Transmission Elasticity in Communities Hyperendemic for Tuberculosis

Pieter Uys,1 Ben J. Marais,2 Simon Johnstone-Robertson,1,3 John Hargrove,1 and Robin Wood3,4

1Department of Science and Technology/National Research Foundations Centre of Excellence in Epidemiological Modeling and Analysis, and 2Children’s Hospital at Westmead and Sydney Emerging Infectious Diseases and Biosecurity Institute, Sydney Medical School, University of Sydney, Sydney, Australia; 3Desmond Tutu HIV Centre, Institute of Infectious Diseases, Molecular Medicine, and 4Department of Medicine, University of Cape Town, Cape Town, South Africa

Background. Despite consistently meeting international performance targets for tuberculosis case detection and treatment success, areas where tuberculosis is hyperendemic fail to achieve the predicted epidemiological impact. In this article, we explore the anomalous relationship between defined performance targets and actual reduction in tuberculosis transmission.

Methods. In areas where tuberculosis is endemic, poorly ventilated social gathering places such as shebeens (informal alcohol drinking places), minibus taxis, and clinic waiting rooms are all potential transmission hot spots. We modeled the transmission reduction achieved by removal of infectious persons in settings with different tuberculosis prevalence rates to demonstrate the concept of transmission elasticity. We then applied this concept to real-life data from a hyperendemic community in Cape Town, South Africa.

Results. In a hyperendemic area, reducing the number of infectious people by a given percentage results in a smaller percentage decrease in the annual risk of infection (ARI) compared with a nonendemic area; for example, removing 10% of infectious persons could result in as little as a 5% reduction in the ARI. With use of real-life data and removal of 60% of infectious individuals with tuberculosis, as would be achieved by meeting current performance targets of 70% case detection and 85% cure, the estimated ARI reduction is 50%.

Conclusions. The relationship between the number of infectious people removed and the decrease in ARI is nonlinear. The concept of transmission elasticity has important implications for the formulation of universal performance targets, since hyperendemic areas would require more stringent targets to achieve comparable transmission reduction.

Tuberculosis is caused by Mycobacterium tuberculosis, which spreads by aerosol transmission. Individuals with active lung disease expel infectious droplets into the ambient air, where they remain suspended for a considerable period in the absence of adequate ventilation. Progression to disease is uncommon among immunocompetent people following inhalation of an infectious droplet. However, high levels of ongoing transmission and sufficient numbers of vulnerable people sustain the epidemic in areas where tuberculosis is endemic [1–3]. Molecular epidemiology studies have established that in areas where tuberculosis is endemic, disease results from recent infection (primary or reinfection) in the majority of cases [1–3]. This underscores the importance of reducing ongoing transmission in order to establish epidemic control.

The World Health Organization adopted the Directly Observed Therapy Short course (DOTS) strategy together with specified performance targets in order to control the tuberculosis epidemic. Despite curing millions of patients with tuberculosis and saving many lives, the epidemiologic impact of this strategy has been mixed [4, 5]. The global tuberculosis incidence (number of new cases identified each year) shows excessive variability, with the highest rates being recorded in sub-Saharan Africa. Tuberculosis incidence rates in excess of 1,500 cases per 100,000 people have been reported in some hyperendemic communities. Mathematical
models based on data from Western Europe predicted that achieving 70% case detection and 85% cure rates would lead to a decrease of tuberculosis incidence of up to 11% per year [5–7]. However, tuberculosis incidence failed to decrease as predicted in many hyperendemic areas where the DOTS strategy has been fully implemented and performance targets met [5, 8]. This failure could be attributed to inaccurate calculation and/or reporting of estimates or to deficiencies in the mathematical models used to define DOTS-related performance targets. Applying mathematical models in regions other than those from which they were derived is questionable, since these models fail to consider transmission dynamics that may be unique to hyperendemic areas. Typical models assume homogeneity; that is, they assume that infectious people are uniformly distributed over the whole population, but in reality tuberculosis cases are heavily clustered within particular sections of the community. In this article, we explore the potential nonlinearity between the number of infectious tuberculosis cases removed (successfully treated) and the decrease in ongoing tuberculosis transmission, as a novel reason for unreliable predictions by standard models when applied to hyperendemic areas.

