Published in final edited form as:
J Theor Biol. 2011 July 7; 280(1): 88–100. doi:10.1016/j.jtbi.2011.04.011.

Models to understand the population-level impact of mixed strain *M. tuberculosis* infections

Rinat Sergeev[a,b,*], **Caroline Colijn**[c], and **Ted Cohen**[a,d]

[a] Division of Global Health Equity, Brigham and Women's Hospital, 641 Huntington Ave, 02115, Boston, MA, USA

[b] Department of Microelectronics, Ioffe Institute, 26 Polytekhnicheskaya, St Petersburg 194021, Russia

[c] Department of Engineering Mathematics, University of Bristol, University Walk, Bristol, BS8 1TR, United Kingdom

[d] Department of Epidemiology, Harvard School of Public Health, 641 Huntington Ave, 02115, Boston, MA, USA

Abstract

Over the past decade, numerous studies have identified tuberculosis patients in whom more than one distinct strain of *M. tuberculosis* is present. While it has been shown that these mixed strain infections can reduce the probability of treatment success for individuals simultaneously harboring both drug-sensitive and drug-resistant strains, it is not yet known if and how this phenomenon impacts the long-term dynamics for tuberculosis within communities. Strain-specific differences in immunogenicity and associations with drug resistance suggest that a better understanding of how strains compete within hosts will be necessary to project the effects of mixed strain infections on the future burden of drug-sensitive and drug-resistant tuberculosis. In this paper, we develop a modeling framework that allows us to investigate mechanisms of strain competition within hosts and to assess the long-term effects of such competition on the ecology of strains in a population. These models permit us to systematically evaluate the importance of unknown parameters and to suggest priority areas for future experimental research. Despite the current scarcity of data to inform the values of several model parameters, we are able to draw important qualitative conclusions from this work. We find that mixed strain infections may promote the coexistence of drug-sensitive and drug-resistant strains in two ways. First, mixed strain infections allow a strain with a lower basic reproductive number to persist in a population where it would otherwise be outcompeted if has competitive advantages within a co-infected host. Second, some individuals progressing to phenotypically drug-sensitive tuberculosis from a state of mixed drug-sensitive and drug-resistant infection may retain small subpopulations of drug-resistant bacteria that can flourish once the host is treated with antibiotics. We propose that these types of mixed infections, by increasing the ability of low fitness drug-resistant strains to persist, may provide opportunities for compensatory mutations to accumulate and for relatively fit, highly drug-resistant strains of *M. tuberculosis* to emerge.

[*]Corresponding author. Tel.: +1 617 432 1022; fax: +1 617 432 2565, rsergeev@partners.org (R. Sergeev), ccolijn@gmail.com (C. Colijn), tcohen@hsph.harvard.edu (T. Cohen).

Keywords

Epidemiology; Mycobacterium; Mathematical model; Coinfection; Mixed; Infection; Antibiotic resistance

Introduction

Tuberculosis (TB), an infectious disease caused by respiratory transmission of *Mycobacterium tuberculosis* (*M. tb*), was responsible for over 9 million new cases and nearly 2 million deaths in 2008 alone [1]. The global approach for TB control is based on the identification of individuals presenting for medical care with suggestive symptoms (e.g. cough, weight loss), diagnosis by microscopic examination of sputum, and standardized drug regimens that include multiple antibiotics taken for at least six months. While this disease management approach has improved individual patient outcomes and disease control in many settings, the emergence of HIV and the appearance of drug-resistant forms of TB compromise the effectiveness of this strategy and pose additional challenges for TB control [2].

Previous infection and/or disease with TB provides, at best, partial immunity to reinfection with *M. tuberculosis* [3–5]. New molecular methods for strain differentiation have revealed that individuals can be reinfected and harbor more than one distinguishable strain of *M. tb* [6–16]; we will refer to these infections as “mixed infections”. The measured fraction of tuberculosis patients with mixed infections varies markedly between settings; in some studies nearly 1/5 of patients had evidence of multiple strains at the time of diagnosis [9] while in other studies this was a rare finding [14]. Much of this heterogeneity probably reflects true differences between the frequencies of this event in different communities. Since the risk of reinfection (a pre-requisite for a mixed strain infection) is a function of the local prevalence of infectious source cases, mixed infection will be most common in settings with a high disease burden. However, reinfection has also been shown to play an important role in disease dynamics in moderate- [17–18] and low-incidence communities [19]; this may reflect the increased importance of local clustering of disease among household and close social contacts in these lower incidence settings [20]. Heterogeneity in the frequency of detection of mixed strain infections between existing studies may also reflect differences in study participant selection, methods of specimen collection and storage [21], and in approaches for laboratory identification of multiple strains [22]. Furthermore, we note that even those studies with the highest estimates of patients with mixed infections likely underestimate the frequency of this phenomenon since the current methods of detection are insensitive and unable to assess the frequency of mixed infections during the latent stage of infection.

The deleterious effect of mixed strain infections has been shown for individuals harboring both drug-sensitive (DS) and drug-resistant (DR) strains of TB. Van Rie *et al* [11] found that these individuals may respond poorly to treatment; when first-line drugs were administered, the overgrowth of DR strains was observed and when second-line drugs were given to suppress the DR strain, the DS strain was able to reemerge since second line drugs are less effective than the best first-line agents.

While these types of mixed infections compromise the treatment of individuals, the effect of mixed infection on the transmission dynamics of TB in the community is less clear. We propose that it is important to consider what impact mixed infections may have on the long-term distribution of strains present in communities. While early sequencing data suggested limited diversity within the *M. tuberculosis* species [23], new analyses have resulted in an

emerging consensus of a robust phylogenetic structure within the *M. tuberculosis* and meaningful biological differences between strain lineages [24–27]. For example, differences between immunogenic properties of strains [24, 28–30] have been demonstrated repeatedly and there are associations between some strain lineages and drug-resistance in some geographic settings [31–33]. The observation that mixed strain infections do occur frequently in some settings, combined with evidence of different abilities of strains to evade and provoke host immune responses and strain associations with drug resistance, suggest that a better understanding of the population-level impact of these types of complex infections is needed.

Here we describe a modeling framework to evaluate the potential importance of mixed infections on the strain diversity in a population. In this paper, we are interested in how mixed infections may promote the stable coexistence of different strains within populations. In particular, we investigate how the phenomenon of mixed infection may protect strains with lower reproductive numbers from being outcompeted and eliminated. Since our goal is to understand stable coexistence, our analysis focuses on the equilibrium states associated with our model system.

While previous models include competition of the strains within hosts with mixed infections [34], the specific mechanisms of this competition are not well understood. Accordingly, we structured our model to permit the systematic investigation of several prototypical mechanisms by which distinct strains may interact. While strains may vary in other important ways (e.g. strain lineage [24–27]), in this paper, we focus our investigation on mixed infections where the infecting strains differ in terms of their drug susceptibility.

The paper is structured as follows. In Section 1, we present our model framework and describe and classify the mechanisms of within-host competition we study. In Section 2, we analyze the steady states of the model and evaluate how the stable coexistence of DR and DS TB depends upon: 1) the presence of the mixed infection state in the model; 2) the specific mechanisms of within-host competition between strains; and 3) the selective pressure associated with drug treatment. Finally, in Section 3, we discuss the implications of these results, outline important assumptions and limitations of the model, and suggest further studies that can improve our understanding of how mixed infections affect the epidemiology and control of TB.

1 Methods

1.1 Model structure and parameterization

We developed a differential equation model, similar to the model used in [35], for assessing the impact of mixed infections on the stable coexistence of strains. The model includes three general classes of health and disease: Susceptible, Latent infection, and Infectious (Figure 1, see also Appendix A for the full set of equations).

Individuals introduced into the population are susceptible to infection by any circulating strain of *M. tuberculosis* (*S*); this compartment is replenished (e.g. birth or immigration) to maintain a constant population size. The natural history of tuberculosis includes a latent period during which infected persons are completely asymptomatic; the majority (~90%) of immunocompetent (i.e. not HIV-infected) latently infected individuals will never progress to disease. We include compartments representing latent infection with either DS (L_S) or DR (L_R) forms of *M. tb*. In contrast to most other TB models (see, for example, [36–38]), we explicitly represent only the minority of latent infections that will result in disease during the lifetime of an infected host (i.e. productive infections). If those in either the L_S or L_R state are productively infected by the other strain type before progressing to the infectious state,

they enter the Mixed Latent class (M) from which they can progress to any one of the three infectious states: I_S = infectious and only harboring the DS strain; I_{SR} = infectious and harboring a majority of the DS strain with a small subpopulation of DR bacteria; and I_R = infectious and harboring predominantly or only DR strain.

The probabilities of progression from the M state to each of the three infectious states depend on the specific mechanism of within-host competition (described below in 1.2). Since the infectious dose for transmission of TB is very small (perhaps as little as a single bacteria [39]), we assume that infectious individuals transmit only the predominant strain in their infection. Consequently, both individuals in the I_S and I_{SR} states transmit DS TB, while those in the I_R state transmit DR TB. While treatment cures the large majority of those in the I_S state, a small fraction will acquire drug resistance as the result of inadequate treatment. Drug treatment of those in the I_{SR} state results in the rapid selection of the resistant subpopulation resulting in a transition to the I_R state. For simplicity, we assume that individuals cannot have more than one episode of active TB disease and that standard treatment does not cure drug resistant disease (a more complex model is provided in Appendix C, but qualitative results are similar to those presented here).

A list of model parameters and variables is provided in Table 1. To highlight important differences between types of parameters and variables, we present them in three categories: 1) parameters that must be defined (Table 1a); 2) variables that describe characteristics of the epidemiological state of the system (Table 1b); and 3) combined parameters that can be calculated based on other parameters and variables (Table 1c).

Since our focus involves analyses of the outcome of inter-strain competition, we used a variant of a reproductive number to efficiently summarize the reproductive capacity of each strain in the absence of the other (R_S and R_R for the drug-sensitive and drug-resistant strain, respectively). In this model we consider the reproductive number of the strains in the presence of antibiotics to highlight differences between the strains in the presence of interventions. Classical competition models that do not include mixed infected states do not permit stable co-existence unless the strains have equivalent reproductive numbers [40]. The five other parameters defined in Table 1a (φ, f_D, $\theta\delta$, $\eta\delta$, and r/π) describe the mechanisms and strength of competition between the strains. This approach allows us to consider how the presence of these interactions affects the abilities of the strains to coexist even when their reproductive numbers are not similar.

