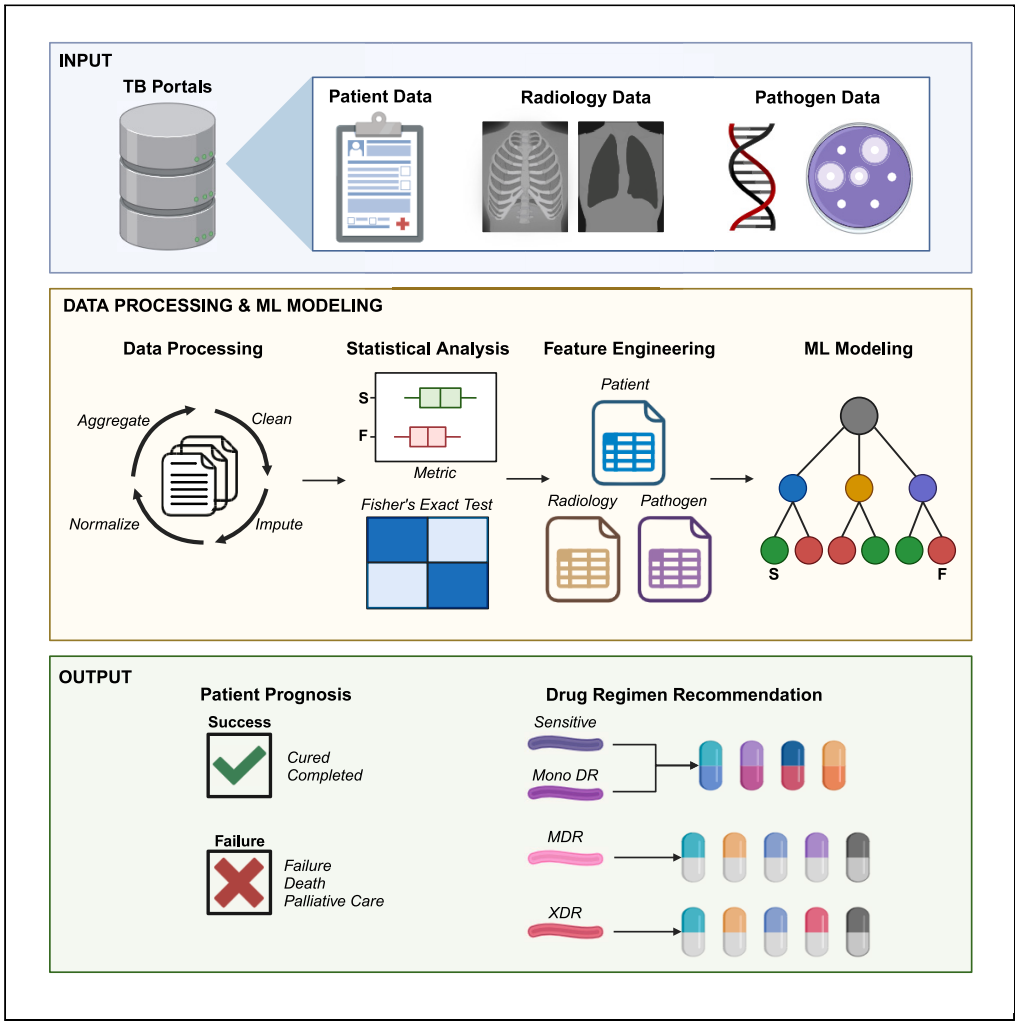
iScience

CellPress
OPEN ACCESS

Article

Integrative analysis of multimodal patient data identifies personalized predictors of tuberculosis treatment prognosis



Awanti Sambarey, Kirk Smith, Carolina Chung, Harkirat Singh Arora, Zhenhua Yang, Prachi P. Agarwal, Sriram Chandrasekaran

csriram@umich.edu

Highlights

Analysis of multimodal data from TB patients with varying levels of drug resistance

Identified clinical, genomic, imaging, and drug features predictive of treatment failure

Uncovered drug regimens effective against specific sets of drug-resistant TB patients

Multimodal ML model predicts prognosis and can lead to personalized TB treatment

Sambarey et al., iScience 27, 109025
February 16, 2024
https://doi.org/10.1016/j.isci.2024.109025

iScience

CellPress
OPEN ACCESS

Article

Integrative analysis of multimodal patient data identifies personalized predictors of tuberculosis treatment prognosis

Awanti Sambarey,[1] Kirk Smith,[1] Carolina Chung,[1] Harkirat Singh Arora,[1] Zhenhua Yang,[2] Prachi P. Agarwal,[3] and Sriram Chandrasekaran[1,4,5,6,*]

SUMMARY

Tuberculosis (TB) afflicted 10.6 million people in 2021, and its global burden is increasing due to multidrug-resistant TB (MDR-TB) and extensively resistant TB (XDR-TB). Here, we analyze multi-domain information from 5,060 TB patients spanning 10 countries with high burden of MDR-TB from the NIAID TB Portals database to determine predictors of TB treatment outcome. Our analysis revealed significant associations between radiological, microbiological, therapeutic, and demographic data modalities. Our machine learning model, built with 203 features across modalities outperforms models built using each modality alone in predicting treatment outcomes, with an accuracy of 83% and area under the curve of 0.84. Notably, our analysis revealed that the drug regimens *Bedaquiline-Clofazimine-Cycloserine-Levofloxacin-Linezolid* and *Bedaquiline-Clofazimine-Linezolid-Moxifloxacin* were associated with treatment success and failure, respectively, for MDR non-XDR-TB. Drug combinations predicted to be synergistic by the INDIGO algorithm performed better than antagonistic combinations. Our prioritized set of features predictive of treatment outcomes can ultimately guide the personalized clinical management of TB.

INTRODUCTION

Tuberculosis (TB), caused by the bacterium *Mycobacterium tuberculosis (Mtb)*, is currently the world's deadliest disease due to a bacterial infection. Of the approximately 10.6 million new cases of TB in 2021, nearly 5% of infections were accounted for by multidrug-resistant tuberculosis (MDR-TB) or extensively drug-resistant (XDR) strains,[1] with the highest burden seen in the World Health Organization (WHO) European Region including Ukraine, Moldova, Belarus, and Russia. The ongoing war and humanitarian crisis in Ukraine and bordering countries are predicted to result in increased MDR-TB cases and disruption of healthcare services.[2] Additionally, the COVID-19 pandemic has resulted in reduced access to TB diagnosis and treatment, reversing decades of progress in disease management globally. The WHO has now called for entirely new strategies to meet the goals for "*End TB*," which aims to reduce TB deaths by 95% by 2035.[3]

Treatments for TB typically involve combination therapy of 4 or more drugs, such as the WHO-recommended regimen of "*isoniazid, rifampicin, pyrazinamide, ethambutol*" (HRZE) for sensitive TB cases, to more complicated regimens that may exceed 9 months for drug-resistant TB cases. Recent updates to TB treatments have included a shorter 6-month BPaLM regimen comprising bedaquiline, pretomanid, linezolid, and moxifloxacin. This regimen may be used for MDR or rifampin-resistant TB patients in place of the previously used 9-month regimen or the longer (≥18 months) regimen.[4] While encouraging, the duration is still long with noticeable side effects which can often lead to treatment non-adherence.[5] Thus there is a need to determine more optimal treatment strategies for TB.[3]

A key challenge has been the lack of integrative algorithms to predict treatment prognosis based on the relationships between different features to identify patients needing tailored treatment approaches.[6,7] Once diagnosis of TB has been confirmed, it becomes vital that clinical healthcare workers make appropriate treatment decisions based on the individual clinical presentation of the disease. Delays in treatment initiation or providing inappropriate treatment to treat drug-resistant strains results in poor prognosis and risk of death.[8] Several factors including presence of comorbidities, patient socioeconomic status, and drug resistance are associated with TB treatment failure. Imaging techniques such as chest X-rays (CXRs) and computed tomography (CT) scans have also been shown to provide high sensitivity as a diagnostic tool and additional insight into TB disease prognosis.[9,10] We hypothesized that with the increasing availability of real-world multimodal patient information, it is now feasible to build prediction models to estimate an individualized probability of a specific treatment endpoint.[11]

[1]Department of Biomedical Engineering, University of Michigan, Ann Arbor, MI 48109, USA
[2]Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI 48109, USA
[3]Department of Radiology, University of Michigan, Ann Arbor, MI 48109, USA
[4]Program in Chemical Biology, University of Michigan, Ann Arbor, MI 48109, USA
[5]Center for Bioinformatics and Computational Medicine, Ann Arbor, MI 48109, USA
[6]Lead contact
*Correspondence: csriram@umich.edu
https://doi.org/10.1016/j.isci.2024.109025

Prior studies have focused on individual data modalities separately to make TB treatment outcome predictions with limited success.[11–14] An unbiased approach that integrates all modalities of host and pathogen data available in the clinical setting is necessary to predict treatment outcomes more accurately.

In this study, we analyze multimodal clinical data from drug-sensitive pulmonary TB, MDR non-XDR TB, and XDR TB patients across different geographical populations and implement machine learning to identify patient, drug, and pathogen features predictive of treatment prognosis. We leverage the National Institute of Allergy and Infectious Diseases (NIAID) TB Portals database, an invaluable resource for multimodal TB data that is continuously updated with new patient information.[15] The database consists of real-world patient data comprising linked socioeconomic/geographic, radiological, clinical, and treatment data as well as genomic information of infecting *Mtb* strains from patients of 10 countries with a high MDR-TB burden, including Ukraine, Moldova, Georgia, India, and Belarus. Further, in contrast to prior predictive models of TB Portals that focus on a single modality and were evaluated using only cross-validation, we conducted additional validation of our multimodal predictions on new unseen patient data that are populated in the TB Portals database, which provides more rigorous evaluation. We also analyze the longitudinal treatment regimens given to patients with different types of drug resistance and determine the drug combinations that are most significantly associated with clinical success in drug-resistant TB at each line of treatment (LoT). Overall, our integrated analysis of clinical, radiological, and genomic features can aid in clinical management of TB by identifying patients at risk of failure.

RESULTS

Statistical exploration of the TB Portals database

The TB Portals database has clinical data from 5,060 patients collected across 10 countries spanning Eastern Europe, Asia, and Africa at the time of analysis (Figures 1A and S1A) and is constantly being populated with new patients from additional countries. Data from Moldova, Georgia, and Belarus are most prevalent, followed by those of Ukraine, Azerbaijan, Romania, and Kazakhstan (Figure 1B). These represent a high burden of drug-resistant TB cases, even though China and India have an overall higher burden of TB.[1]

The dataset had 203 features which we grouped as 3 different modalities: a) patient socio-demographic and clinical characteristics, b) radiological imaging attributes derived from CXRs and CT scans, and c) pathogen drug susceptibility and genomic mutations implicated in resistance to individual drugs. The dataset originally described 5 different treatment outcomes, as listed in Table 1. We pooled the outcomes "*cured*" and "*completed*" as they depicted a "*successful outcome*," while outcomes "*failure*," "*died*," and "*palliative care*" were pooled together as "*failure*," depicting unsuccessful outcome of treatment. After removing samples that had outcomes "*Still on treatment*," "*unknown*," or "*not reported*," there were a total of 4,139 TB patients with these two outcomes of *success* and *failure* that were then considered for further analysis for each modality.

Successful outcomes were present at least 3-fold more than failure. We take this data imbalance into consideration while building and analyzing our models. We also observe a higher male population with TB compared to the female population both overall (Figure 1C), consistent with most TB reports globally. The majority of TB cases in the dataset were new or first instances of TB reported for that patient, while more than 500 patients show instances of relapse, as per the case definition (Figure S1B). There were 6 types of drug susceptibilities observed in the *Mtb* strains namely a) *Sensitive*, implying no resistance to any anti-TB drugs, b) *Mono-DR*, where resistance is seen to one first-line anti-TB drug, c) *Poly-DR*, where resistance is seen to more than one first-line anti-TB drug, d) *MDR non-XDR*, where resistance is seen to at least both isoniazid and rifampin, e) *Pre-XDR*, TB caused by *Mtb* strains that are multidrug resistant and rifampicin resistant (MDR/RR TB) and also resistant to any fluoroquinolone, and f) *XDR-TB* caused by *Mtb* that is resistant to isoniazid and rifampin, plus any fluoroquinolone and at least one of three injectable second-line drugs (i.e., amikacin, kanamycin, or capreomycin).[16,17] The MDR non-XDR type of resistance is observed to be most prevalent across infected populations in the dataset (Figure 1E). Patients with drug-resistant TB show a corresponding increase in treatment times compared to drug-sensitive TB, with the longest average treatment duration seen for patients with MDR non-XDR and XDR TB (Figure S1C).

