

RESEARCH ARTICLE

A Bayesian approach to estimate the probability of resistance to bedaquiline in the presence of a genomic variant

Degefaye Zelalem Anlay^{1,2}, Emmanuel Rivière^{2*}, Pham Hien Trang Tu², Steven Abrams^{2,3}, Annelies Van Rie²

1 School of Nursing, College of Medicine and Health Science, University of Gondar, Gondar, Ethiopia, **2** Global Health Institute, Family Medicine and Population Health, Faculty of Medicine and Health Sciences, University of Antwerp, Antwerp, Belgium, **3** Data Science Institute, Interuniversity Institute for Biostatistics and statistical Bioinformatics, U Hasselt, Diepenbeek, Belgium

* emmanuel.riviere@uantwerpen.be



OPEN ACCESS

Citation: Anlay DZ, Rivière E, Tu PHT, Abrams S, Van Rie A (2023) A Bayesian approach to estimate the probability of resistance to bedaquiline in the presence of a genomic variant. PLoS ONE 18(6): e0287019. <https://doi.org/10.1371/journal.pone.0287019>

Editor: Leonid Chindelevitch, Imperial College London, UNITED KINGDOM

Received: July 14, 2022

Accepted: May 26, 2023

Published: June 14, 2023

Copyright: © 2023 Anlay et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its [Supporting information files](#).

Funding: This work was supported by the Research Foundation Flanders (FWO) [grant numbers G0F8316N and 1S39119N], VLIR-UOS support Degefaye and Trang to follow the master of epidemiology course at university of Antwerp. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Background

Bedaquiline is a core drug for treatment of rifampicin-resistant tuberculosis. Few genomic variants have been statistically associated with bedaquiline resistance. Alternative approaches for determining the genotypic-phenotypic association are needed to guide clinical care.

Methods

Using published phenotype data for variants in *Rv0678*, *atpE*, *pepQ* and *Rv1979c* genes in 756 *Mycobacterium tuberculosis* isolates and survey data of the opinion of 33 experts, we applied Bayesian methods to estimate the posterior probability of bedaquiline resistance and corresponding 95% credible intervals.

Results

Experts agreed on the role of *Rv0678*, and *atpE*, were uncertain about the role of *pepQ* and *Rv1979c* variants and overestimated the probability of bedaquiline resistance for most variant types, resulting in lower posterior probabilities compared to prior estimates. The posterior median probability of bedaquiline resistance was low for synonymous mutations in *atpE* (0.1%) and *Rv0678* (3.3%), high for missense mutations in *atpE* (60.8%) and nonsense mutations in *Rv0678* (55.1%), relatively low for missense (31.5%) mutations and frameshift (30.0%) in *Rv0678* and low for missense mutations in *pepQ* (2.6%) and *Rv1979c* (2.9%), but 95% credible intervals were wide.

Conclusions

Bayesian probability estimates of bedaquiline resistance given the presence of a specific mutation could be useful for clinical decision-making as it presents interpretable probabilities compared to standard odds ratios. For a newly emerging variant, the probability of

Competing interests: The authors have declared that no competing interests exist.

resistance for the variant type and gene can still be used to guide clinical decision-making. Future studies should investigate the feasibility of using Bayesian probabilities for bedaquiline resistance in clinical practice.

Introduction

In 2020, tuberculosis (TB), caused by *Mycobacterium tuberculosis* (*Mtb*), was the 13th leading cause of death and the second leading infectious killer after COVID-19 [1]. The public health threat posed by TB is worsened by the occurrence of drug-resistant TB. In 2019, close to half a million people developed rifampicin-resistant TB (RR-TB) worldwide [2]. In the past decade, important progress has been made, including rapid diagnosis of RR-TB and introduction of short all-oral treatment regimens [3].

Bedaquiline (BDQ) is an essential component of the novel short all-oral RR-TB treatment regimens [4]. Its use is rapidly increasing, with 90 countries using BDQ in 2018 [5]. Soon after its introduction, cases of acquired and primary BDQ resistance and cross-resistance between BDQ and clofazimine (CFZ), another drug frequently used to treat RR-TB, were reported [6–9]. To prevent acquisition and transmission of BDQ resistant TB, the use of BDQ should be accompanied with access to BDQ drug susceptibility tests (DST). Unfortunately, phenotypic DST (pDST) for BDQ, the gold standard, is slow, not yet standardized, and has a low positive predictive value in settings with low prevalence of BDQ resistance [10]. A genotypic DST (gDST) assay could provide accurate and timely information if knowledge on genetic variants and their association with resistance to BDQ is comprehensive [11–13].

Several candidate BDQ resistance genes have been identified [14]. Mutations in the *atpE* gene can cause a loss of binding affinity of BDQ with its drug target, resulting in high-level BDQ resistance *in vitro* [15], but they are rarely observed in clinical isolates [16]. Overexpression of the MmpS5/MmpL5 efflux system caused by mutations in the *Rv0678* gene is the main mechanism of clinical resistance to BDQ [6, 13, 17]. Lastly, mutations in the *pepQ* and *Rv1979c* genes have also been implicated in BDQ resistance, but evidence is lacking [8, 18].

A recent systematic review analyzed the association between phenotypic BDQ resistance and genomic variants in *atpE*, *Rv0678*, *Rv1979c*, and *pepQ* [11]. When applying standard statistical methods, only two out of 313 variants reported in the literature could be statistically associated with resistance [19]. Similarly, no mutations satisfied the criteria for BDQ resistance in the 2021 WHO catalogue of mutations in *Mtb* complex [14]. Consequently, the body of evidence on BDQ genotype-phenotype association cannot inform the use of gDST for patient care.

Bayesian methods can overcome data sparsity by combining data with expert opinion [20]. A Bayesian approach could be of great value in a clinical setting as it is conceptually similar to how clinicians approach the diagnostic process. Furthermore, the credible interval is well suited for clinical practice as it has a more intuitive interpretation [21, 22]. We applied a Bayesian method to estimate the probability of BDQ resistance in the presence of genomic variants.

Materials and methods

Bayesian methods were used to estimate the probability of BDQ resistance in the presence of a mutation. First, a survey was conducted to obtain expert opinion (prior information) on genotype-phenotype associations for BDQ. Next, an update of a recent systematic literature review was performed to obtain the most comprehensive phenotypic-genotypic data [11]. Finally, Bayesian analyses were used to obtain the posterior distribution of the probability of BDQ resistance, the posterior median and its 95% credible interval (CrI) [23].

Survey

Experts were selected based on their active involvement in the field of genotype-phenotype associations in *Mtb* as determined by scientific publications or participation in the development of the WHO catalogue of mutations in *Mtb* [14, 24]. Experts were asked to share contact details of relevant colleagues. The online "Bedaquiline resistance survey" tool (see Survey, S3 File, which includes the consent form and survey presented to experts) focused on the expert rules introduced by Miotto *et al.* [19] and Köser *et al.* [25]. For the *atpE*, *Rv0678*, *Rv1979c*, and *pepQ* genes, experts were asked if they believe that a mutation can cause BDQ resistance. If the expert answered yes, their opinion on the probability of resistance in the presence of different types of mutations (in-frame indel, frameshift indel, synonymous, nonsense, missense, and homoplasmic mutation) was obtained using a six-point Likert scale, except for loss of function mutations (nonsense and frameshift mutation) in the *atpE* gene as these have not been observed in *Mtb* isolates. The survey was piloted by five experts.

Experts' responses were eligible for inclusion in the analysis if the expert responded "yes" to "Do you believe that a mutation in the *atpE* gene can confer resistance to BDQ?", indicated he/she was "sure" or "relatively sure" about their responses, and there was no evidence of poor response quality.

