eight case studies. *J Transl Med* **10**, 39 (2012).
14. Bartelink, I. H. *et al.* New Paradigm for Translational Modeling to Predict Long-term Tuberculosis Treatment Response. *Clin Transl Sci* **10**, 366–379 (2017).
15. Tobias, D. K. *et al.* Second international consensus report on gaps and opportunities for the clinical translation of precision diabetes medicine. *Nat Med* **29**, 2438–2457 (2023).
16. Mosele, M. F. *et al.* Recommendations for the use of next-generation sequencing (NGS) for patients with advanced cancer in 2024: a report from the ESMO Precision Medicine Working Group. *Ann Oncol* **35**, 588–606 (2024).
17. Silva, R. B., Cheng, B., Carvajal, R. D. & Lee, S. M. Dose Individualization for Phase I Cancer Trials With Broadened Eligibility. *Stat Med* **43**, 5534–5547 (2024).
18. Cassidy, J. W., Caldas, C. & Bruna, A. Maintaining Tumor Heterogeneity in Patient-Derived Tumor Xenografts. *Cancer Res* **75**, 2963–2968 (2015).
19. Goh, J. J. N. *et al.* Prospectively predicting BPaMZ phase IIb/III trial outcomes using a translational mouse-to-human platform. *Antimicrob Agents Chemother* **68**, e0061524 (2024).
20. Strydom, N. *et al.* Dose optimization of TBI-223 for enhanced therapeutic benefit compared to linezolid in antituberculosis regimen. *Nat Commun* **15**, 7311 (2024).
21. Zhang, N. *et al.* Mechanistic Modeling of Mycobacterium tuberculosis Infection in Murine Models for Drug and Vaccine Efficacy Studies. *Antimicrob Agents Chemother* **64**, e01727-19 (2020).
22. Magombedze, G., Pasipanodya, J. G. & Gumbo, T. Bacterial load slopes represent biomarkers of tuberculosis therapy success, failure, and relapse. *Commun Biol* **4**, 664 (2021).
23. Solans, B. P., Imperial, M. Z., Olugbosi, M. & Savic, R. M. Analysis of Dynamic Efficacy Endpoints of the Nix-TB Trial. *Clin Infect Dis* **76**, 1903–1910 (2023).
24. Garcia-Cremades, M. *et al.* Risk stratification for MDR/RR-TB: a comprehensive integrated patient-level analysis from TB programs and contemporary clinical trials. *Nat Commun* https://doi.org/10.1038/s41467-026-77931-0 (2026) doi:10.1038/s41467-026-77931-0.

25. Imperial, M. Z. *et al.* A patient-level pooled analysis of treatment-shortening regimens for drug-susceptible pulmonary tuberculosis. *Nat Med* **24**, 1708–1715 (2018).
26. Chang, V. K. *et al.* Risk-stratified treatment for drug-susceptible pulmonary tuberculosis. *Nat Commun* **15**, 9400 (2024).
27. Ernest, J. P. *et al.* Lesion penetration predicts tuberculosis treatment shortening through a validated translational tool to de-risk development. *Nat Commun* https://doi.org/10.1038/s41467-026-78502-z (2026) doi:10.1038/s41467-026-78502-z.
28. Conradie, F. *et al.* Treatment of Highly Drug-Resistant Pulmonary Tuberculosis. *N Engl J Med* **382**, 893–902 (2020).
29. Conradie, F. *et al.* Bedaquiline-Pretomanid-Linezolid Regimens for Drug-Resistant Tuberculosis. *N Engl J Med* **387**, 810–823 (2022).
30. Tweed, C. D. *et al.* A partially randomised trial of pretomanid, moxifloxacin and pyrazinamide for pulmonary TB. *Int J Tuberc Lung Dis* **25**, 305–314 (2021).
31. Cevik, M. *et al.* Bedaquiline-pretomanid-moxifloxacin-pyrazinamide for drug-sensitive and drug-resistant pulmonary tuberculosis treatment: a phase 2c, open-label, multicentre, partially randomised controlled trial. *The Lancet Infectious Diseases* **24**, 1003–1014 (2024).
32. Dorman, S. E. *et al.* Daily rifapentine for treatment of pulmonary tuberculosis. A randomized, dose-ranging trial. *Am J Respir Crit Care Med* **191**, 333–343 (2015).
33. Dorman, S. E. *et al.* Substitution of rifapentine for rifampin during intensive phase treatment of pulmonary tuberculosis: study 29 of the tuberculosis trials consortium. *J Infect Dis* **206**, 1030–1040 (2012).
34. Olugbosi, M. *et al.* Protocol for a phase 2, partially blinded, randomized trial assessing the safety and efficacy of sorfequiline or bedaquiline in combination with pretomanid and linezolid in adult participants with newly diagnosed, drug-sensitive, smear-positive pulmonary tuberculosis (NC-009). *Trials* **27**, 102 (2026).
35. Kurbatova, E. V. *et al.* Phase 2C clinical trial of novel short-course regimens for the treatment of pulmonary tuberculosis: TBTC study 38/CRUSH-TB design. *Contemp Clin Trials* **158**, 108075 (2025).