Drug combinations associated with successful outcomes for drug-resistant TB

In the TB Portals dataset, we observe patients with up to 12 lines of treatment as seen in Figure 2A. We further broke down the cohort to look at individual types of drug resistance, and the regimens most associated with success in each instance. We chose a subset of regimens which were given to a minimum of 10 patients for statistical analysis. Majority of patients received a single line of treatment (LoT 1), and 70.3% of these patients had a successful outcome (Figure 2A).

To determine associations of drug regimens with outcome, we focused on drug regimens given up to 3 LoTs. For each LoT, we calculated the number and prevalence of each treatment regimen in both outcome classes of success and failure. A Fisher's exact test was performed to determine the significant associations between regimen and outcome, broken down by type of resistance using Fisher's exact test (Table 2). The odds ratios are indicative of the likelihood of patients receiving this regimen having a successful outcome, with values > 1 implying likelihood of success, while values < 1 are associated with failure.

As seen in Figure 2B, the number of treatment regimens a patient received varied with the type of resistance of the infecting *Mtb* strain, as a majority of Sensitive TB patients (1,110) and MDR non-XDR TB patients (1,089) received a single treatment regimen (LoT1) (Figure 2B). Of the 1,110 Sensitive TB patients in LoT1, 1,025 patients received the standard-of-care regimen of HRZE, of which 84.2% (864 patients) had successful outcomes. For patients with MDR non-XDR TB who received a single LoT, the treatment regimen "*bedaquiline, clofazimine, cycloserine, levofloxacin, linezolid*" was strongly associated with treatment success, and most patients with MDR non-XDR TB were given this regimen ($p = 0.0004$). Surprisingly, the regimen of "*bedaquiline, clofazimine, linezolid, and moxifloxacin*" showed significant association with failure

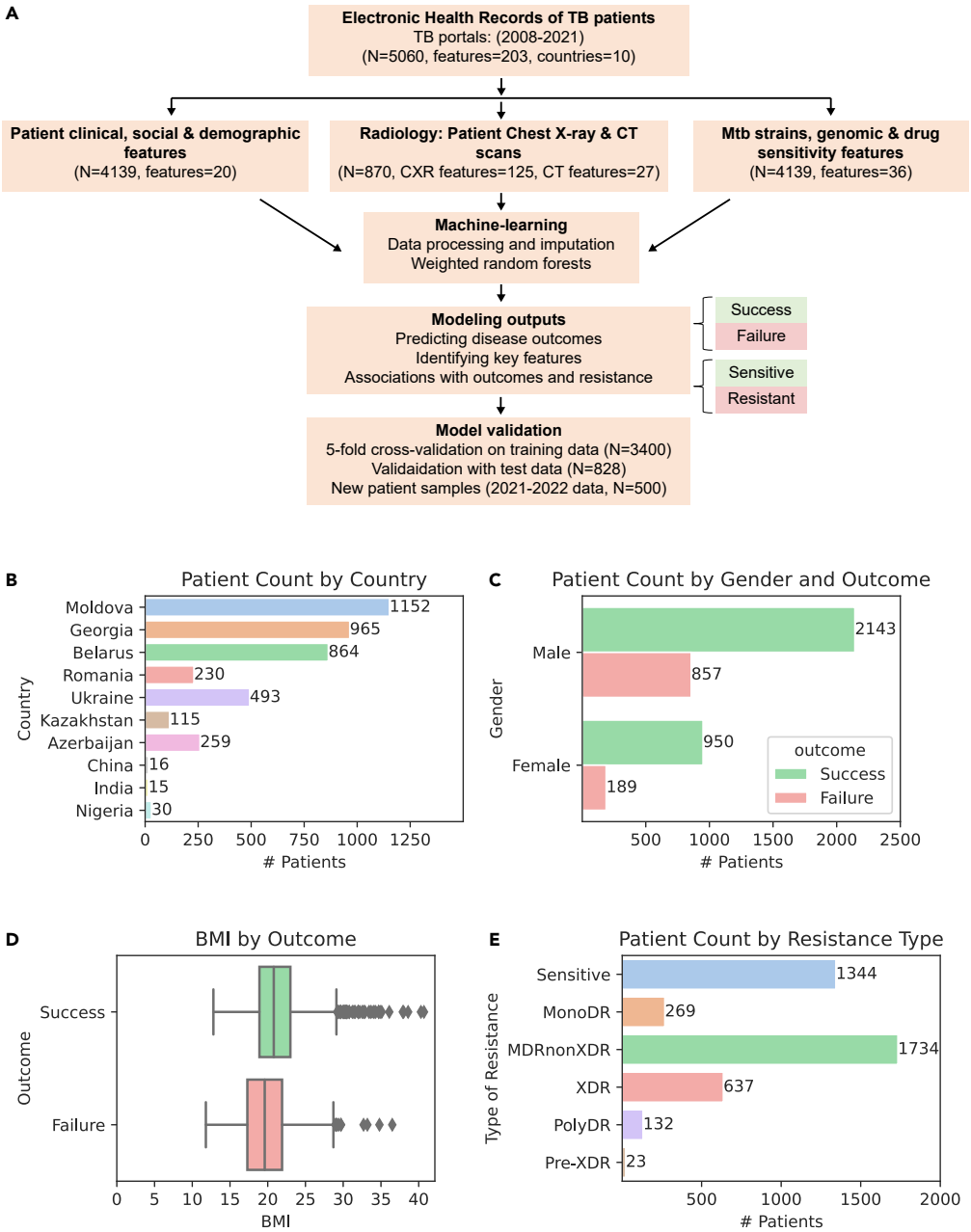


Figure 1. Statistical exploration of the TB portals database

(A) Workflow adopted in this study to analyze multimodal clinical data of TB patients. A total of 5,060 patients from 10 different countries were considered for the analysis, with over 200 features available representing pathogen genomic features, patient clinical and social features, as well as radiological features derived from patient chest X-rays. We analyzed each of these categories separately. Data analysis involved building machine learning models (random forests) and conducting statistical analyses to predict disease outcomes—grouped as Success and Failure, respectively. Feature importance was performed using Shapley analysis as well as hypergeometric tests to determine the predictors associated with both success and failure. A final unified model was built comprising top features across all modalities. All model performances were evaluated by 2 types of validation: a) cross-fold validation (k = 5) where the input data were split into training and test sets and b) predicting outcomes on newer patients that were populated in the TB Portals database.

(B) Overview of the data present in the TB Portals database. Distributions of number of patients per country for all 10 countries present in the TB Portals database.

(C) Distributions of gender by 2 pooled outcomes of success and failure.

(D) Distribution of BMI levels seen across outcomes of success and failure.

(E) Drug susceptibility in *Mtb* strains isolated from patients implicated in resistance to individual drugs.

Table 1. Pooled treatment outcomes considered for this study

Outcome	Definition	Number of patients	Outcome assigned for this study
Cured	Treatment completed as recommended by the national policy without evidence of failure AND three or more consecutive cultures taken at least 30 days apart are negative after the intensive phase	2,556	Success
Completed	Treatment completed as recommended by the national policy without evidence of failure BUT no record that three or more consecutive cultures taken at least 30 days apart are negative after the intensive phase	537	Success
Failure	Treatment terminated or need for permanent regimen change of at least two anti-TB drugs because of a lack of conversion by the end of the intensive phase or a bacteriological reversion in the continuation phase after conversion to negative, or evidence of additional acquired resistance to fluoroquinolones or second-line injectable drug, or adverse drug reactions	481	Failure
Died	A patient who dies for any reason during the course of treatment	565	Failure
Palliative Care	An approach that improves the quality of life of patients and their families facing the problems associated with life-threatening illness, through the prevention and relief of suffering by means of early identification and impeccable assessment and treatment of pain and other problems, physical, psychosocial, and spiritual	101	Failure

Definitions of the outcomes were provided by the TB Portals data dictionary.

for MDR-TB ($p = 0.003$). There were no regimens significantly associated with success or failure for Mono-DR, Pre-XDR, or XDR-TB in the first LoT.

There were only 2 longitudinal regimens with 2 LoTs for MDR non-XDR TB that were given to at least 10 patients. The regimen "*bedaquiline, clofazimine, cycloserine, levofloxacin, linezolid*" followed by "*clofazimine, cycloserine, levofloxacin, linezolid*" was given to 22 patients, and all of them showed a successful treatment outcome ($p = 0.0004$). The regimen "*clofazimine, capreomycin, ethambutol, isoniazid, pyrazinamide, moxifloxacin, prothionamide*" followed by "*clofazimine, ethambutol, moxifloxacin, pyrazinamide*" was also associated with success, although not statistically significant.

There were 193 Mono-DR TB patients who received 2 LoTs. Patients who received *HRZE* followed by *HR* showed significant success ($p < 0.05$). For 147 XDR-TB patients that had 2 LoTs, only 1 regimen *bedaquiline, clofazimine, cycloserine, levofloxacin, linezolid* followed by *bedaquiline, clofazimine, cycloserine, delamanid, linezolid* was given to at least 10 patients all of whom had a successful outcome. There were no regimens significantly associated with success for Poly-DR and Pre-XDR TB at the 2nd LoT.

A total of 326 patients received 3 LoTs. For these patients, there were no regimens that were given to at least 10 patients, so there was not enough statistical power to analyze them. There was 1 regimen given to MDR non-XDR TB patients: *ethambutol, isoniazid, rifampin, pyrazinamide -> cycloserine, ethambutol, kanamycin, levofloxacin, protionamide, pyrazinamide -> cycloserine, ethambutol, levofloxacin, protionamide, pyrazinamide* that was given to 4 patients who all had a successful outcome, and no failures.

Modeling socio-demographic and clinical modalities reveals comorbidities and drug interactions as most predictive of treatment failure

We grouped different features by modality and built machine learning models separately for each modality as illustrated in Figure 1A. There were a total of 20 features describing patient clinical, social, and demographic aspects. After cleaning and imputation of the data, we built a random forest machine learning model. The random forest algorithm was chosen due to its ability to handle mixed data and its interpretability.[18,19] Additionally, random forest performed best in a cross-validation test that we performed on our multimodal data using a variety of machine learning algorithms (Figure S3). We conducted 5-fold cross-validation and hold-out validation on blinded test data as well as on newer patient data populated in the TB Portals database from August 2021 to January 2022. Our model predictions had accuracies of 74.1%,

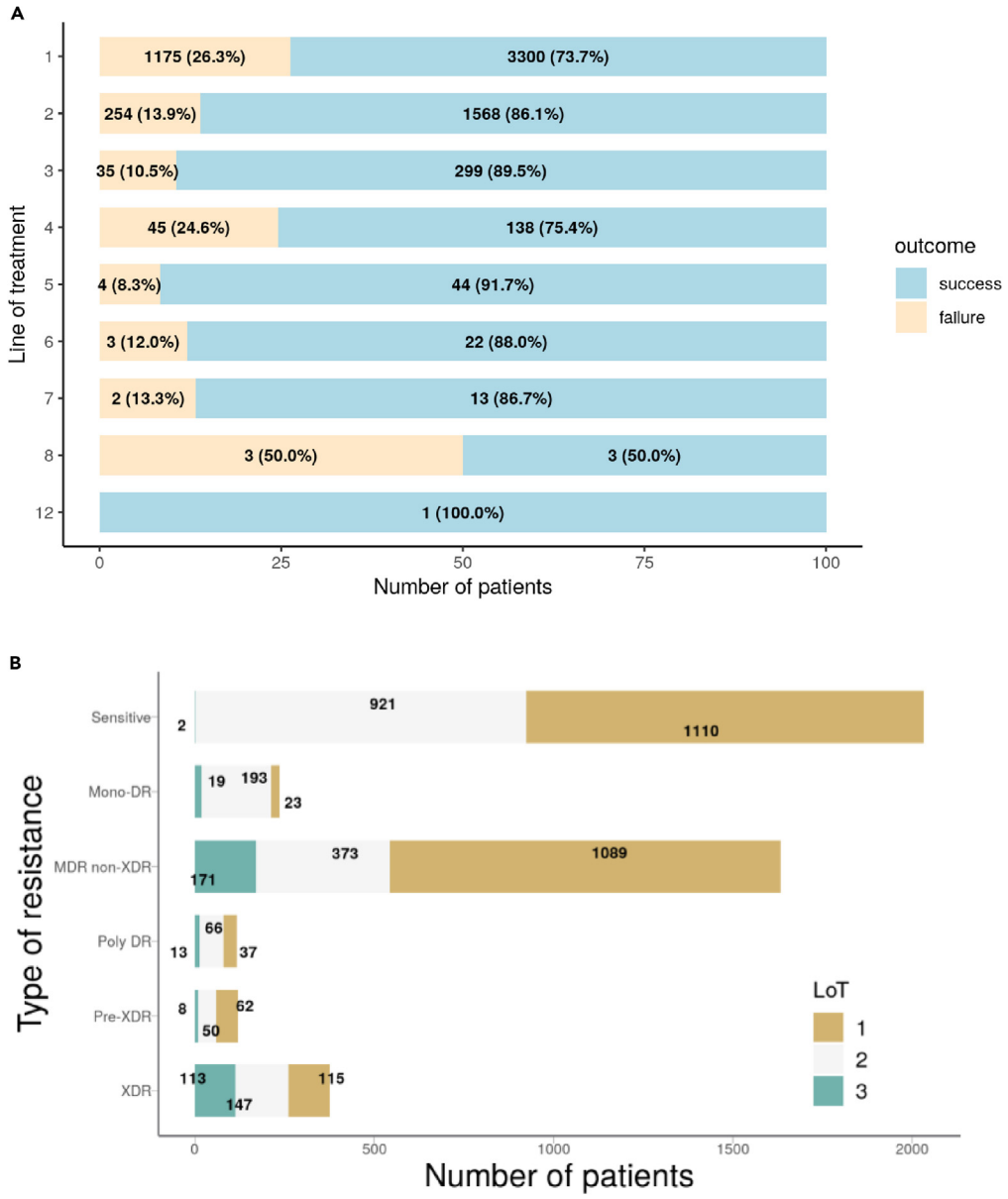


Figure 2. Distribution of treatments across lines of therapy and varying drug resistance

(A) Treatment outcomes for patients who received different lines of therapy (LoT). For each LoT, the number and percentage of patients by outcome are shown.
(B) Patients with different types of resistance and the LoT they received.