Settings that are most conducive to transmission are crowded and poorly ventilated households, workplaces, prisons, and homeless shelters as well as shebeens (informal alcohol drinking places), minibus taxis, clinic waiting rooms, community halls, and churches [9–11]. For convenience, we use the term cohort to mean a group of people who spend time together in such a setting on a regular and frequent basis. For our purposes, the cohort need not comprise the same people each time. Susceptible people may be members of several cohorts; for example, they may undertake a minibus taxi journey to work in the morning (transport cohort), spend several hours at work (work cohort), and then return home in the afternoon with a second transport cohort. In hyperendemic areas around Cape Town, South Africa, many unemployed people spend a great deal of time socializing in shebeens [9, 10]. The essential attributes of a cohort entail spending prolonged time on a regular basis with a group of people, approximately fixed in number, in an enclosed space with inadequate ventilation.

In any given cohort, infectious tuberculosis cases may be present, in which case the cohort would become epidemiologically active. In hyperendemic areas, it is expected that, not infrequently, 2 or more concurrent infectious persons with tuberculosis will form part of a single cohort (doublet cohort), whereas this would be extremely unlikely in low-prevalence areas, where only 1 infectious person would form part of a cohort (singlet cohort). Treating 10% of infectious people in a low-prevalence setting will therefore cause, on average, nearly 10% of the epidemiologically active cohorts to be deactivated so that one can expect a 10% decrease in transmission. This will not be the case in hyperendemic areas, where doublet or multiple exposure cohorts will remain active since they are likely to lose only 1 of the $\geq 2$ infectious members. The expected decrease in transmission will therefore be <10%.

We utilized established transmission theory and real-life data from a well-characterized community to quantify the above considerations and show that in similar settings the relationship between the number of infectious tuberculosis cases successfully treated and the decrease in the annual risk of infection (ARI) is indeed unequal. Thus, reducing the number of infectious people by a given percentage results in a smaller percentage decrease in the ARI than might otherwise be expected. This has important implications for setting optimal performance targets for tuberculosis control programs.

**METHODS**

To quantify the transmission reduction achieved by successfully treating infectious tuberculosis cases, we have to estimate the following: (1) the probability distribution of infectious people among different cohorts, (2) the expected probability of transmission events in cohorts containing variable numbers of infectious people, and (3) for a given cohort, the risk of a susceptible individual becoming infected or reinfected over the course of 1 year, before and after a given percentage reduction in the number of infectious people.

**Probability Distribution of Infectious People Among Cohorts**

The cohorts most conducive to transmission involve prolonged socialization in crowded places with poor ventilation. Any of these cohorts may be shared by 2 or more infectious people. In general, we must find the distribution of infectious people among these cohorts and estimate the proportion of cohorts (of each type) that actually contain $\geq 2$ infectious persons.

Suppose that in a community of $M$ people there are $N$ who have infectious tuberculosis. The probability that any given person in the population is infectious is given by $\rho = N/M$. Assuming independence in the distribution of the infectious people, the probability, $P_k$, that there are $k$ infectious people in a cohort where $X$ people are present can be approximated by a Poisson distribution with parameter $\lambda = \rho X$:

$$P_k = \frac{\lambda^k}{k!} e^{-\lambda}$$  \hspace{1cm} (1)

As an example, we use data from a well-characterized hyperendemic community [12]. Figure 1 reflects the percentage of cohorts likely to contain either 1 or 2 infectious tuberculosis cases in this community, providing for different cohort sizes. For example, with a tuberculosis prevalence of 1,600 cases per 100,000 people and cohort size 20, $\sim 4\%$ of the cohorts will contain 2 infectious people (doublet cohorts) and 24% will
contain only 1 (singlet:doublet ratio of 6:1). With larger cohort sizes the relative number of doublet cohorts increases. The number of cohorts containing ≥3 infectious people seems negligible (with $M = 15,000$ the expected number of triplet cohorts is .8), so we ignored this possibility in our model; triplet cohorts may occur more frequently with exceptional case densities.

