THE ROYAL SOCIETY

doi 10.1098/rspb.2000.1328

Will tuberculosis become resistant to all antibiotics?

Christopher Dye* **and Marcos A. Espinal**

Communicable Diseases, World Health Organization, 1211 Geneva 27, Switzerland

The discovery of high prevalences of antibiotic resistance in some pathogens, in some parts of the world, has provoked fears of a widespread loss of drug efficacy. Here, we use a mathematical model to investigate the evolution of resistance to four major anti-tuberculosis drugs (isoniazid, rifampicin, ethambutol and streptomycin) in 47 sites around the world. The model provides a new method of estimating the relative risk of treatment failure for patients carrying drug-resistant strains and the proportion of patients who develop resistance after failing treatment. Using estimates of these two quantities together with other published data, we reconstructed the epidemic spread of isoniazid resistance over the past 50 years. The predicted median prevalence of resistance among new cases today was 7.0% (range 0.9–64.3%), close to the 6.3% (range 0–28.1%) observed. Predicted and observed prevalences of resistance to isoniazid plus rifampicin (multidrug-resistant or MDR-TB) after 30 years of combined drug use were also similar, 0.9% (0.1–5.9%) and 1.0% (range 0–14.1%), respectively. With current data, and under prevailing treatment practices, it appears that MDR-TB will remain a localized problem, rather than becoming a global obstacle to tuberculosis control. To substantiate this result, further measurements are needed of the relative fitness of drug-resistant strains.

Keywords: tuberculosis; drug resistance; MDR-TB; isoniazid; rifampicin; mathematical modelling

1. INTRODUCTION

Antibiotic resistance threatens to compromise the treatment of 16 million prevalent tuberculosis (TB) cases around the world and to hinder efforts to cut the annual death toll of two million people (Pablos-Mendez *et al.* 1998; Espinal *et al.* 2000*a*,*b*). Expressions of concern and calls for action are running some way ahead of a sufficient understanding of why resistance is spreading. Observations have been made in a few countries, for a few years, on resistance trends, but it remains unclear under what circumstances the prevalence of drug resistance will continue to increase, and to what levels. Our goal in this paper is to define these circumstances more precisely, and hence to specify more accurately epidemiological criteria for the control of drug-resistant TB. The spotlight is on TB, but we make use of fundamental evolutionary principles, taking an approach that could be applied to a wide range of infectious agents.

We first make use of new data from 47 sites around the world (Espinal *et al.* 2000*b*) to evaluate two quantities that influence the rate of spread of resistance. They are (i) the relative risk of treatment failure for patients carrying drug-resistant strains as compared with drug-susceptible strains; and (ii) the proportion of patients who develop resistance on failing treatment (the 'amplifier effect'; Farmer & Kim 1998). Mathematical modelling shows how these two quantities can be derived from the relationship between measured resistance prevalences in new and previously treated TB cases. The second part of the paper examines the epidemiological significance of these parameter estimates by combining them in a full transmission model, and by comparing observed and expected prevalences of drug resistance around the world.

Our analysis leads to the proposition that the global average prevalence of anti-TB drug resistance will remain low. This prediction from modelling, based on new data, is more sanguine than previous analyses (Blower *et al.* 1996; Dye & Williams 2000), and consistent with observed resistance prevalences around the world. It suggests that drug-resistant *Mycobacterium tuberculosis* (MTB), and particularly multidrug-resistant strains of MTB, are not in the process of replacing drug-susceptible strains worldwide.

2. METHODS

(a) *Treatment failure and the selection of drug-resistant strains*

During an outbreak of drug-resistant TB, we first see resistance in patients who have failed treatment. Strains of drug-resistant bacilli, which have arisen by mutation, have an advantage over drug-susceptible strains when the dosage of one or a combination of drugs is too low, i.e. when the treatment regimen is less than recommended (Crofton *et al.* 1997). Patients may receive inadequate regimens because prescriptions are wrong, because they default from treatment or because they take drugs erratically. Pulmonary TB patients carrying resistant bacilli, and with open cavities, can transmit these bacilli to other people who have never before had the disease.

Surveys of antibiotic resistance typically provide data describing the resistance prevalences among new and previously treated cases (Espinal *et al.* 2000*b*). Previously treated resistant cases (P', where the prime denotes 'drug resistant') come from four sources: new resistant cases (I') who have failed treatment; previously treated resistant cases (P') who have failed treatment again; new drug-susceptible cases (I) who have both failed treatment and acquired, through mutation and selection, resistant strains of bacilli; and previously treated drug-susceptible cases (P) who have again failed treatment and become resistant. Thus, the incidence rate of P' (written as $\Delta P'$) after a cycle of treatment, failure and re-treatment (i.e. over the interval t to $t+1$) is

$$\Delta P' = \delta(1-\kappa')(I'_t + P'_t) + \delta(1-\kappa)\rho(I_t + P_t), \quad (1)$$

*Author for correspondence (dyec@who.int).