78.3%, and 81.7% as well as Matthews correlation coefficient (MCC) values ranging from 0.25 to 0.40 for the cross-validation, hold-out validation, and validation on new patient datasets, respectively (Figures 3A and 3B, and Table S5). The MCC is a robust statistical metric which produces a high score only if the predictions have high precision, recall, and accuracy (i.e., all of the confusion matrix categories) and accounts for the biased distribution of success and failure in the dataset.

The most important predictors of failure were determined by Shapley analysis, a game-theoretic approach that measures each feature's marginal contributions to the model.[20,21] Shapley values provide additional insight into the predictors compared to other measures of feature importance. Rather than a global, unidirectional importance score, Shapley values provide local scores at individual observations and inform us which direction the model's prediction was pushed by the feature. For the socio-demographic model, the top predictors based on the Shapley analysis included case definition, BMI, employment, and age of onset (Figures 3C and 3F).

To further inspect what values among these features were truly associated with failure, we analyzed these feature value distributions for patients for each country as well as across countries. Figure 3G shows the Shapley feature value distributions for the 3 most populated

Table 2. Treatment regimens across different lines of therapy and type of drug resistance associated with treatment success

Line of therapy	Type of *Mtb* resistance	Regimen	Success with regimen	Failure with regimen	p value	FDR-corrected p value	Odds ratio	95% confidence interval
1	Sensitive TB	*'Cm', 'E', 'H', 'Lfx', 'Z', 'R'*	6	4	0.05972*	0.0192	1.3444	0.258–4.483
		'E', 'H', 'Lfx', 'Z', 'R'	16	7	0.07698*	0.0192	1.1794	0.432–4.001
		'E', 'H', 'Z', 'R'	864	161	0.04975*	0.0068	1.028	0.706–1.969
	MDR non-XDR TB	*'Bdq', 'Cfz', 'Cs', 'Lfx', 'Lzd'*	399	95	0.00061*	0.0004	1.6083	1.222–2.116
		'E', 'H', 'Z', 'R'	14	20	0.00001*	0.0217	0.2139	0.107–0.429
		'Bdq', 'Cfz', 'Lzd', 'Mfx'	10	15	0.00013*	0.0033	0.2064	0.092–0.464
		'Cm', 'Cs', 'Eto', 'Lfx', 'Z'	2	8	0.00032*	0.048	0.0786	0.017–0.372
2	Sensitive TB	*['E', 'H', 'R', 'S']->['H', 'R']*	45	3	0.7502	1	0.9018	0.271–3.005
		['E', 'H', 'R', 'Z']->['E', 'H', 'R']	21	0	0.6264	1	Inf	NaN - Inf
		['E', 'H', 'R', 'Z']->['H', 'R']	347	21	1	1	0.9985	0.564–1.766
		['E', 'H', 'R']->['E', 'H']	9	1	0.4457	1	0.5394	0.067–4.34
		['E', 'H', 'R']->['H', 'R']	409	25	1	1	0.9794	0.559–1.715
	MDR non-XDR TB	*[Bdq', 'Cfz', 'Cs', 'Lfx', 'Lzd']->[' Cfz', 'Cs', 'Lfx', 'Lzd']*	22	0	0.01114*	0.00046	Inf	NaN - Inf
		['Cfz', 'Cm', 'E', 'H', 'Mfx', 'Pto', 'Z']->['Cfz', 'E', 'Mfx', 'Z']	16	2	0.5469	0.7268	2.1068	0.474–9.368
	MonoDR TB	*['E', 'H', 'R', 'Z']->['E', 'Lfx', 'R', 'Z']*	33	5	0.5444	0.7935	0.6696	0.225–1.991
		['E', 'H', 'R', 'Z']->['H', 'R']	42	3	0.0512*	0.0135	1.7364	0.482–6.254
		['E', 'H', 'R']->['H', 'R']	45	6	0.5951	0.7935	0.7738	0.278–2.158
	XDR TB	*[Bdq', 'Cfz', 'Cs', 'Lfx', 'Lzd']->['Bdq', 'Cfz', 'Cs', 'Dld', 'Lzd'].*	10	0	0.03097*	0.0048	Inf	NaN - Inf

Legend: Significant values (p value < 0.05) are marked with an asterisk. Abbreviations: Bdq - Bedaquiline; Cm - Capreomycin; Cs - Cycloserine; Dld - Delamanid; E– Ethambutol; Eto- Ethionamide; H- Isoniazid; Lfx - Levofloxacin; Lzd - Linezolid; Rifampin - R; Z - Pyrazinamide.

countries in the dataset, namely Belarus, Moldova, and Ukraine. We observe similar trends in the values for top features across the remaining countries (Figure S2) and see similar performance metrics for models built and tested independently on social features for individual countries (Table S6).

BMI ranked as one of the top predictors, with lower values (in blue) associated more with failure compared to higher values. Higher age of onset (red), higher resistance (red/purple), and lower education (blue) are significantly associated with failure. Patients with BMI values less than 18, classified as "*Underweight*" have strong associations with treatment failure ($p = 1.32 \times 10^{-23}$), compared to those that were classified *healthy*, *overweight*, or *obese*. (Figure 3H). We also observe more failure in patients with disability ($p = 3.55 \times 10^{-25}$), lower employment ($p = 2.09 \times 10^{-5}$), and lower education levels ($p = 1.26 \times 10^{-25}$). Prior history of alcohol ($p = 4.46 \times 10^{-34}$), drug abuse ($p = 1.01 \times 10^{-16}$), and smoking ($p = 3.76 \times 10^{-10}$) impacts the treatment outcome as well. Patients with HIV ($p = 3.25 \times 10^{-32}$), anemia ($p = 2.34 \times 10^{-31}$), or hepatitis B/C (2.6×10^{-4}) as comorbidities are also seen to have poorer outcomes, in concordance with what has been formerly observed in the literature.[22] Interestingly, while diabetes is a prevalent comorbidity commonly and has been previously associated with an increased risk of failure and death in TB patients,[23,24] we do not see it significantly associated with poor treatment outcomes in our retrospective analysis.

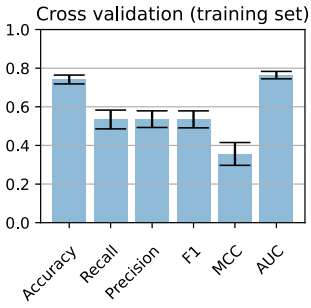
Analysis of pathogen features identifies co-occurring mutations linked with treatment failure

There are 27 different *Mtb* families represented in the data encompassing 3 major genetic lineages L2, L4, and L1, with the L2-Beijing sub-lineage seen to be most prevalent in patients across countries (Table S1). L2 has received much attention due to its high virulence, fast disease progression, and association with antibiotic resistance.[25] The *Mtb* strains show varying drug susceptibilities to the 28 different drugs in the clinical dataset. The Beijing, H3, and T1 families were most prevalent in all infections observed, and strains from these families show all types of drug susceptibilities, from Sensitive TB to XDR TB (Figures 4A–4C).

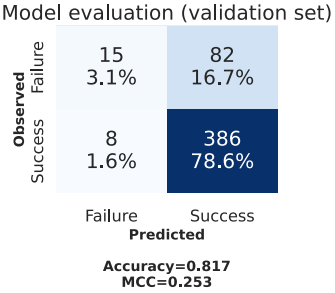
Our machine learning analysis using the pathogen genomics and drug resistance modality could predict treatment failure with accuracies of 72%, 74%, and 74% using cross-validation, hold-out validation, and the new patient data, respectively (Figures 4D and 4E, and Table S5). We observe that the prediction accuracy was relatively lower using pathogen features alone, compared to using clinical features. Feature analysis revealed that higher numbers of *Mtb* colonies as determined by culture and the presence of mycobacterial growth were strongly associated

Socio-demographic model: All patients

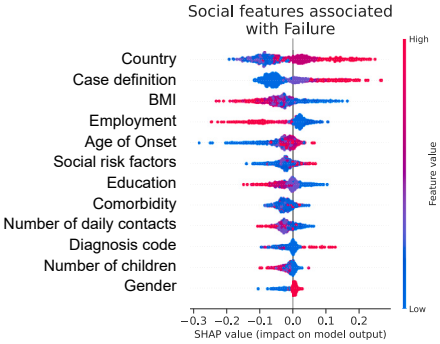
A



B

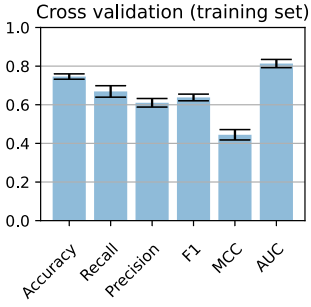


C

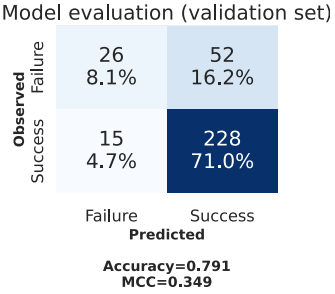


Socio-demographic model: Drug-resistant patients

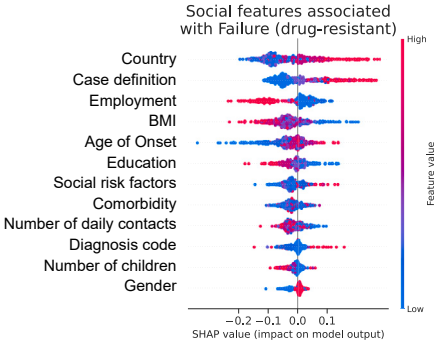
D



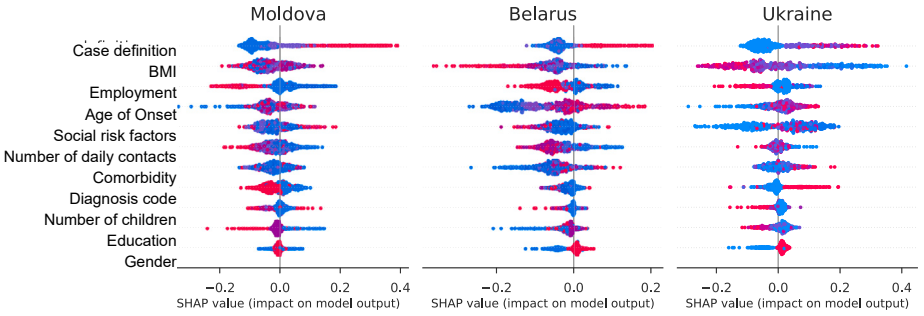
E



F



G **Shapley values for 3 highest populated countries**



H

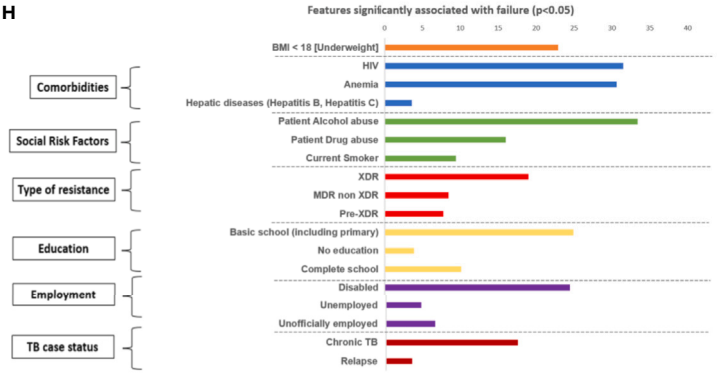


Figure 3. Modeling patient socio-demographic and clinical features

(A) All patients: Model predictions on 80% training data using 5-fold cross-validation with model evaluation metrics of accuracy, recall, precision, F1 score, MCC, and AUC values (error bars represent 95% CI); top features predictive of the failure outcome.

