Tuberculosis epidemics driven by HIV: is prevention better than cure?

Christine S. M. Currie\textsuperscript{a}, Brian G. Williams\textsuperscript{b}, Russell C. H. Cheng\textsuperscript{a} and Christopher Dye\textsuperscript{b}

Objective: To compare the benefits of tuberculosis (TB) treatment with TB and HIV prevention for the control of TB in regions with high HIV prevalence.

Design and methods: A compartmental difference equation model of TB and HIV has been developed and fitted to time series and other published data using Bayesian methods. The model is used to compare the effectiveness of TB chemotherapy with three strategies for prevention: highly active antiretroviral therapy (HAART), the treatment of latent TB infection (TLTI) and the reduction of HIV transmission.

Results: Even where the prevalence of HIV infection is high, finding and curing active TB is the most effective way to minimize the number of TB cases and deaths over the next 10 years. HAART can be as effective, but only with very high levels of coverage and compliance. TLTI is comparatively ineffective over all time scales. Reducing HIV incidence is relatively ineffective in preventing TB and TB deaths over 10 years but is much more effective over 20 years.

Conclusions: In countries where the spread of HIV has led to a substantial increase in the incidence of TB, TB control programmes should maintain a strong emphasis on the treatment of active TB. To ensure effective control of TB in the longer term, methods of TB prevention should be carried out in addition to, but not as a substitute for, treating active cases.

\textit{AIDS} 2003, 17:2501–2508

Keywords: tuberculosis, HIV, epidemiology, mathematical modelling, Bayesian methods, antiretroviral therapy, DOTS, treatment of latent TB infection

Introduction

\textit{Mycobacterium tuberculosis} and HIV are the leading causes of death from infectious diseases among adults [1,2]. The spread of HIV infection has already led to a dramatic increase in tuberculosis (TB) in eastern and southern Africa [3] and threatens to do so elsewhere. The DOTS strategy of the World Health Organization (WHO) is the internationally recommended approach to TB control [3,4] based on diagnosis by smear microscopy, a mechanism to ensure drug supplies, a standardized recording and reporting system, standardized short-course chemotherapy with direct observa-

From the \textsuperscript{a}Faculty of Mathematical Studies, University of Southampton, Southampton, UK and \textsuperscript{b}Communicable Diseases, World Health Organization, Geneva, Switzerland.

Requests for reprints to: Dr C. Dye, Communicable Diseases, World Health Organization, 20 Avenue Appia, P Box PA/OMS/CDS Geneva, Switzerland.

Received: 6 September 2002; revised: 30 January 2003; accepted: 19 March 2003.

DOI: 10.1097/01.aids.0000096903.73209.ac
boosting patients’ immunity by treating them with highly active antiretroviral therapy (HAART) [6] or by treating latent TB infection (LTBI), usually through 6–9 months’ treatment with isoniazid [7]. Previous studies have attempted to calculate the number of TB cases and deaths that can be averted by finding and treating active TB during the course of HIV epidemics [8–10], but none has evaluated the curative approach against the three principal means of prevention. This study compares the effectiveness of preventive and curative methods for the control of TB in high HIV prevalence settings using a dynamical model of TB driven by HIV. Bayesian methods are used to fit the model to the available data on TB and HIV and to determine confidence limits (CL) for parameter estimates and projections. The goal is to compare the benefits of TB treatment with TB and HIV prevention and so provide practical guidance to those responsible for averting the 300,000 or more HIV-related TB deaths that occur each year [1].

Methods

Mathematical model of TB–HIV

An earlier compartmental model of TB–HIV epidemiology [8] was reduced to a single age class (adults 15–49 years), but the options for TB control were extended by including three preventive methods as well as case detection and cure. The model was written in Visual Basic and combined a dynamical model of TB progression with a statistical model of HIV prevalence. Figure 1 illustrates the general structure of the simulation model, which is described in full in the supplementary material (www.who.int/gtb and www.AIDSonline.com) where parameter values and data sources are also given.

