**Title** The distribution of fitness costs of resistance-conferring mutations is a key determinant for the future burden of drug-resistant tuberculosis: a model-based analysis

**Authors** Gwenan M. Knight1, Caroline Colijn2, Sourya Shrestha3, Mariam Fofana3, Frank Cobelens4,5, Richard G. White1, David W. Dowdy3, Ted Cohen6

**Author Affiliations**

1 TB Modelling Group, Centre for the Mathematical Modelling of Infectious Diseases, TB Centre, Department of Infectious Disease Epidemiology, Faculty of Epidemiology and Public Health, London School of Hygiene and Tropical Medicine, United Kingdom

2 Department of Mathematics, Imperial College London, United Kingdom

3 Department of Epidemiology, Johns Hopkins School of Public Health, Baltimore, Maryland;

4 Amsterdam Institute for Global Health and Development, Academic Medical Center, Amsterdam, The Netherlands

5 KNCV Tuberculosis Foundation, The Hague, The Netherlands

6 Department of Epidemiology of Microbial Diseases, School of Public Health, Yale University, New Haven, Connecticut.

**Keywords:** tuberculosis,antibiotic resistance, fitness costs, mathematical modelling

**Running title:** Impact of fitness spread on resistant TB

**Corresponding author:** Ted Cohen, Department of Epidemiology of Microbial Diseases, School of Public Health, Yale University, New Haven, Connecticut. Telephone: (203) 785-5126 Email: theodore.cohen@yale.edu

**Alternate corresponding author:** Gwenan Knight, IDE, EPH, London School of Hygiene and Tropical Medicine, WC1E 7HT United Kingdom. Telephone: +44 (0)20 7636 2407. Email: gwen.knight@lshtm.ac.uk

**Summary:** Different resistance-conferring mutations are associated with different effects on *M. tuberculosis* fitness. By incorporating these differences within a model of TB transmission, we show that both the mean and the variance of these costs drive drug resistant TB prevalence.

**Abstract**

***Background.*** Drug resistance poses a serious challenge for the control of tuberculosis in many settings. It is well established that the expected future trend in resistance depends on the reproductive fitness of drug-resistant *Mycobacterium tuberculosis*. However the variability in fitness between strains with different resistance-conferring mutations has been largely ignored when making these predictions.

***Methods.*** We developed a novel approach for incorporating the variable fitness costs of drug resistance-conferring mutations and for tracking this distribution of fitness costs over time within a transmission model. We used this approach to describe the effects of realistic fitness cost distributions on the future prevalence of drug-resistant tuberculosis.

***Results.*** The shape of the distribution of fitness costs was a strong predictor of the long-term prevalence of resistance. While, as expected, lower average fitness costs of drug resistance-conferring mutations were associated with more severe epidemics of drug-resistant tuberculosis, fitness distributions with greater variance also led to higher levels of drug resistance. For example, compared to simulations in which the fitness cost of resistance was fixed, introducing a realistic amount of variance resulted in a 40% increase in prevalence of drug-resistant tuberculosis after 20 years.

***Conclusions.*** The differences in the fitness costs associated with drug resistance-conferring mutations are a key determinant of the future burden of drug-resistant tuberculosis. Future studies that can better establish the range of fitness costs associated with drug resistance-conferring mutations will improve projections and thus facilitate better public health planning efforts.

**Introduction**

Drug-resistant forms of tuberculosis (DR-TB) are a persistent threat to effective control of tuberculosis in many settings and, by any method of accounting, exact a substantial global health and economic toll [1, 2]. Currently, data to evaluate trends in the burden of drug-resistant TB are limited: in most countries, sufficiently robust surveillance is not available to evaluate whether the incidence of DR-TB is increasing or decreasing. This lack of trend information makes it difficult to assess whether the epidemiology of DR-TB is changing over time, to determine whether interventions have been effective at controlling DR-TB, and to make appropriate plans for future resources needs.