Received 7 June 2000 *Accepted* 14 September 2000

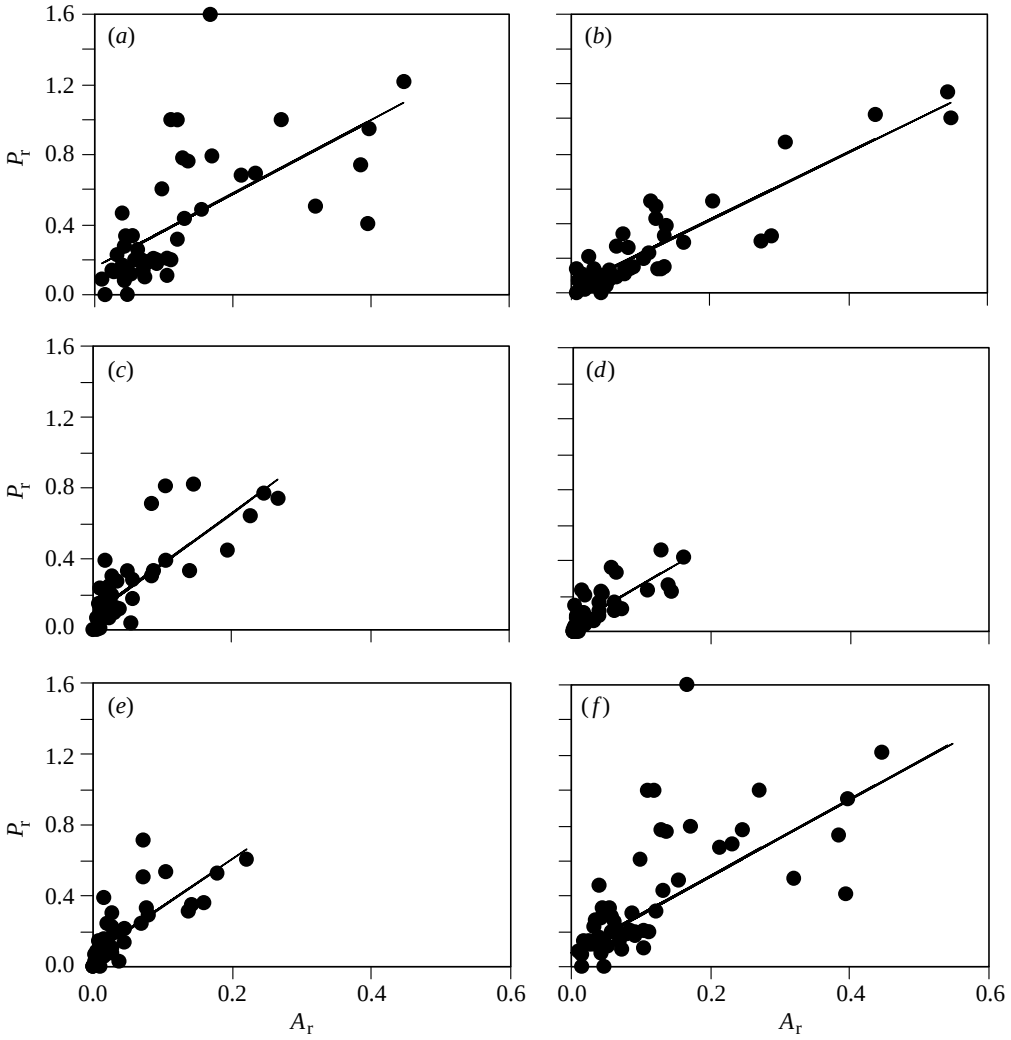

Figure 1. Regressions of P_r on A_r for the four principal anti-tuberculosis drugs, separately, (*a*) isoniazid, (*b*) streptomycin, (*c*) rifampicin, (*d*) ethambutol, and combined (*f*), plus MDR (*e*). Data from 47 countries and areas within countries are from Espinal *et al*. (2000*b*).

where δ is the proportion of cases detected and hence treated per unit time, κ is the proportion of drug-susceptible cases who are cured, κ' is the proportion of drug-resistant cases cured and ρ is the proportion of patients who become resistant to any given drug on failing treatment (the proportion who acquire resistance). The incidence of previously treated drug-susceptible cases is

$$\Delta P = \delta(1-\kappa)(1-\rho)(1_t + P_t), \tag{2}$$

and the ratio of the two incidence rates, $\Delta P'/\Delta P = P_r$, is

$$P_r = \frac{\left[\dfrac{(1-\kappa')(I'_t + P'_t)}{(1-\kappa)(I_t + P_t)} + \rho\right]}{1-\rho}. \tag{3}$$

P_r can be measured from the ratio of incident drug-resistant to incident drug-susceptible cases, who have been previously treated. Similarly, $(I'_t + P'_t)/(I_t + P_t)$ (let this be A_r) can be measured from the ratio of all incident-resistant cases to all incident-susceptible cases, if the ratio of prevalences is the same as the ratio of incidences. This assumes that resistant and susceptible cases are detected at the same rate per capita, and hence exist for the same length of time. Finally, the ratio $(1-\kappa')/(1-\kappa)$ (let this be ϕ) can be interpreted as the relative risk of treatment failure among drug-resistant as compared with drug-susceptible cases. Equation (3) shows that parameters ϕ and ρ can be estimated from a linear regression of P_r on A_r:

$$P_r = a + bA_r, \tag{4}$$

where the intercept $a = \rho/(1-\rho)$ and the gradient $b = \phi/(1-\rho)$, so that $\rho = a/(1+a)$ and $\phi = b/(1+a)$. In this formulation, the regression variables are not completely independent—both include the number of previously treated cases—and will therefore be correlated. However, we are not concerned here with testing statistically for $b > 0$; rather, we want to be sure that b and a lead to unbiased estimates of ϕ and ρ. We performed a set of Monte Carlo simulations to confirm that b and a are indeed unbiased estimators. Our source of data for estimating ϕ and ρ is Espinal *et al*. (2000*b*), who provide matching data on P_r and A_r from 47 countries, or parts of countries, around the world.

(b) *Observed and expected prevalence of drug resistance in populations*

Equations (1) and (2) describe only part of the transmission cycle of drug-susceptible and drug-resistant TB. A model of the

Table 1. *Regression statistics (slope b and intercept a) for the data in figure 1, with estimates of parameters ϕ and ρ, and standard errors (s.e.)*

drug	*b* (s.e.)	*a* (s.e.)	ϕ(s.e.)	ρ(s.e.)
isoniazid	2.104 (0.380)	0.155 (0.062)	1.82 (0.34)	0.134 (0.048)
rifampicin	2.849 (0.296)	0.087 (0.025)	2.62 (0.28)	0.080 (0.021)
ethambutol	2.278 (0.256)	0.040 (0.013)	2.19 (0.25)	0.038 (0.012)
streptomycin	1.946 (0.135)	0.031 (0.023)	1.89 (0.14)	0.030 (0.022)
all four drugs	2.167 (0.130)	0.079 (0.017)	2.01 (0.12)	0.073 (0.015)
MDR	2.678 (0.329)	0.075 (0.021)	2.49 (0.31)	0.070 (0.018)

full cycle, modified from Dye & Williams (2000), is in Appendix A. Parameter estimates for the model, with ranges and distributions, are as previously described, except for the following new information (see also table A1).