Systematic literature review

To obtain the most comprehensive data on BDQ genotype-phenotype, the individual isolate systematic review (publications between 2008 and October 2020) was updated using the same search engines (Europe PubMed Central and Scopus), search terms and methodology for the period of November 2020 to December 30, 2021 [11]. to identify studies that reported genotypic and phenotypic bedaquiline resistance for individual clinical or non-clinical *Mtb* isolates. New records of individual isolates were added to the review database (see S2 File, which includes the screened literature, all extracted data, the meta-analysis, expert survey responses, and Bayesian analysis).

Statistical analysis

All analyses were done using R statistical software (version 4.0.5) (see S1 File). The correlation between survey responses for the same mutation types in different genes was assessed using the Spearman correlation coefficient [26].

Construction of prior distributions

Experts' responses on probability of BDQ resistance in the presence of a variant were translated into prior distributions. For the genes where all experts agreed on its potential role in BDQ resistance, a beta distribution (with shape parameters denoted as α and β) was used to parametrically model the prior distribution for the probability of BDQ resistance for each type of mutation (synonymous mutation, nonsense mutation, frameshift indel, missense mutation, inframe indel, homoplasmic mutation). In order to obtain maximum likelihood (ML) estimates for the α - and β -parameters of the prior distribution for each mutation type, the likelihood function was maximized using the *optim* function (in the *stats* R package) with default Nelder-Mead line search method [27]. To deal with the interval-censored nature of the Likert scale data (i.e., $x_i = [ll_i, ul_i]$, for $i = 1, \dots, n$), the (log)likelihood function, denoted hereunder as $LL(x | \alpha, \beta)$, for the estimation of the beta density was adjusted as follows [28]:

$$LL(x | \alpha, \beta) = \sum_{i=1}^n \log[F(ul_i | \alpha, \beta) - F(ll_i | \alpha, \beta)]$$

where $x = (x_1, \dots, x_n)$ represents the vector of reported probability intervals of BDQ resistance, given the presence of a specific type of mutation, for all experts and $F(s | \alpha, \beta)$ is the cumulative distribution function of a beta distributed random variable with parameters α and β evaluated in s , and ul_i and ll_i the upper and lower limits, respectively, of the probability interval associated with the Likert scale category chosen by expert i .

For genes where experts expressed uncertainty about the role in BDQ resistance, a mixture distribution was used as prior. When all three answers (“yes”, “no”, “I do not know”) were observed for the question “can mutations in gene X confer resistance to BDQ”, a three-component mixture prior distribution was defined using empirical probability weights w_j :

$$f(p_i) = \sum_{j=1}^3 w_j f_j(p_i),$$

where $f_j(p_i)$ represents the probability density for mixture distribution $j = 1, 2, 3$ and nonnegative weights w_j such that $0 < w_j < 1$ for all j and $\sum_{j=1}^3 w_j = 1$ [29]. More specifically, f_1 is the probability density function (pdf) of a beta distribution with shape parameters α_1 and β_1 , f_2 is the discrete probability distribution of a degenerate distribution at zero (i.e., $f_2(p_i) = Pr(P_i = p_i) = I(p_i = 0)$ for $0 < p_i < 1$, where $I(\cdot)$ is the indicator-function) and f_3 is the pdf of a uniform distribution on the unit interval, which is the same as a beta distribution with $\alpha_3 = \beta_3 = 1$ [30]. The mixture proportions of each component were set equal to the observed probability of responding “yes”, “no” or “I do not know”, respectively.

When only two answers were observed, a two-component mixture prior distribution was used:

$$f(p_i) = \sum_{j=1}^2 w_j f_j(p_i).$$

To capture the experts’ opinion on the assumption that the effect of mutations on the phenotype that occurs in laboratory experiments can be extrapolated to the phenotype of clinical isolates containing the same variant, a single unconditional prior distribution was constructed combining the information from experts answering “yes” and the uncertainty from experts answering “I don’t know” or “no.” [31]: More specifically, $f(p_i) = \sum_{j=1}^2 f(p_i | x_j) f(x_j)$ where $f(p_i | x_j)$ represents the conditional distribution of $(P_i | X = x_j)$, $j = 1, 2$, for experts answering “yes” ($j = 1$) and “I don’t know” or “no” ($j = 2$), respectively. Finally, f represents the discrete probability distribution for the aforementioned response categories, i.e., $f(x_j) = Pr(X = x_j)$.

Data likelihood

The genotype-phenotype individual isolate analysis of the updated systematic literature was used for the construction of the data likelihood. Observing phenotypic BDQ resistance ($Z_k = 1$) or susceptibility ($Z_k = 0$) in the presence of a specific mutation (for $k = 1, \dots, m$) can be viewed as a sequence of independent Bernoulli trials with the same probability of resistance, thereby implying that the number of resistant isolates, denoted by $Y = \sum_{k=1}^m Z_k \sim B(m, \pi)$, follows a binomial distribution with m the number of times a specific genetic variant in the target gene of interest was reported in the literature, and π the probability of phenotypic BDQ resistance of an isolate containing such variant.

Posterior distribution and Bayesian inference

To estimate the probability of BDQ resistance and its 95% CrI, inferences were made based on the posterior distribution. Dependent Markov Chain Monte Carlo (MCMC) sampling of the posterior is performed using a Metropolis-Hastings (MH) algorithm. Convergence of the

MCMC chains was checked by inspection of the trace plots and formal Gelman-Rubin convergence statistics, denoted by R . Three chains with starting values 0.25, 0.5 and 0.75 for the resistance probability were run and convergence was declared when R was less than 1 [32–34]. An autocorrelation plot was used to evaluate the dependence of the parameter to the Markov process after the burn-in period [35]. Thinning of the chain was performed every 5th, 10th, or 15th iteration, depending on the results of the autocorrelation plot, and the iteration number was increased after thinning. The sufficiency of the iteration number to estimate the 2.5% and 97.5% quantiles was checked by the Raftery diagnostic [36, 37].

Ethical approval

The study has been reviewed and approved by the Research Ethics Committee of the University Hospital of the University of Antwerp (REF number: 21/06/093). The survey was anonymous, all participated experts gave written informed consent.

Results

Expert survey results and estimation of the prior

Of the 120 invited researchers, 46 (38.3%) completed the survey. Thirteen (28.3%) respondents were excluded from the analysis because they did not fulfil the predefined inclusion criteria: 11 participants reported they were not sure about their response and the responses by 2 participants were of poor quality (same answers to all questions and completion of the full survey in less than 5 minutes). Of the 33 experts whose results could be included, most ($n = 25$, 75.8%) were laboratory researchers and almost all ($n = 27$, 81.8%) resided in a high-income country (Table 1).

Regarding general expert rules on genotype-phenotype associations, most experts believed that resistance-conferring mutations occurring in isolation also confer resistance when occurring in combination with other mutations ($n = 26$, 78.8%), that any increase above the critical concentration is clinically relevant ($n = 20$, 60.6%), and that the phenotypic effect of a mutation in a laboratory strain can be extrapolated to the effect in a clinical isolate ($n = 22$, 66.7%) (Table 2).

Regarding the role of specific genes in BDQ resistance, all experts believed that a variant in *atpE* or *Rv0678* can cause resistance to BDQ. For the *pepQ* and *Rv1979c* genes, 51.5% (17/33) and 78.1% (25/32) were uncertain whether a variant can cause resistance to BDQ, respectively.

Table 1. Characteristics of experts contributing data to the expert opinion analysis ($n = 33$).