In the absence of robust data on trends, mathematical models have served as a tool to help guide our thinking about how DR-TB epidemics may progress over time and which factors may influence these trends [3-6]. One of the most important determinants of DR-TB projections is the reproductive number of drug-resistant forms of TB, defined as the expected number of secondary cases of DR-TB that are attributable to a single patient infectious with DR-TB. When the reproductive number exceeds the critical threshold of one, each existing case of DR-TB will cause, on average, at least another case of DR-TB through transmission and the DR-TB epidemic will not be contained. The reproductive number depends on pathogen biological factors, factors impacting the duration of the infectious period, and the degree of vulnerability of the population in which the pathogen is being spread [7].

Drug resistance arises initially in the bacterium that causes TB disease, *Mycobacterium tuberculosis*, via chromosomal mutations [8]; these rare sporadic mutants may be selected by sub-optimal treatment leading to acquired resistance. After drug resistance emerges among individuals receiving ineffective treatment, these forms of resistant *M. tuberculosis* may be transmitted directly to others leading to primary (or transmitted) resistance.

 Multiple studies have shown that different mutations can confer similar resistance phenotypes, but may be associated with very different effects on the reproductive capacity (“fitness”) of strains [9-12]. For example, resistance to rifampicin can be encoded by several different mutations in the *rpoB* gene [13], each of which has a different effect on *in vitro* growth rates [9-11]. These experimental measures of fitness often correlate well with the apparent reproductive fitness in clinical populations – in several settings those mutations that are least costly are those that are preferentially transmitted [9, 11, 14, 15]. Worryingly, recent genomic studies have found that in several settings there is already a dominance of multidrug-resistant (MDR) TB strains with these lowest cost mutations [16-18], as well as strains with compensatory mutations that can partially ameliorate the initial fitness cost to resistance [19].

 Determining the speed with which these different DR-TB strains arise and spread is vital for understanding the epidemic potential of DR-TB. Several previous mathematical models have investigated DR-TB spread [3, 5, 20, 21]; however, most models assume a single reproductive fitness level for DR-TB strains. This assumption does not allow for the possibility that some strains of DR-TB will have less costly mutations and may be preferentially transmitted, leading to changes in the mean and distribution of fitness costs within the population of resistant strains over time. A few models have allowed for a small number of fitness levels of DR-TB strains [22, 23], but consideration of realistic distributions of fitness costs associated with drug resistance-conferring mutations has not been incorporated into a simple modelling framework.

Here we use experimental data to parameterise the distribution of fitness costs at resistance acquisition, and introduce a novel method for dynamically tracking changes in fitness within a DR-TB population. Using this new model, we illustrate how heterogeneity in the fitness costs of mutation impacts the expected future burden of DR-TB.

**Methods**

We expanded a standard model for TB transmission to include a function that tracks the distribution of reproductive fitness costs of DR forms of *M. tuberculosis* over time. We considered the effect of realistic distributions of fitness costs on the projected burden of resistance to a new drug over a 20-year time horizon.

***TB transmission model***

We modified a previously published model [20], which is structured similarly to other TB models [4, 24]. The model includes three basic health states: TB-uninfected, latent TB infection (LTBI) and active (infectious) TB disease (Figure 1). LTBI is modelled as an asymptomatic and non-infectious state that persists throughout an individual’s life and may reactivate to active (symptomatic, infectious) TB disease at any time. We also allow for rapid progression of disease upon initial infection, reflecting the fact that the majority of individuals who develop active TB do so within 5 years of their initial infection [25]. The strains causing infection and disease are classified by resistance phenotype to the new drug as either drug-susceptible (DS-TB) or drug-resistant (DR-TB). DR-TB strains appear first via acquired resistance (i.e. sporadic mutation and subsequent selection among individuals ineffectively treated for active DS-TB). These resistant strains may then be transmitted at a rate determined by the reproductive fitness associated with the specific mutation responsible for the resistant phenotype.

 We calibrated the model by altering the *M. tuberculosis* transmission rate to reach a base case steady-state TB prevalence of 150 per 100,000 prior to new drug introduction. The probability of acquiring resistance was benchmarked to a baseline scenario of current rifampicin resistance levels and the level of treatment success for DR-TB set to that for MDR-TB [20]. A table with all parameter values is available in the Supplementary materials (Supplemental Table 1). It was assumed that a new drug was introduced at time zero. The outputs were the prevalence of active cases with resistance to this new drug (DR-TB) per 100,000 at 5 and 20 years from the time of drug introduction.