Transmission Risk Within a Cohort Containing a Given Number of Infectious People

The circumstances conducive to tuberculosis transmission are those of close confinement in poorly ventilated areas for prolonged periods. Being a member of an active cohort implies tuberculosis exposure and the possibility of acquiring infection. The number of transmission events within each cohort can be estimated using the Wells-Riley equation [13–15] or the Gammaitoni and Nucci model [16]. Under the conditions we are considering here there is little difference in the predictions made by these 2 models [17, 18], and we elect to use the Wells-Riley model:

$$C/S = (1 - e^{-kptq/Q})$$

In this equation, $C$ is the number of transmission events, $S$ is the number of susceptible people, $k$ is the number of infectious people in each cohort, $T$ is the time of exposure, $q$ is the infectivity of the infectious people, $p$ is the susceptible respiration rate, and $Q$ is the germ-free ventilation.

As an example, using typical data from our well-defined study community [12], $S = 18$ (for 2 infectious people in a cohort of size 20), $q = 1$, $T = t \times D$, where $t = 1$ h/d and $D = 60$ d is the period of infectiousness, and $P = 14,400$ L/d. Figure 2 reflects the probability of transmission within cohorts, depending on the number of infectious cases and various ventilation rates ($Q$). For a $Q$ value of 2 m$^3$/h, the probability of transmission when 1 infectious person is present in a cohort is .46, whereas for 2 infectious people in a cohort, the probability is .71.

Transmission Reduction Resulting From Removal of Infectious Cases

We assume passive case finding, in which cases emerge from the community in a random way with equal chances of treatment success. We calculate the proportional reduction in transmission events after the removal, in a random manner, of a given proportion of infectious cases as follows.

The expected annual risk of infection within a cohort is given by the expected proportion of the year during which 1 infectious person is present in the cohort multiplied by the probability of transmission when 1 infectious person is present plus the similar factor for 2 infectious people. This annual risk is therefore specified by the following formula:

$$R = \sum_{k=1}^{\lambda} \frac{\lambda^k}{k!} \exp(-\lambda)(1 - \exp\left(-\frac{kpTq}{Q}\right))$$

The probability that 3 or more infectious people are present is negligible and is therefore ignored.

For convenience, we write equation (3) as follows:

$$R = \sum_{k=1}^{\lambda} \frac{\lambda^k}{k!} \exp(-\lambda)(1 - \exp(-k\phi))$$

We require the proportional effect on $R$ when the number of infectious people (ie, the prevalence) is reduced by a given proportion. This corresponds exactly to the concept of price elasticity of demand in economics, and we therefore find it appropriate to use the prevalence elasticity of annual risk of infection or transmission elasticity given by the following:

$$\eta = \frac{dR}{d\rho} \frac{\rho}{R}$$

Applying equation (5) to equation (4) and using $\lambda = \rho X$ (equation [1]) we find the following:

$$\eta = \frac{(1-\lambda)+(\lambda-\lambda^2)\theta}{1+\phi\theta}$$

where $\theta = 1 + e^{-\phi}$ with $\phi = pTq/Q$ as above.

Note that $\eta = dR/d\rho \cdot p/R$ (equation [5]) is the limiting value of $\Delta R/\Delta \rho \cdot p/R$, where $\Delta R$ and $\Delta \rho$ are (small) increments.
in \( R \) and \( \rho \), respectively. We observe that \( \Delta R/\Delta \rho \cdot \rho/R \) can be written as

\[
\frac{100(\Delta R)}{R} \cdot \frac{\rho}{\Delta \rho}
\]

where the minus signs indicate decreases rather than increases. Thus,

\[
\left(100\frac{\Delta R}{R}\right) \sim \eta \left(100\frac{\Delta \rho}{\rho}\right),
\]

that is, the percentage reduction in the ARI is approximately \( \eta \) times the percentage reduction in the prevalence.