Two recent studies of the molecular epidemiology of TB clusters (two or more cases with identical restriction fragment length polymorphism patterns, presumed to be related by transmission) provide estimates of the relative fitness (the ratio of basic case reproduction numbers) of drug-resistant as compared with drug-susceptible TB. First, D. van Soolingen and collaborators (personal communication) found that MTB strains carrying the AA315 mutation, which is associated with isoniazid resistance, did not generate significantly fewer clusters of secondary cases than isoniazid-susceptible strains (relative risk, RR = 0.8, 95% CI 0.6–1.2). They previously found that isoniazid-resistant strains did generate fewer clusters of cases (RR = 0.7, 95% CI 0.5–0.9; Van Soolingen *et al.* 1999), though results of the two studies are not significantly different. With this uncertainty, we allow the relative fitness of izoniazid-resistant strains to vary uniformly between 0.5 and 1.0. In the second study, Garcia-Garcia *et al.* (2000) reported that cases resistant to both isoniazid and rifampicin (MDR-TB) generated only 16% (95% CI 4–60%) as many clusters of cases as drug-susceptible strains. Based on this result, we allow the relative fitness of MDR strains to vary uniformly between 0.04 and 0.6, where the uniform distribution allows maximal uncertainty from minimal data. These estimates of relative fitness are based on the relative numbers of clusters of cases generated by different strains. They are not directly dependent on the absolute number of clusters observed in these studies, which is sensitive to the completeness of sampling schedules (Glynn *et al.* 1999).

The relative fitness of drug-resistant strains must be calculated over the entire life cycle of MTB, including transmissibility, the duration of infectiousness and human susceptibility to disease. Clustering studies do sample over a full generation of MTB. For mathematical modelling, we adjusted the relative fitness by changing the transmission or contact rate, c, with which the basic case reproduction number, R_0, varies in direct proportion (see Appendix A).

To reconstruct epidemics of drug resistance over the past few decades, we first used equations (A1)–(A8) (with time-step 0.1 year) to establish a model population with stable, drug-susceptible disease only. Case detection (δ) and cure rates (κ) were set at 50% per year and 70%, respectively, generating an equilibrium prevalence of infection of *ca.* 30%, and a steady incidence rate of new infectious cases of *ca.* 50 per 100 000 per year. These rates are typical of the many highly endemic countries (Dye *et al.* 1999). Cases of either isoniazid resistance or MDR-TB were then introduced to this population by allowing resistance to arise among treatment failures ($\rho > 0$, estimated as above). The epidemic of isoniazid resistance has been underway for 50 years (since the drug was first introduced) and the epidemic of MDR-TB resistance for 30 years (since rifampicin was introduced). We compared the predicted prevalences of isoniazid resistance and MDR-TB after 50 and 30 years with the prevalences observed by Espinal *et al.* (2000*b*). Uncertainty and sensitivity analyses were carried out by Monte Carlo simulation using Palisade *@Risk* (Dye & Williams 2000) and parameter values with ranges in table A1.

3. RESULTS

(a) *Treatment failure and the selection of drug-resistant strains*

Regressions of P_r on A_r for isoniazid, rifampicin, ethambutol and streptomycin, and for MDR, are shown in figure 1. The regression statistics and the resulting estimates of parameters ϕ and ρ are in table 1.

Figure 1 and table 1 together illustrate four main results. First, the range of ratios of resistant to susceptible cases is greatest for isoniazid and streptomycin. These are the drugs that have been used longest—streptomycin since the 1940s and isoniazid since the 1950s. Ethambutol (1960s) was introduced before rifampicin (1970s), but has been used less intensively. Second, estimates of ϕ are greater than one, indicating, as expected, that treatment failure has been more frequent for drug-resistant as compared with drug-sensitive strains. Third, the largest estimate of ϕ is for rifampicin (table 1). Rifampicin resistance is strongly associated with MDR, for which ϕ is also relatively high (2.49). Fourth, all estimates of ρ are positive, and all are significantly greater than zero ($t \geqslant 2.8$, $p < 0.01$), with the exception of that for streptomycin. On average, 7.3% of treatment failures develops resistance ($\rho = 0.073$). The other estimates of ρ imply that isoniazid resistance is more easily selected than resistance to other drugs when treatment fails, but the differences between drugs are not statistically significant.

(b) *Observed and expected prevalence of drug resistance in populations*

We can explore the epidemiological significance of estimates for ϕ and ρ by putting them in the full model (see Appendix A), and then solving to forecast the spread of resistance from the point at which drugs were introduced. A complete set of parameter estimates is available for isoniazid-resistant and MDR strains.