1.2 Modeling within-host competition between DR and DS strains

Within-host competition between DR and DS strains is introduced into the model depicted in Figure 1 through the set of the probabilities (f_S, f_{SR}, f) that determines the risk of progression from the mixed infection class M to I_S, I_{SR}, or I_R, respectively. We express these probabilities by specifying the chance that a particular strain "wins" the within-host competition (f for DR and 1-f for DS), and the chance for a subdominant strain to remain within the Infectious host (θ) as follows:

$$\begin{aligned} f_S &= (1-\theta)(1-f) \\ f_{SR} &= \theta(1-f). \end{aligned} \tag{1}$$

The probabilities f and θ cannot be directly measured since the type of infections present within a host cannot be determined during latency using existing diagnostic tools. However, the properties that govern the probability f can be categorized into two distinct classes: 1) strain-independent properties (i.e. factors that apply equally to either strain and are related to either the order in which infections occurred, the time delay between the infections, or the

anatomic location in the lung in which the infection was seeded) and 2) strain-dependent properties (i.e. factors that reflect inherent differences between the strains' ability to expand within a host and evade immune responses). We introduce the overall probability f as the weighted average of two probabilities, f_N and f_D, representing these strain-independent and strain-dependent mechanisms, respectively:

$$f=(1-\varphi)f_N+\varphi f_D. \tag{2}$$

The weighting factor $\varphi \in [0,1]$ expresses the relative importance of these factors to the outcome of the competition.

Under the strong assumption that the competing strains elicit similar immunological host responses and are identical except for their response to antibiotics (i.e. $\varphi=0$), the overall probability f is equal to f_N. That is, the outcome of the competition is determined only by the number of reinfection events and the timing and anatomic location in which infections occur. Since the frequency of infection with each strain type is dependent on the prevalence of infectious source cases with that type of disease, this probability varies with the epidemiological state of the system. As we show in detail in Appendix B, at the equilibrium f_N is related to the variables v (i.e. the fraction of the total infection force due to the DR strain) and n_λ (i.e. the expected number of reinfection events occurring during latency) by a simple expression:

$$f_N=\frac{1+vn_\lambda}{2+n_\lambda}. \tag{3}$$

In contrast, if we assume that inherent differences between the strains is the sole determinant of the outcome of within-host competition (i.e. $\varphi=1$), the overall probability f is equal to f_D. In this case, the outcome of strain competition depends only on the ability of DR strain to compete with a DS strain within a host, and not, as above, on the epidemiological state of the system.

In reality, we expect the outcome of within-host competition of strains to depend on both strain-independent and strain-dependent factors (i.e. $0<\varphi<1$). The overall probability f in this case is a dynamic variable, described in equilibrium by the following expression:

$$f=(1-\varphi)\frac{1+vn_\lambda}{2+n_\lambda}+\varphi f_D. \tag{4}$$

2 Results

2.1 Steady state analysis

The set of differential equations, describing the dynamics of the seven (of eight) variables, listed in Table 1b, is derived in Appendix A-1 (see equation (A6)). Together with the equation (4) for the variable f, it forms the full set of equations (A7). Three elementary equilibria (disease free: ***0;** exclusively DS strain: ***S**; and exclusively DR strain: ***R**) are presented in Appendix A-2 as the time-independent solutions of (A7). Here we focus on the conditions when a fourth equilibrium, designated by *****, is possible: the coexistence of DS and DR strains.

The equations (A7) cannot be completely resolved analytically for the ***** equilibrium (see equations (A15–A16) of Appendix A-2 for details). Nevertheless, they allow us to determine combinations of strain reproductive numbers (R_S and R_R) that allow for coexistence (i.e. when the equilibrium value of the fraction of the total force of infectious due the DR strain is between 0 and 1: $0<v^*<1$).

Thus, analysis of the limit $v^*\to 1$ in equations (A16) yields the condition for R_S that permits the DS strain to invade the population already saturated by a DR strain:

$$\begin{gathered} R_S > \frac{R_R}{1+n_\lambda^{*R}\varphi\left(1-\frac{2+n_\lambda^{*R}}{1+n_\lambda^{*R}}f_D\right)} \\ R_R \geq 1 \\ n_\lambda^{*R} \equiv \frac{r}{\pi}(R_R-1). \end{gathered} \tag{5}$$

Here n_λ^{*R} is the value of n_λ at ***R** (equilibrium with only DR TB) and depends only on the reproductive number of the DR strain (equation (A12) of Appendix A-2).

In the absence of acquired resistance ($\eta\delta=0$), analysis of the limit $v^*\to 0$ in (A16) yields the similar condition for a DR strain to invade a population in which the DS strain exists alone at steady state:

$$\begin{gathered} R_R > \frac{R_S}{1+n_\lambda^{*S}\left((1-\theta\delta)\varphi\left(\frac{2+n_\lambda^{*S}}{1+n_\lambda^{*S}}f_D-\frac{1}{1+n_\lambda^{*S}}\right)+\theta\delta\right)} \\ R_S \geq 1 \\ n_\lambda^{*S} \equiv \frac{r}{\pi}(R_S-1). \end{gathered} \tag{6}$$

Here n_λ^{*S} is the value of n_λ in ***S** (equilibrium with only DS TB) and depends only on the reproductive number of the DS strain (equation (A9) of Appendix A-2). If we allow acquired resistance ($\eta\delta>0$), the DR strain will always be present at some frequency when DS TB is circulating and treatment is available ($v^*>0$, see Appendix A-2). Nevertheless, even in this case, condition (6) allows us to determine the parametric area where the transmission of DR strain is self-sustaining.

The inequalities (5–6) allow us to determine the conditions under which coexistence is expected for strains with different reproductive numbers. If all the inequalities in (5–6) are simultaneously satisfied, the two strains can coexist at equilibrium ($0<v^*<1$). Conversely, if, for a particular set of the parameters ($\theta\delta$, φ, f_D, π/r), inequalities (5–6) cannot be resolved with respect to R_S and R_R, stable coexistence is not possible.

The substitution of equalities into the first lines of (5) and (6) allows us to examine the conditions when the coexistence area degenerates. We found only two cases when this occurs: when there is no disease at all (a trivial case) and when the strains are completely indistinguishable:

$$\begin{cases} R_S=R_R=1 & \textit{disease free, or} \\ \theta\delta=\varphi=0 \\ R_S=R_R & \textit{indistinguishable strains} \end{cases}. \tag{7}$$

Both these cases represent the extreme edges of the parametric space and, at reasonable values of the parameters, stable coexistence is always possible. In the other words, for every $R_S>1$ and when strain-dependent differences ($\varphi>0$) or drug treatment ($\theta\delta>0$) is present, there is a finite range of R_R over which DR and DS strains can coexist.

2.2 The coexistence of DS and DR strains

By examining inequalities (5–6), we investigate the influence of different factors on the shape of the coexistence area.

Mixed infection—As our defined state of mixed-strain infection can occur only after at least two independent infection events, the frequency of mixed-infections should depend on the number of reinfection events. The expected number of reinfections by each strain, according to the expressions for the single-strain equilibria (n_λ^{*R} in (5) and n_λ^{*S} in (6)), is strongly dependent on the strain reproductive numbers. Thus, the overall effect of the state of mixed infection is expected to increase in settings where the reproductive numbers of each of the strains is large. This effect is clearly demonstrated when we simplify the conditions (5–6) for the symmetrical case, substituting $f_D=1/2$ (i.e. each strain equally likely to win the within-host competition) and $\theta=0$ (i.e. no remaining subpopulation of DR strains once a host progresses from mixed infection to DS disease). For the criteria of coexistence, $0<v^*<1$, we find:

$$\begin{aligned} R_R &> \frac{1}{1+\varphi(n_\lambda^{*S})^2/2(1+n_\lambda^S)} R_S \\ R_S &> \frac{1}{1+\varphi(n_\lambda^{*R})^2/2(1+n_\lambda^R)} R_R \\ \min(R_S, R_R) &> 1. \end{aligned} \tag{8}$$

Expanding the expressions in (8) at the limit of small prevalences (R_S, $R_R\to1$) we find that for any value of R_S there is a finite range of R_R that allows for coexistence. This range is a square function of the number of reinfections and also has a strong dependence on the mechanism of competition (φ) (see equations (A19–A21) in Appendix A-3 for more details):

$$|R_R - R_S| < \tfrac{\varphi}{2}(n_\lambda^{*S})^2 = \tfrac{\varphi}{2}\tfrac{r^2}{\pi^2}(R_S - 1)^2. \tag{9}$$

The Figure 2a shows the results of the numerical solution of (8) for arbitrary reproductive number values, and demonstrates how the width of the coexistence area increases with the reproductive abilities of the strains.

Within-host competition—Equation (9) shows that in the absence of the drug treatment ($\delta=0$), the width of the coexistence area is linear with φ. The coexistence area is smallest if the competition of the strains within the host is determined by strain-independent factors and does not include strain-specific differences ($\varphi=0$) and largest when inherent differences between strains entirely determine the outcome of competition ($\varphi=1$).

The role of strain-specific competitive ability (f_D) can be demonstrated if we simplify the equilibrium equations at the low-prevalence limit ($R_S-1\ll1$), where we assume that strain competition is entirely strain-dependent ($\varphi=1$), and there is no drug treatment ($\delta=0$). If the competitive abilities of the strains are not equal ($f_D\neq0.5$), the range of R_R that allows for coexistence is a linear function of the number of reinfections and is shifted proportionally to ($1–2\ f_D$) (see equations (A22–A24) in Appendix A-3 for more details):

$$R_S+(1-2f_D)n_\lambda^{*S}<R_R<R_S+\frac{(1-2f_D)}{1-\frac{r}{\pi}(1-2f_D)}n_\lambda^{*S}. \tag{10}$$

Equation (10) shows that a strain may have a strong enough within-host competitive ability to allow it to stably persist in populations even when it has a lower reproductive number than the competing strain. As shown in Figure 2b, this within-host competitive ability can have a particularly pronounced effect in settings where reinfections occur frequently.

Drug treatment—Individuals progressing to DS disease from mixed infection may retain small subpopulations of DR bacteria. Among those individuals (I_{SR}) drug treatment may rapidly select for resistance. This effect is represented in the equilibrium equations by the parameter combination $\theta\delta$. This mechanism provides the DR strain an opportunity to dominate in populations even in the absence of strain-dependent differences in within-host competition. This selective effect of drugs can be demonstrated most clearly in the case of strain-independent competition, (i.e. if we substitute φ=0 into (5–6)), where, according to the condition (7), we do not expect coexistence of the strains in the absence of drug treatment (i.e. when δ=0). The presence of drugs (δ>0) alters the balance and allows for coexistence (see equations (A25–A29) in Appendix A-3 for more details):

$$\frac{1}{1+n_\lambda^{*S}\theta\delta}<R_R/R_S<1. \tag{11}$$

Figure 2c shows that drug treatment makes coexistence possible at lower DR strain relative fitness than is possible without treatment. This effect is more pronounced where the expected number of reinfections is larger. We emphasize that this result describes only the impact of drug treatment on those progressing to disease from the mixed infection class. In reality, drug treatment will also affect the reproductive number of the drug sensitive strain by limiting the expected duration of infectivity, which favors the DR strain as well. Thus, the total competitive benefit which drug treatment provides for the DR strain is the sum of these two effects.