(B) Model evaluation on 20% hold-out validation data.

(C) Top features predictive of failure for all patients.

(D) Drug-Resistant TB patients: Model predictions on 80% training data using 5-fold cross validation with model evaluation metrics of accuracy, recall, precision, F1 score, MCC and AUC values (error bars represent 95% CI).

(E) Model evaluation on 20% hold-out validation data for Drug resistant TB patients.

(F) Top features predictive of failure for drug resistant TB patients.

(G) Top features and their Shapley values across the 3 most populated countries in TB Portals. Each dot represents a single patient, with blue color representing lower values for each feature. The X axis represents the impact of the feature value driving model outcomes, with values above 0 associated with failure, and those below 0 associated with success.

(H) For 6 features identified to be significantly associated with failure ($p < 0.05$), the feature category is broken down to depict significant associations between feature subtypes and failure, highlighted by different colors. All p values are significant ($p < 0.05$) after multiple hypothesis correction (FDR <0.1). The scale represents negative log p values determined by hypergeometric tests.

with treatment failure. Of the 28 different drugs present in treatment combinations across all regimens, resistance to rifampin, isoniazid, kanamycin, streptomycin, ethambutol, capreomycin, and amikacin was more significantly predictive of treatment failure as determined by Shapley feature analysis. It is important to note that the TB Portals data only report mutation at the gene level and not the overall variant level. Notably, co-occurring mutations in Mtb genes gyrA, inhA, katG, rpoB, and rpsL were seen to be strongly linked with unfavorable treatment outcomes (odds ratio 8.6, Table S3). Mutations in the gyrA gene, particularly at positions 90, 91, and 94, have been frequently reported among fluoroquinolone-resistant *Mtb* (FQr-MTB) isolates.

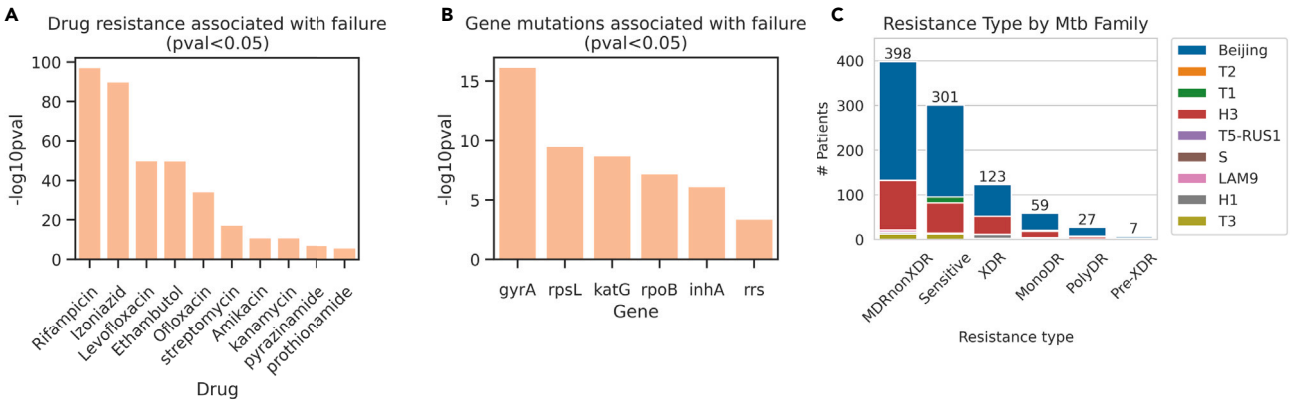
For the different drug regimens assigned per patient (1,084 unique regimens in total), drug interaction scores were calculated using the machine learning tool INDIGO-MTB.[25] These scores provide a quantitative description of the nature of interaction between drugs (synergy, additivity, and antagonism), with lower scores associated with strong synergy among the drugs in each regimen. Here, a drug combination predicted as synergistic implies that the same amount of growth inhibition would be achieved with a lower dose when the drugs are combined compared to them acting individually. We observe that regimens used to treat drug-sensitive TB have the strongest synergy and correspondingly lower treatment times (Figures S1C and S1D), while drug combinations used to treat resistant TB cases show a wider range including weak synergistic and antagonistic interactions. These interactions may lead to increased treatment times for drug-resistant TB.

Lung volume, pleural effusion, and bronchial obstruction are significantly predictive of treatment failure and associated with drug-resistant TB

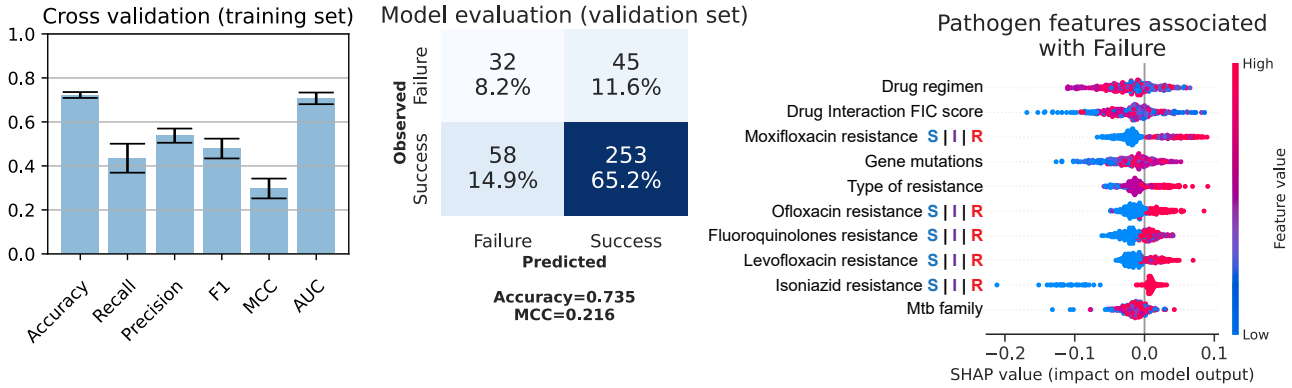
Radiological imaging using CXR and/or CT scans are typically used to aid clinicians in reaching a diagnosis of TB and monitoring clearance of infection. They complement *Mtb* culturing and symptoms. Imaging can reveal TB lesions of differing size, shapes, and characteristics (e.g., cavitation) occurring anywhere in the lungs. This dataset had 406 patients with at least 1 X-ray available, with an average of 3 X-rays per patient taken over their course of treatment. We chose the CXR taken closest to the date of treatment initiation to assess if there were any features that would help indicate treatment prognosis at that time. We validated our model on 289 new patients with CXR available in the January 2022 dataset. As only 59 additional patients had CT scans available, we conducted validation with new patient entries from TB Portals using CXR data alone.

Machine learning models built on CXR and CT data individually could predict treatment prognosis reasonably well within the training data (accuracies of 74% and 83%, respectively) (Figures 5A–5C and Table S5). Validation with new patient CXR data showed an accuracy of 75% and MCC of 0.22 (Figure S3). Despite the CT model predicting well within cross-validation, it performed very poorly in the external validation set (accuracy 80%, MCC = 0). This is most likely due to the small number of observations and poor imputation of missing data. The most significant TB-related manifestations that are predictive of both treatment outcome and resistance are shown in Figures 5C and Table 3.

The TB Portals database annotates each CXR into sextants to highlight regions that were afflicted in the patient's lungs viz. upper left (UL), upper right (UR), middle left (ML), middle right (MR), lower left (LL), and lower right (LR). TB is typically manifested as an upper respiratory infection, but it has been shown to infect different parts of the lung based on disease severity.[13] Our analysis indicated that the overall affected abnormal volume across both lungs was more significant in predicting poor disease prognosis rather than individual affected sextants. The presence of bronchial obstruction, pleuritis (inflammation of the tissues that line the lungs and chest cavity), a decrease in lung capacity, and the presence of lung opacities due to nodules, nodes (seen in shadow patterns), and airspace disease (infiltrates) were strongly predictive of both treatment failure and were significantly associated with drug-resistant TB. It is important to note that we analyzed the clinical presentation of drug resistance as seen through imaging data for our study, as drug resistance can only be predicted directly by culturing the infecting *Mtb* strain. Interestingly, the presence of calcified or partially calcified nodules across most sextants, nodes larger than 10 mm, and the presence of post-TB residuals were seen to be associated with drug resistant TB (DR-TB) cases, but they were not especially predictive of treatment failure (Table 3). The overall percentage of abnormal lung volume, which is a quantification of TB severity, was most predictive of treatment failure. Similarly, lymph node enlargement (lymphadenopathy)[14,26] and collapse of the upper or middle lung were indicative of a poor treatment prognosis. Chest imaging usually results in several hundred features (152 imaging features in this dataset), and our analysis identified 21 TB manifestations to be most clinically predictive.



D Pathogen model: All patients



E Pathogen model: Drug-resistant patients

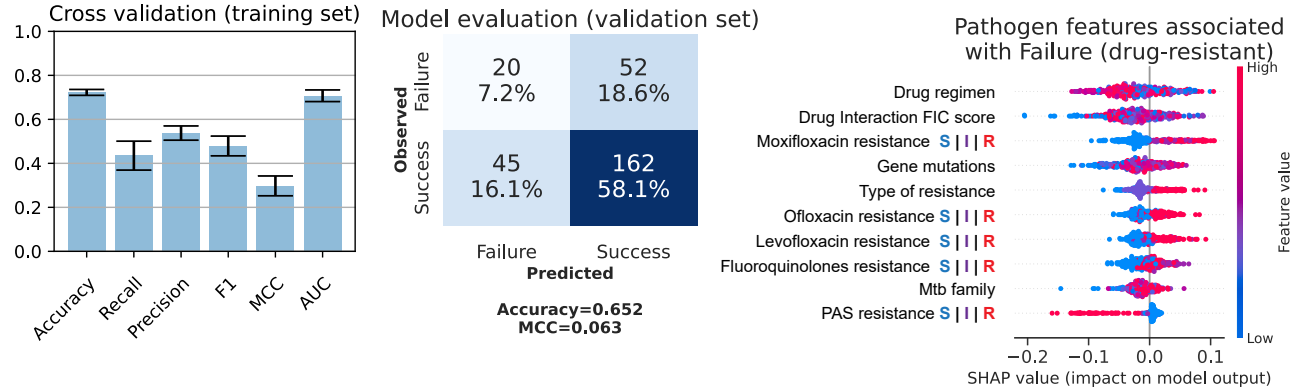


Figure 4. Modeling pathogen genomic features and drug susceptibilities

(A) Distribution of *Mtb* families across different types of resistance observed in the TB Portals data.

(B) Resistance to individual drugs associated with failure determined by hypergeometric tests.

(C) Gene mutations associated with failure. All p values are significant ($p < 0.05$) after multiple hypothesis correction (FDR <0.1).

(D) All patients: model predictions on 80% training data using 5-fold cross-validation with model evaluation metrics of accuracy, recall, precision, F1 score, MCC, and AUC values (errors bars represent 95% CI); model predictions on 20% hold-out validation data; top features predictive of the failure outcome.

(E) Model evaluation and top features predictive of failure for drug-resistant patients only.