36. Harrison, L. J. *et al.* ACTG A5409 (RAD-TB): Study protocol for a phase 2 randomized, adaptive, dose-ranging, open-label trial of novel regimens for the treatment of pulmonary tuberculosis. *Trials* **26**, 291 (2025).
37. Dorman, S. E. *et al.* Four-Month Rifapentine Regimens with or without Moxifloxacin for Tuberculosis. *N Engl J Med* **384**, 1705–1718 (2021).
38. Ngara, B. *et al.* Translational Modeling of BTZ-043 to Predict Phase 2A Efficacy and Evaluate Drug-Drug Interactions With Bedaquiline, Pretomanid, and Linezolid in Murine Models. *J Infect Dis* **231**, e901–e911 (2025).
39. Ernest, J. P. *et al.* Translational predictions of phase 2a first-in-patient efficacy studies for antituberculosis drugs. *Eur Respir J* 2300165 (2023) doi:10.1183/13993003.00165-2023.
40. Mahmood, I. Application of allometric principles for the prediction of pharmacokinetics in human and veterinary drug development. *Adv Drug Deliv Rev* **59**, 1177–1192 (2007).
41. Rowland, M., Peck, C. & Tucker, G. Physiologically-based pharmacokinetics in drug development and regulatory science. *Annu Rev Pharmacol Toxicol* **51**, 45–73 (2011).
42. Li, S.-Y. *et al.* Next-Generation Diarylquinolines Improve Sterilizing Activity of Regimens with Pretomanid and the Novel Oxazolidinone TBI-223 in a Mouse Tuberculosis Model. *Antimicrob Agents Chemother* **67**, e0003523 (2023).
43. WHO consolidated guidelines on tuberculosis: module 4: treatment and care. *World Health Organization* https://www.who.int/publications/i/item/9789240107243 (2025).
44. Cadena, A. M., Fortune, S. M. & Flynn, J. L. Heterogeneity in tuberculosis. *Nat Rev Immunol* **17**, 691–702 (2017).
45. Lenaerts, A., Barry, C. E. & Dartois, V. Heterogeneity in tuberculosis pathology, microenvironments and therapeutic responses. *Immunol Rev* **264**, 288–307 (2015).
46. Bustion, A. E. *et al.* The kinetics of bedaquiline diffusion in tuberculous cavities open a window for emergence of resistance. *J Infect Dis* jiaf303 (2025) doi:10.1093/infdis/jiaf303.

47. Bowness, R. *et al.* The relationship between Mycobacterium tuberculosis MGIT time to positivity and cfu in sputum samples demonstrates changing bacterial phenotypes potentially reflecting the impact of chemotherapy on critical sub-populations. *J Antimicrob Chemother* **70**, 448–455 (2015).
48. Gammino, V. M. *et al.* Bacteriologic monitoring of multidrug-resistant tuberculosis patients in five DOTS-Plus pilot projects. *Int J Tuberc Lung Dis* **15**, 1315–1322 (2011).
49. Eddy, D. M. *et al.* Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. *Med Decis Making* **32**, 733–743 (2012).
50. Terranova, N. & Venkatakrishnan, K. Machine Learning in Modeling Disease Trajectory and Treatment Outcomes: An Emerging Enabler for Model-Informed Precision Medicine. *Clin Pharmacol Ther* **115**, 720–726 (2024).
51. Huang, Z., Denti, P., Mistry, H. & Kloprogge, F. Machine Learning and Artificial Intelligence in PK-PD Modeling: Fad, Friend, or Foe? *Clin Pharmacol Ther* **115**, 652–654 (2024).
52. Target regimen profiles for tuberculosis treatment, 2023 update. https://www.who.int/publications/i/item/9789240081512.
53. Jindani, A. *et al.* Four-Month High-Dose Rifampicin Regimens for Pulmonary Tuberculosis. *NEJM Evidence* **2**, EVIDoa2300054 (2023).
54. Phillips, P. P. J. *et al.* Limited role of culture conversion for decision-making in individual patient care and for advancing novel regimens to confirmatory clinical trials. *BMC Med* **14**, 19 (2016).
55. Pappalardo, F., Russo, G., Tshinanu, F. M. & Viceconti, M. In silico clinical trials: concepts and early adoptions. *Brief Bioinform* **20**, 1699–1708 (2019).
56. Baaz, M., Cardilin, T., Lignet, F. & Jirstrand, M. Optimized scaling of translational factors in oncology: from xenografts to RECIST. *Cancer Chemother Pharmacol* **90**, 239–250 (2022).
57. Dewhirst, M. W. & Secomb, T. W. Transport of drugs from blood vessels to tumour tissue. *Nat Rev Cancer* **17**, 738–750 (2017).
58. Donath, M. Y., Dinarello, C. A. & Mandrup-Poulsen, T. Targeting innate immune mediators in type 1 and type 2 diabetes. *Nat Rev Immunol* **19**, 734–746 (2019).