3 Discussion

While recent studies reveal that TB patients may simultaneously harbor multiple distinct strains of TB [6–16] and find that these complex infections may complicate treatment and result in worse outcomes for patients infected by both DR and DS strains [11], it is not yet clear what impact these mixed infections have on the transmission dynamics or control of disease within communities. The fact that even relatively crude methods of strain differentiation have detected that in some high incidence settings nearly 1/5 of incident TB cases may have multiple strains suggests that the population-level impact of mixed infections requires additional investigation. In this manuscript, we develop models to assess the effects of mixed infections on strain ecology and disease dynamics. The model provides a systematic framework for considering the mechanisms by which strains compete within an individual-host, investigating how mixed infections can affect the dynamics of drug-resistant TB, and identifying priority questions for future research.

Our models demonstrate how mixed strain infections may promote the long-term coexistence of strains within a host population. That is, models that do *not* allow for individual hosts to simultaneously harbor multiple strains exhibit competitive exclusion of strains under a wider range of parameter values than models that allow for mixed strain

infections [35]. This is similar to the extinction of less competitive species in Lotka-Volterra systems if inter-specific (between species) competition is equal or greater than intra-specific (within species) one [40–41]. This means that models where hosts can harbor mixed infections allow the strains with lower basic reproductive numbers to persist in conditions where they may otherwise be eliminated.

With respect to TB, we find that mixed strain infections may increase the opportunity for DR strains to persist in areas where they might otherwise be outcompeted [42]. Once DR strains are present within populations, our models find that, even in the absence of antibiotic pressure, DR strains with a relatively low basic reproductive number (compared with DS strains) may nonetheless persist, especially if they are relatively successful competitors within coinfected hosts. The long-term persistence of these DR strains is worrisome since this persistence allows for opportunity for evolution [43]. Previous research has shown that many TB drug-resistance conferring mutations result in substantial fitness costs [44], but that these initial costs may be compensated by secondary mutations [45–50]. Accordingly, mixed infections, by allowing for longer persistence of low fitness strains, provide the bacteria with a better chance to accumulate these additional mutations.

Anti-TB antibiotics further increase the competitive ability of DR strains (i.e. lowers the DR invasion threshold) in several ways. First, imprudent antibiotic use may select for resistant strains in those with DS infections (i.e. acquired drug resistance). This mechanism is how DR initially arises in most settings, although importation of DR strains is also possible. Second, antibiotics selectively remove DS competitors and increase the relative reproductive number of the DR strain. We note that neither of these first two mechanisms depends on the presence of mixed strain infections. Mixed strain infections offer a novel third mechanism by which antibiotics facilitate the emergence of DR strains. Since hosts with apparent DS disease may have actually progressed from a mixed infection and retained a small subpopulation of DR mutants, antibiotics can act to select previously outcompeted DR bacteria. This is essentially the phenomenon that van Rie *et al* describe [11]. In addition to offering another route by which DR can emerge, this third mechanism is problematic for the evaluation of control programs because we will misclassify cases of transmitted resistance as acquired drug resistance. This misclassification will make treatment programs appear worse (i.e. overestimate rates of acquired resistance) and make it difficult to judge the performance of existing interventions and adequately plan new strategies.

As with all models, we make important simplifying assumptions about the natural history of disease. In Appendix C we present models with more complex representations of the natural history of TB; we find that our qualitative results apply to this more complicated models and we describe how each of these alternative models can be reduced to the model depicted in Figure 1 by redefining of parameters.

While we can gain some broad qualitative insight into the potential population-level impact of mixed strain infections through this modeling framework, our ignorance of important parameter values limits the conclusions that we can currently make. In particular, there are few data that can currently be used to understand how different strains of TB may compete within a single host. There are intriguing data to suggest that some strain lineages provoke different levels and types of host immune response (i.e. Beijing-lineage strains [24, 28–30]) and that these strains may, in some circumstances, have an increased likelihood of having a drug resistant phenotype [31–33]. Experiments in which animals were simultaneously and/or sequentially infected with different strain lineages would greatly improve our understanding of these mechanisms of within-host competition.

Acknowledgments

The work is supported by NIH grants DP2OD006663 and U54GM088558. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Office of the Director of the National Institute of Health and the National Institute of General Medical Sciences of the National Institutes of Health.

References

1. Global tuberculosis control: a short update to the 2009 report. World Health Organization; 2009.
2. Multidrug and extensively drug-resistant TB (M/XDR-TB): 2010 global report on surveillance and response. World Health Organization; 2010.
3. Lambert ML, Hasker E, van Deun A, Roberfroid D, Boelaert M, van der Stuyft P. Recurrence in tuberculosis: relapse or reinfection? Lancet Infectious Diseases. 2003; 3:282–7. [PubMed: 12726976]
4. Gomes MGM, Franco AO, Gomes MC, Medley GF. The reinfection threshold promotes variability in tuberculosis epidemiology and vaccine efficacy. Proceedings of the Royal Society B. 2004; 271:617–23. [PubMed: 15156920]
5. Chiang CY, Riley LW. Exogenous reinfection in tuberculosis. Lancet Infectious Diseases. 2005; 5:629–36. [PubMed: 16183517]
6. Niemann S, Richter E, Rusch-Gerdes S, Schlaak M, Greinert U. Double infection with a resistant and a multidrug-resistant strain of Mycobacterium tuberculosis. Emerging Infectious Diseases. 2000; 6:548–51. [PubMed: 10998389]
7. Richardson M, Carroll NM, Engelke E, van der Spuy GD, Salker F, Munch Z, Gie RP, Warren RM, Beyers N, van Helden PD. Multiple Mycobacterium tuberculosis strains in early cultures from patients in a high-incidence community setting. Journal of Clinical Microbiology. 2002; 40:2750–4. [PubMed: 12149324]
8. Shamputa IC, Rigouts L, Eyongeta LA, Aila NAE, van Deun A, Salim AH, Willery E, Locht C, Supply P, Portaels F. Genotypic and phenotypic heterogeneity among Mycobacterium tuberculosis isolates from pulmonary tuberculosis patients. Journal of Clinical Microbiology. 2004; 42:5528–36. [PubMed: 15583277]
9. Warren RM, Victor TC, Streicher EM, Richardson M, Beyers N, van Pittius NG, van Helden PD. Patients with active tuberculosis often have different strains in the same sputum specimen. American Journal of Respiratory and Critical Care Medicine. 2004; 169:610–614. [PubMed: 14701710]
10. Baldeviano-Vidalon GC, Quispe-Torres N, Bonilla-Asalde C, Gastiaburu-Rodriguez D, Pro-Cuba JE, Llanos-Zavalaga F. Multiple infection with resistant and sensitive M. tuberculosis strains during treatment of pulmonary tuberculosis patients. International Journal of Tuberculosis and Lung Disease. 2005; 9:1155–60. [PubMed: 16229228]
11. van Rie A, Victor TC, Richardson M, Johnson R, van der Spuy GD, Murray EJ, Beyers N, van Pittius NCG, van Helden PD, Warren RM. Reinfection and Mixed Infection Cause Changing Mycobacterium tuberculosis Drug-Resistance Patterns. American Journal of Respiratory and Critical Care Medicine. 2005; 172:636–642. [PubMed: 15947286]
12. Shamputa IC, Jugheli L, Sadradze N, Willery E, Portaels F, Supply P, Rigouts L. Mixed infection and clonal representativeness of a single sputum sample in tuberculosis patients from a penitentiary hospital in Georgia. Respiratory Research. 2006; 7:99. [PubMed: 16846493]
13. Dickman KR, Nabyonga L, Kateete DP, Katabazi FA, Asiimwe BB, Mayanja HK, Okwera A, Whalen C, Joloba ML. Detection of multiple strains of Mycobacterium tuberculosis using MIRU-VNTR in patients with pulmonary tuberculosis in Kampala, Uganda. BMC Infectious Diseases. 2010; 10:349. [PubMed: 21143966]
14. Mallard K, McNerney R, Crampin AC, Houben R, Ndlovu R, Munthali L, Warren RM, French N, Glynn JR. Molecular detection of mixed infections of Mycobacterium tuberculosis strains in sputum samples from patients in Karonga District, Malawi. Journal of Clinical Microbiology. 2010; 48:4512–8. [PubMed: 20962138]

15. Cohen T, Wilson D, Wallengren K, Samuel EY, Murray M. Mixed-Strain Mycobacterium tuberculosis Infections among Patients Dying in a Hospital in KwaZulu-Natal, South Africa. Journal of Clinical Microbiology. 2011; 49:385–8. [PubMed: 20980576]

16. Wang JY, Hsu HL, Yu MC, Chiang CY, Yu FL, Yu CJ, Lee LN, Yang PC. the TAMI Group. Mixed infection with Beijing and non-Beijing strains in pulmonary tuberculosis in Taiwan: prevalence, risk factors, and dominant strain. Clinical Microbiology and Infection. 2011 in press.

17. Caminero JA, Pena MJ, Campos-Herrero MI, Rodriguez JC, Afonso O, Martin C, Pavon JM, Torres MJ, Burgos M, Cabrera P, Small PM, Enarson DA. Exogenous reinfection with tuberculosis on a European island with a moderate incidence of disease. American Journal of Respiratory and Critical Care Medicine. 2001; 163:717–20. [PubMed: 11254530]

18. Garcia de Viedma D, Marin M, Hernangomez S, Diaz M, Serrano MJR, Alcala L, Bouza E. Tuberculosis recurrences - Reinfection plays a role in a population whose clinical/epidemiological characteristics do not favor reinfection. Archives of Internal Medicine. 2002; 162:1873–9. [PubMed: 12196086]

19. Dobler CC, Crawford ABH, Jelfs PJ, Gilbert GL, Marks GB. Recurrence of tuberculosis in a low-incidence setting. European Respiratory Journal. 2009; 33:160–7. [PubMed: 18829676]

20. Cohen T, Colijn C, Finklea B, Murray M. Exogenous re-infection and the dynamics of tuberculosis epidemics: local effects in a network model of transmission. Journal of the Royal Society Interface. 2007; 4:523–31.