Before making comparisons of observed and expected prevalences of drug resistance (the fraction of incident cases that is drug resistant), we first make some general

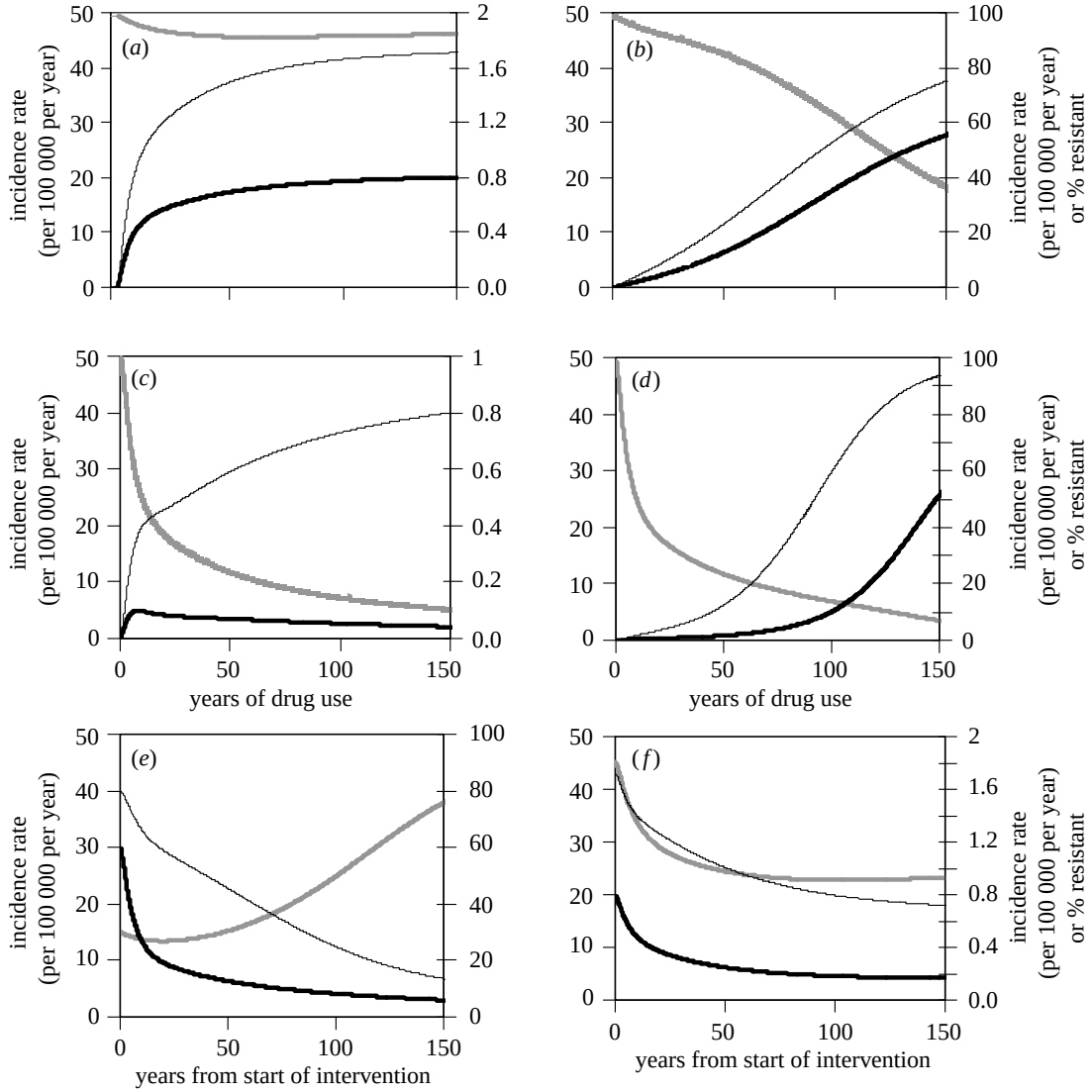

Figure 2. Possible trajectories in the numbers of drug-susceptible (grey lines, left axis) and MDR-TB cases (heavy lines, right axis), and in the fraction of cases resistant (light lines, right axis). (*a–d*) Dynamics of the epidemic of drug resistance. (*a*) Poor control of drug-susceptible cases ($\delta = 0.5$, $\kappa = 0.7$), low transmissibility of resistant strains ($c'/c = 0.4$); (*b*) as (*a*) but high transmissibility of resistant strains ($c'/c = 0.9$); (*c*) good control of drug-susceptible cases ($\delta = 0.7$, $\kappa = 0.9$), low transmissibility of resistant strains ($c'/c = 0.4$); (*d*) as (*c*) but high transmissibility of resistant strains ($c'/c = 0.9$). (*e*,*f*) Interventions to cut the prevalence of drug resistance. (*e*) Good control of resistant strains ($\delta' = 0.5$, $\kappa' = 0.9$), with high transmissibility ($c'/c = 0.9$); (*f*) good control of susceptible strains ($\delta = 0.7$, $\kappa = 0.9$), low transmissibility of resistant strains ($c'/c = 0.4$). Other parameters as for MDR-TB.

observations on the relationship between prevalences and numbers of resistant cases. When a new drug is used to treat TB patients, both the number of resistant cases and the prevalence of resistance will at first increase, provided some patients fail treatment (figure 2*a–d*). What happens thereafter depends on the absolute and relative magnitudes of the basic case reproduction numbers of drug-susceptible (R_{0S}) and drug-resistant disease (R_{0R}). Consider a new drug that is used poorly so that case detection and cure rates remain unaltered after its introduction. If $R_{0S} > 1$ and $R_{0S} > R_{0R}$, the incidence and prevalence of resistance increase to a steady state in which fewer than 100% of new cases are resistant (figure 2*a*). If resistant strains have the biggest reproduction number ($R_{0R} > R_{0S} > 1$), drug-resistant disease will eventually replace drug-susceptible disease altogether (figure 2*b*). Expectations for the performance of a new drug would, however, be much higher. If the case detection and cure rates of drug-susceptible disease rise so that R_{0S} falls below 1, the incidence of drug-susceptible disease will decline towards elimination. If $R_{0R} < 1$ too, because, for example, the resistant strain has low transmissibility, the number of resistant cases will first rise and then fall (figure 2*c*). But if the resistant strain is highly transmissible ($R_{0R} > 1$ while $R_{0S} < 1$), its incidence will increase as drug-susceptible disease proceeds to extinction (figure 2*d*). With these various combinations of reproduction numbers the prevalence of drug resistance always increases through time, but the numbers of resistant cases may go up and/or down.