59. Dingemanse, J. & Appel-Dingemanse, S. Integrated pharmacokinetics and pharmacodynamics in drug development. *Clin Pharmacokinet* **46**, 713–737 (2007).

60. Nair, A. T. N. *et al.* Heterogeneity in phenotype, disease progression and drug response in type 2 diabetes. *Nat Med* **28**, 982–988 (2022).

Figure Legends

Figure 1. Overview of the translational meta-model framework for *in silico* prediction of dose-ranging Phase 2B tuberculosis regimen outcomes. The translational meta-model framework integrates mechanistic preclinical models of host immunity, pharmacokinetics (PK), pharmacodynamics (PD), and lesion coverage with historical clinical data-driven models of biomarker conversion and patient phenotypes to quantitatively predict Phase 2B regimen outcomes. Abbreviations: CFU, colony-forming units; $E_{Metabolite}$, metabolite-driven bacterial killing rate; E_{Parent}, parent-driven bacterial killing rate; E_{Total}, total drug-induced bacterial killing rate; $f_{immune}(B,T)$, immune function dependent on bacterial load and incubation time; f_u, fraction unbound; K_{death}, natural bacterial death rate; K_{growth}, bacterial growth rate; MR, metabolic ratio; PD, pharmacodynamics; PK, pharmacokinetics; TSCC, time to stable culture conversion; TTP, time to positivity; ULOQ, upper limit of quantification

Figure 2. Prospective application of the translational meta-model framework for blinded prediction of dose-ranging SPaL regimen outcomes in the Phase 2 NC-009 trial. a, Workflow of the translational meta-model framework applied to blinded prediction of SPaL regimen outcomes. The lesion coverage component (dashed outline) was applied as a sensitivity analysis. **b,** Representative outputs from individual framework components. Solid lines represent predicted medians, and shaded areas represent the 2.5th to 97.5th percentiles of the predicted medians. Abbreviations: CFU, colony-forming units; L, linezolid; Pa, pretomanid; PD, pharmacodynamics; PK, pharmacokinetics; S, sorfequiline; TSCC, time to stable culture conversion; TTP, time to positivity

Figure 3. Back-translation, forward validation, and prospective application of the translational meta-model framework for prediction of Phase 2B tuberculosis regimen outcomes. a, Iterative learning from clinically evaluated regimens and prospective application to investigational regimens. b, Retrospective validation through back-translation of three regimens evaluated in historical trials to finalize Iteration 1 of the translational meta-model framework: BPaL in the Nix-TB (NCT02333799) and ZeNix (NCT03086486) trials, PaMZ in the STAND trial (NCT02342886), and BPaMZ in the SimpliciTB trial (NCT03338621). Predicted proportions of patients with stable culture conversion at weeks 4 and 8 were compared with observed clinical

trial outcomes. c, Predicted hazard ratios for stable culture conversion by week 8 relative to the HRZE control, compared with observed hazard ratios in the STAND and SimpliciTB trials. Predicted hazard ratios were estimated against baseline-matched observed outcomes for HRZE from TBTC Studies 29 and 29X (NCT00694629). d, Forward validation of blinded predictions of dose-ranging SPaL regimen outcomes in the Phase 2 NC-009 trial (NCT06058299). Predictions were generated before trial readout using Iteration 1 of the translational meta-model framework and subsequently compared with observed clinical outcomes once NC-009 results became available. e, Prospective prediction of dose-ranging TBAJ-587-PaL regimen outcomes. Predictions were generated using Iteration 2 of the translational meta-model framework, updated through learnings from the NC-009 trial. Boxes and whiskers represent the 2.5th to 97.5th percentiles and the range of the predicted medians, respectively. Solid lines represent predicted medians, and shaded areas represent the 2.5th to 97.5th percentiles of the predicted medians. Open diamonds and dashed lines represent observed clinical trial data. In c, filled circles and open diamonds represent predicted and observed hazard ratios, respectively, with error bars indicating 95% confidence intervals. Abbreviations: B, bedaquiline; E, ethambutol; H, isoniazid; L/LZD, linezolid; M, moxifloxacin; Pa/PMD, pretomanid; Pa_{100}, pretomanid 100 mg; Pa_{200}, pretomanid 200 mg; R, rifampin; S, sorfequiline; TBTC, Tuberculosis Trials Consortium; Z, pyrazinamide