The model is made up of two submodels, the first describing the transitions between states for individuals who are not infected with HIV, or who are in the early stages of HIV (stages 1 and 2 of the WHO staging system [11,12]) and the second describing transitions for those in the later stages of HIV. Active TB can arise through any of three mechanisms. Those who acquire a new TB infection either develop progressive primary disease within 1 year or enter a latent state from which TB can arise by reactivation or reinfection. A proportion of individuals who are latently infected can also develop TB within 1 year of reinfection. Active TB may be infectious or non-infectious. A separate HIV model determines the incidence of HIV in each time step and is used to determine the rate at which people move from one TB submodel to another. After approximately 4 years, the time lag between HIV infection and late-stage HIV (defined here as WHO stage 3 [12,13]), this number of individuals move from the first submodel to the corresponding state (TB uninfected, latently infected) in the second submodel. During the later stages of HIV, coinfection leads to a greatly increased risk of developing TB, though a smaller fraction of those with active TB become infectious. Individuals with late-stage HIV infections (WHO stages 3 and above) also have higher death rates, with and without active TB. Reasonably good data are available for the prevalence of HIV infection over time in the countries of interest but the future course of the epidemic is much less certain. The available data for the HIV prevalence over time was fitted to a double logistic function (see supplementary material at www.who.int/gtb and www.AIDSonline.com), which allows one to vary the initial rate of increase, the peak prevalence, the final steady-state prevalence and the rate of convergence to the steady state. The fitted prevalence is then used to estimate the incidence of HIV. The flexibility of the HIV model allows a range of scenarios to be explored for the future course of the HIV epidemic.

Fitting the model to data

A Bayesian methodology was used to fit the model to the available data [14,15]. Prior estimates of the distribution of each parameter are combined with the distribution of the likelihood function (obtained by fitting the model to the available data) to give the posterior distribution. Prior distributions for the parameters describing transitions between TB states were

![Fig. 1. Outline of the tuberculosis (TB) submodel. In the full model [see Methods, and supplementary material (www.who.int/gtb and www.AIDSonline.com)], active TB may be infectious or non-infectious, with movement allowed from active non-infectious disease to active infectious disease. An identical submodel, with different parameter values, describes those with HIV. Death can occur in any state, but death rates are higher for patients with active disease.](image-url)
or in the early stages of HIV infection, remain infected assumed that cured TB patients uninfected with HIV, are at no additional risk of dying from TB. It is those who are treated that become non-infectious and a given time period. The cure rate is the proportion of active cases that are found and begin treatment during

The TB case detection rate is the proportion of new, interventions considered in this study.

Although a simpler model may still have captured the TB–HIV model as well as the parameter values. The accuracy of the results depends on the structure of

Partial rank correlation coefficients show that the parameters responsible for most of the uncertainty in model outputs are those for which there is least information, i.e. those describing the effect of HIV infection on the course of TB. They are, for those with HIV, the rate of progression from confection to active TB, the proportion infectious of those with active TB, the death rate of those with TB and the relapse rate to active TB among those who have failed treatment.

The accuracy of the results depends on the structure of the TB–HIV model as well as the parameter values. Although a simpler model may still have captured the main features of the data, the model structure is the simplest that could be used to explore all of the interventions considered in this study. **Interventions**

The TB case detection rate is the proportion of new, active cases that are found and begin treatment during a given time period. The cure rate is the proportion of those who are treated that become non-infectious and are at no additional risk of dying from TB. It is assumed that cured TB patients uninfected with HIV, or in the early stages of HIV infection, remain infected with TB; those that have late-stage HIV infections

The main effect of TLTI is to eliminate the chance of developing active TB for 70% of infected people who receive it; the other 30% are assumed to receive no benefit [21,22]. Ideally, TLTI is given only to those who are already infected with TB (and never to those with active disease); however, TB infections cannot always be identified by tuberculin skin-testing, especially in anergic subjects coinfected with HIV [7]. Therefore HIV infection was used as the criterion for the administration of TLTI, and coverage was measured as the fraction of patients that receive one course of treatment between initial HIV infection and death. It was assumed that those given TLTI are protected from TB infection for the duration of treatment [23]. Treatment is either for 6 months or for life; patients treated for 6 months return to the state whence they came, either latent or uninfected.