Transmission elasticity reflects the potential effectiveness of attempts to reduce transmission by case detection and treatment. In the ideal situation where \( \eta = 1 \), the percentage reduction in ARI is exactly equal to the percentage reduction in prevalence. This is implicitly assumed in most mathematical models where the law of mass action is applied to estimate the rate of transmission (rate of transmission = force of infection \( \times \) number of infectious people \( \times \) number of susceptible people). We are concerned here with the possibility of settings where \( \eta < 1 \). At a low prevalence the elasticity is close to 1, but in hyperendemic communities this may drop significantly below 1. With a value of .6, the percentage reduction in transmission (ARI) will be only 60% of the percentage reduction in prevalence, since ARI reduction is \( \eta \) times the percentage reduction in prevalence.

RESULTS

Figure 3 shows variable rates of transmission reduction (transmission elasticity) at different tuberculosis prevalences within cohorts of size 20 and 40. Figure 3 and formula (6) show that \( \eta \rightarrow 1 \) as \( \lambda \rightarrow 0 \); that is, at low prevalences there is parity between transmission reduction and the percentage reduction in prevalence. Transmission elasticity is demonstrated at high prevalences, where parity with the percentage reduction in prevalence is no longer maintained. At a tuberculosis prevalence of 2,000 cases per 100,000 people the ratio (reflected on the y-axis) drops to <.5, which corresponds to a proportional reduction in ARI that is less than half of the proportional reduction in tuberculosis prevalence.

In order to estimate the effect on the ARI within a community where tuberculosis cases are successfully treated, it is necessary to supply parameter values for formula (6). These parameter values are specific to the community under consideration and also vary from one cohort type to another. We consider 2 population components in order to accommodate variable exposure risk. The high-risk component includes people who spend a minimum of 1 h/d in the company of \( \geq 20 \) people in circumstances conducive to \( M. \) tuberculosis transmission. Examples of such conditions include shebeens, taxis, clinics, and community halls; it is estimated that 39% of the study population would fall into this high-risk component (Robin Wood, personal communication). Restriction fragment length polymorphism data suggest that the majority of transmission occurs within this high-risk component where the tuberculosis prevalence is estimated to be >3,200 cases per 100,000 people [12]. Household transmission is relatively less common than in low-prevalence areas [12, 19], and household cohorts were excluded from this analysis. For simplicity we assumed zero infection risk in the low-risk component, which should result in more conservative estimates for the community as a whole.

We assigned the following typical values: \( q = 1 \), \( T = t \times D \), where \( t = 1 \) is the minimum time expressed in hours per day and \( D = 60 \) d is the period of infectiousness, \( P = 14,400 \) L/d, and \( Q = 4 \) m\(^3\)/h (96,000 L/d) [12]. We then calculated the risk of infection for the high-risk component using formula (4). The risk for the remainder of the population is, as explained above, assumed to be 0. The risk of infection for these 2 components...
combined is then the weighted sum of these 2 risks with the weightings given by the total proportions that each of these groups form within the community as a whole, namely, 39% and 61%. We now estimate the transmission elasticity in the population as a whole by considering this weighted sum of the risks.

To illustrate the result, we consider the effect on transmission, in the community as a whole, of a 10% reduction in prevalence. With the conservative assumptions made, a 10% case reduction at a prevalence of 1,600 tuberculosis cases per 100,000 people, estimated transmission would be reduced by 8.3% (Figure 4). Given current performance targets (70% case detection and 85% treatment completion), it is of interest to estimate what the transmission impact would be if these targets were achieved and 60% (.7 × .85 = .595) of tuberculosis cases cured. Figure 5 demonstrates the variable ARI reductions achieved given different case densities (initial prevalence rates).