The prevalence of resistance will only decline if a control programme succeeds in cutting the incidence of drug-resistant cases more than the incidence of drug susceptibles. This can be done by intervening during an

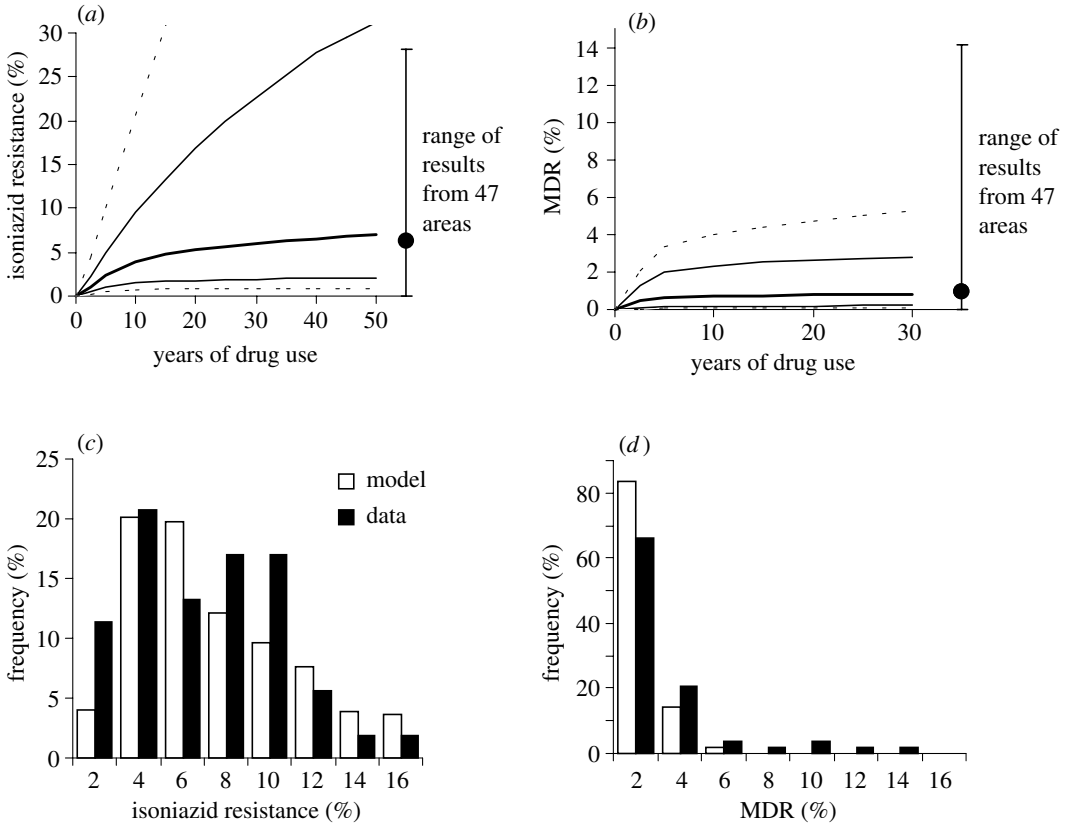

Figure 3. Predicted and observed changes in the proportions of new cases that are isoniazid resistant (*a*,*c*) and MDR (*b*,*d*). (*a*,*b*) Trajectories of the median epidemics (thick lines) for (*a*) isoniazid and (*b*) MDR, with fifth and 95th centiles (thin lines), and lower and upper limits (dotted lines), derived from 1000 Monte Carlo simulations. The data are summarized as medians (solid dot), with lower and upper limits (±bars), taken from the same source as figure 1. (*c*,*d*) Frequency distributions of predicted (open bars) and observed (filled bars) fractions of cases resistant to (*c*) isoniazid and (*d*) MDR after 50 and 30 years of drug use, respectively.

epidemic to cure a higher proportion of drug-resistant cases, thereby preventing transmission (figure 2*e*). It can also be done by curing a higher proportion of drug-susceptible cases (preventing the development of resistance), but only if R_{0R} is not too large (figure 2*f*).

The model forecasts a median prevalence of isoniazid resistance of 7.0% (range 0.9–64.3%) after 50 years of drug use. This is about the same as the observed median of 6.3% (0–28.1%) reported from the 47 sites described by Espinal *et al.* (2000*b*) (figure 3*a*,*c*). The forecast for MDR-TB is 0.9% (0.1–5.9%) after 30 years of drug use, close to the observed median of 1.0% (0–14.1%) (figure 3*b*,*d*). Most strikingly, the forecasts in figure 3 suggest that, on average, the prevalences of isoniazid resistance and MDR have already been close to saturation for decades and are likely to remain below 5%. The model is less likely, *a priori*, to represent accurately variation around the median; nonetheless, the frequency distributions of observations were similar for model and data (figure 3*c*,*d*).

However, in some countries (and parts of countries) the prevalence of MDR is much higher than 5%, e.g. 11% in Henan Province, China, 14% in Estonia, 9% in Latvia and 9% in Ivanovo Oblast, Russia (Espinal *et al.* 2000*b*). These rates exceed the upper bounds of model forecasts in figure 3 and require explanation.

There are three main reasons why MDR rates can be high. A sensitivity analysis of the results in figure 3 shows

Table 2. *Sensitivity analysis of results in figure 3: parameters most responsible for variation in the prevalence of isoniazid resistance and MDR among new TB cases, as measured by partial rank correlation coefficients (PRCC)*

parameter	PRCC (MDR)	PRCC (isoniazid)
c'/c	0.847	0.456
κ	−0.356	−0.486
ρ	0.322	0.665
p	0.141	0.073
δ	0.108	0.192
ν_f	0.104	0.154
c	0.041	0.087
ϕ	0.030	0.035

that rates of isoniazid resistance after 50 years, and of MDR after 30 years, are most influenced by uncertainty in the relative transmission rate of drug-resistant strains, c'/c, the proportion of drug-susceptible cases cured, κ, and the proportion acquiring resistance, ρ (table 2). The relative transmission or contact rate is the most important of these parameters because of its influence on the basic case reproduction number, R_0. If c'/c is about 0.5, a high estimate according to Garcia-Garcia *et al.* (2000), the MDR rate saturates at under 5% (figure 4*a*). But if c'/c is close to 1, MDR is still spreading quickly after 30 years, and would eventually replace all other strains of *M. tuberculosis*.