Figure 4. Comparison of Phase 2B regimen efficacy to inform treatment-shortening potential. Predictions were generated based on participant-level baseline characteristics from the NC-009 trial (NCT06058299) using Iteration 2 of the translational meta-model framework. Regimen ranking also included prospectively predicted outcomes for investigational regimens under evaluation in the Phase 2 TBTC Study 38/CRUSH-TB (NCT05766267) and ACTG A5409/RAD-TB (NCT06192160) trials. Circles represent the median predicted proportions of patients with stable culture conversion at week 8, and horizontal lines represent the ranges of the predicted medians. Diamonds represent baseline-matched observed proportions of patients with stable culture conversion at week 8 for HPZM from TBTC Study 31/ACTG A5349 (NCT02410772) and HRZE from TBTC Studies 29 and 29X (NCT00694629). The shaded area represents the Phase 2B efficacy benchmark (0.76-1.00), defined by the predicted BPaMZ and observed HPZM outcomes, both having with demonstrated treatment-shortening efficacy in late-phase trials and provides a quantitative reference for assessing the treatment-

shortening potential of investigational regimens. BPaL, a clinically successful 6-month regimen for drug-resistant tuberculosis, was excluded from the treatment-shortening benchmark. Dose labels indicate sorfequiline doses for SPaL, TBAJ-587 doses for TBAJ-587-PaL, sutezolid doses for BPaU, and TBI-223 doses for BPaO. Abbreviations: ACTG, Advancing Clinical Therapeutics Globally; B, bedaquiline; D, delamanid; E, ethambutol; H, isoniazid; L, linezolid; M, moxifloxacin; O, TBI-223; P, rifapentine; Pa, pretomanid; R, rifampin; Rb, rifabutin; S, sorfequiline; TBTC, Tuberculosis Trials Consortium; U, sutezolid; Z, pyrazinamide