21. Lumb R, Ardian M, Waramori G, Syahrial H, Tjitra E, Maguire GR, Anstey NM, Kelly RM. An alternative method for sputum storage and transport for Mycobacterium tuberculosis drug resistance surveys. International Journal of Tuberculosis and Lung Disease. 2006; 10:172–7. [PubMed: 16499256]

22. Stavrum R, Mphahlele M, Ovreas K, Muthivhi T, Fourie PB, Weyer K, Grewal HMS. High Diversity of Mycobacterium tuberculosis Genotypes in South Africa and Preponderance of Mixed Infections among ST53 Isolates. Journal Of Clinical Microbiology. 2009; 47:1848–56. [PubMed: 19386854]

23. Musser JM, Amin A, Ramaswamy S. Negligible genetic diversity of Mycobacterium tuberculosis host immune system protein targets: Evidence of limited selective pressure. Genetics. 2000; 155:7–16. [PubMed: 10790380]

24. Lopez B, Aguilar D, Orozco H, Burger M, Espitia C, Ritacco V, Barrera L, Kremer K, Hernandez-Pando R, Huygen K, van Soolingen D. A marked difference in pathogenesis and immune response induced by different Mycobacterium tuberculosis genotypes. Clinical and Experimental Immunology. 2003; 133:30–7. [PubMed: 12823275]

25. Tsolaki AG, Gagneux S, Pym AS, de la Salmoniere Y-OLG, Kreiswirth BN, Soolingen DV, Small PM. Genomic Deletions Classify the Beijing/W Strains as a Distinct Genetic Lineage of Mycobacterium tuberculosis. Journal of Clinical Microbiology. 2005:3185–3191. [PubMed: 16000433]

26. de Jong BC, Hill PC, Aiken A, Awine T, Antonio M, Adetifa IM, Jackson-Sillah DJ, Fox A, DeRiemer K, Gagneux S, Borgdorff MW, McAdam KPWJ, Corrah T, Small PM, Adegbola RA. Progression to active tuberculosis, but not transmission, varies by Mycobacterium tuberculosis lineage in the Gambia. Journal of Infectious Diseases. 2008; 198:1037–1043. [PubMed: 18702608]

27. Thwaites G, Caws M, Chau TTH, D'Sa A, Lan NTN, Huyen MNT, Gagneux S, Anh PTH, Tho DQ, Torok E, Nhu NTQ, Duyen NTH, Duy PM, Richenberg J, Simmons C, Hien TT, Farrar J. Relationship between Mycobacterium tuberculosis genotype and the clinical phenotype of pulmonary and meningeal tuberculosis. Journal of Clinical Microbiology. 2008; 46:1363–8. [PubMed: 18287322]

28. van Crevel R, Nelwan RHH, de Lenne W, Veeraragu Y, van der Zanden AG, Amin Z, van der Meer JWM, van Soolingen D. Mycobacterium tuberculosis Beijing genotype strains associated with febrile response to treatment. Emerging Infectious Diseases. 2001; 7:880–3. [PubMed: 11747703]

29. Tanveer M, Hasan Z, Kanji A, Hussain R, Hasan R. Reduced TNF-α and IFN-γ responses to Central Asian strain 1 and Beijing isolates of Mycobacterium tuberculosis in comparison with

H37Rv strain. Transactions of the Royal Society of Tropical Medicine and Hygiene. 2009; 103:581–7. [PubMed: 19375139]

30. Parwati I, van Crevel R, van Soolingen D. Possible underlying mechanisms for successful emergence of the Mycobacterium tuberculosis Beijing genotype strains. Lancet Infect Dis. 2010; 10:103–11. [PubMed: 20113979]
31. Glynn JR, Whiteley J, Bifani PJ, Kremer K, van Soolingen D. Worldwide occurrence of Beijing/W strains of Mycobacterium tuberculosis: A systematic review. Emerging Infectious Diseases. 2002; 8:843–9. [PubMed: 12141971]
32. Iwamoto T, Yoshida S, Suzuki K, Wada T. Population Structure Analysis of the Mycobacterium tuberculosis Beijing Family Indicates an Association between Certain Sublineages and Multidrug Resistance. Antimicrobial Agents and Chemotherapy. 2008; 52:3805–9. [PubMed: 18694954]
33. Niemann S, Diel R, Khechinashvili G, Gegia M, Mdivani N, Tang YW. Mycobacterium tuberculosis Beijing Lineage Favors the Spread of Multidrug-Resistant Tuberculosis in the Republic of Georgia. Journal of Clinical Microbiology. 2010; 48:3544–50. [PubMed: 20702677]
34. Alavez-Ramirez JA, Castellanos JRA, Esteva L, Flores JA, Fuentes-Allen JL, Garcia-Ramos G, Gomez G, Lopez-Estrada J. Within-host population dynamics of antibiotic-resistant M. tuberculosis. Mathematical Medicine and Biology. 2007; 24:35–56. [PubMed: 17060491]
35. Colijn C, Cohen T, Murray M. Latent Coinfection and the Maintenance of Strain Diversity. Bulletin of Mathematical Biology. 2009; 71:247–263. [PubMed: 19082663]
36. Blower SM, Small PM, Hopewell PC. Control strategies for tuberculosis epidemics: new models for old problems. Science. 1996; 273:497–500. [PubMed: 8662538]
37. Dye C, Williams BG. Criteria for the control of drug-resistant tuberculosis. Proc Natl Acad Sci USA. 2000; 97:8180–5. [PubMed: 10859359]
38. Iannelli M, Martcheva M, Li XZ. Strain replacement in an epidemic model with super-infection and perfect vaccination. Mathematical Biosciences. 2005; 195:23–46. [PubMed: 15922003]
39. Jacobs AL. Infective dose in pulmonary tuberculosis. Tubercle. 1941; 22:266–71.
40. Armstrong RA, McGehe R. Competitive Exclusion. The American Naturalist. 1980; 115:151–70.
41. Zeeman ML. Extinction In Competitive Lotka-Volterra Systems. Proceedings of the American Mathematical Society. 1995; 123:87–96.
42. Strauss OJ, Warren RM, Jordaan A, Streicher EM, Hanekom M, Falmer AA, Albert H, Trollip A, Hoosain E, van Helden PD, Victor TC. Spread of a Low-Fitness Drug-Resistant Mycobacterium tuberculosis Strain in a Setting of High Human Immunodeficiency Virus Prevalence. Journal of Clinical Microbiology. 2008; 46:1514–6. [PubMed: 18272712]
43. Antia R, Regoes RR, Koella JC, Bergstrom CT. The role of evolution in the emergence of infectious diseases. Nature. 2003; 426:658–61. [PubMed: 14668863]
44. Billington OJ, McHugh TD, Gillespie SH. Physiological cost of rifampin resistance induced in vitro in Mycobacterium tuberculosis. Antimicrob Agents Chemother. 1999; 43:1866–9. [PubMed: 10428904]
45. Bottger EC, Springer B, Pletschette M, Sander P. Fitness of antibioticresistant microorganisms and compensatory mutations. Nature Medicine. 1998; 4:1343–4.
46. Bjorkman J, Nagaev I, Berg OG, Hughes D, Andersson DI. Effects of environment on compensatory mutations to ameliorate costs of antibiotic resistance. Science. 2000; 287:1479–82. [PubMed: 10688795]
47. Cohen T, Sommers B, Murray M. The effect of drug resistance on the fitness of Mycobacterium tuberculosis. Lancet Infect Dis. 2003; 3:13–21. [PubMed: 12505028]
48. Gagneux S, Long CD, Small PM, Van T, Schoolnik GK, Bohannan BJM. The Competitive Cost of Antibiotic Resistance in Mycobacterium tuberculosis. Science. 2006; 312:1944–6. [PubMed: 16809538]
49. Shcherbakov D, Akbergenov R, Matt T, Sander P, Andersson DI, Bottger EC. Directed mutagenesis of Mycobacterium smegmatis 16S rRNA to reconstruct the in vivo evolution of aminoglycoside resistance in Mycobacterium tuberculosis. Molecular Microbiology. 2010; 77:830–40.
50. Luciani F, Sisson SA, Jiang H, Francis AR, Tanaka MM. The epidemiological fitness cost of drug resistance in Mycobacterium tuberculosis. PNAS. 2009; 106:14711–5. [PubMed: 19706556]

51. Cohen T, Colijn C, Murray M. Modeling the effects of strain diversity and mechanisms of strain competition on the potential performance of new tuberculosis vaccines. Proceedings of the National Academy of Sciences. 2008; 105:16302–7.

52. Lipsitch M, Colijn C, Cohen T, Hanage WP, Fraser C. No coexistence for free: Neutral null models for multistrain pathogens. Epidemics. 2009; 1:2–13. [PubMed: 21352747]

Appendix A. The equilibrium analysis

A1. Dynamic equations

The differential equations corresponding to the model in Figure 1 are as follows:

$$\begin{aligned}
\tfrac{d}{dt}S &= r-(r+\lambda_S+\lambda_R)S\\
\tfrac{d}{dt}L_S &= \lambda_S S-(\pi+\lambda_R)L_S\\
\tfrac{d}{dt}L_R &= \lambda_R S-(\pi+\lambda_S)L_R\\
\tfrac{d}{dt}M &= \lambda_R L_S+\lambda_S L_R-\pi M\\
\tfrac{d}{dt}I_S &= \pi(L_S+f_S M)-(\mu+d)I_S\\
\tfrac{d}{dt}I_{SR} &= \pi f_{SR}M-(\mu+d)I_{SR}\\
\tfrac{d}{dt}I_R &= \pi(L_R+fM)+d(\eta I_S+I_{SR})-\mu I_R.
\end{aligned} \tag{A1}$$

In order to rewrite the equations for the independent set of the variables from Table 1b, we use the following substitution steps. First, we link the frequency of those in the each of the Infectious classes with the infectious force, expressing the infectivity coefficients as a function of strain reproductive numbers:

$$\begin{aligned}
I_S &= \tfrac{1-\varepsilon}{(\mu+d)R_S}\lambda_S\\
I_{SR} &= \tfrac{\varepsilon}{(\mu+d)R_S}\lambda_S\\
I_R &= \tfrac{1}{\mu R_R}\lambda_R.
\end{aligned} \tag{A2}$$

Second, we express the strain infectious forces though the relative fraction of the infection force due to DR strain (v) and the expected number of reinfections (n_λ) (as, the total infectious force, $\lambda_R+\lambda_S$, provides the overall rate of reinfection, and $1/\pi$ provides the duration of the Latency period):

$$\begin{aligned}
\lambda_S &= (1-v)\pi n_\lambda\\
\lambda_R &= v\pi n_\lambda.
\end{aligned} \tag{A3}$$

Third, we express the drug treatment rate as a function of the treatment coverage δ:

$$d = \mu\delta/(1-\delta). \tag{A4}$$

And finally, we use the expressions (1) for the probabilities f_S and f_{SR}:

$$\begin{aligned}
f_S &= (1-\theta)(1-f)\\
f_{SR} &= \theta(1-f).
\end{aligned} \tag{A4}$$

Applying the substitutions (A2–5) to the dynamic set (A1) we obtain the following system, describing the dynamics of all the variables from Table 1b (with the exception of f):

$$\begin{aligned}
\tfrac{1}{r}\tfrac{d}{dt}S &= 1-(1+\tfrac{\pi}{r}n_\lambda)S\\
\tfrac{1}{\pi}\tfrac{d}{dt}L_S &= (1-\nu)n_\lambda S-(1+\nu n_\lambda)L_S\\
\tfrac{1}{\pi}\tfrac{d}{dt}L_R &= \nu n_\lambda S-(1+(1-\nu)n_\lambda)L_R\\
\tfrac{1}{\pi}\tfrac{d}{dt}M &= \nu n_\lambda L_S+(1-\nu)n_\lambda L_R-M\\
\tfrac{(1-\delta)}{\mu R_S}(1-\nu)n_\lambda\tfrac{d}{dt}\varepsilon &= \theta(1-f)M-\varepsilon(L_S+(1-f)M)\\
\tfrac{(1-\delta)}{\mu R_S}\tfrac{d}{dt}(1-\nu)n_\lambda &= (L_S+(1-f)M)-\tfrac{1}{R_S}(1-\nu)n_\lambda\\
\tfrac{1}{\mu R_R}\tfrac{d}{dt}\nu n_\lambda &= (L_R+fM)+\tfrac{\delta}{R_S}(\eta+(1-\eta)\varepsilon)(1-\nu)n_\lambda-\tfrac{1}{R_R}\nu n_\lambda.
\end{aligned} \tag{A6}$$

Here we list the equations in a form that keeps the right side of the equations as simple as possible to facilitate our subsequent investigation of equilibrium conditions.