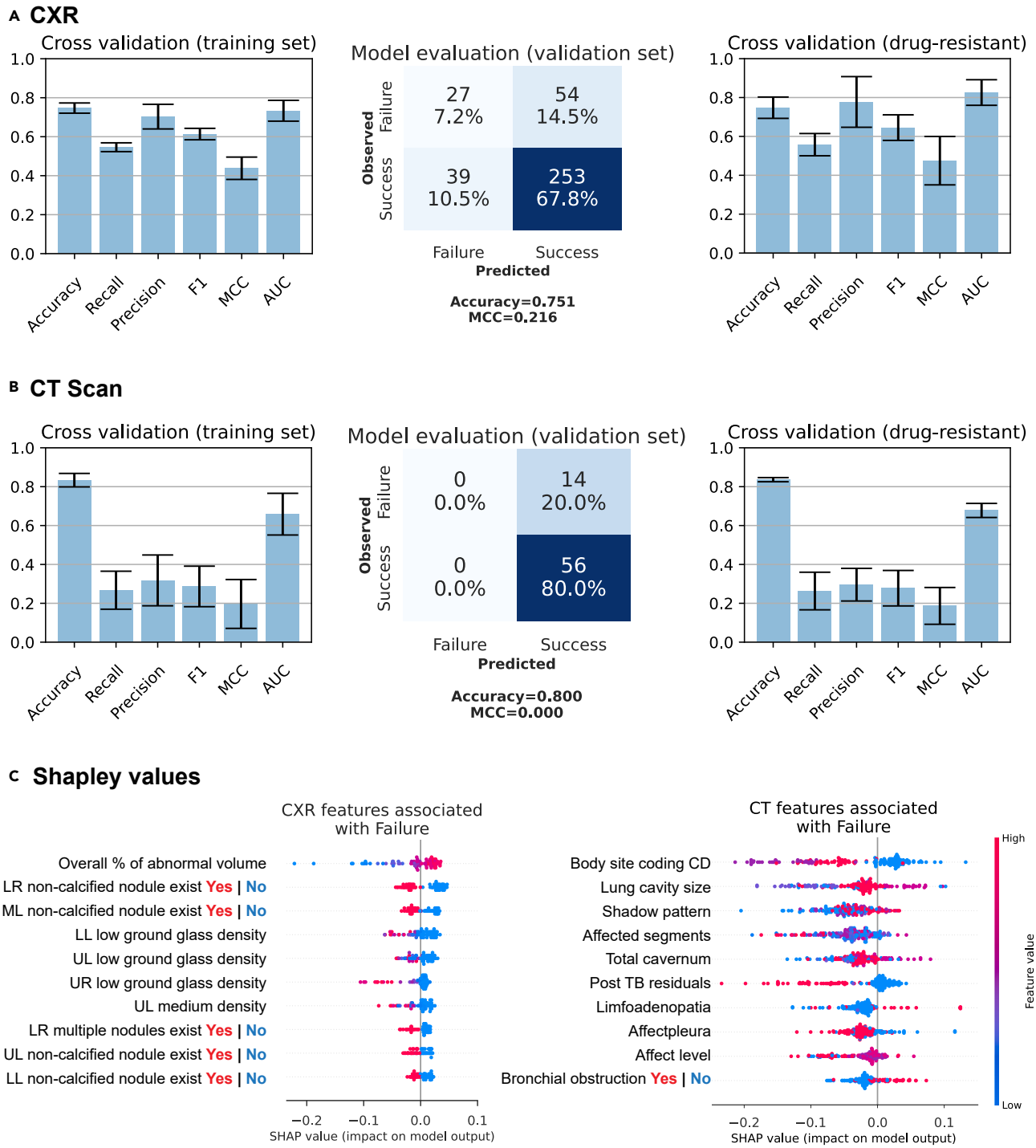


Figure 5. Modeling imaging modalities (CXR and CT data)

(A) CXR data: model predictions on 80% training data using 5-fold cross-validation with model evaluation metrics of accuracy, recall, precision, F1 score, MCC, and AUC values and model predictions on 20% hold-out validation data.

(B) CT scans data: model predictions on 80% training data using 5-fold cross-validation with model evaluation metrics of accuracy, recall, precision, F1 score, MCC, and AUC values and model predictions on 20% hold-out validation data; top features predictive of the failure outcome.

(C) Top features predictive of the failure outcome identified for both CXR and CT data.

Table 3. Imaging features associated with treatment failure and drug-resistant TB

Chest imaging features	Association with failure ($p < 0.05$)	Association with resistance ($p < 0.05$)
Body site: Both lungs	4.16e-06	9.09e-04
Pleuritis: Yes	4.15e-05	2.11e-60
Bronchial obstruction: Yes	7.33e-04	0.04
Lung capacity decrease: Yes	0.0017	4.35–04
Opacity due to nodule, node, and infiltrate	0.0068	3.22e-04
Overall percent of abnormal volume	3.1e-013	$p > 0.05$
Small cavities: UL|UR	4.2e-11	$p > 0.05$
Small cavities: ML|MR	5.4e-06	$p > 0.05$
Collapse: UL|UR	7.5e-05	$p > 0.05$
Enlarged lymph nodes: Yes	4.75e-04	$p > 0.05$
Nodes more than 10 mm	$p > 0.05$	7.16e-011
Affect pleura: Yes	$p > 0.05$	0.022
Post TB residuals: Yes	$p > 0.05$	0.0028
LL| LR calcified or partially calcified nodules: Yes	$p > 0.05$	6.03e-05
ML|MR calcified or partially calcified nodules: Yes	$p > 0.05$	3.12e-08
UL| UR calcified or partially calcified nodules: Yes	$p > 0.05$	7.00e-11

Chest imaging features that are significantly associated with both resistance and treatment failure ($p < 0.05$) and those associated with either failure or drug resistance.

Integrated multimodal analysis outperforms model predictions of individual modalities

A multimodal machine learning model was trained using the socio-demographic, pathogen, and CXR datasets. CT data were excluded due to the low number of overlapping patients. In the primary model, we prioritized sample size and therefore used patients in any of the three datasets, which led to imputing a large amount of missing X-ray data. Using this model, which included 163 features after data cleaning, treatment failure was predicted with an accuracy of 79% (area under the curve [AUC] of 0.826 and MCC of 0.495) in the cross-validation, and an accuracy of 83.2% (AUC of 0.84 and MCC of 0.451) on an external validation set with new patients (Figure 6A and Table S5). A secondary model, which prioritized data completeness, used only patients present in all three datasets and performed almost as well, despite the much smaller sample size (Figure S4). For comparison, the primary integrated model performed better in the external patient predictions than the individual models for clinical (AUC 0.797, MCC 0.253), pathogen (AUC 0.704, MCC 0.216), CXR (AUC 0.645, MCC 0.216), and CT (AUC 0.504, MCC 0.0) feature models (Table S5). These results are comparable to or better than those of prior models in the literature that predict TB treatment failure (AUC = 0.70 [28], 0.74 [29], and 0.79 [13]).

For model interpretation, a simplified decision tree model was created to visualize the typical prediction path that a patient observation follows when its clinical outcome is being decided. The complexity of the tree was constrained by limiting the tree depth to enable ease of interpretation and visualization. The random forest model is made of an ensemble of such decision trees. The simplified decision tree (Figure 6E) used in the integrated model is helpful to understand which features are instrumental in differentiating between success and failure. For example, following the rightmost path through the tree, patients that are unemployed (employment field <2), older (age >57), prescribed a mostly non-synergistic drug regimen with INDIGO score >0.88, and a smoker (social risk factor score = 1) are at a high risk of failure (3-fold increase in odds).

We next analyzed correlations among the top 30 features based on Shapley values (Table 4), from the integrated model (Figure 6D). We find high correlations ($p < 0.05$) among features within each modality, as well as between features across modalities (Table S4). The type of resistance is significantly associated with the drug regimen and corresponding drug interaction score determined by INDIGO-MTB. We also observe that higher drug interaction scores are more associated with drug regimens used to treat drug-resistant TB in comparison with drug-sensitive TB (as seen in Figure S1D), implying lesser synergy among combination therapies used to treat drug-resistant TB. Interestingly, we observe higher INDIGO scores (i.e., more antagonistic interactions) correlated with worse clinical presentation of TB, as indicated by overall percentage of abnormal lung volume. Resistance to isoniazid, rifampicin, and fluoroquinolones emerges as the top feature among the pathogen modality (Table 4).[27] The type of resistance is also correlated with the imaging attributes of increased bronchial obstruction and affected lung segments as seen in the clinical presentation of drug-resistant TB. We speculate that this is likely due to the association between chronic obstructive pulmonary disease (COPD) and drug-resistant TB.[28] COPD patients also have weakened immunity which makes them more susceptible to MDR-TB. Disease severity as reported by diagnosis code correlated with the number of *Mtb* colonies and the affected lung segments. Social risk factors and comorbidities show associations with gender and education; employment and BMI are correlated, highlighting the socioeconomic impact in TB prognosis and management.[29,30]

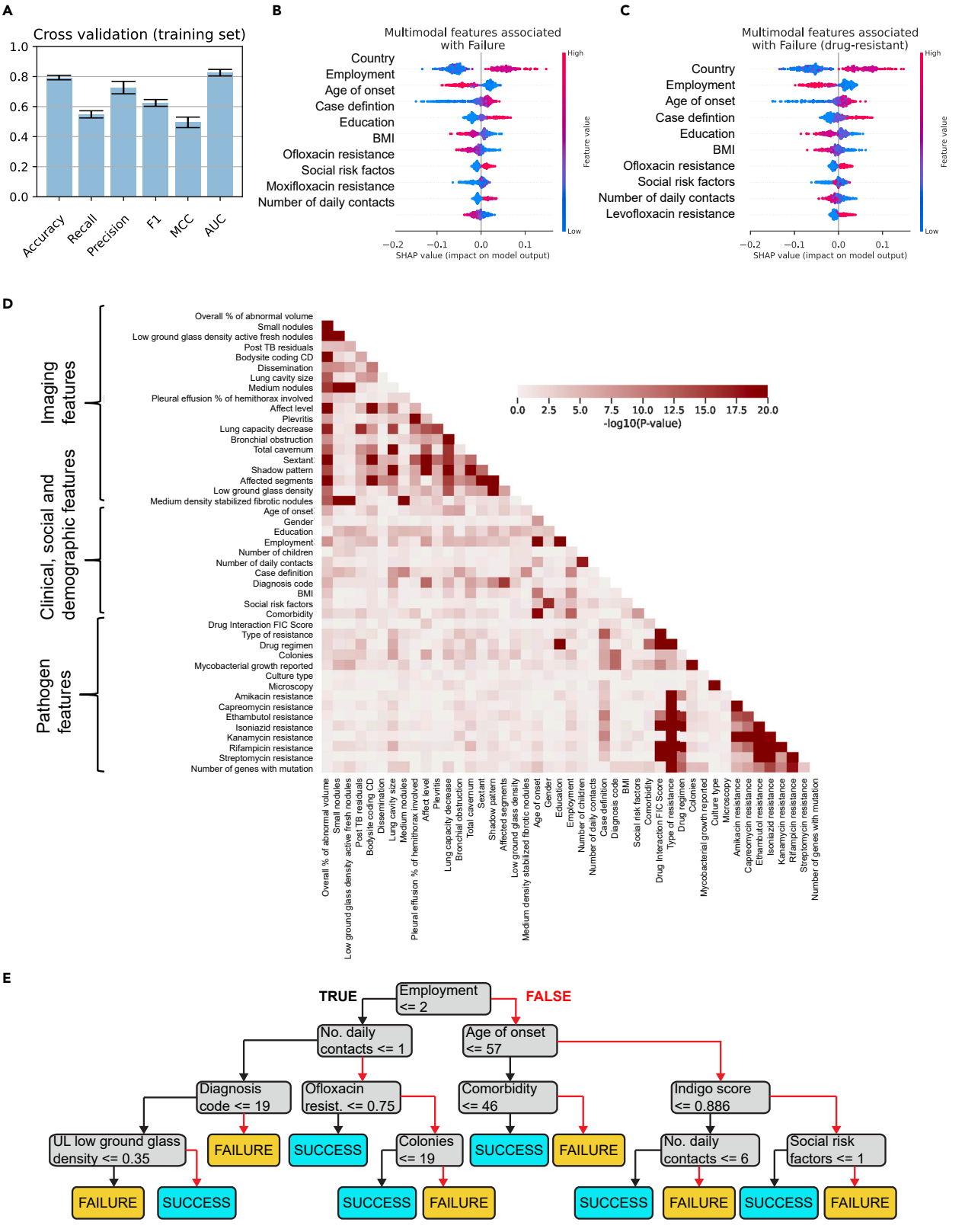

A
Cross validation (training set)
1.0
0.8
0.6
0.4
0.2
0.0
Accuracy
Recall
Precision
F1
MCC
AUC
B
Multimodal features associated with Failure
Country
Employment
Age of onset
Case defintion
Education
BMI
Ofloxacin resistance
Social risk factos
Moxifloxacin resistance
Number of daily contacts
High
Feature value
Low
−0.2
−0.1
0.0
0.1
SHAP value (impact on model output)
C
Multimodal features associated with Failure (drug-resistant)
Country
Employment
Age of onset
Case defintion
Education
BMI
Ofloxacin resistance
Social risk factors
Number of daily contacts
Levofloxacin resistance
High
Feature value
Low
−0.2
−0.1
0.0
0.1
SHAP value (impact on model output)
D
Imaging features
Clinical, social and demographic features
Pathogen features
Overall % of abnormal volume
Small nodules
Low ground glass density active fresh nodules
Post TB residuals
Bodysite coding CD
Dissemination
Lung cavity size
Medium nodules
Pleural effusion % of hemithorax involved
Affect level
Plevritis
Lung capacity decrease
Bronchial obstruction
Total cavernum
Sextant
Shadow pattern
Affected segments
Low ground glass density
Medium density stabilized fibrotic nodules
Age of onset
Gender
Education
Employment
Number of children
Number of daily contacts
Case definition
Diagnosis code
BMI
Social risk factors
Comorbidity
Drug Interaction FIC Score
Type of resistance
Drug regimen
Colonies
Mycobacterial growth reported
Culture type
Microscopy
Amikacin resistance
Capreomycin resistance
Ethambutol resistance
Isoniazid resistance
Kanamycin resistance
Rifampicin resistance
Streptomycin resistance
Number of genes with mutation
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
-log10(P-value)
E
TRUE
Employment <= 2
FALSE
No. daily contacts <= 1
Age of onset <= 57
Diagnosis code <= 19
Ofloxacin resist. <= 0.75
Comorbidity <= 46
Indigo score <= 0.886
UL low ground glass density <= 0.35
FAILURE
SUCCESS
Colonies <= 19
SUCCESS
FAILURE
No. daily contacts <= 6
Social risk factors <= 1
FAILURE
SUCCESS
SUCCESS
FAILURE
SUCCESS
FAILURE
SUCCESS
FAILURE

Figure 6. Integrated multimodal model

(A) For the integrated model, the model predictions on 80% training data using 5-fold cross-validation with model evaluation metrics of accuracy, recall, precision, F1 score, MCC, AUC values and correlations are shown.