Figures

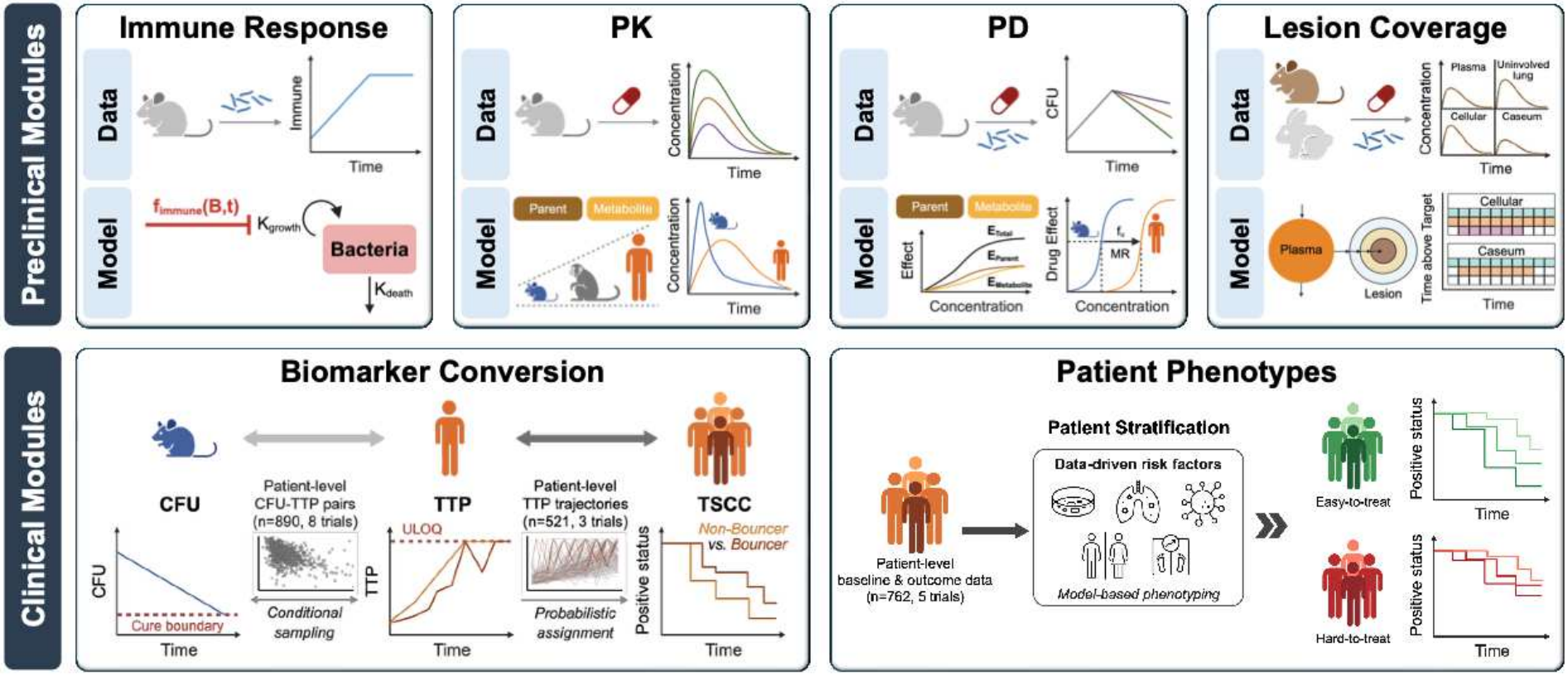


Figure 1. Overview of the translational meta-model framework for *in silico* prediction of dose-ranging Phase 2B tuberculosis regimen outcomes.

The translational meta-model framework integrates mechanistic preclinical models of host immunity, pharmacokinetics (PK), pharmacodynamics (PD), and lesion coverage with historical clinical data-driven models of biomarker conversion and patient phenotypes to quantitatively predict Phase 2B regimen outcomes.

Abbreviations: CFU, colony-forming units; $E_{Metabolite}$, metabolite-driven bacterial killing rate; E_{Parent}, parent-driven bacterial killing rate; E_{Total}, total drug-induced bacterial killing rate; $f_{immune}(B,T)$, immune function dependent on bacterial load and incubation time; f_u, fraction unbound; K_{death}, natural bacterial death rate; K_{growth}, bacterial growth rate; MR, metabolic ratio; PD, pharmacodynamics; PK, pharmacokinetics; TSCC, time to stable culture conversion; TTP, time to positivity; ULOQ, upper limit of quantification