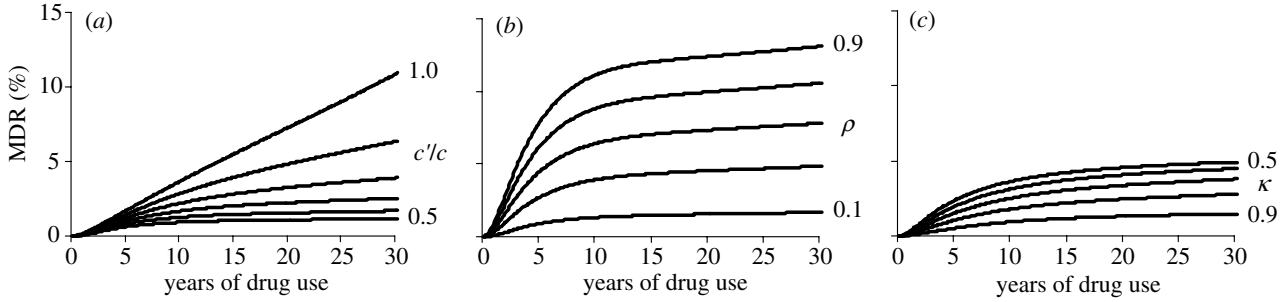

Figure 4. Growth of MDR-TB among new TB cases in relation to (*a*) the relative transmissibility of drug-resistant strains (c'/c varying from 0.5 to 1.0, in steps of 0.1), (*b*) the proportion of treatment failures that develops resistance (ρ from 0.1 to 0.9 in steps of 0.2), (*c*) the cure rate of drug-susceptible disease (κ from 0.5 to 1.0 in steps of 0.1). In all cases, $\delta = 0.7$ and $\kappa = 0.8$, so that $R_0 < 1$ for drug-susceptible disease, which is therefore in long-term decline. The relative transmissibility of drug-resistant strains, c'/c, is 0.7 in (*b*) and (*c*).

The effect of increasing ρ is quite different: the MDR rate saturates below 15%, even with ρ as high as 0.9 (figure 4*b*). A high proportion acquiring resistance has the same effect as a low cure rate of drug-susceptible cases; in figure 4*c*, the MDR rate reaches a maximum of less than 15%, even with κ as low as 0.6.

4. DISCUSSION

Our main result, based on the newest data and prevailing treatment practices, is that TB resistant to isoniazid, or to isoniazid plus rifampicin (MDR-TB), is likely to remain at low levels in most parts of the world. This is consistent with the findings of the most recent global review, in which significant upward trends in isoniazid resistance were reported in only four out of 25 countries studied, and a rise in the prevalence of MDR in only one country, Estonia (Espinal *et al.* 2000*b*). The predicted, stable coexistence between drug-resistant and drug-susceptible disease is also consistent with previous theoretical work (Castillo-Chavez & Feng 1997). Our forecasts suggest that drug-resistant TB, and particularly multidrug-resistant strains of TB, are not in the process of replacing drug-susceptible strains globally. Although TB cases are continuously emigrating from countries where the prevalence of drug resistance is high (and we have not investigated the consequences of such movement), it seems unlikely that MDR-TB, if left untreated, will 'affect tens of millions worldwide' (Heymann *et al.* 1999).

However, we need more information to be confident about this prediction. It is vital to understand why the prevalence of drug resistance has already reached exceptionally high levels in some countries, such as the Baltic States and parts of India, Russia and China. In the present analysis, the most important unknown is the relative fitness (here represented by relative transmissibility or contact rate, c'/c) of resistant as compared with susceptible strains. If the prevalence of resistance and the number of resistant cases are increasing because resistant strains are almost as transmissible as susceptible strains ($c' \approx c$), then the task is to interrupt a sustainable transmission cycle. In the language of deterministic epidemic theory, the goal is to reduce the basic case reproduction number of drug-resistant disease below 1 ($R_{0R} < 1$). Our previous analysis of the drug-resistance problem (Dye & Williams 2000) allowed for the possibility that MDR strains do indeed have high relative fitness (so that $R_{0R} > 1$). Erring towards caution, we concluded that controlling the epidemic would require higher cure rates for resistant cases. The method of control for MDR-TB would be to draw on the small armoury of expensive and relatively toxic second-line drugs.

This new analysis, based on new data, offers a more benign view of the epidemic of antibiotic resistance. In particular, if MDR strains have low relative fitness (Garcia-Garcia *et al.* 2000) we must look for other explanations for high resistance prevalences. The main contenders are that the cure rate of drug-susceptible disease (and hence, probably, of drug-resistant disease) has been low (small κ), or that treatment failure readily selects for resistance (large ρ). In either case, the principal remedy will be to use first-line drugs to their full potential. Short-course chemotherapy can cure over 85% of pan-susceptible cases and up to 80% of cases resistant to either isoniazid or rifampicin (Espinal *et al.* 2000*a*). A high cure rate for drug-susceptible cases, coupled with a high rate of case detection, will force down the incidence of drug-resistant cases and, provided R_{0R} is not large, the prevalence of drug resistance among all cases. Under these circumstances, second-line drugs would be needed as a back-up to cure individual drug-resistant cases, but not to contain the epidemic.

We have proposed in this paper a new method for estimating both ρ, the fraction of patients who acquire drug resistance after failing treatment, and ϕ, the relative risk of treatment failure in patients with drug-resistant disease. The method is indirect, but has some advantages over cohort studies: information is already freely available for many patients in many parts of the world; and the estimates reflect the outcome of common treatment practices rather than the results of tightly controlled trials. Moreover, simulation studies suggest that the method gives unbiased estimates of ρ and ϕ. One disadvantage is that parameter estimates are averages, which conceal important variation, for example, between countries.