Characteristic		Number (%) of experts
Expert type	Mainly clinical research	4 (12.1%)
	Mainly laboratory research	25 (75.8%)
	Mainly epidemiological and public health research	4 (12.1%)
Years of experience in tuberculosis research	1 to 5 years	5 (15.1%)
	6 to 10 years	9 (27.3%)
	11 to 15 years	8 (24.2%)
	>15 years	11 (33.3%)
Income status of respondent's country	High-income country	27 (81.8%)
	Middle-income country	3 (9.1%)
	Low-income country	3 (9.1%)

<https://doi.org/10.1371/journal.pone.0287019.t001>

Table 2. Expert (n = 33) opinion on general rules regarding genotype-phenotype associations for BDQ resistance in *Mtb*.

General expert rule [Miotto <i>et al.</i> [19] and Köser <i>et al.</i> [24]]	Response	Frequency (%)
Rule 1. A mutation that confers phenotypic resistance to BDQ when it occurs on its own also causes resistance when it occurs in combination with other mutations in the same gene or mutations elsewhere in the genome.	Always	3 (9.1%)
	Most of the time	23 (69.7%)
	Sometimes	6 (18.2%)
	Never	0
	I do not know	1 (3.0%)
Rule 2. The critical concentration (cc) for BDQ (as defined by WHO or EUCAST) corresponds to a clinical breakpoint, and any minimal inhibitory concentration increase above the cc is clinically significant (i.e., both low level and high-level resistance are clinically relevant for BDQ).	Yes	20 (60.6%)
	No	3 (9.1%)
	I do not know	10 (30.3%)
Rule 3. The effect of a specific mutation on the phenotype observed in laboratory experiments can be extrapolated to the effect that same mutation would have in clinical isolates. Results of laboratory experiments can, therefore, be used as a source of evidence.	Yes	22 (66.7%)
	No	9 (27.3%)
	I do not know	2 (6.0%)

BDQ = bedaquiline; Mtb = Mycobacterium tuberculosis

<https://doi.org/10.1371/journal.pone.0287019.t002>

The probability that a synonymous mutation can confer resistance was assumed low (never or rarely) for all genes by most experts (*atpE* (84.4%), *Rv0678* (68.8%), *pepQ* (90.0%), *Rv1979c* (100.0%)) (Table 3). Most experts believed that a nonsense mutation frequently to always confers resistance (*Rv0678* (84.8%), *pepQ* (60.0%), *Rv1979c* (50.0%)) and that frameshift indels frequently to always confer resistance when occurring in *Rv0678* (93.8%) or *pepQ* (50.0%) but not when occurring in *Rv1979c* (0.0%). Opinions on the probability of resistance in the presence of a missense mutation or inframe indel varied. The opinion of individual experts on the role of types of mutations in conferring BDQ resistance was correlated for the *atpE* and *Rv0678* genes (synonymous mutations ($r = 0.38$, $p = 0.03$), missense mutations ($r = 1$, $p < 0.001$)).

The prior median probability of BDQ resistance was lowest in the presence of a synonymous mutation (0.1% for *atpE*, 0.2% for *Rv0678*, 26.0% for *pepQ*), higher for inframe indels (32.8% for *pepQ*, 39.6% for *atpE*, 45.9% for *Rv0678*) and missense mutations (38.0% for *pepQ*, 60.9% for *atpE*, 66.9% for *Rv0678*) and highest for frameshift (46.1% for *pepQ*, 85.2% for *Rv0678*) and nonsense mutations (47.9% for *pepQ*, 90.0% for *Rv0678*). For the *Rv1979c* gene, the prior median probability of resistance was 42.0% for all types of mutations (Table 4; S1 and S2 Tables, showing more shape parameters).

Updated systematic literature review results

The updated search identified 503 additional studies of which five studies [38–42] were eligible for inclusion in the analysis, bringing the total number of included studies to 46 (see S1 Fig, showing the literature update flowchart). The majority of records were removed because they were reviews ($n = 140$), did not report on BDQ ($n = 152$), or studied non-tuberculous mycobacteria ($n = 47$). In the 46 included studies, 313 unique variants (223 from the published systematic review and 90 additional unique variants) were reported in 756 isolates. Most variants (92.3%, $n = 289$) occurred in clinical isolates of which 98 were reported in isolates phenotypically resistant to BDQ. Of the 189 (25.0%) BDQ resistant isolates, 142 (75.1%) were clinical isolates. Among these 142 phenotypically resistant clinical isolates, the presence of one or more

Table 3. Expert opinion on the probability of BDQ resistance given the presence of a specific class of mutations.

Mutation type	Probability of phenotypic BDQ resistance						
	Count ^a	Rarely (<5%)	Occasionally 5–24%	Sometimes 25–49%	Frequently 50–74%	Very frequently (75–94%)	Almost always (≥95%)
<i>atpE</i>							
Synonymous	32	27	4	1	0	0	0
Missense	30	2	4	5	8	7	4
Inframe indel	31	5	4	10	5	5	2
Homoplastic	33	4	3	3	4	10	9
<i>Rv0678</i>							
Synonymous	32	22	2	2	3	0	3
Nonsense	33	0	1	4	5	10	13
Frameshift	32	1	1	0	8	13	9
Missense	30	2	4	5	8	7	4
Inframe indel	30	3	2	11	9	5	0
Homoplastic	33	1	3	8	5	8	8
<i>pepQ</i>							
Synonymous	10	9	0	0	1	0	0
Nonsense	10	1	2	1	2	2	2
Frameshift	10	2	1	2	2	1	2
Missense	10	0	4	4	1	1	0
Inframe indel	9	3	1	2	2	1	0
Homoplastic	10	1	1	3	1	2	2
<i>Rv1979c</i>							
Synonymous	2	2	0	0	0	0	0
Nonsense	2	0	1	0	1	0	0
Frameshift	2	0	1	1	0	0	0
Missense	2	0	2	0	0	0	0
Inframe indel	1	1	0	0	0	0	0
Homoplastic	2	0	0	1	0	0	1

^a Count For *atpE* and *Rv0678*, all 33 experts believed these genes can play a role in BDQ resistance, but some experts did not complete all questions; for *pepQ* 12 experts believed this gene can play a role but 2 to 3 experts did complete all questions; for *Rv1979c*, only 2 experts believed this gene can play a role in BDQ resistance.

BDQ = bedaquiline

<https://doi.org/10.1371/journal.pone.0287019.t003>

Rv0678 variant was reported in 106 (74.6%) isolates. The frequency of unique variants ranged from 1 to 41, with *Rv0678* 192_193insG being the most frequent variant (Tables 5–7).

Posterior median probability of BDQ resistance by gene and variant type

For the *atpE* gene, prior (expert opinion) and posterior (prior plus literature review data) probabilities were similar for synonymous mutations (prior median: 0.1%; posterior median: 0.1%) and missense mutations (prior median: 61.0%; posterior median: 60.8%) but 95% CrIs were wide (Table 8). For the *Rv0678* gene, posterior median probabilities were lower than prior median probabilities for frameshift mutations (prior median: 85.2%; posterior median: 30.0%), inframe indels (prior median: 46.0%; posterior median: 29.6%), missense mutations (prior median: 67.0%; posterior median: 31.5%) and nonsense mutations (prior median: 90.3%; posterior median: 55.1%), and similar for synonymous mutations (prior median: 0.2%; posterior median: 3.3%), but all with wide 95% CrIs. For the *pepQ* and *Rv1979c* genes, the posterior median probabilities were lower than the prior median probabilities for frameshift

Table 4. Prior probability distribution for different types of variants by gene of interest.