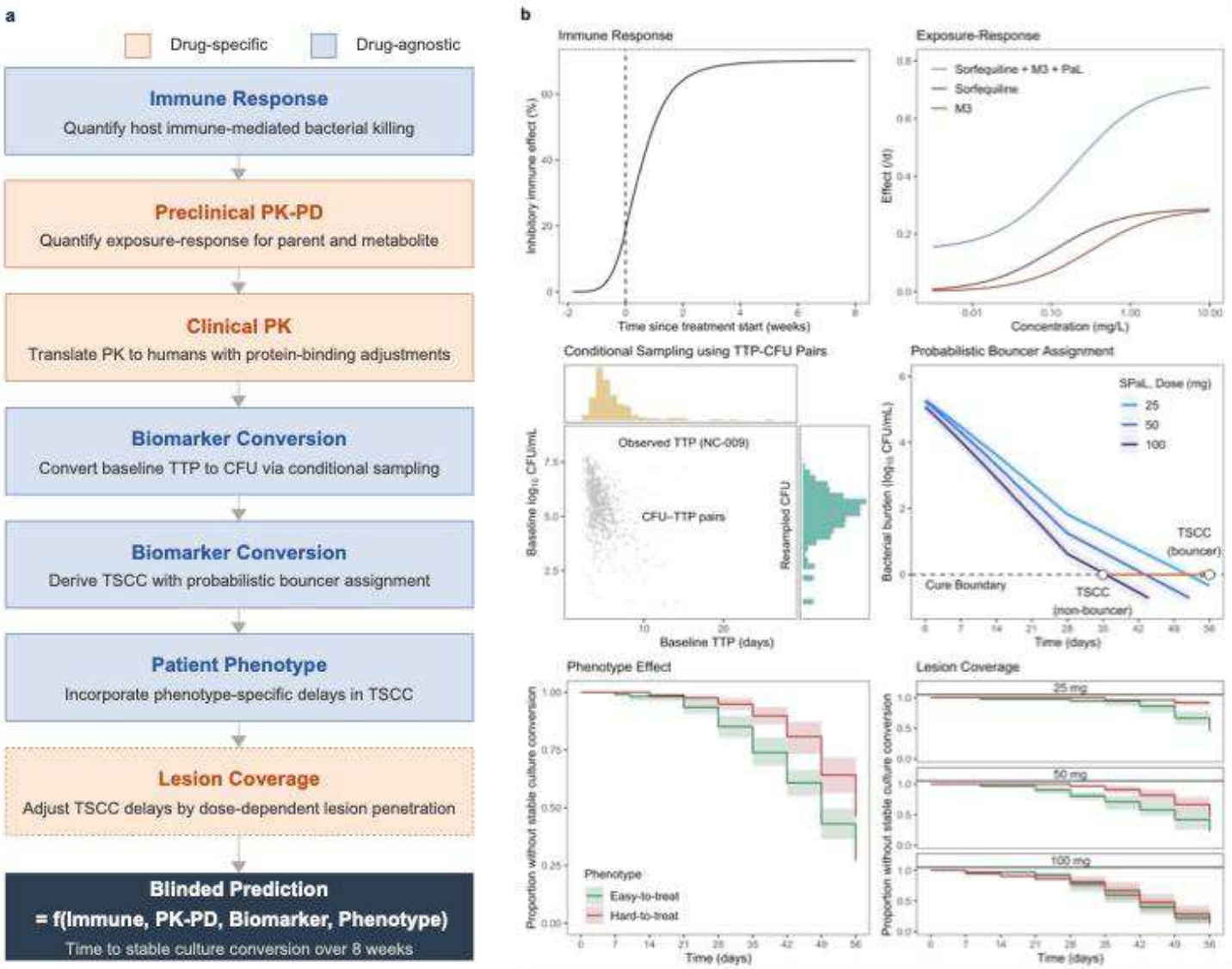


Figure 2. Prospective application of the translational meta-model framework for blinded prediction of dose-ranging SPaL regimen outcomes in the Phase 2 NC-009 trial. a, Workflow of the translational meta-model framework applied to blinded prediction of SPaL regimen outcomes. The lesion coverage component (dashed outline) was applied as a sensitivity analysis. **b,** Representative outputs from individual framework components. Solid lines represent predicted medians, and shaded areas represent the 2.5th to 97.5th percentiles of the predicted medians.

Abbreviations: CFU, colony-forming units; L, linezolid; Pa, pretomanid; PD, pharmacodynamics; PK, pharmacokinetics; S, sorfequiline; TSCC, time to stable culture conversion; TTP, time to positivity