There are presently few direct estimates of the relative treatment failure (ϕ) and the proportion of cases acquiring resistance (ρ) that can be compared with the indirect estimates presented here. Espinal *et al.* (2000*b*) reported average relative failure rates for isoniazid, rifampicin and MDR-TB of 1.4, 1.6 and 2.1, as compared

with outcomes for drug-susceptible cases. These are similar to our own indirect estimates of 1.1, 2.1 and 2.2.

By contrast, direct estimates of the proportion of patients becoming resistant do not always correspond with our indirect estimates. L. Rigouts and colleagues (personal communication) found in a small study that ten out of 12 TB patients carrying isoniazid-resistant strains also acquired rifampicin resistance after failing treatment, thereby becoming multidrug resistant. This gives $\rho = 0.83$, which is more than ten times higher than our estimate of $\rho = 0.070$. A study by the Indian Council for Medical Research (personal communication) also obtained a high value of $\rho = 32/60 = 0.53$ for the proportion of cases that fail treatment and acquire MDR-TB (rifampicin resistance adding to isoniazid resistance). But estimates of ρ for isoniazid and rifampicin resistance in patients carrying fully susceptible strains were only $5/25 = 0.2$ and $1/25 = 0.04$, respectively, which are close to our own estimates of 0.13 and 0.08. The rate of development of monoresistance may be relatively low because patients are more likely to fail for reasons other than drug resistance. Discrepancies in the comparisons for MDR remain to be explained. The differences could be resolved by testing the assumptions of our model (e.g. that the detection and treatment rates of drug-susceptible and drug-resistant cases are the same) and by making further direct observations on the proportion of patients who acquire resistance having failed treatment.

Last, our calculations warn against defining the course of drug-resistance epidemics, or progress in their control, solely in terms of prevalence (i.e. the fraction of cases that is resistant, often called a 'rate'). It has been noted that the prevalence of resistance can remain high, even in countries that have achieved high rates of case detection and cure (Horsburgh 2000). But we have shown here how, in principle, the prevalence of resistance can go up while, more importantly, the number of resistant cases is going down.

In conclusion, the crux of the drug-resistance problem is to determine whether drug-resistant strains, including MDR-TB, can persist in self-sustaining transmission cycles. If they cannot, as suggested by this analysis, containment will require high cure rates for susceptible cases only. If they can, containment will require high cure rates for resistant cases too. Molecular epidemiological studies of relative fitness (Van Soolingen *et al.* 1999; Garcia-Garcia *et al.* 2000) are one approach to finding out. Another is to closely observe control programmes that have adopted best practices in standard short-course chemotherapy, without introducing second-line drugs to treat MDR-TB cases (as recommended by Crofton *et al.* 1997), to see if the incidence of drug-resistant as well as drug-susceptible TB is in decline.

Thanks to Paul Nunn and Mario Raviglione for innumerable discussions on antibiotic resistance, and for helpful comments on a draft.

APPENDIX A. MATHEMATICAL MODEL FOR THE EVOLUTION OF DRUG RESISTANCE

The following system of difference equations, simplified from Dye & Williams (2000), describes the dynamics of infection and drug-susceptible pulmonary disease in adults.

Uninfected proportion:

$$S_{t+1} = S_t - c(I_t + P_t)S_t + \mu(1 - S_t) + \mu_i I_t + \mu_{if} P_t. \quad \text{(A1)}$$

Latent (slow breakdown to disease):

$$L_{s,t+1} = L_t + c(1-p)(I_t + P_t)S_t - (cpx(I_t + P_t) + \nu_s + \mu)L_{s,t}. \quad \text{(A2)}$$

Latent (fast breakdown to disease):

$$L_{f,t+1} = L_{f,t} + cp(I_t + P_t)S_t + cpx(I_t + P_t)L_{s,t} - (\nu_f + \mu)L_{f,t}. \quad \text{(A3)}$$

New infectious:

$$I_{t+1} = I_t + f(\nu_s L_{s,t} + \nu_f L_{f,t} + r_n C_{n,t}) - (\delta + \mu + \mu_i + n)I_t. \quad \text{(A4)}$$

Non-infectious:

$$N_{t+1} = N_t + (1 - f)(\nu_s L_{s,t} + \nu_f L_{f,t}) - \mu N_t. \quad \text{(A5)}$$

Previously treated, infectious:

$$P_{t+1} = P_t + \delta(1 - \kappa)(I_t + P_t) - (\delta + \mu + \mu_{if})P_t. \quad \text{(A6)}$$

Self-cured:

$$C_{n,t+1} = C_{n,t} + nI_t - (r_n + \mu)C_{n,t}. \quad \text{(A7)}$$

Cured (by treatment):

$$C_{t+1} = C_t + \delta\kappa(I_t + P_t) - \mu C_t. \quad \text{(A8)}$$

Definitions are in table A1. If there are two kinds of pathogen instead of one—drug resistant as well as drug susceptible—they will interact, and the above model needs to be expanded as follows. Drug-resistant cases first arise through treatment failure. Thus, incidence rate $\delta(1 - \kappa)(I_t + P_t)$ in equation (A6) is actually the sum of $\delta(1 - \kappa)\rho(I_t + P_t)$ and $\delta(1 - \kappa)(1 - \rho)(I_t + P_t)$, as shown in equations (1) and (2). Treatment failures carrying resistant strains can transmit to uninfected people in equation (A1), who will be challenged at rate $c'(I'_t + P'_t)S_t$, as well as at rate $c(I_t + P_t)S_t$. Moreover, infected individuals in equation (A2) can be reinfected at rate $c'px(I'_t + P'_t)L_{s,t}$, from which follows rapid breakdown to disease (equation (A3)). Individuals infected with drug-resistant strains can, reciprocally, be reinfected with drug-susceptible strains. The relative fitness of resistant as compared with susceptible strains is represented in this analysis by c'/c, i.e. through differential transmission; the breakdown rates from infection to disease, ν_f and ν_s, are assumed to be the same for susceptible and resistant strains. The dynamics of non-infectious disease are represented very simply in equation (A5) (e.g. excluding treatment) because our focus here is on transmission and the rate of spread of drug resistance, and not on the total TB burden.