Gene	Mutation type	n ^a	Mean	Median	Variance	IQR
<i>atpE</i>	Synonymous mutation	32	2.9%	0.1%	0.5	1.9%
	Inframe indel	31	43.4%	39.6%	10.2	58.7%
	Missense mutation	30	57.0%	61.0%	9.9	57.4%
	Homoplastic mutation	33	62.4%	74.6%	12.9	68.9%
<i>Rv0678</i>	Synonymous mutation	32	19.8%	0.2%	11.4	24.0%
	Nonsense mutation	33	80.0%	90.3%	5.4	29.5%
	Frameshift mutation	32	76.0%	85.0%	5.9	34.6%
	Inframe indel	30	46.9%	46.0%	7.4	45.6%
	Missense mutation	30	63.1%	67.0%	6.6	41.4%
	Homoplastic mutation	33	65.5%	74.0%	9.5	52.9%
<i>pepQ</i>	Synonymous mutation	33	27.6%	26.0%	2.6	26.1%
	Nonsense mutation	33	46.2%	47.9%	3.9	51.1%
	Frameshift mutation	33	45.2%	46.1%	3.9	51.0%
	Inframe indel	33	36.7%	32.9%	3.4	43.3%
	Missense mutation	33	38.9%	38.0%	2.8	37.5%
	Homoplastic mutation	33	46.3%	48.2%	3.8	50.2%
<i>Rv1979c</i>	Synonymous mutation	32	42.0%	42.0%	5.8	42.0%
	Nonsense mutation	32	42.0%	42.0%	5.8	42.0%
	Frameshift mutation	32	42.0%	42.0%	5.8	42.0%
	Inframe indel	32	42.0%	42.0%	5.8	42.0%
	Missense mutation	32	42.0%	42.0%	5.8	42.0%
	Homoplastic mutation	32	42.0%	42.0%	5.8	42.0%

^a The number of experts who responded to the survey; IQR = interquartile range

<https://doi.org/10.1371/journal.pone.0287019.t004>

mutations in *pepQ* (prior median: 46.1%; posterior median: 13.7%), missense mutations in *pepQ* (prior median: 26.0%; posterior median: 2.6%), missense mutations in *Rv1979c* (prior median: 42.0%; posterior median: 25.0%) and synonymous mutations in *Rv1979c* (prior median: 42.0%; posterior median: 2.9%), with mostly wide 95% CrIs (Table 8).

Posterior median probability of BDQ resistance for specific mutations

For single nucleotide polymorphisms (SNPs), the probability of resistance ranged from 0.4% for the silent mutations (*Rv0678_C108T*, *Rv0678_T798G* and *atpE_G183A*) to $\geq 95\%$ for the nonsense mutation *Rv0678_C400T* and missense mutation *atpE_G187C*. However, some missense and nonsense mutations had a substantially lower median probability, while some silent mutations had a much higher probability of resistance (Tables 5 and 7). For SNPs, the probability of resistance ranged from 0.3% to 97.1% in *atpE*, from 0.4% to 96.9% in *Rv0678*, from 22.1% to 28.7% in *pepQ*, and from 2.9% to 29.6% in *Rv1979c*. For indels, the probability of resistance ranged from 10.0% to 98.2% in *Rv0678*, and from 18% to 27% in *pepQ* (Tables 6 and 7). The 95% CrIs were wide for almost all variants.

Discussion

Accurate knowledge of the presence of BDQ resistance is important when initiating a BDQ-containing treatment regimen or changing a BDQ-containing regimen when a patient does not respond to their RR-TB treatment. Because pDST requires a lengthy culture process, there

Table 5. Posterior median probability of BDQ resistance given a specific single nucleotide polymorphism in the *Rv0678* gene.

Rv0678 SNPs	pDST										OR estimate (95% CI)	Posterior probability estimate (%) (95% CrI)						
	All					Clinical							Non-clinical					
	R	S	R	S	R	R	S	R	S	R			R	S	R	S		
C400T (nonsense)	3	0	1	0	2	0	0	0	0	0	0	0	0	0	0	0	∞ (1.6-∞)	96.9 (60.6-99.9)
G61T, T276A (nonsense)	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	∞ (0.1-∞)	94.2 (39.2-99.9)
C189A	3	0	0	0	3	0	0	0	3	0	0	0	0	0	0	0	∞ (1.6-∞)	82.9 (46.0-99.6)
A152C, C158G, G287A	2	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	∞ (0.7-∞)	84.2 (37.2-99.4)
G361A	3	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	∞ (1.6-∞)	82.9 (46.0-99.6)
C403G, G281A	2	0	0	0	2	0	0	0	2	0	0	0	0	0	0	0	∞ (0.7-∞)	84.2 (37.2-99.4)
A97G	2	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	∞ (0.7-∞)	84.2 (37.2-99.4)
C148T	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	∞ (0.1-∞)	78.1 (24.6-99.3)
C280T	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	∞ (0.1-∞)	78.1 (24.6-99.3)
A271C, A413G, A65T, G120C, G197A, T407C	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	∞ (0.1-∞)	78.1 (24.6-99.3)
A263G, C155T, C257T, G215A, G326C, T124C, T332A A199C, C247T, C286G, G203A, T118A	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	∞ (0.1-∞)	78.1 (24.6-99.3)
T200G	2	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	∞ (0.7-∞)	84.2 (37.2-99.4)
C286T	2	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	∞ (0.1-∞)	84.2 (37.2-99.4)
G5T	2	1	2	1	0	0	0	0	0	0	0	0	0	0	0	0	7.6 (0.4-450.5)	67.1 (25.9-94.5)
T157C	5	1	5	1	0	0	0	0	0	0	0	0	0	0	0	0	19.3 (2.1-910.4)	79.7 (45.7-96.9)
T131C	1	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0.0 (0.0-206.3)	58.7 (16.1-92.8)
C403T, A202G	1	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	3.8 (0.1-298.8)	58.7 (16.1-92.8)
C107T, C158T, C176T, C185T, C296T, G193A, T59G, T323C, C143T, T2C	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	3.8 (0.1-298.8)	58.7 (16.1-92.8)
C343T(nonsense)	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0.0 (0.0-148.0)	53.9 (7.3-95.5)
C214T	1	2	0	0	2	1	0	0	0	0	0	0	0	0	0	0	1.9 (0.0-36.9)	46.2 (11.5-83.9)
A152G	2	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	3.8 (0.3-52.8)	55.6 (19.2-87.9)
A14G, A331G, C-11A, C-53A, C279A, C406G, G109A, G149C, G194C, G196T, G245A, G259A, G352A, G358A, G393C, G421C, G61A, T-20A, T240G, T254G, T274G, T437G, T469G, T93G, A223G, A322G, A44G, A442C, C302A, C319T, C40T, C41T, C47A, C86T, G149A, G184A, G266T, G326A, G467A, T154C	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0.0 (0.0-148.0)	44.2 (5.6-89.7)
T239G	2	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	3.8 (0.3-52.8)	55.6 (19.2-87.9)
G287T	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0.0 (0.0-148.0)	58.7 (16.1-92.8)
A208G	2	1	2	1	0	0	0	0	0	0	0	0	0	0	0	0	7.6 (0.4-450.5)	67.1 (25.9-94.5)
T59C	0	2	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0.0 (0.0-20.3)	32.7 (4.0-78.2)
T416C	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	3.8 (0.1-298.8)	58.7 (16.1-92.8)
C265T (silent)	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	3.8 (0.1-298.8)	43.1 (2.4-94.4)
G337A	1	3	1	3	0	0	0	0	0	0	0	0	0	0	0	0	1.3 (0.0-15.8)	38.6 (9.3-76.5)
T78G, C112T(nonsense)	0	2	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0.0 (0.0-20.3)	37.6 (4.4-83.5)
A-4T, A172C, C247G, G253T, G362A, G7A, T236C, T278C, G417A, G61A, G250A, T119C	0	2	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0.0 (0.0-20.3)	32.7 (4.0-78.2)
G120C	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0.0 (0.0-20.3)	32.7 (4.0-78.2)
G106A, T365C	0	2	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0.0 (0.0-20.3)	32.7 (4.0-78.2)
A187G, T425G, A67G, C220G, C251T	0	2	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0.0 (0.0-20.3)	32.7 (4.0-78.2)
T248C	1	2	1	2	0	0	0	0	0	0	0	0	0	0	0	0	1.9 (0.0-36.8)	46.2 (11.5-83.9)
T136C	1	2	1	2	0	0	0	0	0	0	0	0	0	0	0	0	1.9 (0.0-36.8)	46.2 (11.5-83.9)
T350G	2	5	2	5	0	0	0	0	0	0	0	0	0	0	0	0	1.5 (0.1-9.4)	36.9 (11.6-67.8)
C6A, G485A, G307A	0	3	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0.0 (0.0-9.2)	26.1 (2.9-68.5)
G259C, A292G	0	4	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0.0 (0.0-5.8)	21.7 (2.4-61.3)
T341C	1	4	1	4	0	0	0	0	0	0	0	0	0	0	0	0	1.0 (0.0-9.7)	32.8 (7.5-68.4)
C268T	2	8	2	8	0	0	0	0	0	0	0	0	0	0	0	0	1.0 (0.1-4.8)	27.6 (8.2-54.8)
A11C, A165C	0	5	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0.0 (0.0-4.2)	18.2 (2.0-54.4)
A67C	0	6	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0.0 (0.0-3.2)	16.3 (1.8-50.0)
T254C, T118G	0	7	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0.0 (0.0-2.6)	14.4 (1.6-44.2)
T119C	0	8	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0.0 (0.0-20.3)	13.2 (1.4-40.8)