The equations (A6), together with the equation (4) for f, provide the full set of equalities for the values of all the variables (S, L_S, L_R, M, ε, n_λ,ν, and f) from Table 1b in equilibrium:

$$\begin{aligned}
S &= \tfrac{1}{1+n_\lambda\pi/r}\\
L_S &= (1-\nu)\tfrac{n_\lambda}{1+\nu n_\lambda}S\\
L_R &= \nu\tfrac{n_\lambda}{1+(1-\nu)n_\lambda}S\\
M &= \nu(1-\nu)\left(\tfrac{{n_\lambda}^2}{1+\nu n_\lambda}+\tfrac{{n_\lambda}^2}{1+(1-\nu)n_\lambda}\right)S\\
\varepsilon &= \tfrac{\theta(1-f)M}{L_S+(1-f)M}\\
(1-\nu)\tfrac{n_\lambda}{R_S} &= L_S+(1-f)M\\
\nu\tfrac{n_\lambda}{R_R}-(1-\nu)\tfrac{\delta((1-\eta)\varepsilon+\eta)n_\lambda}{R_S} &= L_R+fM\\
f &= (1-\varphi)\tfrac{1+\nu n_\lambda}{2+n_\lambda}+\varphi f_D.
\end{aligned} \tag{A7}$$

A2. System equilibria

The dynamic system (A6) has at most 4 equilibria (as the solutions of (A7)), namely: Disease free; DS strain only; DR strain only; and DR and DS coexistence. In order to distinguish the epidemiological state variables with respect to the state we label them in superscripts with the marker of the state: ***0** (disease free), ***S** (DS only), ***R** (DR only), and ***** (coexistence) correspondingly.

a) Disease free (*0; $n_\lambda^{*0}\equiv 0$)

$$(S^{*0},L_S^{*0},L_R^{*0},M^{*0},I_S^{*0},I_{SR}^{*0},I_R^{*0})=(1,0,0,0,0,0,0). \tag{A8}$$

This equilibrium is stable only if both R_R, R_S<1, and unstable otherwise.

b) DS strain only (*S; $\nu^{*S}\equiv 0$)

$$\begin{aligned}
(S^{*S},L_S^{*S},L_R^{*S},M^{*S},I_S^{*S},I_{SR}^{*S},I_R^{*S}) &= \left(\tfrac{1}{R_S},\tfrac{n_\lambda^{*S}}{R_S},0,0,\tfrac{\pi}{\mu+d}\tfrac{n_\lambda^{*S}}{R_S},0,0\right)\\
n_\lambda^{*S} &\equiv \tfrac{r}{\pi}(R_S-1).
\end{aligned} \tag{A9}$$

This equilibrium is not possible unless R_S>1 and $\eta\delta$=0, and is stable only if the invasion reproductive number for the DR strain is less than 1. The latter condition can be tested by

calculating the secondary DR cases (K) a single infectious DR individual would produce by adding up all the individual routes of reproduction:

$$\begin{gathered}K_{S\to L_R\to I_R}+K_{S\to L_R\to M\to I_R}+K_{S\to L_S\to M\to I_R}+K_{S\to L_R\to M\to I_{SR}\to I_R}+K_{S\to L_S\to M\to I_{SR}\to I_R}<1\\ K_{S\to L_R\to I_R}=\frac{R_R}{R_S}\frac{1}{1+n_\lambda^{*S}}\\ K_{S\to L_R\to M\to I_R}=f^{*S}\frac{R_R}{R_S}\frac{n_\lambda^{*S}}{1+n_\lambda^{*S}}\\ K_{S\to L_S\to M\to I_R}=f^{*S}\frac{R_R}{R_S}n_\lambda^{*S}\\ K_{S\to L_R\to M\to I_{SR}\to I_R}=\theta\delta(1-f^{*S})\frac{R_R}{R_S}\frac{n_\lambda^{*S}}{1+n_\lambda^{*S}}\\ K_{S\to L_S\to M\to I_{SR}\to I_R}=\theta\delta(1-f^{*S})\frac{R_R}{R_S}n_\lambda^{*S}\\ f^{*S}=(1-\varphi)\frac{1}{2+n_\lambda^{*S}}+\varphi f_D\end{gathered}\tag{A10}$$

Substituting the expressions on the lines 2–6 of (A10) into the inequality on line 1, we obtain the relationship between the strain reproductive numbers (R_S and R_R) that provides the stability of ***S** (compare with (6) from main text):

$$R_R/R_S<\frac{1}{1+n_\lambda^{*S}\left((1-\theta\delta)\varphi\left(\frac{2+n_\lambda^{*S}}{1+n_\lambda^{*S}}f_D-\frac{1}{1+n_\lambda^{*S}}\right)+\theta\delta\right)}.\tag{A11}$$

c) DR strain only (*R; $v^{*R}\equiv 1$)

$$\begin{gathered}(S^{*R},L_S^{*R},L_R^{*R},M^{*R},I_S^{*R},I_{SR}^{*R},I_R^{*R})=\left(\frac{1}{R_R},0,\frac{n_\lambda^{*R}}{R_R},0,0,0,\frac{\pi}{\mu}\frac{n_\lambda^{*R}}{R_R}\right)\\ n_\lambda^{*R}\equiv\frac{r}{\pi}(R_R-1).\end{gathered}\tag{A12}$$

This equilibrium is not possible unless $R_R>1$, and it is stable only if the invasion reproductive number for DS strain is below 1. This condition can be tested by analogy with (A10):

$$\begin{gathered}K_{S\to L_S\to I_S}+K_{S\to L_S\to M\to(I_S+I_{SR})}+K_{S\to L_R\to M\to(I_S+I_{SR})}<1\\ K_{S\to L_S\to I_S}=\frac{R_S}{R_R}\frac{1}{1+n_\lambda^{*R}}\\ K_{S\to L_S\to M\to(I_S+I_{SR})}=(1-f^{*R})\frac{R_S}{R_R}\frac{n_\lambda^{*R}}{1+n_\lambda^{*R}}\\ K_{S\to L_R\to M\to(I_S+I_{SR})}=(1-f^{*R})\frac{R_S}{R_R}n_\lambda^{*R}\\ f^{*R}=(1-\varphi)\frac{n_\lambda^{*R}}{2+n_\lambda^{*R}}+\varphi f_D.\end{gathered}\tag{A13}$$

Summation of (A13), yields the relation between the strain reproductive numbers that provide the stability of ***R** (compare with (5) from main text):

$$R_R/R_S>1+n_\lambda^{*R}\varphi\left(1-\frac{2+n_\lambda^{*R}}{1+n_\lambda^{*R}}f_D\right).\tag{A14}$$

d) Mixed-strain: the coexistence of DS and DR strains (*; 0<v^*<1)

The sizes of all the 7 compartments in this equilibrium can be expressed by just two values: the equilibrium fraction of the infection force due to DR strain (v^*) and the number of reinfections (n_λ^*):

$$\begin{aligned}
S^* &= \frac{1}{1+n_\lambda^*\pi/r} \\
L_S^* &= \frac{(1-v^*)n_\lambda^*}{1+v^*n_\lambda^*}S^* = \frac{(1-v^*)n_\lambda^*}{1+v^*n_\lambda^*}\frac{1}{1+n_\lambda^*\pi/r} \\
L_R^* &= \frac{v^*n_\lambda^*}{1(1-v^*)n_\lambda^*}S^* = \frac{v^*n_\lambda^*}{1+(1-v^*)n_\lambda^*}\frac{1}{1+n_\lambda^*\pi/r} \\
M^* &= n_\lambda^*S^* - L_S^* - L_R^* = \frac{v^*(1-v^*)(2+n_\lambda^*)n_\lambda^{*2}}{v^*(1-v^*)n_\lambda^{*2}+n_\lambda^*+1}\frac{1}{1+n_\lambda^*\pi/r} \\
I_S^* &= (1-\varepsilon^*)\frac{\pi}{(\mu+d)R_S}(1-v^*)n_\lambda^* \\
I_{SR}^* &= \varepsilon^*\frac{\pi}{(\mu+d)R_S}(1-v^*)n_\lambda^* \\
I_R^* &= \frac{\pi}{\mu R_R}v^*n_\lambda^* \\
\varepsilon^* &\equiv \frac{\theta}{1+\frac{1}{v^*n_\lambda^*}\frac{1}{1+\varphi\left(\frac{1+v^*n_\lambda^*}{1+(1-v^*)n_\lambda^*}-f_D\frac{2+n_\lambda^*}{1+(1-v^*)n_\lambda^*}\right)}}.
\end{aligned} \tag{A15}$$

For the values v^* and n_λ^* we have the following two equations:

$$\begin{aligned}
&\frac{R_S}{R_R} = \eta\delta\frac{(1-v^*)}{v^*} + \frac{1+(1-v^*)n_\lambda^*\left\langle\left\langle\frac{v^*n_\lambda^*}{1+n_\lambda^*}\middle|\frac{2+n_\lambda^*}{1+n_\lambda^*}f_D-\frac{1}{1+n_\lambda^*}\right\rangle_\varphi\middle|1\right\rangle_{\theta\delta}}{1+v^*n_\lambda^*\left\langle\frac{(1-v^*)n_\lambda^*}{1+n_\lambda^*}\middle|1-\frac{2+n_\lambda^*}{1+n_\lambda^*}f_D\right\rangle_\varphi} \\
&(1-v^*)\left(\frac{1}{R_S}-\frac{\delta((1-\eta)\varepsilon^*+\eta)}{R_R}\right)+v^*\frac{1}{R_R} = \frac{1}{1+n_\lambda^*\pi/r} \equiv S^*.
\end{aligned} \tag{A16}$$

The brackets denote an average, with the appropriate weighting shown in subscript:

$\langle g|h\rangle_\alpha \equiv (1-\alpha)g+\alpha h$. For ease of presentation visualization, we also retain ε^* in the above equation; ε^* is a complicated combination of v^* and n_λ^*.