By reducing TB prevalence among HIV-positive individuals, the death rate of those in late-stage HIV was effectively reduced, thereby increasing the late-stage HIV population. It is assumed, however, that this has a negligible effect on HIV transmission and we do not include a corresponding rise in HIV incidence.

In our model, HAART prevents patients from moving from early- to late-stage HIV infection, effectively preventing a patient’s HIV infection from progressing so long as they continue to take the appropriate combination of drugs. Since the increase in life expectancy of patients taking HAART (as currently formulated) is estimated to be 5–7 years [24–26], or less [27], this is an optimistic view of the effectiveness of HAART. As yet, there are few data on compliance with HAART; to explore the full potential of antiretroviral therapy, plus one scenario that is more realistic, annual dropout rates of 0% or 20% have been used [28]. There is no explicit allowance for the emergence of drug resistance under HAART, and it is assumed that HAART has no impact on HIV transmission. The coverage of interventions that do reduce HIV transmission (condoms, change of sexual behaviour, etc.) is expressed in terms of the effects on HIV incidence (e.g., reducing the annual HIV incidence rate by 1% from the point of intervention onwards).

The impact of interventions were measured over and above present levels of coverage. For the TB treatment measures, it was assumed that 50% of the new infectious TB cases that arise each year are detected obtained from published studies (see supplementary material at www.who.int/gtb and www.AIDSonline.com), while prior distributions of the parameters describing the HIV epidemic were obtained by fitting the HIV model to HIV prevalence data from antenatal clinics in the relevant countries [16]. The likelihood function was determined by fitting the model output to estimates of TB incidence from each country [3]. Posterior probability distributions for the parameters were obtained using importance sampling, together with a standard optimization routine [15,17–19].

With the Bayesian approach, if more information is available from the prior distribution than from the new data for a particular parameter, the posterior distribution will depend mainly on the prior information. Conversely, if the prior information on a particular parameter is weak and the data constrain the parameter to a relatively small range of values, the posterior distribution will depend mainly on the likelihood function.

Confidence intervals around each series of projected TB incidence and death rates were obtained by carrying out 1000 simulations using parameter values randomly chosen from the posterior distributions. Uncertainty analysis was used to identify the components of the model that most influenced our results, judging sensitivity from partial rank correlation coefficients calculated between each outcome measure and each of the parameters in the model [20].

Partial rank correlation coefficients show that the parameters responsible for most of the uncertainty in model outputs are those for which there is least information, i.e. those describing the effect of HIV infection on the course of TB. They are, for those with HIV, the rate of progression from confection to active TB, the proportion infectious of those with active TB, the death rate of those with TB and the relapse rate to active TB among those who have failed treatment.

The accuracy of the results depends on the structure of the TB–HIV model as well as the parameter values. Although a simpler model may still have captured the main features of the data, the model structure is the simplest that could be used to explore all of the interventions considered in this study.
and 70% of these are cured, which is thought to be typical for sub-Saharan Africa [3]. For HAART, coverage is measured as the fraction of HIV-infected persons progressing to AIDS that receive antiretroviral drugs. Similarly, the coverage of TLTI is measured as the fraction of HIV-infected persons (including all TB and HIV coinfected persons) given one course of treatment between HIV infection and death. For condoms and other measures designed to prevent infection, coverage was expressed in terms of its effect on HIV incidence, applying a fixed percentage reduction in annual HIV incidence from the point of intervention onwards. For all three of the preventive measures, it was assumed that coverage was negligible prior to the modelled interventions. Consequently, there is great potential to improve on prevention, much less to improve on cure.

Analysing the model output
The model was first used to generate a reference TB epidemic for a country with a high prevalence of HIV infection. Kenya was chosen because reasonably good data are available both for HIV and TB, and because the epidemic is more advanced in Kenya than in some African countries (such as South Africa), but less advanced than in others (such as Uganda). It is not known, and the data do not show, whether the prevalence of HIV will continue to rise, remain steady or fall, so three different underlying HIV epidemics have been considered in which prevalence, in the absence of any further intervention, levels off at its current value, increases by half or falls by half.