(Continued)

Table 5. (Continued)

R0678 SNPs	All						pDST						OR estimate (95% CI)	Posterior probability estimate (%) (95% CrI)
	R		S		Clinical		Non-clinical		R		S			
	1	2	1	2	R	S	R	S	R	S	R	S		
T437C	0	1	1	12	1	12	0	0	0	0	0	0	0.0 (0.0–2.1)	15.1 (3.2–37.7)
C225T, A318G, C127T, C384A, C460T, C81T, G168A, G363A, G390C, G408A, G441A, T297G (Silent)	0	1	0	1	0	1	0	0	0	0	0	0	0.0 (0.0–148.0)	0.5 (0.0–68.0)
C447T, G195A (silent)	0	2	0	2	0	2	0	0	0	0	0	0	0.0 (0.0–20.3)	0.5 (0.0–46.0)
G108T, T798G (silent)	0	3	0	3	0	3	0	0	0	0	0	0	0.0 (0.0–9.2)	0.4 (0.0–34.0)
C305T, T299A, A172C, C201G, T56C, G122T, C143A, A380C, A38C	1	0	1	0	1	0	0	0	0	0	0	0	∞ (0.1–∞)	78.1 (24.6–99.3)
G137A	0	1	0	1	0	1	0	0	0	0	0	0	0.0 (0.0–148.0)	44.2 (5.6–89.7)

BDQ = bedaquiline; SNP = Single Nucleotide Polymorphism; pDST = phenotypic drug susceptibility test; R = resistant, S = susceptible, OR = odds ratio, CI = confidence interval, CrI = credible interval

<https://doi.org/10.1371/journal.pone.0287019.t005>

Table 6. Posterior median probability of BDQ resistance of Rv0678 indels.

Rv0678 indels	pDST										OR estimate (95% CI)	Posterior probability estimate (%) (95% CrI)
	All		Clinical				Non-clinical					
	R	S	R	S	R	S	R	S	R	S		
198_199insG	3	0	3	0	0	0	0	0	0	0	0	95.7 (57.4–100.0)
139_141insTG, 145_147indel, 16_17delGG, 172_173insS6110, 18_19delGG, 19delG, 212delC, 330delA, 107_108insG, 15delC, 325_326insG, 464_466insGC	1	0	1	0	0	0	0	0	0	0	0	91.1 (36.1–99.9)
259_260insG, 272_273insS6110, 334_335insS6110, 349_350insS6110, 38_39insA, 65_66insS6110, 94_95insS6110	1	0	0	0	0	0	1	0	0	0	0	91.1 (36.1–99.9)
138_139insG	4	3	4	3	0	0	0	0	0	0	0	62.3 (29.6–88.4)
107delG, 138_140insGG, 138_140insGA, 140_141insG, 142_143delCT, 142_143insC, 142delC, 15DelG, 176_177delCG, 176_178insGG, 192_194insCG, 193_194insG, 214delC, 262_263insA, 274_283delTATTCCCGT, 288delG, 318_320insCG, 335delC, 43_44insA, 437_438insT, 457delG, 46indel?, 211delG, 28_29insT, 291delC, 29delT, 329delC, 382delG, 461delT, 492_493insGA	0	1	0	1	0	0	0	0	0	0	0	51.7 (6.6–94.0)
193delG	2	6	0	6	2	0	0	0	0	0	0	34.2 (10.7–65.3)
140_141insC, 291_292insA, 29delG, 359_360insG, 464_465insC, 136_137insG, 198delG	0	2	0	2	0	0	0	0	0	0	0	36.2 (4.1–82.6)
400delCGACGGCTGGCG (inframe indel)	0	1	0	1	0	0	0	0	0	0	0	29.7 (1.6–82.2)
133_134insTG, 184_185insC, 435delT, 465_466insC	0	3	0	3	0	0	0	0	0	0	0	30.8 (3.1–72.3)
137_138insG	1	3	1	3	0	0	0	0	0	0	0	41.4 (10.1–79.6)
139_140insG	4	3	4	3	0	0	0	0	0	0	0	62.3 (29.6–88.4)
192delG, 288delC	1	6	1	6	0	0	0	0	0	0	0	26.6 (5.7–59.4)
274_275insA	2	6	2	6	0	0	0	0	0	0	0	34.2 (10.7–65.3)
434delT	0	4	0	4	0	0	0	0	0	0	0	22.6 (2.6–64.1)
141_142insC	4	19	4	19	0	0	0	0	0	0	0	21.3 (8.4–40.0)
16delG	1	5	1	5	0	0	0	0	0	0	0	30.3 (6.8–65.2)
418_419insG	0	9	0	9	0	0	0	0	0	0	0	11.8 (1.3–39.5)
192_193insG	3	38	3	38	0	0	0	0	0	0	0	10.0 (3.4–20.9)
11_63del, 430_431insGC, 466_467insGC, 141_142insTC, 142_143insGATC, 95delT	1	0	1	0	0	0	0	0	0	0	0	91.1 (36.1–99.9)
del, 181_182insG, 150_151insG	2	0	2	0	0	0	0	0	0	0	0	93.1 (47.9–99.9)
144_145insC	10	0	10	0	0	0	0	0	0	0	0	98.2 (79.9–100.0)
ins_156110	1	0	1	0	0	0	0	0	0	0	0	91.1 (36.1–99.9)

BDQ = bedaquiline; pDST = phenotypic drug susceptibility test, R = resistant, S = susceptible, OR = odds ratio, CI = confidence interval, CrI = credible interval

<https://doi.org/10.1371/journal.pone.0287019.t006>

Table 7. Posterior probability of BDQ resistance given a specific mutation in *atpE*, *pepQ* and *Rv1979c* genes.