(B) Top features identified by Shapley values predictive of failure for all patients.

(C) Top features identified by Shapley values predictive of failure for drug-resistant patients.

(D) Feature associations between top features across all modalities. The color scale is based on -log(p values) for correlations between these features.

(E) Decision tree for the integrated model showing features differentiating between success and failure at each leaf of the tree. The black arrows indicate if the condition in the leaf is true, and the red arrow implies that the condition is false.

DISCUSSION

TB remains a significant challenge globally, especially in the aftermath of the COVID-19 pandemic as TB services have been disproportionately affected.[31,32] Of particular concern are the regions in Eastern Europe which have a high incidence of TB, TB-HIV co-infection, and MDR-TB (MDR non-XDR TB).[33,34] The ongoing war and humanitarian crisis in Ukraine have also affected the healthcare systems in the area. Given these scenarios, it becomes especially important to analyze available clinical data to aid in TB prognosis and guide optimal treatment decisions for each patient.[8,35–38]

We performed a novel multimodal data integration from over 5,000 patients across 10 countries from the NIH NIAID TB Portals. In contrast to black box machine learning models, we implemented a "transparent" machine learning model to identify patient, drug, and pathogen features predictive of drug resistance and treatment prognosis in individual patients. Our analysis of over 200 different features across different host and pathogen modalities revealed several significant predictors associated with successful clinical outcomes. While multimodal integration has recently been applied to other diseases and has shown promising results,[39] this is the first such analysis for TB and our study integrates a wide range of features and data types.

Multimodal analysis also highlights the relative importance of features, which can help prioritize public health efforts to mitigate TB. For example, we observe that BMI is more important than drug regimens in our Shapley importance analysis. Features related to nutrition, particularly lower BMI, and the presence of HIV and anemia are significantly associated with failure. TB is typically more prevalent in the undernourished population.[40,41] Our findings are thus consistent with these reports.

Multimodal analysis also revealed novel associations that have not been previously reported, such as between imaging modalities and type of resistance and drug combinations. Here we discovered that the type of resistance is correlated with the imaging attributes of increased bronchial obstruction and affected lung segments. Further, drug combinations with antagonistic drug interactions predicted by INDIGO correlated with worse clinical presentation of TB, as indicated by overall percentage of abnormal lung volume.

Our study provides three key insights related to predicting TB prognosis and determining optimal treatment regimens. Firstly, multimodal integration outperforms models using a single data modality (such as imaging or pathogen features alone). Secondly, analysis of different patient populations grouped by type of resistance and treatment regimens revealed drug combinations that are associated with clinical success in drug-resistant TB. Thirdly, drug interactions are predictive of treatment success with INDIGO-predicted synergistic drug combinations outperforming antagonistic combinations. These insights are discussed in the following.

Our integrated multimodal machine learning model could predict treatment outcomes with an accuracy of 83.2% and an MCC of 0.451. These results are especially encouraging, as data were collected across several hospitals within each country, each of which had their own data collection and reporting criteria. As a result, and as is common with real-world patient data, there were several redundancies, missing information, and poor labeling, which made the data noisy. Despite these limitations, our final integrated model proved to be quite robust when predicting on unseen patient data. We also demonstrated that the model performed well at various levels of missing data imputation, as we achieved a similar accuracy in our secondary model that prioritized data completeness over sample size. The final model developed in our study can help predict TB prognosis and determine optimal treatment regimens based on real world multimodal data.

Notably, we identified drug combinations that are associated with clinical success in drug-resistant TB. Interestingly, the regimen "*bedaquiline, clofazimine, cycloserine, levofloxacin, linezolid*" as 1st LoT showed significant success in MDR-TB patients. This is concordant with recent reports on new regimens approved by the Food and Drug Administration (FDA) involving bedaquiline and levofloxacin.[42] In addition, for those MDR patients that progressed to an additional round of treatment, the same regimen followed by "*clofazimine, cycloserine, levofloxacin, linezolid*" showed a higher success rate. Recent studies have shown improved outcomes after additions of *bedaquiline* and *clofazimine* to combination therapies for treating MDR-TB,[43] and these results further underscore the clinical importance of these treatments. Surprisingly, we observed the regimen "*bedaquiline, clofazimine, linezolid, and moxifloxacin*" to be associated with failure for MDR non-XDR TB. Prior studies have shown that moxifloxacin-containing regimens for TB were mostly antagonistic,[44] and that *levofloxacin*-based regimens have better outcomes compared to *moxifloxacin*-based combination treatments for MDR-TB.[45] Further, cycloserine is known to be effective in MDR cases and is recommended by the WHO for treating MDR-TB, and its presence in the first regimen may increase treatment success.[46,47]

For XDR TB we observe that patients initially received "*bedaquiline, clofazimine, cycloserine, levofloxacin, linezolid*" similar to the MDR non-XDR TB patients, and, as they progressed, they were prescribed "*bedaquiline, clofazimine, cycloserine, delamanid, linezolid.*" Here, the only difference with MDR treatment is the inclusion of *delamanid* in place of *levofloxacin*, which showed high success for XDR TB. This is consistent with delamanid's effectiveness against XDR-TB, which are resistant to fluoroquinolones like levofloxacin.[48,49]

Table 4. Top features determined across modalities

Socio-demographic and clinical features	Pathogen genomic features and drug susceptibilities	Imaging modalities (CXR and CT)
Country	Ofloxacin resistance	Overall percentage of abnormal volume
Employment	Moxifloxacin resistance	ML calcified nodules
Age of onset	Levofloxacin resistance	ML_smallnodules
Case definition	Fluoroquinolones resistance	MR_non-calcified nodules
Education	Isoniazid resistance	UR_largecavities
BMI	Drug regimen	LR_non-calcified nodules
Social risk factors	Indigo_score	UR_lowgroundglassdensity
Number of daily contacts	Rifampicin resistance	
Number of children	Type_of_resistance	
Comorbidity	Gene mutations	
Gender	No. of genes with mutations	
	Mtb Colonies	

Swapping levofloxacin with delamanid is also predicted to make the combination slightly more synergistic (INDIGO interaction score = 0.87 with bedaquiline, clofazimine, cycloserine, delamanid, linezolid vs. 0.94 for bedaquiline, clofazimine, cycloserine, levofloxacin, linezolid).

Interestingly, we discovered that synergistic drug combinations were associated with treatment success. The drug interaction score was a top predictor according to the Shapley importance analysis. Currently, the role of drug interactions in treatment outcome is unclear. Prior studies have found that drug interaction scores correlate with TB clinical trial outcomes and mouse models. For example, Ma et al.[50] performed a meta-analysis of clinical trials of drug combinations against drug-sensitive TB and found a significant correlation between drug interaction scores and percentage of patients with negative sputum clearance. In this study, we observe a similar trend but with a larger number of drug combinations and with DR-TB patients. The standard 4-drug regimen of HRZE was synergistic (drug interaction score = 0.8) and associated with treatment success of Sensitive and Mono-DR TB cases. Treatment regimens involving the combination of *bedaquiline, levofloxacin, clofazimine, and Linezolid* (drug interaction score = 0.87) see more success in treating MDR and XDR TB. Drug interaction scores for all regimens predicted by INDIGO are included in Table S2.

In sum, our study analyzes multi-domain information from patients across geographical regions and focuses on the urgent need to improve TB clinical management particularly in the face of increasing drug resistance. The findings of our study are especially important given various humanitarian crises worldwide to meet the WHO's goals to "End TB" by 2035. This study can serve as a framework for managing other drug-resistant infections as well and presents a way for integrating large clinical and bio-molecular datasets[51] to further translational medicine.

Limitations of the study

There are some limitations to our study. As with any real-world patient data, the TB Portals dataset has several missing and noisy information, since it collects information from multiple hospitals across 10 countries, each with their own collection and reporting protocol. Further, there is a significant imbalance in the outcomes, with success more prevalent than failure. We address these limitations by considering random forests with an inverse weighting approach, with failure assigned higher weight than success. Random forests perform better than other modeling approaches as they can work with mixed data types and with missing values.[11] Despite the missing information, our models were robust and performed significantly well statistically. While the accuracy is not high enough for clinical implementation, this study represents a significant improvement over prior studies that utilize limited data types. Future models may incorporate data from more countries and also additional data within each country to build separate models and improve accuracy further.

Other limitations include that the drug interaction scores determined by INDIGO-MTB are computed based on *in vitro* interaction data, with the assumption that drug regimens are provided simultaneously. It would be more clinically relevant to calculate sequential drug interaction scores for longitudinal treatments[52] and account for pharmacokinetics and dynamics.[44,52] Further, genomic information is only available for a few patients, based on sequencing availability at the time of collection. Whole-genome sequencing is expensive in high-burden TB countries and may not always be feasible. For this study, we analyzed the prevalence of mutations at the gene level for the 8 genes associated with drug resistance, and in future studies we plan to analyze the individual DNA sequences of infecting *Mtb* strains where available to look at individual pathogenic variants and co-mutation prevalence and their associations with treatment outcomes. Of particular value is the CRyPTIC consortium to further analyze *Mtb* genetic variants and associations with resistance as a next step.[53] Finally, we used a transparent machine learning approach here to enable both interpretation and prediction of TB treatment outcomes. The use of black box methods like deep neural networks may lead to models with higher predictive accuracy. Nevertheless, our multimodal machine learning model's performance is better than prior models' in the literature that were built with fewer modalities.[13,54,55]

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 MODELS AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Data compilation
  - Data processing
  - Culture and microscopy
  - Drug sensitivity test (DST)
  - Drug-interaction scores
  - *Mtb* genotype, lineage, and families
  - Chest X rays and CT data
  - Outcomes
  - Data imputation
  - Data modalities
  - Machine learning
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Data visualization

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.109025.

ACKNOWLEDGMENTS

We are grateful to the University of Michigan Precision Health, Office of Vice-Provost for Research, the Michigan Institute of Clinical and Health Research (MICHR), Michigan Medicine Pandemic Research Recovery program, and MCUBED and NIH grants R56AI150826 and R01AI150826 for funding this study.