 Both a deterministic and stochastic version were implemented in R [26]. The stochastic model allows for exploration of chance events: in particular, this model variant allows us to include the effect that small drug-resistant sub-populations may die out by chance, even when the effective reproductive number exceeds unity. Full details of model implementation are available in the Supplementary materials.

***Modelling the fitness of new resistance mutations***

To parameterize the shape of the distribution of fitness costs associated with various drug resistance-conferring mutations, we used previously published data for rifampicin on the frequency of each resistance-conferring mutation in the set of spontaneous mutants derived *in vitro* and their relative fitness levels from growth competition experiments. The data were available from two experimental studies [9, 10] and are shown in Figure 2a & Table 1. This acquisition distribution had a mean relative fitness (versus susceptible strains) of 0.87, with most mutations clustered around a relative fitness of 0.86, and all above 0.5.

 In our models, we explored how several beta distributions (examples in Figure 2b) of similar shape to the empirical rifampicin data affected the projected trajectory of DR-TB over time. These distributions are bounded between 0 and 1 and parameterised by two shape parameters, which we selected to produce a range of mean fitness levels from 0.5-0.9 and a variance between 0.004 and 0.032.

***Tracking the distribution of fitness costs among drug-resistant strains over time***

To capture the effect of the natural history dynamics (Figure 1) on fitness, we developed a function that tracks the proportion of active and latent cases with DR-TB strains at each level of relative fitness over time (see Supplementary materials). This function accounts for the distribution of fitness costs associated with new mutations (the acquisition distribution, captured with different beta distributions) and the preferential transmission of strains with higher fitness. At each time, the function returns a mean relative fitness of extant DR-TB strains among active (i.e. infectious) TB cases that is then used in the dynamic transmission model to determine the number of subsequent infections (Figure 1). Hence relative fitness is here defined as relative ability to transmit (rather than, for example, relative ability to cause disease after transmission).

***Impact of background force of TB infection***

To test the impact of differing forces of TB infection on projections of drug resistance, we performed analyses where we assumed lower (50/100,000) and higher (1000/100,000) prevalence of TB than in our base case scenario.

**Results**

***The projected burden of DR-TB is dependent on the fitness cost of resistance***

As expected, if we assume that resistance is associated with a single, fixed fitness cost, the projected level of DR-TB at 5 and 20 years (circles, Figure 3a&b) is strongly dependent on this cost. For example, if resistance-conferring mutations confer a 50% fitness cost, then the projected prevalence of DR-TB 20 years after drug introduction is 3 DR-TB cases per 100,000 (mean of acquisition distribution: 0.5; Figure 3b). If they confer only a 10% cost then the projected prevalence of DR-TB is more than five times higher at 17 DR-TB cases per 100,000 (mean of acquisition distribution: 0.9; Figure 3b).

***The projected burden of DR-TB is also strongly dependent on the variance of fitness costs of resistance***

When we include a distribution of costs associated with resistance-conferring mutations (e.g. those in Figure 2b), we find that the projected prevalence of DR-TB is dependent on both the mean and variance of this distribution (Figures 3a&b, Supplementary Figure 3).

While this dependence on variation is less evident at the 5-year time horizon, after 20 years, the projected prevalence of resistance is clearly affected by the variance in costs of resistance, especially at intermediate values of the mean fitness cost (Figure 3b). For example, simulations for which we assume a distribution of fitness costs to resistance with a mean fitness cost of 20% (mean of acquisition distribution: 0.8) and a variance of 0.03 produces a 41% higher prevalence at 20 years from drug introduction than simulations in which we assume a constant fixed mean fitness cost of 20% (Figure 3b).