The basic case reproduction number, R_0, is the number of secondary infectious cases arising from one primary case introduced into a fully susceptible population (Anderson & May 1991). This quantity is central to our analysis because it determines whether or not drug-resistant TB is maintained by a self-sustaining transmission

Table A1. *Definitions and values of parameters and variables in model (A1)–(A8)*

parameter	definition	lower	mode or point[a]	upper
c	per-capita contact rate of infectious, drug-susceptible cases per year	7.0	9.8	12.6
c'/c	relative contact rate of infectious, MDR-TB (upper) or isoniazid-resistant cases (lower)	0.04 0.5	0.32[a] 0.75[a]	0.6 1.0
x	fraction of reinfected persons who develop TB at rate ν_f	0.1	0.35	0.6
p	fraction of newly infected persons who develop primary progressive TB	0.08	0.14	0.25
f	fraction of TB cases that are infectious	0.5	0.65	0.65
n	per-capita rate at which infectious cases self-cure per year	0.15	0.2	0.25
r_n	per-capita rate of relapse to infectious TB following self-cure per year	0.02	0.03	0.04
μ	per-capita death rate from causes other than TB per year	—	0.015	—
μ_i	per-capita death rate from untreated TB per year; μ_{if}, for treatment failures, assumed to be the same	0.2	0.3	0.4
ν_f	per-capita rate of breakdown to progressive primary TB per year	0.76	0.88	0.99
ν_s	per-capita rate of breakdown to TB by endogenous reactivation per year	0.0001	0.000 11	0.0003
δ	per-capita detection and treatment rate of new and previously treated cases per year, drug susceptible or drug resistant	0.4	0.5[a]	0.6
κ	proportion of drug-susceptible cases cured; κ' set to 62% of these values	0.6	0.7[a]	0.8

[a] Midpoints of ranges, with values uniformly distributed across ranges in uncertainty and sensitivity analyses. Other central measures are modes of triangular distributions. Rates are given in units of 'per year'.

cycle. For a population obeying a system of equations such as (A1)–(A8), $R_0 = bc\tau$, where b is the proportion of infections that leads to infectious cases and τ is the average duration of infectiousness, allowing for the effect of chemotherapy. Methods for calculating b and τ, both for drug-susceptible and for drug-resistant disease, are discussed by Dye & Williams (2000).

REFERENCES

Anderson, R. M. & May, R. M. 1992 *Infectious diseases of humans: dynamics and control*, 2nd edn. Oxford University Press.

Blower, S., Small, P. M. & Hopewell, P. C. 1996 Control strategies for tuberculosis epidemics: new models for old problems. *Science* **273**, 497–500.

Castillo-Chavez, C. & Feng, Z. 1997 To treat or not to treat: the case of tuberculosis. *J. Math. Biol.* **35**, 629–656.

Crofton, J., Chaulet, P. & Maher, D. 1997 *Guidelines for the management of drug-resistant tuberculosis*. Geneva: World Health Organization (WHO/TB/96.210).

Dye, C. & Williams, B. G. 2000 Criteria for the control of drug-resistant tuberculosis. *Proc. Natl Acad. Sci. USA* **97**, 8180–8185.

Dye, C., Scheele, S., Dolin, P., Pathania, V. & Raviglione, M. C. 1999 Global burden of tuberculosis: estimated incidence, prevalence and mortality by country. *J. Am. Med. Assoc.* **282**, 677–686.

Espinal, M., Kim, S. J., Suarez, P. G., Kam, K. M., Khomenko, A. G., Migliori, G. B., Baez, J., Kochi, A., Dye, C. & Raviglione, M. C. 2000*a* Standard short-course chemotherapy for drug resistant tuberculosis: treatment outcomes in six countries. *J. Am. Med. Assoc.* **283**, 2586–2612.

Espinal, M. (and 11 others) 2000*b* *Anti-tuberculosis drug resistance in the world*, report no. 2. Geneva: World Health Organization (WHO/CDS/TB/2000.278).

Farmer, P. & Kim, J. Y. 1998 Community-based approaches to the control of multidrug resistant tuberculosis: introducing 'DOTS-plus'. *Br. Med. J.* **317**, 671–674.

Garcia-Garcia, M. de L. (and 11 others) 2000 Clinical consequences an transmissibility of drug-resistant tuberculosis in southern Mexico. *Arch. Intern. Med.* **160**, 630–636.

Glynn, J. R., Vynnycky, E. & Fine, P. E. M. 1999 Influence of sampling on estimates of clustering and recent transmission of *Mycobacterium tuberculosis* derived from DNA fingerprinting techniques. *Am. J. Epidemiol.* **149**, 366–371.

Heymann, S. J., Brewer, T. F., Wilson, M. E. & Fineberg, H. V. 1999 The need for global action against multi-drug resistant tuberculosis. *J. Am. Med. Assoc.* **281**, 2138–2140.

Horsburgh Jr, C. R. 2000 The global problem of multidrug-resistant tuberculosis: the genie is out of the bottle. *J. Am. Med. Assoc.* **283**, 2575–2576.

Pablos Mendez, A. (and 10 others) 1998 Global surveillance for antituberculosis drug resistance, 1994–1997. *N. Engl. J. Med.* **338**, 1641–1649.

Van Soolingen, D., Borgdorff, M. W., de Haas, P., Sebek, M. M. G. G., Veen, J., Dessens, M., Kremer, K. & Van Embden, J. D. A. 1999 Molecular epidemiology of tuberculosis in the Netherlands: a nationwide study from 1993 through 1997. *J. Infect. Dis.* **180**, 726–736.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.