AUTHOR CONTRIBUTIONS

A.S. and S.C. conceptualized the study, gained access to the NIAID TB Portals data, analyzed the data, and wrote the manuscript. A.S., K.S., C.C., and H.S.A. conducted statistical and machine learning analyses for the data and prepared the figures. Z.Y., P.A., and S.C. supervised the study, designed the research, and edited the manuscript. All authors read, contributed to, and approved of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: September 29, 2023
Revised: December 8, 2023
Accepted: January 22, 2024
Published: January 29, 2024

REFERENCES

1. Chakaya, J., Khan, M., Ntoumi, F., Aklillu, E., Fatima, R., Mwaba, P., Kapata, N., Mfinanga, S., Hasnain, S.E., Katoto, P.D.M.C., et al. (2021). Global Tuberculosis Report 2020 - Reflections on the Global TB burden, treatment and prevention efforts. Int. J. Infect. Dis. *113*, S7–S12. https://doi.org/10.1016/j.ijid.2021.02.107.
2. Holt, E. (2022). Tuberculosis services disrupted by war in Ukraine. Lancet Infect. Dis. *22*, e129.
3. Dartois, V.A., and Rubin, E.J. (2022). Anti-tuberculosis treatment strategies and drug development: challenges and priorities. Nat. Rev. Microbiol. *20*, 685–701.
4. World Health Organization (2022). Rapid Communication: Key Changes to the Treatment of Drug-Resistant Tuberculosis. https://apps.who.int/iris/bitstream/handle/10665/353743/WHO-UCN-TB-2022.2-eng.pdf?sequence=1.
5. Fox, W.S., Strydom, N., Imperial, M.Z., Jarlsberg, L., and Savic, R.M. (2023). Examining nonadherence in the treatment of tuberculosis: The patterns that lead to failure. Br. J. Clin. Pharmacol. *89*, 1965–1977.
6. Mdluli, K., Kaneko, T., and Upton, A. (2015). The Tuberculosis Drug Discovery and Development Pipeline and Emerging Drug Targets. Cold Spring Harb. Perspect. Med. *5*, a021154. https://doi.org/10.1101/cshperspect.a021154.
7. Zumla, A., Chakaya, J., Centis, R., D'Ambrosio, L., Mwaba, P., Bates, M., Kapata, N., Nyirenda, T., Chanda, D.,

Mfinanga, S., et al. (2015). Tuberculosis treatment and management—an update on treatment regimens, trials, new drugs, and adjunct therapies. Lancet Respir. Med. *3*, 220–234.
8. Lino Ferreira da Silva Barros, M.H., Oliveira Alves, G., Morais Florêncio Souza, L., da Silva Rocha, E., Lorenzato de Oliveira, J.F., Lynn, T., Sampaio, V., and Endo, P.T. (2021). Benchmarking Machine Learning Models to Assist in the Prognosis of Tuberculosis. Informatics *8*, 27.
9. Ordonez, A.A., Wang, H., Magombedze, G., Ruiz-Bedoya, C.A., Srivastava, S., Chen, A., Tucker, E.W., Urbanowski, M.E., Pieterse, L., Fabian Cardozo, E., et al. (2020). Dynamic imaging in patients with tuberculosis reveals heterogeneous drug exposures in pulmonary lesions. Nat. Med. *26*, 529–534.
10. Huang, Q., Yin, Y., Kuai, S., Yan, Y., Liu, J., Zhang, Y., Shan, Z., Gu, L., Pei, H., and Wang, J. (2016). The value of initial cavitation to predict re-treatment with pulmonary tuberculosis. Eur. J. Med. Res. *21*, 20.
11. Peetluk, L.S., Ridolfi, F.M., Rebeiro, P.F., Liu, D., Rolla, V.C., and Sterling, T.R. (2021). Systematic review of prediction models for pulmonary tuberculosis treatment outcomes in adults. BMJ Open *11*, e044687.
12. Bui, V.C.B., Yaniv, Z., Harris, M., Yang, F., Kantipudi, K., Hurt, D., Rosenthal, A., and Jaeger, S. (2023). Combining Radiological and Genomic TB Portals Data for Drug Resistance Analysis. IEEE Access *11*, 84228–84240.
13. Koo, H.-K., Min, J., Kim, H.W., Lee, J., Kim, J.S., Park, J.S., and Lee, S.-S. (2020). Prediction of treatment failure and compliance in patients with tuberculosis. BMC Infect. Dis. *20*, 622.
14. Rosenfeld, G., Gabrielian, A., Meyer, A., and Rosenthal, A. (2022). Radiologist observations of chest X-rays (CXR) predict sputum smear microscopy status in TB Portals, a real-world database of tuberculosis (TB) cases. Preprint at bioRxiv. https://doi.org/10.1101/2022.04.21.22273975.
15. Rosenthal, A., Gabrielian, A., Engle, E., Hurt, D.E., Alexandru, S., Crudu, V., Sergueev, E., Kirichenko, V., Lapitskii, V., Snezhko, E., et al. (2017). The TB Portals: an Open-Access, Web-Based Platform for Global Drug-Resistant-Tuberculosis Data Sharing and Analysis. J. Clin. Microbiol. *55*, 3267–3282.
16. Johnson, R., Streicher, E.M., Louw, G.E., Warren, R.M., van Helden, P.D., and Victor, T.C. (2006). Drug resistance in Mycobacterium tuberculosis. Curr. Issues Mol. Biol. *8*, 97–111.
17. Mitchison, D.A. (2005). Drug resistance in tuberculosis. Eur. Respir. J. *25*, 376–379.
18. Genuer, R., Poggi, J.-M., Tuleau-Malot, C., and Villa-Vialaneix, N. (2017). Random Forests for Big Data. Big Data Res. *9*, 28–46.
19. Mughal, H., Bell, E.C., Mughal, K., Derbyshire, E.R., and Freundlich, J.S. (2022). Random Forest Model Predictions Afford Dual-Stage Antimalarial Agents. ACS Infect. Dis. *8*, 1553–1562.
20. Merrick, L., and Taly, A. (2020). The Explanation Game: Explaining Machine Learning Models Using Shapley Values. In Machine Learning and Knowledge Extraction (Springer International Publishing), pp. 17–38.
21. Winter. (2002). The Shapley Value. Handbook of Game Theory with Economic Applications.
22. Johnson, J.M., Mohapatra, A.K., Velladath, S.U., and Shettigar, K.S. (2022). Predictors of treatment outcomes in drug resistant tuberculosis-observational retrospective study. Int. J. Mycobacteriol. *11*, 38–46.
23. Baker, M.A., Harries, A.D., Jeon, C.Y., Hart, J.E., Kapur, A., Lönnroth, K., Ottmani, S.-E., Goonesekera, S.D., and Murray, M.B. (2011). The impact of diabetes on tuberculosis treatment outcomes: a systematic review. BMC Med. *9*, 81.
24. Adane, H.T., Howe, R.C., Wassie, L., and Magee, M.J. (2023). Diabetes mellitus is associated with an increased risk of unsuccessful treatment outcomes among drug-susceptible tuberculosis patients in Ethiopia: A prospective health facility-based study. J. Clin. Tuberc. Other Mycobact. Dis. *31*, 100368.
25. Gagneux, S. (2012). Host–pathogen coevolution in human tuberculosis. Philos. Trans. R. Soc. Lond. B Biol. Sci. *367*, 850–859.
26. Ahmed, H.G.E., Nassar, A.S., and Ginawi, I. (2011). Screening for tuberculosis and its histological pattern in patients with enlarged lymph node. Pathol. Res. Int. *2011*, 417635.
27. Sarathy, J., Blanc, L., Alvarez-Cabrera, N., O'Brien, P., Dias-Freedman, I., Mina, M., Zimmerman, M., Kaya, F., Ho Liang, H.-P., Prideaux, B., et al. (2019). Fluoroquinolone Efficacy against Tuberculosis Is Driven by Penetration into Lesions and Activity against Resident Bacterial Populations. Antimicrob. Agents Chemother. *63*, e02516-18. https://doi.org/10.1128/AAC.02516-18.
28. Zhao, J.-N., Zhang, X.-X., He, X.-C., Yang, G.-R., Zhang, X.-Q., Xin, W.-G., and Li, H.-C. (2015). Multidrug-Resistant Tuberculosis in Patients with Chronic Obstructive Pulmonary Disease in China. PLoS One *10*, e0135205.
29. Santos, M.d.L.S.G., Vendramini, S.H.F., Gazetta, C.E., Oliveira, S.A.C., and Villa, T.C.S. (2007). Poverty: socioeconomic characterization at tuberculosis. Rev. Latino-Am. Enferm. *15*, 762–767.
30. Gupta, D., Das, K., Balamughesh, T., Aggarwal, N., and Jindal, S.K. (2004). Role of socio-economic factors in tuberculosis prevalence. Indian J. Tubercul. *51*, 27–32.
31. 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.
32. Zhou, S., Van Staden, Q., and Toska, E. (2020). Resource reprioritisation amid competing health risks for TB and COVID-19. Int. J. Tubercul. Lung Dis. *24*, 1215–1216.
33. Acosta, C.D., Dadu, A., Ramsay, A., and Dara, M. (2014). Drug-resistant tuberculosis in Eastern Europe: challenges and ways forward. Public Health Action *4*, S3–S12.
34. Marais, B.J., Lönnroth, K., Lawn, S.D., Migliori, G.B., Mwaba, P., Glaziou, P., Bates, M., Colagiuri, R., Zijenah, L., Swaminathan, S., et al. (2013). Tuberculosis comorbidity with communicable and non-communicable diseases: integrating health services and control efforts. Lancet Infect. Dis. *13*, 436–448. https://doi.org/10.1016/s1473-3099(13)70015-x.
35. Gröschel, M.I., Walker, T.M., van der Werf, T.S., Lange, C., Niemann, S., and Merker, M. (2018). Pathogen-based precision medicine for drug-resistant tuberculosis. PLoS Pathog. *14*, e1007297.
36. Falzon, D., Jaramillo, E., Gilpin, C., and Weyer, K. (2018). The Role of Novel Approaches and New Findings in the Pharmacology of Tuberculosis Medicines in Improving Treatment Outcomes. Clin. Infect. Dis. *67*, S365–S367.
37. Lange, C., Aarnoutse, R., Chesov, D., van Crevel, R., Gillespie, S.H., Grobbel, H.-P., Kalsdorf, B., Kontsevaya, I., van Laarhoven, A., Nishiguchi, T., et al. (2020). Perspective for Precision Medicine for Tuberculosis. Front. Immunol. *11*, 566608.
38. Olaru, I.D., Lange, C., and Heyckendorf, J. (2016). Personalized medicine for patients with MDR-TB. J. Antimicrob. Chemother. *71*, 852–855.
39. Kline, A., Wang, H., Li, Y., Dennis, S., Hutch, M., Xu, Z., Wang, F., Cheng, F., and Luo, Y. (2022). Multimodal machine learning in precision health: A scoping review. NPJ Digit. Med. *5*, 171.
40. Papathakis, P., and Piwoz, E. (2008). Nutrition and Tuberculosis: A Review of the Literature and Considerations for TB Control Programs (United States Agency for International Development, Africa's Health 2010 Project), p. 1.
41. VanValkenburg, A., Kaipilyawar, V., Sarkar, S., Lakshminarayanan, S., Cintron, C., Prakash Babu, S., Knudsen, S., Joseph, N.M., Horsburgh, C.R., Sinha, P., et al. (2022). Malnutrition leads to increased inflammation and expression of tuberculosis risk signatures in recently exposed household contacts of pulmonary tuberculosis. Front. Immunol. *13*, 1011166.
42. Burki, T. (2019). BPaL approved for multidrug-resistant tuberculosis. Lancet Infect. Dis. *19*, 1063–1064.
43. Yao, G., Zhu, M., Nie, Q., Chen, N., Tu, S., Zhou, Y., Xiao, F., Liu, Y., Li, X., and Chen, H. (2023). Improved outcomes following addition of bedaquiline and clofazimine to a treatment regimen for multidrug-resistant tuberculosis. J. Int. Med. Res. *51*. 3000605221148416.
44. Cicchese, J.M., Sambarey, A., Kirschner, D., Linderman, J.J., and Chandrasekaran, S. (2021). A multi-scale pipeline linking drug transcriptomics with pharmacokinetics predicts *in vivo* interactions of tuberculosis drugs. Sci. Rep. *11*, 5643.
45. Sidamo, T., Shibeshi, W., Yimer, G., Aklillu, E., and Engidawork, E. (2021). Explorative Analysis of Treatment Outcomes of Levofloxacin- and Moxifloxacin-Based Regimens and Outcome Predictors in Ethiopian MDR-TB Patients: A Prospective Observational Cohort Study. Infect. Drug Resist. *14*, 5473–5489.
46. Li, Y., Wang, F., Wu, L., Zhu, M., He, G., Chen, X., Sun, F., Liu, Q., Wang, X., and Zhang, W. (2019). Cycloserine for treatment of multidrug-resistant tuberculosis: a retrospective cohort study in China. Infect. Drug Resist. *12*, 721–731.
47. He, Y., and Li, X. (2022). The treatment effect of Levofloxacin, Moxifloxacin, and Gatifloxacin contained in the conventional therapy regimen for pulmonary tuberculosis: Systematic review and network meta-analysis. Medicine *101*, e30412.
48. Skripconoka, V., Danilovits, M., Pehme, L., Tomson, T., Skenders, G., Kummik, T., Cirule, A., Leimane, V., Kurve, A., Levina, K., et al. (2013). Delamanid improves outcomes and reduces mortality in multidrug-resistant tuberculosis. Eur. Respir. J. *41*, 1393–1400.
49. Nasiri, M.J., Zangiabadian, M., Arabpour, E., Amini, S., Khalili, F., Centis, R., D'Ambrosio, L., Denholm, J.T., Schaaf, H.S., van den Boom, M., et al. (2022). Delamanid-containing regimens and multidrug-resistant

tuberculosis: A systematic review and meta-analysis. Int. J. Infect. Dis. *124* (*Suppl 1*), S90–S103.