***The mean relative fitness increases over time***

When including a distribution of fitness costs, the mean relative fitness of DR-TB strains circulating in the population increases over time (Supplementary Figure 3b). The rate of increase was faster when there was a higher variance in the distribution of fitness costs to resistance. This means that it is possible that resistance-conferring mutations associated with high average fitness costs may nonetheless lead to high levels of DR-TB when associated variation around this average cost is large. For example, the levels of DR-TB 20 years after drug introduction achieved from an acquisition distribution with a mean fitness cost of 25% (mean of acquisition distribution: 0.75), can exceed the levels from an acquisition distribution with a mean fitness cost of 20% (mean of acquisition distribution: 0.8) when the variance associated with the greater average cost of mutation is higher (blue diamond vs. purple cross, Figure 3b). Worryingly, the prevalence of resistance achieved from an acquisition distribution with a given mean fitness cost with no variance is comparable to the prevalence achieved when the mean is 10% smaller but has a reasonable degree of variance.

***Stochastic effects slightly reduce the expected levels of resistance, but exhibit wide divergence***

As in the deterministic model, the long-term levels of resistance achieved in a stochastic model were dependent on both the mean and variance of the acquisition distribution (Supplementary Figure 4). While the stochastic model allows for chance die out of individual resistant strains resulting in a slightly lower mean projected DR-TB level (Supplementary Figure 5), elimination of resistance is unlikely due to the continued acquisition of resistance during treatment of drug susceptible disease [3]. The stochastic model results are highly divergent and illustrate that while the expected levels of resistance are lower than in the deterministic projections, chance events could also promote even higher levels of resistance (Supplementary Figure 5).

***The dependence of projections in DR-TB on variability of fitness costs is maintained at very different forces of TB infection***

Our results supporting the importance of variability of fitness costs are maintained at both much higher (1000 cases/100,000 population) and lower levels of TB transmission (50 cases/100,000 population) (Supplementary Figures 6&7).

**Discussion**

Improved projections of the spread of drug-resistant tuberculosis (DR-TB) must account for the fact that not all strains of DR-TB have the same epidemic potential. Multiple studies have demonstrated that the same phenotypic resistance can be conferred by different mutations, each of which may be associated with different effects on reproductive fitness [27]. Our results demonstrate that not only the absolute (mean) magnitude of these fitness costs, but also the variation in fitness cost between strains, is an important determinant of future epidemic trajectories of DR-TB. By better understanding the relationship between mutations and fitness costs, we can improve predictions of future levels of DR-TB and facilitate enhanced public health planning efforts. Specifically, with this model framework, we can combine laboratory data on fitness and observational studies on the distribution of fitness costs associated with resistance conferring mutations (such as [28]) with snapshots of fitness from population studies of clinical isolates (such as [29]), to better understand the threat of on-going transmission of resistance. By extending cross-sectional data accordingly, this modelling framework can inform better forecasts of resistance levels and predictions of the impact of interventions for control.

 Here, we developed a new approach to model the effects of variation in fitness costs of drug resistance-conferring mutations on short and longer-term prevalence of DR-TB. Our model suggests that the shape of this distribution of fitness costs is a key contributor to resistance levels over time; by considering such variation we find that the projected burden of DR-TB could be nearly 50% higher after 20 years compared to scenarios in which such variation is ignored. Similar to an earlier model [22], we find that wide distributions in the costs of resistance-conferring mutations allow for increasingly frequent generation of relatively fit resistant strains that can be transmitted and subsequently jeopardize DR-TB control even if the current average DR-TB fitness within a population is low.

More generally, our finding that a wide variance of fitness costs associated with resistance is associated with greater epidemic potential is closely related to Fisher’s fundamental theorem of natural selection which states that “the rate of increase in fitness of any organism at any time is equal to it genetic variance in fitness at that time” [30]. The link between variation in fitness and the rate of change of a fitness-associated trait in a population has been made more formally by Price [31, 32] and previously applied to models of parasite evolution [33].