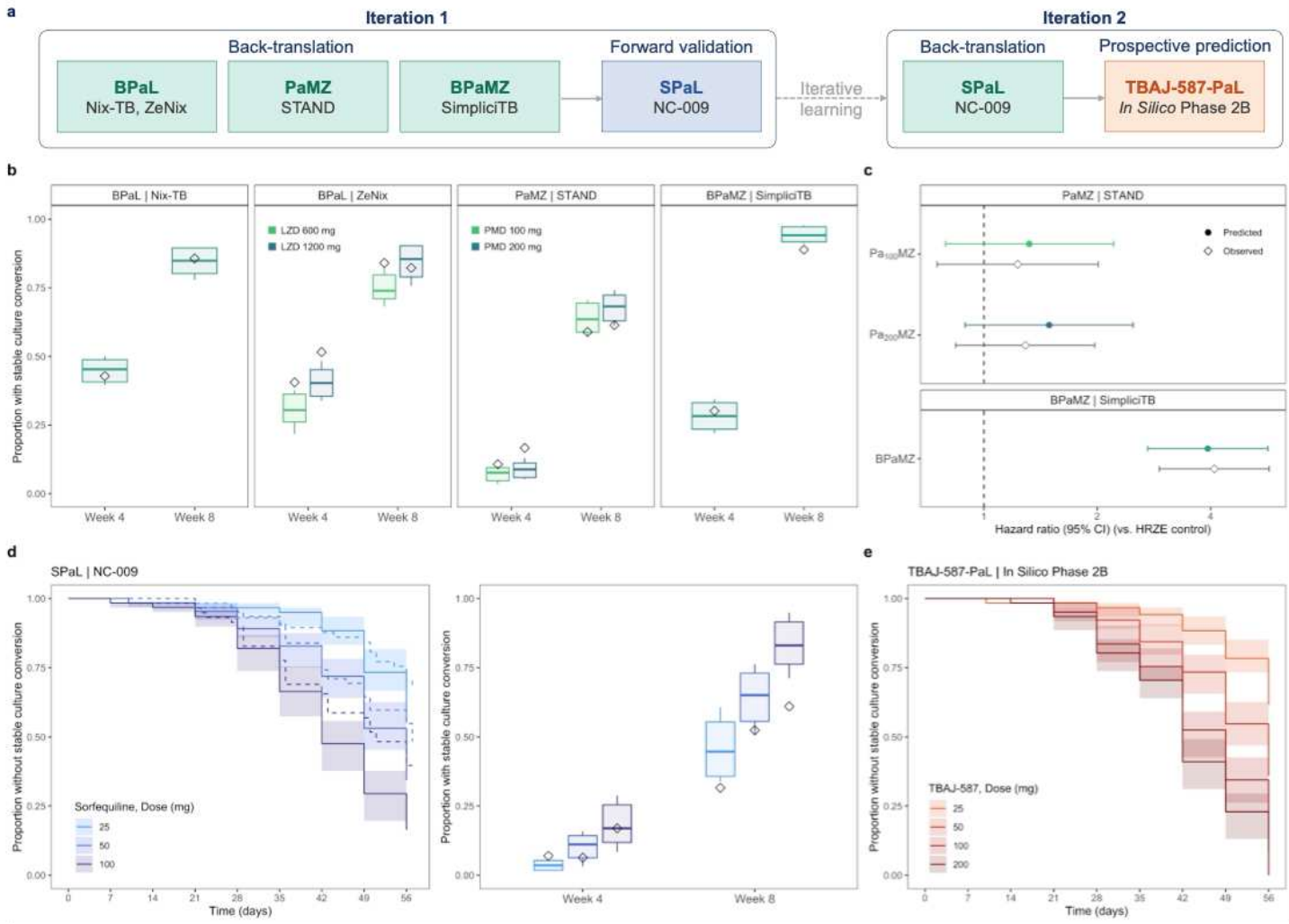


Figure 3. Back-translation, forward validation, and prospective application of the translational meta-model framework for prediction of Phase 2B tuberculosis regimen outcomes. a, Iterative learning from clinically evaluated regimens and prospective application to investigational regimens. **b,** Retrospective validation through back-translation of three regimens evaluated in historical trials to finalize Iteration 1 of the translational meta-model framework: BPaL in the Nix-TB (NCT02333799) and ZeNix (NCT03086486) trials, PaMZ in the STAND trial (NCT02342886), and BPaMZ in the SimpliciTB trial (NCT03338621). Predicted proportions of patients with stable culture conversion at weeks 4 and 8 were compared with observed clinical trial outcomes. **c,** Predicted hazard ratios for stable culture conversion by week 8 relative to the HRZE control, compared with observed hazard ratios in the STAND and SimpliciTB trials. Predicted hazard ratios were estimated against baseline-matched observed outcomes for HRZE from TBTC Studies 29 and 29X (NCT00694629). **d,** Forward validation of blinded predictions of dose-ranging SPaL regimen outcomes in the Phase 2 NC-009 trial (NCT06058299). Predictions were generated before trial readout using Iteration 1 of the translational meta-model framework and subsequently compared with observed clinical outcomes once NC-009 results became available. **e,** Prospective prediction of dose-ranging TBAJ-587-PaL regimen outcomes. Predictions were generated using Iteration 2 of the translational meta-model framework, updated through

learnings from the NC-009 trial. Boxes and whiskers represent the 2.5th to 97.5th percentiles and the range of the predicted medians, respectively. Solid lines represent predicted medians, and shaded areas represent the 2.5th to 97.5th percentiles of the predicted medians. Open diamonds and dashed lines represent observed clinical trial data. In **c**, filled circles and open diamonds represent predicted and observed hazard ratios, respectively, with error bars indicating 95% confidence intervals.

Abbreviations: B, bedaquiline; E, ethambutol; H, isoniazid; L/LZD, linezolid; M, moxifloxacin; Pa/PMD, pretomanid; Pa_{100}, pretomanid 100 mg; Pa_{200}, pretomanid 200 mg; R, rifampin; S, sorfequiline; TBTC, Tuberculosis Trials Consortium; Z, pyrazinamide