The first equation of (A16) is derived by dividing the line 6 in (A7) by the line 7 and simplifying the result by substituting the expressions for L_S^*, L_R^*, and M^* from (A15). This equation shows how the relative fitness of the strains (left-side of equation) are linked with the mechanisms of interaction between the strains (described by parameters on right-side of the equation).

The second equation is obtained by summation of the lines 6 and 7 in (A7) and the following simplification with the use of expressions for S^*, L_S^*, L_R^*, and M^* in (A15). This equation shows how the two strains together deplete the pool of susceptible individuals.

The conditions when ***** equilibrium exists can be obtained by analyzing the behavior of the first equation in (A16) at the limit of either $v^*\to 0$ or $v^*\to 1$, substituting the corresponding value of n_λ^* from the second equation. For non-zero acquired resistance in the presence of treatment ($\eta\delta>0$), we always have $v^*>0$, and the only boundary of the coexistent area corresponds to $v^*=1$ (the boundary between ***** and ***R**) (compare with (A14)):

$$1<R_R<R_S\left(1+n_\lambda^{*R}\varphi\left(1-\frac{2+n_\lambda^{*R}}{1+n_\lambda^{*R}}f_D\right)\right). \tag{A17}$$

If we ignore acquired resistance, then we also have a boundary between ***** and ***S** with an additional condition, corresponding to $v^*=0$ (compare with (A11)):

$$1<R_S<\frac{R_R}{1+n_\lambda^{*S}\left\langle\varphi\left(\frac{2+n_\lambda^{*S}}{1+n_\lambda^{*S}}f_D-\frac{1}{1+n_\lambda^{*S}}\right)\middle|1\right\rangle_{\theta\delta}}. \tag{A18}$$

Conditions allowing for the existence of the equilibrium ***** map 1:1 to the conditions when each of the other equilibria (***0**, ***S**, and ***R**) are unstable.

A3. Mixed-strain equilibrium analysis

Despite the complexity of the equations (A16), there are several important cases when the system can be significantly simplified and analyzed.

a) Symmetric reproduction ($\eta\delta=\theta\delta=0$; $f_D=1/2$)

If we assume that the model has symmetrical structure in respect to both the strains and the strains differ only by their reproductive numbers, the equations (A16) can be rewritten as:

$$\frac{R_S}{R_R}=\frac{1+\frac{\varphi}{2}(1-v^*)\frac{n_\lambda^{*2}}{1+n_\lambda^*+(1-\varphi)v^*(1-v^*)n_\lambda^{*2}}}{1+\frac{\varphi}{2}v^*\frac{n_\lambda^{*2}}{1+n_\lambda^*+(1-\varphi)v^*(1-v^*)n_\lambda^{*2}}}$$
$$\frac{1-v^*}{R_S}+\frac{v^*}{R_R}=\frac{1}{1+n_\lambda^*\pi/r}. \tag{A19}$$

In the case of low prevalences ($n_\lambda^*\to 0$) the equations (A19) can be resolved:

$$v^*=\frac{1}{2}-\frac{1}{\varphi}\frac{\pi^2}{r^2}\frac{R_S-R_R}{(R_S-1)}$$
$$n_\lambda^*=\frac{r}{\pi}(R_S-1). \tag{A20}$$

As the fraction of infectious force in mixed-equilibrium is restricted to $0<v^*<1$, the coexistence is allowed if R_R falls into the range:

$$R_S-\frac{\varphi}{2}\frac{r^2}{\pi^2}(R_S-)^2<R_R<R_S+\frac{\varphi}{2}\frac{r^2}{\pi^2}(R_S-1)^2, \tag{A21}$$

yielding the equation (9) from the main text.

b) Strain-dependent within-host competition ($\varphi=1$)

If the chance of the bacteria to win within-host competition is entirely determined by strain-specific factors, and there are no effects of the acquirement of resistance and within-host drug selection ($\eta\delta=\theta\delta=0$), the equations (A16) can be significantly simplified:

$$\frac{R_S}{R_R}=\frac{1+n_\lambda^*+(1-v^*)n_\lambda^*((2+n_\lambda^*)f_D-1)}{1+n_\lambda^*+v^*n_\lambda^*(1+n_\lambda^*-(2+n_\lambda^*)f_D)}$$
$$\frac{1-v^*}{R_S}+\frac{v^*}{R_R}=\frac{1}{1+n_\lambda^*\pi/r}. \tag{A22}$$

In the case of low prevalences ($n_\lambda^*\to 0$) and different abilities of strains to within-host competition ($f_D\neq 0.5$) the equations (A22) can be resolved:

$$n_\lambda^*=\frac{r}{\pi}((1-v^*)R_S+v^*R_R-1)=\frac{R_S-R_R}{2f_D-1}$$
$$v^*=\frac{R_S-1}{R_S-R_R}-\frac{\pi}{r}\frac{1}{2f_D-1}. \quad (A23)$$

Substituting the last equation of (A23) into to the condition $0<v^*<1$ yields the equation (10) from the main text for the range of R_R allowing coexistence:

$$R_S+\frac{r}{\pi}(1-2f_D)(R_S-1)<R_R<R_S+\frac{r}{\pi}\frac{(1-2f_D)}{1-\frac{r}{\pi}(1-2f_D)}(R_S-1). \quad (A24)$$

c) Strain-independent within-host competition (φ=0)

If within-host dynamics of the bacteria are similar for both strains, the equations (A16) can be rewritten as:

$$\frac{R_S}{R_R}=1+\eta\delta\frac{(1-v^*)}{v^*}+\theta\delta(1-v^*)n_\lambda^*\frac{1+(1-v^*)n_\lambda^*}{1+n_\lambda^*+v^*(1-v^*)n_\lambda^{*2}}$$
$$(1-v^*)\left(\frac{1}{R_S}-\frac{\theta\delta(1-\eta)v^*n_\lambda^*/(1+v^*n_\lambda^*)+\eta\delta}{R_R}\right)+v^*\frac{1}{R_R}=\frac{1}{1+n_\lambda^*\pi/r}. \quad (A25)$$

The first equation in (A25) shows that coexistence may occur even if the reproductive numbers R_S and R_R differ from each other when either acquired resistance ($\eta\delta>0$) or selection of DR subpopulations by drugs ($\theta\delta>0$) is possible. In the absence of these two effects ($\eta\delta=\theta\delta=0$) the system (A25) reduces to:

$$R_S=R_R$$
$$S^*=1/R_S, \quad (A26)$$

yielding the fundamental result that only the two strains with equal reproductive numbers may coexist without interaction using the same susceptible pool.

If acquired resistance can occur but there is no selection of DR subpopulations ($\eta\delta>0$; $\theta\delta=0$), the system (A25) can be resolved as:

$$v^*=\frac{1}{1+(R_S-R_R)/(\eta\delta R_R)}$$
$$n_\lambda^*=\frac{r}{\pi}\left(R_S\frac{1+\eta\delta R_R/(R_S-R_R)}{1-\eta\delta {R_S}^2/(R_R(R_S-R_R))}-1\right). \quad (A27)$$

If, alternatively, acquired resistance does not occur but DR subpopulations occurring among those progressing from mixed infections can be selected by treatment ($\eta\delta=0$; $\theta\delta>0$), the system (A25) can be rewritten as:

$$\frac{R_S}{R_R}=1+\theta\delta(1-v^*)n_\lambda^*\frac{1+(1-v^*)n_\lambda^*}{1+n_\lambda^*+v^*(1-v^*)n_\lambda^{*2}}$$
$$(1-v^*)\left(\frac{1}{R_S}-\frac{v^*n_\lambda^*}{1+v^*n_\lambda^*}\frac{\theta\delta}{R_R}\right)+v^*\frac{1}{R_R}=\frac{1}{1+n_\lambda^*\pi/r}. \quad (A28)$$

The analysis of the limits $v^* \to 0$ and $v^* \to 1$ in (A28) yields the equation (11) from the main text for the range of R_R, allowing the coexistence of the strains specifically because of within-host drug selection effect:

$$R_S/(1+\theta\delta\tfrac{r}{\pi}(R_S-1)) < R_R < R_S. \tag{A29}$$

Links to previous models—The equations (A16) allow us advance the results of earlier studies of mixed infection with DR and DS strains of TB [35, 51]. These previous models are "nested" within our current approach and represent the setting in which $\varphi=1$, $f_D=1/2$ and $\theta=0$, resulting in simplified equations:

$$\begin{gathered}\frac{R_S}{R_R}=\eta\delta\frac{(1-v^*)}{v^*}+\frac{1+(1-v^*)n_\lambda^{*2}/2(1+n_\lambda^*)}{1+v^*n_\lambda^{*2}/2(1+n_\lambda^*)}\\ \frac{1-v^*}{R_S}+\frac{v^*}{R_R}=\frac{1}{1+n_\lambda^*\pi/r},\end{gathered} \tag{A30}$$

(see Figure 2a). Our model allows us to include the effects of both strain-dependent and strain-independent mechanisms of within-host competition ($\varphi<1$ in equation (11); $f_D\neq 1/2$, Figure 2b) as well as the potential for drug treatment to select for minority drug-resistant strains present in individuals with apparent drug sensitive disease ($\theta=1$, the effect of alteration $\delta=0\to\delta=1$ on Figure 2c).

Appendix B. Strain-independent competition within the host

Since we are interesting in studying coexistence, we must first confirm that our model does not incorporate artificial (i.e. "built in") mechanisms that support the stable coexistence of strains. We test the "neutrality" of the model by considering two strains that are identical except for some neutral marker. According to [52], a neutral model must meet two criteria: "ecological neutrality" and "population genetic neutrality". Ecological neutrality requires the dynamics of the ecological variables (the number of uninfected hosts, the number infected with 0, 1, 2, … strains, *etc*) to depend on the ecological state variables but, given these, to be independent of the identity of the particular strains involved. Population genetic neutrality means that there should be no stable equilibrium frequency of the strains in the model; rather, it should be possible to choose initial conditions to guarantee an arbitrary frequency of strains that remains constant for all time. Application of both the criteria to our model allows us to determine f_N (and, using (2), for probability f) for the important cases we need.