 Our results suggest that the fraction of resistance mutations that harbour minimal fitness costs (i.e. those in the upper tail of a highly variable fitness cost distribution, approaching the fitness of DS-TB) is an important determinant of the epidemic potential of DR-TB. Once strains with mutations that confer resistance without substantial fitness costs appear and are selected for by ineffective treatment, they will become the preferentially transmitted resistant strains and will contribute to increases in mean fitness of DR-TB over time. DR strains with mutations that confer large fitness costs may also accumulate secondary mutations that compensate or ameliorate these initial fitness costs [19, 27, 34], although we have not considered such effects here. Fitness costs of resistance-conferring mutations are conditional on strain genetic background [8, 35], which could influence the relative prevalence of specific lineages under the selective pressure of TB drug treatment [36, 37]. TTTThese mechanisms suggest that the mean fitness of DR-TB may increase in the long term (>5 years). Hence population-based studies that estimate relative fitness should regard their estimates as specific to a particular moment in time [29]. This increase in fitness may have already occurred in several settings where multidrug-resistant (MDR) strains of TB appear to be readily transmitted [15-18]. This increases urgency for TB control programmes to improve the detection and treatment of MDR- and extensive drug resistance (XDR-) TB [15, 38] to minimize the probability that strains with low cost mutations appear, are selected for, and subsequently spread.

 While our model was designed to investigate the impact of variation in fitness costs of resistance-conferring mutations on short and longer-term trends on DR-TB, there are important determinants of the future burden of DR-TB beyond biological fitness. Most importantly, as TB control programs improve their ability to rapidly detect and effectively treat individuals with DR-TB, the duration of infectiousness with DR-TB strains, and hence the reproductive number of DR-TB, will be reduced. We did not consider such improvements to TB control programs. Furthermore, in the interest of simplicity, we aggregated resistance into a single phenotype in the model, which does not reflect the heterogeneity in resistance patterns observed clinically. In addition, we have not considered host susceptibility factors, such as co-infection with HIV, which have complex and time-varying effects on the incidence of TB and the spread of DR-TB [23]. For these reasons, the projection of trends in DR-TB should not be viewed as quantitative predictions of expected levels of drug resistance in the future. Despite these caveats, our results strongly support the need for additional research to better understand the likelihood of emergence of relatively fit strains of DR-TB, whether these occur through the sporadic appearance of low-cost resistance-conferring mutations or because of the accumulation of compensatory mutations.

This model suggests that even if the mean fitness cost associated with resistance-conferring mutations is large, if a subset of strains have much smaller fitness costs or harbour costly mutations that can subsequently be compensated, these strains will be preferentially transmitted. This process skews the range of resistance mutations observed and suggests that those mutations with the lowest fitness cost should be prioritised for molecular drug resistance tests if the goal of such testing is to provide an early warning for risk of transmitted resistance. It should be emphasized here that due to the dynamic nature of fitness, the most prevalent mutations in the population may not be associated with the smallest fitness cost – this distribution will depend on time since drug introduction.

 In conclusion, we found that, in addition to the mean fitness cost associated with drug resistance appearance, the variance in fitness costs of specific drug resistance-conferring mutations is a key determinant of future trends of DR-TB. Our results are important both to understand the factors affecting the useful lifespan of existing anti-TB drugs, but also for projections about the speed at which we expect to observe resistance to new anti-TB drugs in the development pipeline [39, 40]. Given the importance of the distribution in fitness costs, it would be valuable, though challenging [41], to design additional studies aiming to estimate the ranges of such fitness costs at resistance emergence and in DR-TB populations over time. To guard against the appearance and continued selection of fit drug-resistant strains, further investment to improve the capacity of TB programs to detect and effectively treat individuals with DR-TB is essential.

**Funding**

This work was supported by the Bill and Melinda Gates Foundation (Work Order 10). R.G.W. is additionally funded by the Medical Research Council (UK) (MR/J005088/1), the Bill and Melinda Gates Foundation (Grants 21675/OPP1084276 and 19790.01), CDC/PEPFAR via the Aurum Institute (U2GPS0008111), and USAID/IUTLD/The Union North America (TREAT TB: Technology, Research, Education, and Technical Assistance for Tuberculosis). C.C. is additionally funded by the Engineering and Physical Sciences Research Council (EPSRC EP/K026003/1).

**Acknowledgement**

G.M.K. would like to thank Rein Houben for useful discussions.

References

1. Zignol M, van Gemert W, Falzon D, et al. Surveillance of anti-tuberculosis drug resistance in the world: an updated analysis, 2007-2010. Bulletin of the World Health Organization **2012**; 90(2): 111-9D.