Several points should be noted:

1. The proportional reductions calculated are for the infection rate, not the disease rate. However, the risk to progress from infection to disease should remain constant in the short term, and thus the proportional reduction in disease rate will be similar. However, if the population profile changes considerably over time (reducing overall vulnerability) then correspondence may no longer apply. These factors should be accounted for in mathematical models of the longer-term epidemiological impact where, in the typical scenario, the state variables are updated using successive time-step iterations. These changing values imply a possible change in risk to progress from infection to disease. The changes may be small at each time step, but the cumulative effect could be substantial.

2. Our calculations reflect the direct impact on transmission; they do not consider secondary or more dynamic longer-term effects. When modeling longer-term impact, various delays have to be factored in; delays that facilitate ongoing transmission will only accentuate the elasticity effect.

3. In the present study we investigated the situation in a particular high-prevalence community. However the methods could be applied to low-burden settings as well, where infectious people are heterogeneously distributed. Within particular transmission hot spots such as prisons, homeless shelters, or immigrant clusters, prevalences of >500 cases per 100,000 people would suggest that transmission elasticity becomes relevant. If averaged over the entire low-burden setting, the effect may not be observable.

**DISCUSSION**

In an environment where tuberculosis is hyperendemic, it is likely that susceptible individuals may be simultaneously exposed to >1 infectious person with tuberculosis. Previous mathematical models failed to consider the potential importance of this transmission overlap, which has particular relevance in hyperendemic communities. Reducing the tuberculosis prevalence will reduce ongoing transmission, but our data suggest that this reduction is highly variable (transmission elasticity) depending on the initial prevalence rate and likelihood of transmission overlap. Transmission elasticity may partly explain why the transmission impact observed in some communities that achieved and maintained global performance targets failed to meet expectations. Although the effect seems modest at an averaged population level, it is likely that transmission elasticity is greatly enhanced within particular transmission hot spots where high-risk individuals are likely to congregate.

Only 2 avenues exist to gain control of the global tuberculosis epidemic: (1) reduce host vulnerability at the population level and/or (2) limit ongoing *M. tuberculosis* transmission. The challenge is greatly increased by factors such as the coexistent human immunodeficiency virus (HIV) infection epidemic,
drug-resistant tuberculosis, and the multiple social determinants of tuberculosis [4, 5]. Our work focused exclusively on transmission dynamics and the development of mathematical models that are more robust and predictive in hyperendemic areas. Although we utilized data from a particular setting, we believe the underlying principles are generalizable and transmission elasticity should be considered in future attempts to model tuberculosis epidemiology in hyperendemic settings. Simplified models are likely to overestimate the epidemiological impact that will be achieved in areas with significant transmission overlap, as is likely to occur within transmission hot spots. A remaining shortcoming of current transmission models is their failure to consider the impact of diagnostic delay and transmission saturation [20]. With delayed diagnosis, the transmission impact of curative treatment is greatly reduced [21]. Current targets do not reflect the importance of limiting diagnostic delay, which is essential to reduce ongoing transmission.

Transmission elasticity demonstrates that any control policy is likely to be less effective than predicted by current models especially in high-prevalence settings. Dowdy and Chaisson [22] showed that once case detection rates stabilize at any constant level <80%, no further reduction in tuberculosis incidence occurs. Adding the transmission elasticity effect to this finding provides further support for efforts to optimize case detection and treatment success, irrespective of whether existing targets have been met. Delays in accurate diagnosis and institution of effective treatment erode the effectiveness of target interventions, which has particular relevance for the spread of drug-resistant tuberculosis [23, 24].

In conclusion, the incorporation of transmission elasticity in future mathematical models supports the identification of more stringent performance targets for global tuberculosis control. Tuberculosis control programs should vigorously pursue improvements in case detection, reductions in diagnostic delay, and reductions in time to effective treatment, particularly in hyperendemic areas where transmission elasticity is likely to be most relevant.

Acknowledgments

We thank Paul Helden, who helped to stimulate thinking on the topic and encouraged us to take it forward.

Potential conflicts of interest. All authors: No reported conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed in the Acknowledgements section.

References