The equation set (A1) can be rewritten for the case of two identical strains via the substitution of $d=\varphi=\theta=0$ and setting $I_{SR}=0$. For the remaining compartments we have:

$$\begin{gathered}\tfrac{d}{dt}S=r-(r+\lambda_1+\lambda_2)S\\ \tfrac{d}{dt}L_1=\lambda_1 S-(\pi+\lambda_2)L_1\\ \tfrac{d}{dt}L_2=\lambda_2 S-(\pi+\lambda_1)L_2\\ \tfrac{d}{dt}M=\lambda_2 L_1+\lambda_1 L_2-\pi M\\ \tfrac{d}{dt}I_1=\pi(L_1+(1-f_N)M)-\mu I_1\\ \tfrac{d}{dt}I_2=\pi(L_2+f_N M)-\mu I_2.\end{gathered} \tag{B1}$$

Here we are using the subscripts '1' and '2', instead of 'S' and 'R', in order to underscore that the two strains are indistinguishable.

a) Ecological neutrality

According to this criterion, the ecological dynamics of the infection must be independent of the relative strain frequencies. Thus, introducing L, I and λ for the combined latent class, infectious class, and force of infection:

$$\begin{aligned} L&=L_1+L_2+M \\ I&=I_1+I_2 \\ \lambda&=\lambda_1+\lambda_2, \end{aligned} \tag{B2}$$

and summarizing 2–4 and 5–6 equations in (B1) we obtain:

$$\begin{aligned} \tfrac{d}{dt}S&=r-(r+\lambda)S \\ \tfrac{d}{dt}L&=\lambda S-\pi L \\ \tfrac{d}{dt}I&=\pi L-\mu I. \end{aligned} \tag{B3}$$

The absence of any dependence of the totals on the strains composition proves that the system (B1) satisfies the ecologic neutrality criteria regardless of the value of f_N.

b) Population genetic neutrality

b1) Criterion formulation

According to the population genetic neutrality criterion, for every value of the relative strain frequency there is a choice of the initial conditions possible that keeps this value constant regardless of the subsequent dynamics of the system. In the other words, if we introduce the time-dependent frequency of the second strain as:

$$v=\lambda_2/(\lambda_1+\lambda_2)=I_2/I, \tag{B4}$$

and postulate that the relation between the infection forces remained the same over the past:

$$v|_{t<0} \equiv const, \tag{B5}$$

we must find such an expression for f_N, that provides:

$$\tfrac{d}{dt}v\Big|_{t=0}=0, \tag{B6}$$

regardless of the dynamics of S, L and I.

It is challenging to provide a full analysis of the strain frequency dynamics for the whole system (B1), thus we break this analysis into several steps. Applying the definition (B4) to the last equation of (B1), and using the equation (B3) for the total Infectious derivative, we first obtain:

$$\tfrac{d}{dt}v=\frac{\pi}{I}(L_2+f_N M-vL). \tag{B7}$$

Now, using the criteria (B6), we obtain the general expression for the competition probability that forces the model to satisfy the population genetic neutrality criteria:

$$f_N = v\frac{L}{M} - \frac{L_2}{M}. \tag{B8}$$

To derive the ratios L/M and L_2/M, we require insight on the dynamics of the latency compartments L, L_2, and M.

b2) Criterion for the steady state

We first assume, that the system is in the ecological steady state and:

$$\frac{d}{dt}S = \frac{d}{dt}L = \frac{d}{dt}I = 0. \tag{B9}$$

According to genetic neutrality criteria, the frequencies of each of the latency compartments must also remain:

$$\begin{aligned} \frac{1}{\pi}\frac{d}{dt}L_2 &= vL - (1+(1-v)n_\lambda)L_2 = 0 \\ \frac{1}{\pi}\frac{d}{dt}M &= vn_\lambda L + (1-2v)n_\lambda L_2 - (1+vn_\lambda)M = 0. \end{aligned} \tag{B10}$$

Here the number of reinfections n_λ and the fraction of the infection force due to the second strain v are introduced in the same way as in equation (A3) of the Appendix A:

$$\begin{aligned} \lambda_1 &= (1-v)\pi n_\lambda \\ \lambda_2 &= v\pi n_\lambda. \end{aligned} \tag{B11}$$

Solving (B10) in respect to L/M and L_2/M, and substituting them to (B8), we obtain a probability of the strain-independent competition probability in equilibrium:

$$f_N = \frac{1+vn_\lambda}{2+n_\lambda}. \tag{B12}$$

Together with the expression (2) from the main text, (B12) reproduces the expression (4) for overall probability f. (B12) shows that the strain-independent probability f_N is a function of the number of reinfection events and strain frequencies. We note the probability (B12) approaches 1/2 if reinfection is rare ($n_\lambda \ll 1$) or if the strain infection forces are equal ($v=1/2$). The other important conclusion from the analysis of ((B12) is that the strain-independent probability of the strain to win the competition is monotonically increasing with the relative frequency of the strain in population. In other words, in order to preserve neutrality, the more frequent strain has to show a competitive advantage. In contrast, when we use strain-dependent probability ($\varphi=1 \rightarrow f=f_D$), we implicitly provide the less frequent strain with a competitive boost.

b3) Criterion in the generalized case

There are no simple expressions for M/L and L_2/L ratios in generalized dynamical case, but there is still some insight we can gain. We describe the current distribution of the strains in the population within the following assumptions:

1. Among those in the Latent class, there is a known distribution g_n of the number of reinfection events (n) the individuals experience by the moment of developing the disease:

$$\sum_{n\geq 0} g_n \equiv 1. \tag{B13}$$

2. Each strain has been responsible for the same fraction of the total infection force in the past (v_t=v=*const*).
3. The probability that an infection event is due to a particular strain is equal to its relative frequency (i.e. an assumption of homogenous mixing).

In such a case, the probability that a particular latently infected individual has been reinfected n times, and $m \in [0..\, n{+}1]$ of the total number of infections are due to the second strain, is:

$$\xi_n^m = g_n \begin{pmatrix} m \\ n+1 \end{pmatrix} v^m (1-v)^{n+1-m}; \qquad \sum_{m\leq n+1} \xi_n^m \equiv g_n. \tag{B14}$$

With the use of (B13), for the ratios M/L and L_2/L we have:

$$\begin{gathered} \frac{M}{L} = \sum_{n\geq 1} \sum_{0<m<n+1} \xi_n^m = \sum_{n\geq 1} g_n (1 - v^{n+1} - (1-v)^{n+1}) \\ \frac{L_2}{L} = \sum_{n\geq 0} \chi_n^{n+1} = \sum_{n\geq 0} g_n v^{n+1}. \end{gathered} \tag{B15}$$

Substituting (B15) to (B8) we obtain the expression for the strain-independent competition probability:

$$f_N = \frac{v - \sum_{n\geq 0} g_n v^{n+1}}{\sum_{n\geq 1} g_n (1 - v^{n+1} - (1-v)^{n+1})}. \tag{B16}$$

Accordingly, different assumptions about the distribution g_n result in different expressions for the probability f_N.

b4) Reinfection number distribution examples

Here we calculate the expression (B16) for a few distributions of the number of reinfection events.

1. Geometric distribution case—Let the value g_n to have a geometric distribution with the mean n_λ:

$$g_n = \frac{1}{1+n_\lambda} \left(\frac{n_\lambda}{1+n_\lambda} \right)^n, \tag{B17}$$

relevant to the equilibrium case with a normal distribution for the disease-progression times. In this case, the summation of (B16) results into familiar:

$$f_N = \frac{1+\nu n_\lambda}{2+n_\lambda}, \tag{B18}$$

reproducing (B12).

2. Poisson distribution case—Let the value g_n to have a Poisson distribution with the mean n_λ:

$$g_n = \frac{{n_\lambda}^n}{n!}\exp(-n_\lambda), \tag{B19}$$

relevant to the equilibrium case, when the duration of latency is set to be the same for all hosts. In this case, the summation of (B16) results into:

$$f_N = \frac{\nu(1-\exp(-(1-\nu)n_\lambda))}{1-\nu\exp(-(1-\nu)n_\lambda)-(1-\nu)\exp(-\nu n_\lambda)}. \tag{B20}$$

3. Two-point distribution—Let the value g_n to have the distribution (B17), but with non-zero values only for n=0,1:

$$g_0 = \frac{1}{1+n_\lambda};\ g_1 = \frac{n_\lambda}{1+n_\lambda};\ g_{n>1} \equiv 0. \tag{B21}$$

This is possible if only a single reinfection event can occur. In this case, the summation of (B16) results in a simple expression:

$$f_N \equiv \frac{1}{2}. \tag{B22}$$

Appendix C. An expanded model

Here we expand the model depicted in Figure 1 to include a more detailed representation of the natural history of tuberculosis. The results of our simpler model presented in the main text are supported by this analysis and similar relationships are found through the redefinition of model parameters.

The model shown in Figure C1 is based on the model on Figure 1 but with several new details considered. We allow for a different susceptibility to reinfection as compared to infection (by the factor κ), allow for different mortality from each of the compartments (μ_S, μ_L, and μ_I for Susceptible, Latent and Infectious compartments correspondingly), allow for those in the latent compartment the opportunity to arrest progression (with the rate γ), provide those with DR and DS disease treatments of different efficacy (with the rates d_R and d_S), and allow those with Infectious disease to self-cure at the rate h.

The full system of the dynamic equations of the Model on Figure A1 is:

$$\begin{aligned}
\tfrac{d}{dt}S &= \mu_S + \gamma(L_S + L_R + M) - (\mu_S + \lambda_R + \lambda_S)S \\
\tfrac{d}{dt}L_S &= \lambda_S S - (\mu_L + \gamma + \pi + \kappa\lambda_R)L_S \\
\tfrac{d}{dt}L_R &= \lambda_R S - (\mu_L + \gamma + \pi + \kappa\lambda_S)L_R \\
\tfrac{d}{dt}M &= \kappa\lambda_R L_S + \kappa\lambda_S L_R - (\mu_L + \gamma + \pi)M \\
\tfrac{d}{dt}I_S &= \pi(L_S + f_S M) - (\mu_I + h + d_S)I_S \\
\tfrac{d}{dt}I_{SR} &= \pi f_{SR} M - (\mu_I + h + d_S)I_{SR} \\
\tfrac{d}{dt}I_R &= \pi(L_R + fM) + d_S(\eta I_S + I_{SR}) - (\mu_I + h + d_R)I_R.
\end{aligned} \tag{C1}$$

The system (C1) is distinct from (A1) and has different dynamical behavior. However, in the following, we show that at equilibrium, the system described by (C1), can be reduced to that of (A1) through simple parameter renormalization.