Gene	Mutation	pDST						OR estimate (95% CI)	Posterior probability estimate (%) (95% CrI)
		All		Clinical		Non-clinical			
		R	S	R	S	R	S		
<i>atpE</i>	G187C	11	0	1	0	10	0	∞ (9.9- ∞)	97.1 (78.2-99.9)
	A83T	2	0	0	0	2	0	∞ (0.7- ∞)	88.2 (34.9-99.9)
	A83C, T-72C	1	0	0	0	1	0	∞ (0.1- ∞)	81.5 (19.0-99.9)
	A83G	3	1	0	1	3	0	11.7 (0.9-615.6)	73.0 (30.8-96.8)
	C188T, G183T, A196G, G73A, G-53A	0	1	0	1	0	0	0.0 (0.0-148.0)	29.5 (0.6-88.4)
	C198G	1	0	1	0	0	0	∞ (0.1- ∞)	81.5 (19.0-99.9)
	G183C	1	0	1	0	0	0	∞ (0.1- ∞)	81.5 (19.0-99.9)
	T-138C	0	3	0	3	0	0	0.0 (0.0-9.2)	13.9 (0.3-60.0)
	G82A	0	4	0	4	0	0	0.0 (0.0-5.8)	11.0 (0.2-52.3)
	G183A (silent mutations)	0	1	0	1	0	0	0.0 (0.0-148.0)	0.3 (0.0-23.0)
<i>pepQ</i>	T131C, A324G, A706G, C-31T, C114G, C206T, C269T, C371T, C483A, C7T, C914T, G-12C, G1108A, G274A, G278A, G454A, G500T, G640T, T641C	0	1	0	0	0	1	0.0 (0.0-206.3)	28.7 (1.8-80.2)
	811delC	0	1	0	0	0	1	0.0 (0.0-206.3)	27.0 (0.6-85.0)
	A1021G, A407G, C233T, G20A, G347T, G925A	0	2	0	2	0	0	0.0 (0.0-28.3)	22.1 (1.3-66.0)
	424delC	0	2	0	0	0	2	0.0 (0.0-28.3)	18.0 (0.4-70.8)
	A155C, G1216A, G724A, T1057G, T311C, T824C, A733C, C562T	0	1	0	1	0	0	0.0 (0.0-285.9)	29.6 (1.3-83.0)
<i>Rv1979c</i>	T798G (silent)	0	2	0	2	0	0	0.0 (0.0-39.3)	20.8 (0.9-70.2)
	A1403G, G20A, A187G, C520T, C1432G	0	2	0	2	0	0	0.0 (0.0-39.3)	20.8 (0.9-70.2)
	G114T/C	0	8	0	8	0	0	0.0 (0.0-4.3)	7.4 (0.3-33.3)
	G114C	1	23	1	23	0	0	0.3 (0.0-2.0)	6.6 (1.0-20.2)
	G1226A	2	37	2	37	0	0	0.4 (0.0-1.5)	6.5 (1.6-16.8)
	T151A	0	14	0	14	0	0	0.0 (0.0-2.2)	5.1 (0.2-23.2)
	A857G	0	22	0	22	0	0	0.0 (0.0-1.3)	2.9 (0.1-14.9)

BDQ = bedaquiline; pDST = phenotypic drug susceptibility test, R = resistant, S = susceptible, OR = odds ratio, CI = confidence interval, CrI = credible interval

<https://doi.org/10.1371/journal.pone.0287019.t007>

Table 8. Prior and posterior probability of BDQ resistance by mutation type and gene of interest.

Gene	Mutation type	Prior		pDST		Posterior		95% CrI
		Mean probability	Median probability	R	S	Mean probability	Median probability	
<i>atpE</i>	Synonymous	2.9%	0.1%	0	1	2.4%	0.1%	0.0–19.7%
	Missense	57.0%	61.0%	20	13	60.5%	60.8%	44.0–76.1%
<i>Rv0678</i>	Synonymous	19.8%	0.2%	1	23	4.5%	3.3%	0.2–15.4%
	Nonsense	80.0%	90.3%	5	5	54.8%	55.1%	27.3–80.7%
	Frameshift	76.0%	85.2%	69	164	30.0%	30.0%	24.2–36.1%
	Inframe indel	46.9%	46.0%	0	1	33.0%	29.6%	1.7–81.8%
	Missense	63.1%	67.0%	91	200	31.5%	31.5%	26.3–36.8%
	Frameshift	45.2%	46.1%	0	3	18.1%	13.7%	0.2–59.0%
<i>pepQ</i>	Missense	38.8%	26.0%	0	33	3.4%	2.6%	0.1–11.7%
<i>Rv1979c</i>	Synonymous	42.0%	42.0%	0	2	20.8%	25.0%	0.9–70.2%
	Missense	42.0%	42.0%	3	122	3.2%	2.9%	0.9–6.9%

BDQ = bedaquiline; CrI = Credible interval; pDST = phenotypic drug susceptibility; R = Resistant; S = Susceptible

<https://doi.org/10.1371/journal.pone.0287019.t008>

is great interest in a genotypic assay. Unfortunately, developing such assays is challenging given the existence of multiple BDQ resistance candidate genes, the occurrence of variants across the genes, and the difficulty to confidently assign a BDQ phenotype to genetic variants using standard statistical methods due to the low prevalence of variants [11, 43, 44].

To overcome these challenges, we combined expert opinions and published genotype-phenotype data to estimate the Bayesian probability of BDQ resistance in the presence of a genomic variant. Our results showed that many experts were uncertain about the role of *pepQ* and *Rv1979c* variants in BDQ resistance and that experts overestimated the probability of resistance for most variant types, resulting in lower posterior (median) probabilities as compared to prior estimates. As expected, the posterior probability of BDQ resistance was low for synonymous mutations in *atpE* and *Rv0678* (<5%) and relatively high for missense mutations in *atpE* (60.8%) and nonsense mutations in *Rv0678* (55.1%). However, some missense and nonsense mutations had a substantially lower median probability, while some synonymous mutations had a much higher probability of resistance. This could either suggest that the presence of a specific type of mutation is not sufficient to determine the resistance phenotype or that the contradictory findings are a consequence of limited data availability. Surprisingly, the posterior probability of BDQ resistance was relatively low for missense mutations and frameshift in *Rv0678* ($\leq 35\%$) and low for missense mutations in *pepQ* and *Rv1979c* (<3%). For specific SNPs and indels, the probability of resistance varied greatly between variants of the same type, even in the same gene, and 95% CrIs were wide due to sparse data.

In our analysis, the use of a Bayesian approach did not substantially improve the ability to predict bedaquiline resistance in the presence of specific variants. Nevertheless, we believe this is an important step forward for two reasons. First, the Bayesian approach is flexible and allows inclusion of different types of data. When, for example, data in the effect of variants on protein structure becomes available, this information could be used to improve the Bayesian prediction model. Second, for clinicians, the interpretation of a probability may be more intuitive than the interpretation of an odds ratio. For example, for variant *Rv0678_A202G*, the odds ratio of 3.8 (95% CI of 0.05–298.8) is interpreted as: the odds of resistance to BDQ for an *Mtb* isolate with a A202G variant is 3.8 times that of isolates without the A202G variant, independent of the presence of wild type or any other variant, and if we performed the analysis 100 times using different datasets (based on different samples), the estimate of the odds ratio

would lie between the lower and upper limit of the constructed CI in 95% of the analyses. In contrast, the interpretation of the probability of resistance to BDQ for an *Mtb* isolate with a A202G variant derived from a posterior distribution with median (or mean) of 58.7% (95% CrI 16.1–92.8%) is more intuitive: the probability of resistance to BDQ for the isolate containing the A202G variant is estimated to be 58.7% and we are 95% certain that the true probability of resistance lies between 16.1% and 92.8%, given the data at hand.