To explore the generality of these findings, the model was also fitted to data from Uganda and South Africa [16], where the HIV epidemics are, respectively, more and less advanced.

Results
Figure 2a gives three possible projections for the HIV epidemic in Kenya, each of which is consistent with the available data and corresponds to the HIV prevalence levelling off at 50%, 100% or 150% of current levels. If HIV prevalence declines in Kenya, even without additional interventions, TB incidence is expected to fall (Fig. 2b), with a time lag equal to the estimated time lag between HIV infection and progression to WHO stage 3 (or late-stage HIV), approxi-
If HIV prevalence stabilizes or continues to increase, then the number of TB cases is also expected to increase (Fig. 2c,d), by about 20% for constant HIV prevalence and by approximately 60% for a 50% increase in HIV prevalence.

Figure 3a shows the impact on TB incidence and Figure 4a the impact on TB deaths of a 10% increase in coverage of each intervention for the epidemic in which HIV prevalence stabilizes at its current level. Note, in particular, that while reducing HIV incidence has little impact on TB incidence in the short term it has the greatest impact in the long term.

The relative effectiveness of the different interventions is judged first by applying and maintaining the same, small improvements in coverage and recording consequent reductions, over 10 and 20 years, in the numbers of new TB cases (Fig. 3b,d) and TB deaths (Fig. 4b,d). Larger HIV epidemics generate larger burdens of TB, and more cases (Fig. 3) and deaths (Fig. 4) are averted by each intervention. In all scenarios considered, the most effective way to reduce TB incidence is by increasing TB case detection and cure rates (Fig. 3b,d); however, over 20 years, reducing HIV incidence is not significantly worse than improving case detection and cure rates for TB. The most effective way to avert TB deaths (Fig. 4b,d) is by improving case detection. For both TB cases and TB deaths, TLTI is relatively ineffective, even with lifelong treatment. HAART is also relatively ineffective, especially when compliance is less than 100%. While the relative impact of the different interventions is similar over 10 and 20 years, the benefits of reducing HIV incidence become more pronounced over the longer time scale, i.e. after a time delay exceeding the time spent in the early stages (WHO stages 1 and 2) of HIV (Figs 3a and 4a). Nonetheless, reducing HIV incidence is not expected to have as great an impact on TB incidence or deaths within the next 20 years as finding and treating those with active TB (Figs 3d and 4d).

Since unit changes in the coverage of very different interventions are unlikely to be equally feasible or equally costly, their impacts were also compared when each is exploited to its full potential and coverage is increased to 100%. Under these circumstances, HAART has the greatest impact on TB incidence (Fig. 3c) and is as good as 100% case detection in preventing deaths (Fig. 4c). TLTI remains relatively ineffective (Figs 3c and 4c).

---

**Fig. 3. Projected changes in tuberculosis (TB) incidence in Kenya when HIV prevalence stabilizes at its current value.** (a) The impact of interventions on TB incidence. The coverage of each intervention was increased by 10% from the start of 2002 onwards. (b) Expected reductions in the number of TB cases over 10 years when the coverage of interventions is increased by 1%, and when HIV prevalence stabilizes at 50% (red), 100% (blue) or 150% (green) of its present value. (c) Expected reductions in the number of TB cases over 10 years when coverage is increased from baseline to 100%. (d) As (b) but with impact measured over 20 years. HAART, highly active antiretroviral therapy; TLTI, treatment of latent TB infection. Error bars give 95% confidence limits for the point estimates.
A third way of comparing interventions is to ask what improvement in coverage would be needed to match the impact of a 5% increase in the case detection rate, over the baseline level of 50%. The same reduction in the number of TB cases over 10 years could be obtained by any of the following means: reduce HIV incidence by 24% (95% CL, 15–40); increase the coverage of HAART from 0 to 21% (95% CL, 13–29), assuming 20% dropout each year; provide 6 months’ TLTI to 33% (95% CL, 17–55) of all HIV-infected persons; or increase the TB cure rate from 70 to 76% (95% CL, 75–78). Therefore, all interventions, except augmenting the TB cure rate, require relatively large increases in coverage to compete with a 5% improvement in case detection.