By analogy with derivations (A1–7), applying the following substitutions:

$$\begin{aligned}
I_S^* &= \frac{\pi^2}{(\mu_L + \gamma + \pi)(\mu_I + h + d_S)} \frac{1-\varepsilon^*}{R_S}(1-\kappa^*)n_\lambda^* \\
I_{SR}^* &= \frac{\pi^2}{(\mu_L + \gamma + \pi)(\mu_I + h + d_S)} \frac{\varepsilon^*}{R_S}(1-\kappa^*)n_\lambda^* \\
I_R^* &= \frac{\pi^2}{(\mu_L + \gamma + \pi)(\mu_I + h + d_R)} \frac{1}{R_R} v^* n_\lambda^* \\
d_S &= (\mu_I + h)\frac{\delta}{1-\delta} \\
f_S^* &= (1-\theta)(1-f^*) \\
f_{SR}^* &= \theta(1-f^*),
\end{aligned} \tag{C2}$$

and setting the time-derivatives to zero, yields the following equations for the equilibrium values of the variables of the Table 1b:

$$\begin{aligned}
\mu_S + \gamma(L_S^* + L_R^* + M^*) &= (\mu_S + \pi n_\lambda^*)S^* \\
(1-v^*)n_\lambda^* S^* &= (\tfrac{\mu_L + \gamma + \pi}{\pi} + \kappa v^* n_\lambda^*)L_S^* \\
v^* n_\lambda^* S^* &= (\tfrac{\mu_L + \gamma + \pi}{\pi} + \kappa(1-v^*)n_\lambda^*)L_R^* \\
\kappa(v^* n_\lambda^* L_S^* + (1-v^*)n_\lambda^* L_R^*) &= \tfrac{\mu_L + \gamma + \pi}{\pi} M^* \\
\varepsilon^* &= \theta(1-f^*)M^*/(L_S^* + (1-\theta)(1-f^*)M^*) \\
(1-v^*)n_\lambda^*/R_S &= \tfrac{\mu_L + \gamma + \pi}{\pi}(L_S^* + (1-f^*)M^*) \\
v^* n_\lambda^*/R_R - (1-v^*)\delta((1-\eta)\varepsilon^* + \eta)n_\lambda^*/R_S &= \tfrac{\mu_L + \gamma + \pi}{\pi}(L_R^* + f^* M^*)
\end{aligned} \tag{C3}$$

Summarizing the equation lines 2–4 in (C3), we obtain:

$$L_S^* + L_R^* + M^* = \tfrac{\pi}{\mu_L + \gamma + \pi} n_\lambda^* S^*. \tag{C4}$$

Substituting (C4) into the first equation of (C3) and using the following redefinition of the model parameters and variables:

$$\begin{aligned}
\tilde{r} &= \tfrac{\mu_S \pi}{\mu_S + \pi}\kappa \\
\tilde{n}_\lambda^* &= \tfrac{\pi}{\mu_L + \gamma + \pi}\kappa n_\lambda^* \\
\tilde{L}_S^* &= \kappa L_S^* \\
\tilde{L}_R^* &= \kappa L_R^* \\
\tilde{M}^* &= \kappa M^*,
\end{aligned} \tag{C5}$$

we reduce our system of equations to one that is identical to (A7) (and, therefore, according to Appendix B, obtaining the same expression for the equilibrium value of *f*):

$$
\begin{aligned}
S^* &= \frac{1}{1+\tilde{n}_\lambda^* \pi/\tilde{r}} \\
\tilde{L}_S^* &= \frac{(1-\nu^*)\tilde{n}_\lambda^*}{1+\nu^*\tilde{n}_\lambda^*} S^* \\
\tilde{L}_R^* &= \frac{\nu^*\tilde{n}_\lambda^*}{1+(1-\nu^*)\tilde{n}_\lambda^*} S^* \\
\tilde{M}^* &= \nu^*(1-\nu^*)\left(\frac{\tilde{n}_\lambda^{*2}}{1+\nu^*\tilde{n}_\lambda^*} + \frac{\tilde{n}_\lambda^{*2}}{1+(1-\nu^*)\tilde{n}_\lambda^*}\right) S^* \\
\varepsilon^* &= \frac{\theta(1-f^*)\tilde{M}^*}{\tilde{L}_S^* + (1-f^*)\tilde{M}^*} \\
(1-\nu^*)\frac{\tilde{n}_\lambda^*}{R_S} &= \tilde{L}_S^* + (1-f^*)\tilde{M}^* \\
\nu^*\frac{\tilde{n}_\lambda^*}{R_R} - (1-\nu^*)\frac{\delta((1-\eta)\varepsilon^*+\eta)\tilde{n}_\lambda^*}{R_S} &= \tilde{L}_R^* + f^*\tilde{M}^* \\
f^* &= (1-\varphi)\frac{1+\nu^*\tilde{n}_\lambda^*}{2+\tilde{n}_\lambda^*} + \varphi f_D.
\end{aligned}
\tag{C6}
$$

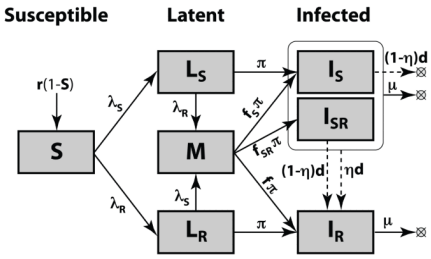

Figure 1. Model structure

Compartment and parameters are defined in Table 1. The dashed arrows show the effects of drug treatment. The population size is normalized to 1.

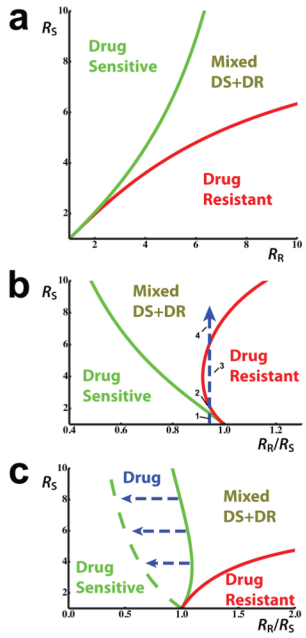

Figure 2. Coexistence analysis

The effect of (a) reproductive numbers, (b) relative fitness, and (c) drug treatment on the coexistence area of drug-sensitive and drug-resistant strains

a) The coexistence area as a function of DS and DR reproductive numbers (see equation (8)). The following parameters have been chosen for visualization here: φ=1 and f_D=1/2 (equivalent to f_S+f_{SR}=f=1/2); δ;=0; η=0; and r=0.2 π.

b) The coexistence area for the case where one of the strains has a higher competitive ability (see equation (10)). Here φ=1 and f_D=0.7 (equivalent to f_S+f_{SR}=0.3 and f=0.7); δ=0; η=0; and r=0.2 π. The blue dashed arrow shows the variation of the steady state frequencies of the strains in a hypothetical scenario of increasing prevalence. In this case, a DR strain with a lower effective reproductive number, but higher within-host competitive ability, may 1) be excluded from the population, 2) coexist with the DS strain, 3) exclude the DS strain, 4) and again coexist with the DS strain as prevalence increases.

c) Drug treatment may expand the area of coexistence by selecting for minority DR strains in those progressing to sensitive disease from mixed infection (the flow I_{SR}→I_R on figure 1, see also equation (11)). The solid lines are the borders of coexistence area for $\theta\delta$=0, whereas the dashed lines show $\theta\delta$=1. The other parameters are: φ=1 and f_D=0.3 (equivalent to f_S+f_{SR}=0.7 and f=0.3); η=0; and r=0.2 π.

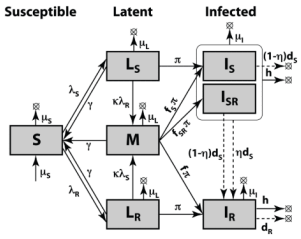

Figure C1. Expanded model structure

Additional parameters for this model (beyond the model shown in Figure 1) include κ (susceptibility to reinfection compared to infection); μ_S, μ_L, and μ_I (mortality rates for Susceptible, Latent and Infectious compartments, respectively); γ (rate of reversion from latency); h (rate of TB self-cure); d_S and d_R (drug cure rates for drug-sensitive and drug-resistant TB).

Table 1

State variables and parameters

a) Model parameters that must be provided. We define logical ranges and identify values used to demonstrate model behavior in Figure 2.
b) Model variables that characterize the epidemiologic state of the system. The final column references the equations that allow us to determine equilibrium values of variables.
c) Combined parameters that are combinations of the variables and parameters of Table 1a–b. The corresponding expressions are provided in final column.

a) Independent Parameters

Parameter	Description	Logical range	Selected for demonstration
R_S	Basic reproductive number of DS strain including the effect of antibiotics (referred in the text as "reproductive number of DS strain")	[0..∞)	[1..10]
R_R	Basic reproductive number of DR strain including the effect of antibiotics (referred in the text as "reproductive number of DR strain")	[0..∞)	[1..10]
φ	Fraction of the outcome of within-host competition that is dependent on strain-specific factors	[0..1]	0, 1
f_D	Probability for a DR strain to outcompete a DS strain within a co-infected host if the outcome depends entirely on strain-specific factors	[0..1]	0.3, 0.5, 0.7
θ	Probability of DR strain to persist in a host if outcompeted by DS strain	[0..1]	$\theta\delta$ = 0, 1
δ	Proportion of TB cases receiving antibiotic treatment (i.e. treatment coverage)	[0..1]	$\theta\delta$ = 0, 1
η	Probability of acquiring drug resistance on treatment[1]	[0..1]	$\eta\delta$ = 0
r	Susceptible recruitment rate[2]	(0..1) π	0.2 π
π	Progression rate from Latent infection to TB disease[3]		
μ	Rate of self-cure from TB[3]		

b) Variables of Epidemiologic State

Variable[4]	Description	Equation(s), determining the value in Equilibria
S	Frequency of Susceptibles (S=1 for disease-free equilibrium)	A8; A9; A12; A15
L_S	Frequency of Latent infection with DS strain only	A8; A9; A12; A15
L_R	Frequency of Latent infection with DR strain only	A8; A9; A12; A15
M	Frequency of Mixed Latent infection with both DS and DR strains	A8; A9; A12; A15
n_λ	The expected number of reinfections the host will experience during latency period under the current infectious force (referred in the text as "number of reinfections").	7 or A9; 8 or A12; A16
v	Fraction of total infection force due to the DR strain	A16
ε	Fraction of those with DS TB that have some remaining DR bacteria	A15
f	Overall probability that DR strain to outcompete a DS strain within a co-infected host	4

c) Combined Parameters

Parameter[4]	Description	Combination
d	Rate at which those with TB are detected and treated	$\mu\delta(1-\delta)$
f_N	Probability that DR strain to outcompete a DS strain within a co-infected host if the outcome does *not* depend on strain-specific factors	$(f-\varphi f_D)/(1-\varphi)$
f_S	Probability that host with mixed infection develops DS TB with *no* remaining DR bacteria	$(1-\theta)(1-f)$
f_{SR}	Probability that a host with mixed infection develops DS TB with *some* remaining DR bacteria	$\theta(1-f)$

c) Combined Parameters

Parameter[4]	Description	Combination
λ_R	Force of infection due to the DR strain	$vn_\lambda\pi$
λ_S	Force of infection due to the DS strain	$(1-v)n_\lambda\pi$
I_S	Frequency of DS TB without subpopulation of DR bacteria	$(1-\varepsilon)\lambda_S/(\mu+d)\,R_S$
I_{SR}	Frequency of DS TB with a subpopulation of DR bacteria	$\varepsilon\lambda_S/(\mu+d)\,R_S$
I_R	Frequency of DR TB	$\lambda_R/\mu R_R$

[1] The value of the acquired resistance is set to zero for the results in Figure 2 in order to show the DR strain invasion threshold

[2] $r<\pi$, as the duration of latency cannot exceed the lifespan

[3] The value of the parameter is not needed for the results shown in Figure 2

[4] The equilibrium values for the dynamic variables or their combinations are denoted by corresponding superscripts: ***0** (disease free), ***S** (DS only), ***R** (DR only), and ***** (mixed-strain)