The results of our study should be interpreted in light of some limitations. First, while the most accurate priors would be obtained by limiting the survey to the most knowledgeable experts, we deliberately extended the survey to obtain a broad and globally representative perspective. While this could have caused the wide range of opinions observed, the responses on the probability of resistance by mutation type in the *Rv0678* and *atpE* genes were similar for experts with a laboratory versus clinical, epidemiological, or public health background. To limit the impact of extreme values in the survey responses, we used the median instead of the mean to summarize the prior probabilities. In the face of discrepancies, determining whether the prior or the data is correct is challenging [23]. Second, in the construction of the prior distributions, we did not account for the uncertainty around the estimated shape parameters of the beta distributions. In order to do so, one would need to specify hyperprior distributions for α and β thereby increasing posterior uncertainty further. Third, the published expert rules were developed for resistance to other drugs used to treat TB and may not fully apply to BDQ. For example, while *Miotto et al* assumed that nonsense and frameshift mutations can be predicted with confidence to cause a loss-of-function phenotype [19, 25], we found that the posterior probability of BDQ resistance in the presence of a nonsense mutation in *Rv0678* was only 55.1% and the probability of resistance for frameshift mutations in *Rv0678* was only 29.9%. Fourth, the phenotype-genotype reported in studies included in the analysis may not always be accurate due to lack of standardization of BDQ pDST and the use of the current critical concentration to determine the BDQ resistance phenotype [24]. This may especially be a problem for mutations that confer a modest increase in BDQ MIC, which overlaps with the wild-type MIC distribution [41]. Fifth, to maximally use all reported data on BDQ resistance given a specific mutation, data on clinical and non-clinical isolates was combined and analyzed. Furthermore, the published data may not reflect the actual sampling distributions of the phenotype for a specific variant. Finally, posterior distributions can only be estimated for variants that have been reported in the literature. Given that the introduction of BDQ is recent, it is possible that many variants have not yet been observed or reported in the literature. The strength of the Bayesian approach is that it also provides a resistance probability for different variant types in the candidate resistance genes, which can give guidance to the clinician when confronted with a novel variant.

Probabilistic approaches to predict phenotypic resistance to BDQ from genotypic data could be an alternative approach to the standard statistical methods and help guide treatment decisions as it presents physicians with more intuitively interpretable probabilities (and 95% CrI) as compared to odds ratios (and corresponding 95% CI). For a variant not previously reported, the probability of resistance for a type of variant in a specific gene can still be used to guide clinical decision making. Since a probabilistic approach is new for clinicians, studies should investigate how feasible it is for physicians to use probabilities of BDQ resistance in clinical practice.

Supporting information

S1 Fig. Study selection after the published systematic review and individual meta-analysis. (DOCX)

S1 Table. Prior probability distribution parametrised by beta shape parameters (α and β) based on experts' responses in the three of genes of interest.

(DOCX)

S2 Table. Mixture distribution and mixing proportion for the prior distribution of pepQ and Rv1979c.

(DOCX)

S1 File. Supplementary R code.

(TXT)

S2 File. Supplementary data.

(XLSX)

S3 File. Supplementary questionnaire.

(DOCX)

Acknowledgments

We thank Claudio Köser for his input into the development of the survey tool and the Tuberculosis Omics ResearCH team members for piloting the survey tool. We also thank the scientists who participated in the survey.

Author Contributions

Conceptualization: Degefaye Zelalem Anlay, Pham Hien Trang Tu, Steven Abrams, Annelies Van Rie.

Data curation: Degefaye Zelalem Anlay, Annelies Van Rie.

Formal analysis: Degefaye Zelalem Anlay, Emmanuel Rivière, Pham Hien Trang Tu, Steven Abrams, Annelies Van Rie.

Funding acquisition: Annelies Van Rie.

Investigation: Degefaye Zelalem Anlay, Emmanuel Rivière, Pham Hien Trang Tu, Steven Abrams, Annelies Van Rie.

Methodology: Degefaye Zelalem Anlay, Steven Abrams, Annelies Van Rie.

Project administration: Annelies Van Rie.

Supervision: Steven Abrams, Annelies Van Rie.

Writing – original draft: Degefaye Zelalem Anlay, Emmanuel Rivière, Pham Hien Trang Tu, Steven Abrams, Annelies Van Rie.

Writing – review & editing: Degefaye Zelalem Anlay, Emmanuel Rivière, Pham Hien Trang Tu, Steven Abrams, Annelies Van Rie.

References

1. World Health Organization. Tuberculosis 2021 <https://www.who.int/news-room/fact-sheets/detail/tuberculosis>.
2. World Health Organization. Key fact: Tuberculosis 2020 updated 14 October 2020. <https://www.who.int/news-room/fact-sheets/detail/tuberculosis>
3. World Health Organization. Global tuberculosis report 2021. Geneva: WHO; 2021.

4. Cox E, Laessig K. FDA approval of bedaquiline—the benefit–risk balance for drug-resistant tuberculosis. *New England Journal of Medicine*. 2014; 371(8):689–91. <https://doi.org/10.1056/NEJMp1314385> PMID: 25140952
5. World Health Organization. WHO consolidated guidelines on drug-resistant tuberculosis treatment: World Health Organization; 2019.
6. Andries K, Vilellas C, Coeck N, Thys K, Gevers T, Vranckx L, et al. Acquired resistance of *Mycobacterium tuberculosis* to bedaquiline. *PLoS One*. 2014; 9(7):e102135–e. <https://doi.org/10.1371/journal.pone.0102135> PMID: 25010492
7. Ismail NA, Omar SV, Joseph L, Govender N, Blows L, Ismail F, et al. Defining bedaquiline susceptibility, resistance, cross-resistance and associated genetic determinants: a retrospective cohort study. *EBio-Medicine*. 2018; 28:136–42. <https://doi.org/10.1016/j.ebiom.2018.01.005> PMID: 29337135
8. Nimmo C, Millard J, Van Dorp L, Brien K, Moodley S, Wolf A, et al. Population-level emergence of bedaquiline and clofazimine resistance-associated variants among patients with drug-resistant tuberculosis in southern Africa: a phenotypic and phylogenetic analysis. *The Lancet Microbe*. 2020; 1(4):e165–e74. [https://doi.org/10.1016/S2666-5247\(20\)30031-8](https://doi.org/10.1016/S2666-5247(20)30031-8) PMID: 32803174
9. Villellas C, Coeck N, Meehan CJ, Lounis N, de Jong B, Rigouts L, et al. Unexpected high prevalence of resistance-associated Rv0678 variants in MDR-TB patients without documented prior use of clofazimine or bedaquiline. *Journal of Antimicrobial Chemotherapy*. 2017; 72(3):684–90. <https://doi.org/10.1093/jac/dkw502> PMID: 28031270
10. World Health Organization. Technical report on critical concentrations for TB drug susceptibility testing of medicines used in the treatment of drug-resistant TB. <https://apps.who.int/iris/handle/10665/260470>. 2018.
11. Ismail N, Rivière E, Limberis J, Huo S, Metcalfe JZ, Warren RM, et al. Genetic variants and their association with phenotypic resistance to bedaquiline in *Mycobacterium tuberculosis*: a systematic review and individual isolate data analysis. *The Lancet Microbe*. 2021. [https://doi.org/10.1016/s2666-5247\(21\)00175-0](https://doi.org/10.1016/s2666-5247(21)00175-0) PMID: 34796339
12. McNerney R, Zignol M, Clark TG. Use of whole genome sequencing in surveillance of drug resistant tuberculosis. *Expert review of anti-infective therapy*. 2018; 16(5):433–42. <https://doi.org/10.1080/14787210.2018.1472577> PMID: 29718745
13. Nieto Ramirez LM, Quintero Vargas K, Diaz G. Whole Genome Sequencing for the Analysis of Drug Resistant Strains of *Mycobacterium tuberculosis*: A Systematic Review for Bedaquiline and Delamanid. *Antibiotics*. 2020; 9(3):133. <https://doi.org/10.3390/antibiotics9030133> PMID: 32209979
14. World Health Organization. Catalogue of mutations in *Mycobacterium tuberculosis* complex and their association with drug resistance. 2021.
15. Matteelli A, Carvalho AC, Dooley KE, Kritski A. TMC207: the first compound of a new class of potent anti-tuberculosis drugs. *Future microbiology*. 2010; 5(6):849–58. <https://doi.org/10.2217/fmb.10.50> PMID: 20521931
16. Guo H, Courbon GM, Bueler SA, Mai J, Liu J, Rubinstein JL. Structure of mycobacterial ATP synthase bound to the tuberculosis drug bedaquiline. *Nature*. 2021; 589(7840):143–7. <https://doi.org/10.1038/s41586-020-3004-3> PMID: 33299175
17. Yamamoto K, Nakata N, Mukai T, Kawagishi I, Ato M. Coexpression of MmpS5 and MmpL5 Contributes to Both Efflux Transporter MmpL5 Trimerization and Drug Resistance in *Mycobacterium tuberculosis*. *mSphere*. 2021; 6(1):e00518–20. <https://doi.org/10.1128/mSphere.00518-20> PMID: 33408221
18. Kadura S, King N, Nakhoul M, Zhu H, Theron G, Köser CU, et al. Systematic review of mutations associated with resistance to the new and repurposed *Mycobacterium tuberculosis* drugs bedaquiline, clofazimine, linezolid, delamanid and pretomanid. *J Antimicrob Chemother*. 2020; 75(8):2031–43. <https://doi.org/10.1093/jac/dkaa136> PMID: 32361756
19. Miotto P, Tessema B, Tagliani E, Chindelevitch L, Starks AM, Emerson C, et al. A standardised method for interpreting the association between mutations and phenotypic drug resistance in *Mycobacterium tuberculosis*. *European Respiratory Journal*. 2017; 50(6). <https://doi.org/10.1183/13993003.01354-2017> PMID: 29284687
20. Sadia F, Hossain SS. Contrast of bayesian and classical sample size determination. *Journal of Modern Applied Statistical Methods*. 2014; 13(2):23.
21. Ellison AM. Bayesian inference in ecology. *Ecology letters*. 2004; 7(6):509–20.
22. Hespanhol L, Vallio CS, Costa LM, Saragiotto BT. Understanding and interpreting confidence and credible intervals around effect estimates. *Brazilian journal of physical therapy*. 2019; 23(4):290–301. <https://doi.org/10.1016/j.bjpt.2018.12.006> PMID: 30638956
23. Rindskopf D. Reporting Bayesian Results. *Evaluation Review*. 2020:0193841X20977619. <https://doi.org/10.1177/0193841X20977619> PMID: 33380224