2. WHO. Global Tuberculosis Report, **2014**.

3. Blower SM, Gerberding JL. Understanding, predicting and controlling the emergence of drug-resistant tuberculosis: a theoretical framework. Journal of molecular medicine **1998**; 76(9): 624-36.

4. Blower SM, Small PM, Hopewell PC. Control strategies for tuberculosis epidemics: new models for old problems. Science **1996**; 273(5274): 497-500.

5. Dye C, Williams BG. Criteria for the control of drug-resistant tuberculosis. Proceedings of the National Academy of Sciences of the United States of America **2000**; 97(14): 8180-5.

6. Blower SM, McLean AR, Porco TC, et al. The intrinsic transmission dynamics of tuberculosis epidemics. Nature medicine **1995**; 1(8): 815-21.

7. Cohen T, Dye C, Colijn C, Williams B, Murray M. Mathematical models of the epidemiology and control of drug-resistant TB. Expert review of respiratory medicine **2009**; 3(1): 67-79.

8. Trauner A, Borrell S, Reither K, Gagneux S. Evolution of drug resistance in tuberculosis: recent progress and implications for diagnosis and therapy. Drugs **2014**; 74(10): 1063-72.

9. Gagneux S, Long CD, Small PM, Van T, Schoolnik GK, Bohannan BJ. The competitive cost of antibiotic resistance in *Mycobacterium tuberculosis*. Science **2006**; 312(5782): 1944-6.

10. Mariam DH, Mengistu Y, Hoffner SE, Andersson DI. Effect of *rpoB* mutations conferring rifampin resistance on fitness of *Mycobacterium tuberculosis*. Antimicrobial agents and chemotherapy **2004**; 48(4): 1289-94.

11. Billington OJ, McHugh TD, Gillespie SH. Physiological cost of rifampin resistance induced *in vitro* in *Mycobacterium tuberculosis*. Antimicrobial agents and chemotherapy **1999**; 43(8): 1866-9.

12. Davies AP, Billington OJ, Bannister BA, Weir WR, McHugh TD, Gillespie SH. Comparison of fitness of two isolates of *Mycobacterium tuberculosis*, one of which had developed multi-drug resistance during the course of treatment. The Journal of infection **2000**; 41(2): 184-7.

13. Morlock GP, Plikaytis BB, Crawford JT. Characterization of spontaneous, in vitro-selected, rifampin-resistant mutants of *Mycobacterium tuberculosis* strain H37Rv. Antimicrobial agents and chemotherapy **2000**; 44(12): 3298-301.

14. Luciani F, Sisson SA, Jiang H, Francis AR, Tanaka MM. The epidemiological fitness cost of drug resistance in *Mycobacterium tuberculosis*. Proceedings of the National Academy of Sciences of the United States of America **2009**; 106(34): 14711-5.

15. Ioerger TR, Feng Y, Chen X, et al. The non-clonality of drug resistance in Beijing-genotype isolates of *Mycobacterium tuberculosis* from the Western Cape of South Africa. BMC genomics **2010**; 11: 670.

16. Lanzas F, Karakousis PC, Sacchettini JC, Ioerger TR. Multidrug-resistant tuberculosis in Panama is driven by clonal expansion of a multidrug-resistant *Mycobacterium tuberculosis* strain related to the KZN extensively drug-resistant *M. tuberculosis* strain from South Africa. Journal of clinical microbiology **2013**; 51(10): 3277-85.

17. Marais BJ, Mlambo CK, Rastogi N, et al. Epidemic spread of multidrug-resistant tuberculosis in Johannesburg, South Africa. Journal of clinical microbiology **2013**; 51(6): 1818-25.

18. Casali N, Nikolayevskyy V, Balabanova Y, et al. Evolution and transmission of drug-resistant tuberculosis in a Russian population. Nature genetics **2014**; 46(3): 279-86.

19. Comas I, Borrell S, Roetzer A, et al. Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes. Nature genetics **2012**; 44(1): 106-10.

20. Shrestha S, Knight GM, Fofana M, et al. Drivers and Trajectories of Resistance to New First-Line Drug Regimens for Tuberculosis. Open Forum for Infectious Diseases **2014**; 1(2).