50. Ma, S., Jaipalli, S., Larkins-Ford, J., Lohmiller, J., Aldridge, B.B., Sherman, D.R., and Chandrasekaran, S. (2019). Transcriptomic Signatures Predict Regulators of Drug Synergy and Clinical Regimen Efficacy against Tuberculosis. mBio *10*, e02627-19. https://doi.org/10.1128/mBio.02627-19.
51. Lussier, Y.A., Butte, A.J., Li, H., Chen, R., and Moore, J.H. (2018). Translational informatics of population health: How large biomolecular and clinical datasets unite. In Biocomputing 2019 (WORLD SCIENTIFIC), pp. 455–459.
52. Chung, C.H., and Chandrasekaran, S. (2022). A flux-based machine learning model to simulate the impact of pathogen metabolic heterogeneity on drug interactions. PNAS Nexus *1*, pgac132. https://doi.org/10.1093/pnasnexus/pgac132.
53. The CRyPTIC Consortium (2022). A data compendium associating the genomes of 12,289 Mycobacterium tuberculosis isolates with quantitative resistance phenotypes to 13 antibiotics. PLoS Biol. *20*, e3001721.
54. Kalhori, S.R.N., Nasehi, M., and Zeng, X.J. (2010). A logistic regression model to predict high risk patients to fail in tuberculosis treatment course completion. IAENG Int. J. Appl. Math. *40*, 102–107.
55. Sauer, C.M., Sasson, D., Paik, K.E., McCague, N., Celi, L.A., Sánchez Fernández, I., and Illigens, B.M.W. (2018). Feature selection and prediction of treatment failure in tuberculosis. PLoS One *13*, e0207491.
56. Rosenthal, A., Gabrielian, A., Engle, E., Hurt, D.E., Alexandru, S., Crudu, V., Sergueev, E., Kirichenko, V., Lapitskii, V., Snezhko, E., et al. (2017). The TB Portals: an Open-Access, Web-Based Platform for Global Drug-Resistant-Tuberculosis Data Sharing and Analysis. J. Clin. Microbiol. *55*, 3267–3282. https://doi.org/10.1128/jcm.01013-17.
57. Galkina, K., Nosova, E., and Krasnova, M. (2012). Modern molecular direct tests for rapid identification and drug susceptibility testing of Mycobacterium tuberculosis. Eur. Respir. J. *40*.
58. Beaulieu-Jones, B.K., and Moore, J.H. (2017). MISSING DATA IMPUTATION IN THE ELECTRONIC HEALTH RECORD USING DEEPLY LEARNED AUTOENCODERS. Pac. Symp. Biocomput. *22*, 207–218.
59. Vidakovic, B. (2011). Statistics for Bioengineering Sciences: With MATLAB and WinBUGS Support (Springer Science & Business Media).
60. Akoglu, H. (2018). User's guide to correlation coefficients. Turk. J. Emerg. Med. *18*, 91–93.
61. Ranstam, J. (2016). Multiple P-values and Bonferroni correction. Osteoarthritis Cartilage *24*, 763–764.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Multimodal TB patient data	Rosenthal, A. et al. The TB Portals: an Open-Access, Web-Based Platform for Global Drug-Resistant-Tuberculosis Data Sharing and Analysis. J. Clin. Microbiol. (2017) https://doi.org/10.1128/jcm.01013-17.	https://tbportals.niaid.nih.gov/
Software and algorithms		
Python v. 3.7.14 Shapley package	Python Software Foundation	https://www.python.org/downloads/ https://pypi.org/project/shapley/#files
Matlab R2021b	Mathworks	https://www.mathworks.com/products/matlab.html
RStudio 4.3.0	The R project for statistical computing	https://www.r-project.org/

RESOURCE AVAILABILITY

Lead contact

Requests for further information should be directed to and will be fulfilled by the lead contact, Sriram Chandrasekaran (csriram@umich.edu).

Materials availability

This study did not create new unique reagents.

Data and code availability

- This paper analyzes data from the NIH-NIAID TB portals database for patient cases available from 2008 until August 2021 [*TB_portals_Update_Aug2021*] after signing a data-sharing agreement.
- The code for machine learning and statistical analysis conducted in this study is available on Synapse: TB - syn32884986 - Wiki (synapse.org).
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODELS AND STUDY PARTICIPANT DETAILS

This study did not use experimental models.

METHOD DETAILS

Data compilation

Patient data was obtained from the TB portals database for patient cases available from 2008 until August 2021 [TB_portals_Update_Aug2021] after signing a data-sharing agreement. The patient data is encoded with unique IDs determined by the database, with no disclosure of the individual patient's name. The database includes patients from 10 countries spanning Eastern Europe, Central Asia, and Africa with a heavy burden of drug-resistant TB. It continues to be actively populated with new patient data, including those who are still undergoing treatment. For every patient, we obtained associated information pertaining to patient cases (clinical features), Radiological information (chest X-ray images, CT scans and their annotations), drug regimens, biochemistry, drug sensitivity profile for the pathogen associated with each infection and the corresponding specimen. The data collected comes from a range of sources – clinical trials, research studies, as well as routine collection of atypical patient cases receiving medical care. There is no single identifiable data collection protocol that is uniformly enforced. They utilize a uniform data dictionary with generally accepted medical terminology and data field values. The TB Portals program uses multiple mechanisms to ensure quality of the data submitted and data checks and uniformity are embedded into the system using dropdown lists and ranges of permissible values.[56]

Data processing

The dataset had missing values for several features as well as entries with values 'Not Reported', 'Unknown', 'Others', 'Not Specified'. These were collectively labeled as "NaN" for further processing. Duplicate patient records as determined by patient id numbers were removed.

Entries with conflicting values (eg. entries reporting 'Yes AND No' or 'Resistant And Sensitive') were also reported as NaN for that feature if the entries for other features for that patient were not conflicting. For every patient, there were 203 variables of mixed data types, with 11 numerical and 192 categorical features. The categorical features were encoded into numerical values, with numerical values assigned based upon the number of levels an attribute, starting from 0. For example, the attribute 'education' has six levels, ranging from "no education" - which is assigned 0, to "college and higher" which is encoded as 6. This process is repeated for all categorical attributes accordingly. For the machine-learning models, features with over 90% missingness and patients with over 40% missingness were dropped from the analysis prior to imputation.

Culture and microscopy

For culture types, the first culture report of the number of colonies identified in the specimen were considered. Instances which don't report the number of colonies were mapped to the individual sample results associated with specimen id. Growth of mycobacteria is also reported as either positive, negative or both, as well as reports of non-specific mycobacteria and Mycobacteria other than tuberculosis (MOTT). Microscopy results describe the number of acid resistant bacteria in different fields of view. For entries with multiple codes per cell, only the first entry was considered for analysis.

Drug sensitivity test (DST)

The dataset describes DST results for 24 different drugs. Entries are marked R (resistant), S(sensitive) or I (intermediate) to describe observed DST profiles for each drug. For each drug, the DST results are indicated by up to 5 types of test conducted, namely bactec, hain, Le, GeneExpert and lpa.[57] In cases where the tests report different results for the same strain, we consider a cumulative DST case based on the profile reported by the majority of tests.

Drug-interaction scores

For all combinations of drug regimens given to each patient, we computed a 'drug-interaction score' using the INDIGO-MTB tool,[50] which uses Random forests to assign drug interaction scores that capture the nature of interaction between drugs using individual drug response transcriptomics data. The interactions are considered synergistic (score <0.9), additive (scores 0.9-1.2), and antagonistic (scores >1.2) respectively as used in prior studies.

Mtb genotype, lineage, and families

Several *Mtb* families belonging to multiple lineages are represented in this dataset across different populations. The dataset reports *Mtb* strains from 27 different families in infected patients, with some patients infected with 2 or more families (which is unlikely). The TB portals database refers to these families as lineages, which is inaccurate. Lineage refers to the *Mtb* classification based on Large sequence polymorphism (LSP) or Single nucleotide polymorphisms (SNP). Family/sub-lineages refer to the classifications based on spoligotyping. We determined the prevalence of the reported gene mutations in patients with drug-resistant TB. We analyzed the genes with reported mutations individually and those co-occurring with other gene mutations to be strongly associated with failure by a Fisher's exact test, as seen in Table 2.

Chest X rays and CT data

Chest X rays were available for a subset of the patients, taken across multiple time points, with Day 0 considered as the time TB infection was confirmed and treatment was initiated for the patient. Manual annotations of these X rays performed by radiologists and/or general physicians were made available. We chose the annotation information for X Rays recorded at day 0, and in instances where day 0 was not available, we chose the day closest to Day 0 to capture the early days of infection before treatment takes effect. Similarly, we chose the CT annotations closest to day 0 when available.

Outcomes

The outcomes of infection for each patient was considered as a response variable in our analysis. Outcomes of infection and treatment are described for each patient, with 6 outcomes provided, namely *Cured, Completed, Failure, Died, Palliative Care, Lost to Follow up, Still on treatment and Unknown*. As we are interested in analyzing successful and unsuccessful treatment outcomes, we pooled the outcomes 'Cured' and 'Completed' as ***"Success"***, while outcomes 'Failure', 'Died' and "Palliative Care" were grouped together under the outcome ***"Failure"***, as they all indicate unfavorable treatment outcomes per patient. We did not consider patients with outcomes *Lost to Follow up, Still on treatment and Unknown* for our analysis.

Data imputation

Initially, complete case analysis was performed with only patients that had less than 50% NaN values across all features. We repeated the analysis by imputing missing values using the K-nearest neighbor (KNN) imputation method, with k set to 3. This method was chosen as nearest neighbor imputation methods have been shown to be effective for machine learning with missing data for EHR analysis.[58]

Data modalities

The data was split into 3 categories for analysis describing different aspects of host and pathogen data. These categories were a) patient social, clinical and demographic features b) patient radiological features as determined by chest X rays (CXR) and CT scans and c) pathogen genomic features which include gene mutations and drug sensitivity analyses.

Machine learning

We randomly split the input data into training (80%) and test (20%) samples for each category. The training data was then used to build a Random forest model to determine outcomes of Success and Failure. The function uses RandomizedGridSearch within cross-validation to tune the hyperparameters and estimate model performance. The final model uses the hyperparameters from the best-scoring CV fold and is trained on the entire dataset. To account for the class imbalance in the two outcomes of success and failure, we considered an inverse weighting approach to balance the two classes (e.g. if success represents 80% of the outcomes, weights assigned are 1/0.08 and 1/0.02 for success and failure respectively). We performed a 5-fold cross validation and assessed model performance in predicting disease outcome on the training data. The model performance was further evaluated on the test dataset containing the remaining 20% data. Different metrics were considered to evaluate model performance including Accuracy, Precision, Recall, F1 score, Matthews Correlation coefficient and the Pearson's R, to predict the outcome of infection. After identifying the top features for each category, we further evaluated the associations of different levels within the feature with the outcome by calculating p-values based on the hygecdf function (hypergeometric cumulative distribution)[59] in MATLAB. Additional validation was performed on newer data populated in the TB portals database (January 2022).

QUANTIFICATION AND STATISTICAL ANALYSIS

Pairwise statistical tests were performed for all features in the dataset. For comparing two continuous variables, the Kendall rank correlation was used, which is a non-parametric alternative to Pearson's correlation.[60] For two categorical variables, the Chi-squared test was performed to determine if there is a significant difference between the observed and expected frequencies of the associated contingency table. When comparing continuous and categorical variables, the Kruskal-Wallis test was used, which tests whether there is a difference between the groups of the continuous feature. The p-values from all of these tests were combined into a matrix and plotted as a heatmap. For all comparisons, a Bonferroni correction test[61] was applied.

Data visualization

We use the Python Shapley package for data visualization. All machine learning and statistical analysis was conducted in Python version 3.7.13 and repeated in Matlab version R2021b. Fisher's exact test to determine associations between treatment regimens and gene mutations and outcomes of success/failure were performed in Rstudio using R version 4.3.0 (2023-04-21).