24. Walker TM, Miotto P, Köser CU, Fowler PW, Knaggs J, Iqbal Z, et al. The 2021 WHO catalogue of Mycobacterium tuberculosis complex mutations associated with drug resistance: a genotypic analysis. *The Lancet Microbe*. 2022; 3(4):e265–e73. [https://doi.org/10.1016/S2666-5247\(21\)00301-3](https://doi.org/10.1016/S2666-5247(21)00301-3) PMID: 35373160
25. Köser CU, Cirillo DM, Miotto P. How to optimally combine genotypic and phenotypic drug susceptibility testing methods for pyrazinamide. *Antimicrobial agents and chemotherapy*. 2020; 64(9). <https://doi.org/10.1128/AAC.01003-20> PMID: 32571824
26. Sullivan GM, Artino AR Jr., Analyzing and interpreting data from likert-type scales. *J Grad Med Educ*. 2013; 5(4):541–2. <https://doi.org/10.4300/JGME-5-4-18> PMID: 24454995
27. Nelder JA, Mead R. A Simplex Method for Function Minimization. *The Computer Journal*. 1965; 7(4):308–13.
28. Sun J. *The statistical analysis of interval-censored failure time data*: Springer; 2007.
29. Bouguila N, Ziou D, Monga E. Practical Bayesian estimation of a finite beta mixture through Gibbs sampling and its applications. *Statistics and Computing*. 2006; 16(2):215–25.
30. Lehoczky JP. Distributions, Statistical: Special and Continuous. In: Smelser NJ, Baltes PB, editors. *International Encyclopedia of the Social & Behavioral Sciences*. Oxford: Pergamon; 2001. p. 3787–93.
31. Rosner B. *Fundamentals of biostatistics*: Cengage learning; 2015.
32. Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. *Journal of computational and graphical statistics*. 1998; 7(4):434–55.
33. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. *Statistical science*. 1992; 7(4):457–72.
34. Vats D, Knudson C. Revisiting the gelman-rubin diagnostic. *arXiv preprint arXiv:181209384*. 2018.
35. Link WA, Eaton MJ. On thinning of chains in MCMC. *Methods in ecology and evolution*. 2012; 3(1):112–5.
36. Carlo M. Comment: One long run with diagnostics: Implementation strategies for markov chain. *Statistical Science*. 1992; 7(4):493–7.
37. Raftery AE, Lewis S. How many iterations in the Gibbs sampler?: WASHINGTON UNIV SEATTLE DEPT OF STATISTICS; 1991.
38. Chesov E, Chesov D, Maurer FP, Andres S, Utpatel C, Barilar I, et al. Emergence of bedaquiline-resistance in a high-burden country of tuberculosis. *Eur Respir J*. 2021.
39. Ismail NA, Omar SV, Moultrie H, Bhyat Z, Conradie F, Enwerem M, et al. Assessment of epidemiological and genetic characteristics and clinical outcomes of resistance to bedaquiline in patients treated for rifampicin-resistant tuberculosis: a cross-sectional and longitudinal study. *Lancet Infect Dis*. 2021. [https://doi.org/10.1016/S1473-3099\(21\)00470-9](https://doi.org/10.1016/S1473-3099(21)00470-9) PMID: 34780706
40. Sonnenkalb L, Carter J, Spitaleri A, Iqbal Z, Hunt M, Malone K, et al. Deciphering Bedaquiline and Clofazimine Resistance in Tuberculosis: An Evolutionary Medicine Approach. *bioRxiv*. 2021.
41. Wu SH, Chan HH, Hsiao HC, Jou R. Primary Bedaquiline Resistance Among Cases of Drug-Resistant Tuberculosis in Taiwan. *Front Microbiol*. 2021; 12:754249. <https://doi.org/10.3389/fmicb.2021.754249> PMID: 34745058
42. Zheng H, He W, Jiao W, Xia H, Sun L, Wang S, et al. Molecular characterization of multidrug-resistant tuberculosis against levofloxacin, moxifloxacin, bedaquiline, linezolid, clofazimine, and delamanid in southwest of China. *BMC Infectious Diseases*. 2021; 21(1):330. <https://doi.org/10.1186/s12879-021-06024-8> PMID: 33832459
43. Battaglia S, Spitaleri A, Cabibbe A, Meehan C, Utpatel C, Ismail N, et al. Characterization of genomic variants associated with resistance to bedaquiline and delamanid in naïve Mycobacterium tuberculosis clinical strains. *Journal of clinical microbiology*. 2020; 58(11).
44. Degiacomi G, Sammartino JC, Sinigiani V, Marra P, Urbani A, Pasca MR. In vitro Study of Bedaquiline Resistance in Mycobacterium tuberculosis Multi-Drug Resistant Clinical Isolates. *Frontiers in microbiology*. 2020; 11:2290. <https://doi.org/10.3389/fmicb.2020.559469> PMID: 33042066