21. Dye C, Williams BG, Espinal MA, Raviglione MC. Erasing the world's slow stain: strategies to beat multidrug-resistant tuberculosis. Science **2002**; 295(5562): 2042-6.

22. Cohen T, Murray M. Modeling epidemics of multidrug-resistant *M. tuberculosis* of heterogeneous fitness. Nature medicine **2004**; 10(10): 1117-21.

23. Sergeev R, Colijn C, Murray M, Cohen T. Modeling the dynamic relationship between HIV and the risk of drug-resistant tuberculosis. Science translational medicine **2012**; 4(135): 135ra67.

24. Dye C, Garnett GP, Sleeman K, Williams BG. Prospects for worldwide tuberculosis control under the WHO DOTS strategy. Directly observed short-course therapy. Lancet **1998**; 352(9144): 1886-91.

25. Vynnycky E, Fine PE. The natural history of tuberculosis: the implications of age-dependent risks of disease and the role of reinfection. Epidemiology and infection **1997**; 119(2): 183-201.

26. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing ([http://www.r-project.org/)](http://www.r-project.org/%29), **2005**.

27. Andersson DI, Hughes D. Antibiotic resistance and its cost: is it possible to reverse resistance? Nature reviews Microbiology **2010**; 8(4): 260-71.

28. Salvatore PP, Becerra MC, Zur Wiesch PA, et al. Fitness Costs of Drug-resistance Mutations in Multidrug Resistant M. tuberculosis: A Household-based Case-control Study. The Journal of infectious diseases **2015**.

29. Grandjean L, Gilman RH, Martin L, et al. Transmission of Multidrug-Resistant and Drug-Susceptible Tuberculosis within Households: A Prospective Cohort Study. PLoS medicine **2015**; 12(6): e1001843.

30. Fisher R. The genetical theory of natural selection. Oxford: Clarendon Press, **1930**.

31. Price GR. Extension of covariance selection mathematics. Annals of human genetics **1972**; 35(4): 485-90.

32. Price GR. Selection and covariance. Nature **1970**; 227(5257): 520-1.

33. Gandon S, Day T. Evolutionary epidemiology and the dynamics of adaptation. Evolution; international journal of organic evolution **2009**; 63(4): 826-38.

34. de Vos M, Muller B, Borrell S, et al. Putative compensatory mutations in the rpoC gene of rifampin-resistant *Mycobacterium tuberculosis* are associated with ongoing transmission. Antimicrobial agents and chemotherapy **2013**; 57(2): 827-32.

35. Fenner L, Egger M, Bodmer T, et al. Effect of mutation and genetic background on drug resistance in *Mycobacterium tuberculosis*. Antimicrobial agents and chemotherapy **2012**; 56(6): 3047-53.

36. Ford CB, Lin PL, Chase MR, et al. Use of whole genome sequencing to estimate the mutation rate of *Mycobacterium tuberculosis* during latent infection. Nature genetics **2011**; 43(5): 482-6.

37. Hanekom M, Gey van Pittius NC, McEvoy C, Victor TC, Van Helden PD, Warren RM. *Mycobacterium tuberculosis* Beijing genotype: a template for success. Tuberculosis (Edinb) **2011**; 91(6): 510-23.

38. Falzon D, Jaramillo E, Wares F, Zignol M, Floyd K, Raviglione MC. Universal access to care for multidrug-resistant tuberculosis: an analysis of surveillance data. The Lancet infectious diseases **2013**; 13(8): 690-7.

39. Zumla A, Nahid P, Cole ST. Advances in the development of new tuberculosis drugs and treatment regimens. Nature reviews Drug discovery **2013**; 12(5): 388-404.

40. WHO. Antimicrobial Resistance: Global Report on Surveillance, **2014**.

41. Cohen T, Sommers B, Murray M. The effect of drug resistance on the fitness of *Mycobacterium tuberculosis*. The Lancet infectious diseases **2003**; 3(1): 13-21.