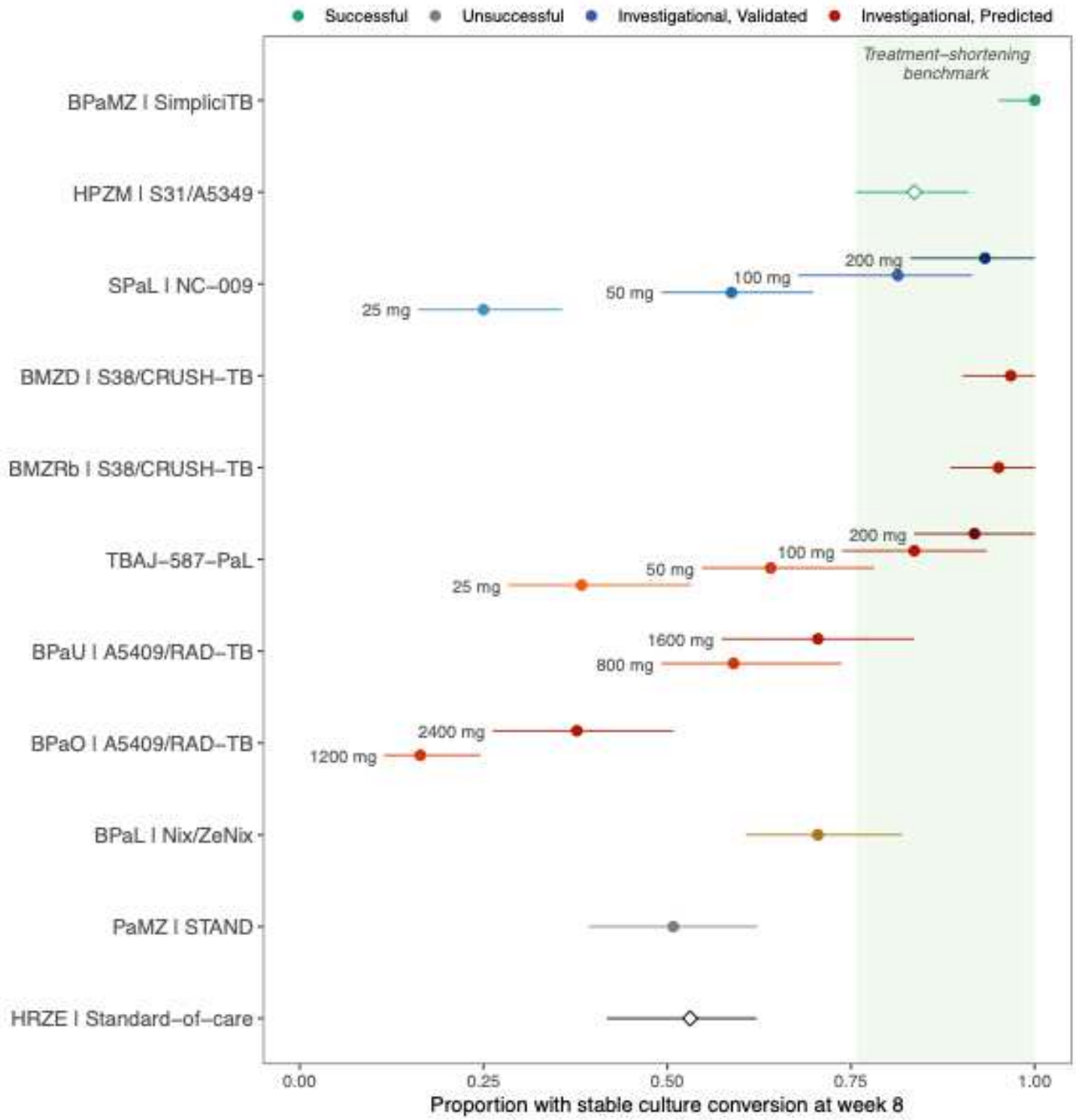


Figure 4. Comparison of Phase 2B regimen efficacy to inform treatment-shortening potential. Predictions were generated based on participant-level baseline characteristics from the NC-009 trial (NCT06058299) using Iteration 2 of the translational meta-model framework. Regimen ranking also included prospectively predicted outcomes for investigational regimens under evaluation in the Phase 2 TBTC Study 38/CRUSH-TB (NCT05766267) and ACTG A5409/RAD-TB (NCT06192160) trials. Circles represent the median predicted proportions of patients with stable culture conversion at week 8, and horizontal lines represent the ranges of the predicted medians. Diamonds represent baseline-matched observed proportions of patients with stable culture

conversion at week 8 for HPZM from TBTC Study 31/ACTG A5349 (NCT02410772) and HRZE from TBTC Studies 29 and 29X (NCT00694629). The shaded area represents the Phase 2B efficacy benchmark (0.76-1.00), defined by the predicted BPaMZ and observed HPZM outcomes, both having with demonstrated treatment-shortening efficacy in late-phase trials and provides a quantitative reference for assessing the treatment-shortening potential of investigational regimens. BPaL, a clinically successful 6-month regimen for drug-resistant tuberculosis, was excluded from the treatment-shortening benchmark. Dose labels indicate sorfequiline doses for SPaL, TBAJ-587 doses for TBAJ-587-PaL, sutezolid doses for BPaU, and TBI-223 doses for BPaO.

Abbreviations: ACTG, Advancing Clinical Therapeutics Globally; B, bedaquiline; D, delamanid; E, ethambutol; H, isoniazid; L, linezolid; M, moxifloxacin; O, TBI-223; P, rifapentine; Pa, pretomanid; R, rifampin; Rb, rifabutin; S, sorfequiline; TBTC, Tuberculosis Trials Consortium; U, sutezolid; Z, pyrazinamide

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- NC009predictionsupplementv2.0ResearchSquare.pdf