A similar exercise was carried out for Uganda, where the HIV epidemic occurred earlier than that in Kenya, assuming that HIV prevalence would decline to 50% of its peak, and for South Africa, with a later epidemic than Kenya, assuming that HIV prevalence would remain constant at its 1999 level; the results are shown in Fig. 5. It was expected that the number of TB cases in Uganda would start to fall in the near future, irrespective of any change in control efforts, because the prevalence of HIV peaked around 1991 and has fallen by about 50% since then [29]. There are also signs that TB incidence is no longer increasing as quickly as in the 1990s (Fig. 5a). Despite the differences between the Ugandan and Kenyan TB and HIV epidemics, the preventive measures are, again, much less effective than the curative measures at averting TB cases and TB deaths (Fig. 5c,e). South Africa appears to be on the threshold of a very large TB epidemic, driven by HIV. With no additional interventions, a twofold increase in TB incidence is forecast before 2010, above a rate that is already very high, whatever the future course of the HIV epidemic. Nonetheless, curative measures are, per unit improvement in coverage, still the best way to diminish TB incidence, as for Uganda and Kenya (Fig. 5d).

**Discussion**

At first sight, these results appear paradoxical. Three different analytical approaches all indicate that the best
A way to manage TB epidemics driven by HIV is to find and treat TB cases rather than to prevent or mitigate the effects of HIV infection. These results are robust to uncertainties in the values of model parameters and are similar for early (Uganda), intermediate (Kenya) and late (South Africa) epidemics. The principal explanation for this finding is that curative measures reduce deaths and decrease transmission in all patients with TB, irrespective of whether patients are infected with HIV. By contrast, the preventive methods are directed at people coinfected with TB and HIV, who typically represent only one-third to one-half of the sources of new TB cases in eastern and southern Africa [1]. In addition, while preventing HIV infection removes the underlying cause of rising TB incidence, the benefits only begin to appear after approximately 4 years [10,11], the time lag between HIV infection and late-stage HIV (WHO stage 3). Taken together, these arguments suggest that preventive measures will work best when intervening early in the course of a large HIV epidemic.

The short-term practical advantage of pursuing curative measures is that national TB control programmes in many African countries are already implementing the WHO DOTS strategy [3]. Coverage can, therefore, be improved by strengthening existing programmes. In Uganda, increasing case detection by only 1% could avert 3200 cases and prevent 2000 deaths from TB over the next 10 years. In Kenya, where the HIV epidemic has not yet started to decline, this could avert 6700 cases and 4900 deaths from TB; in South Africa, where...
HIV is much more prevalent and where the background rate of TB is already high, a 1% increase in case detection could avert approximately 29,000 cases of TB and 22,000 TB deaths over the next 10 years. Greater improvements in case detection would, of course, lead to even greater impacts, and improving cure rates would also increase the impact.

However, even if DOTS is necessary to contain the HIV-related epidemics of TB, it may not be sufficient to bring such epidemics under control for two reasons. First, although curing TB is relatively effective, the results of this analysis suggest that curative programmes on their own will stabilize, but not reverse, TB incidence and deaths. Second, methods for preventing and ameliorating the effects of HIV infection will be essential for tackling AIDS in general, as distinct from HIV-related TB in particular.

We have not attempted here to evaluate the wider benefits of HAART, or of reducing HIV transmission. Rather, our goal has been to reconsider the balance of benefits of HAART, or of reducing HIV transmission. We have not attempted here to evaluate the wider HIV-related TB in particular.

Acknowledgements

We thank G. Elzinga, D. Maher, P. Nunn, P. Onyebujoh, M. Raviglione and F. Scano for numerous discussions that stimulated this work.

References