Table 1: Data on genetic background and mutation, acquisition probability, experimental condition and relative reproductive fitness of *in vitro,* spontaneously acquired rifampicin resistance mutations in *M. tuberculosis* strains.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Strain** **[# colonies selected]**  | **# unique mutations** | **Mutation in *rpoB*** | **Acquisition probability** | **Experimental condition\*\*** | **Relative fitness** | **Notes** | **Reference** |
| Harlingen strain[27] | 3 | S531W | 0.12 | Competition against parental | 0.67 (0.61-0.73) | 1/3 of spontaneous resistances had mutations not in *rpoB* | [10] |
| H526Y | 0.65 | 0.89 (0.84-0.94) |
| S522L | 0.23 | 0.54 (0.51-0.57) |
| “ | “ | Independent | 0.71 (0.62-0.80) |
| 0.86 (0.83-0.89) |
| 0.95 (0.93-0.97) |
| “ | “ | In macrophages | 0.28 (0.22-0.34) |
| 0.63 (0.61-0.65) |
| 0.50 (0.34-0.66) |
| CDC1551 [52]  | 12  | S531L | 0.31 | Competition against parental | 0.91 (0.86-0.97) |  | [9] |
| H526Y | 0.19 | 0.82 (0.75-0.89) |
| H526D | 0.04 | 0.78 (0.73-0.82) |
| S531W | 0.02 | 0.88 (0.78-0.88) |
| H526R | 0.19 | 0.82 (0.75-0.88) |
| S522L | 0.15 | 0.88 (0.80-0.96) |
| Q513L | 0.04 | 0.83 (0.79-0.86) |
| H526P | 0.04 | 0.84 (0.8-0.89) |
| R529Q | 0.02 | 0.58 (0.55-0.61) |
| T85 [63] | 7\* | S531L | 0.46 | Competition against parental | 0.96 (0.93-0.99) |  |
| H526Y | 0.33 | 0.81 (0.78-0.84) |
| H526D | 0.13 | 0.85 (0.82-0.88) |
| S531W | 0.08 | 0.79 (0.75-0.82) |

\* Only four were included in the fitness analysis. \*\* This refers to the experimental condition under which relative fitness was determined.

**Figure legends**

**Figure 1:** **Model outline.** The TB transmission model consists of individuals who are not infected with *Mycobacterium tuberculosis* (*M. tb*) [*M.tb* un-infected] and those who are infected with *M. tb* and in a latent state (non-infectious) [Latent TB infection] or active state (infectious) [Active TB infection]. Drug-resistant *M. tb* appears first through acquired resistance among those with active, drug-susceptible disease. Drug resistant *M. tb* can subsequently be transmitted. The relative transmission potential of strains is dependent on the number of individuals with active disease and the mean fitness of the circulating strains. All susceptible strains have a mean fitness of 1, whilst the resistant strains have a range of relative fitness levels. The changes in this distribution of relative fitness levels are tracked in those with active diseases (as shown) and latent infection (not shown).

**Figure 2: Acquisition distribution.** Distributions of fitness costs associated with new resistance mutations from experiments (a) and generalized via beta distributions as input for the models (b). (a) Distributions of fitness costs from pooled data from Gagneux et al (2006) and Mariam (2004) show that each different mutation within the *rpoB* gene conferring resistance to rifampicin has a different relative fitness level as measured by competitive co-culture with a parental, susceptible strain. The different mutations (shown in different colours) are labeled as original amino acid, codon position of mutation and subsequent new amino acid. The frequency is taken from the number of *in vitro* spontaneous resistance mutations found to have this mutation (Table 1). (b) Examples of the distributions of fitness levels for new mutations that were used as inputs for the deterministic model. Here the examples have three mean fitness values (0.5, 0.75 and 0.9) each with several levels of variance. For example, the two curves with the lowest peaks have means 0.5 and 0.75 have a variance of 0.03. Note that these are beta distributions, which are capped at 1 and have an area under the curve capped at 1.

**Figure 3: Model results.** Fitness distributions with higher variance are associated with higher levels of resistance at 5 (a) and 20 years (b), from the time of drug introduction in the deterministic model. When different levels of variance are included (shape scale), the prevalence of resistance is higher (e.g. compare open circles to crosses at a single mean). This effect is more easily appreciated at the 20-year time point (compare results at 5 (a) and 20 years (b)).