



OPEN

Assortative social mixing and sex disparities in tuberculosis burden

Debebe Shaweno¹, Katherine C. Horton², Richard J. Hayes² & Peter J. Dodd¹✉

Globally, men have higher tuberculosis (TB) burden but the mechanisms underlying this sex disparity are not fully understood. Recent surveys of social mixing patterns have established moderate preferential within-sex mixing in many settings. This assortative mixing could amplify differences from other causes. We explored the impact of assortative mixing and factors differentially affecting disease progression and detection using a sex-stratified deterministic TB transmission model. We explored the influence of assortativity at disease-free and endemic equilibria, finding stronger effects during invasion and on increasing male:female prevalence (M:F) ratios than overall prevalence. Variance-based sensitivity analysis of endemic equilibria identified differential progression as the most important driver of M:F ratio uncertainty. We fitted our model to prevalence and notification data in exemplar settings within a fully Bayesian framework. For our high M:F setting, random mixing reduced equilibrium M:F ratios by 12% (95% CrI 0–30%). Equalizing male case detection then led to a 20% (95% CrI 11–31%) reduction in M:F ratio over 10 years—insufficient to eliminate sex disparities. However, this potentially achievable improvement was associated with a meaningful 8% (95% CrI 4–14%) reduction in total TB prevalence over this time frame.

Considerable sex disparity exists in the burden of tuberculosis (TB) across all regions of the world. Global estimates indicate that around 64% of all TB incidence in adults (aged ≥ 15 years) is among men¹. The ratio of men-to-women for prevalent bacteriologically confirmed pulmonary TB shows wide geographic variations with ratios ranging from 1.2 in Ethiopia to about 5 in Vietnam¹. Similarly, national TB prevalence surveys have found male-to-female ratios in prevalent bacteriologically-confirmed TB (M:F ratios) of around 2.2 across low- and middle-income countries². Even in settings with generalised HIV epidemics, where the prevalence of HIV (a strong risk factor for TB) is higher among women than men, men have higher TB prevalence than women².

The underlying reasons behind excess male burden in different measures are not fully understood. Some studies based on routinely diagnosed TB cases have suggested that the excess notifications among men is a result of systematic underdiagnosis or under reporting of TB in women due to differences in access to care between women and men³. However, prevalence surveys from different settings have usually shown not only a male dominance among undiagnosed cases, but also higher prevalence-to-notification ratios for men, implying a male disadvantage in accessing TB care^{2,4}. In addition, observations that sex disparities in TB rates do not arise before the age of puberty^{3,5} have led to speculations around the role of biological and behavioural factors^{6,7}. Several studies have established a link between TB and behaviours such as smoking and alcohol consumption^{8,9} which are more frequent in men than in women in many high TB burden settings^{1,6,10}, and which are known to enhance the risk of TB progression and transmission^{11,12}.

Sex differences in social roles between men and women can determine the patterns and location of social mixing relevant to transmission of *Mycobacterium tuberculosis* (*M. tb*)⁷. Previous studies have linked substantial TB incidence in adults to their preferential contact with other adults¹³. It has been estimated that 80–90% of infections are acquired from social contacts occurring outside households^{14–16}. Many of these infections likely occur in congregate settings which enhance sex-assortative social mixing such as homeless shelters, prisons, public transports, and bars^{17,18}. Consistent with these reports, a recent systematic review of social contact patterns found widespread sex-assortative social mixing, and also higher rates of workplace contact for men¹⁹. Modelling of the implications of social contact data from Zambia and South Africa for *M. tb* forces-of-infection suggested higher forces-of-infection among men due to sex-assortative mixing and higher male TB prevalence¹³. Sex-assortative mixing therefore provides a mechanism that could amplify disparities in TB burden between men and women arising from other underlying factors.

In this study, we use a sex-stratified TB transmission model to assess the relative role of social mixing, differential progression, and differential case detection rates in establishing and strengthening sex disparities in

¹School of Health and Related Research, University of Sheffield, Sheffield S1 4DA, UK. ²Faculty of Epidemiology and Population Health, Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK. ✉email: p.j.dodd@sheffield.ac.uk

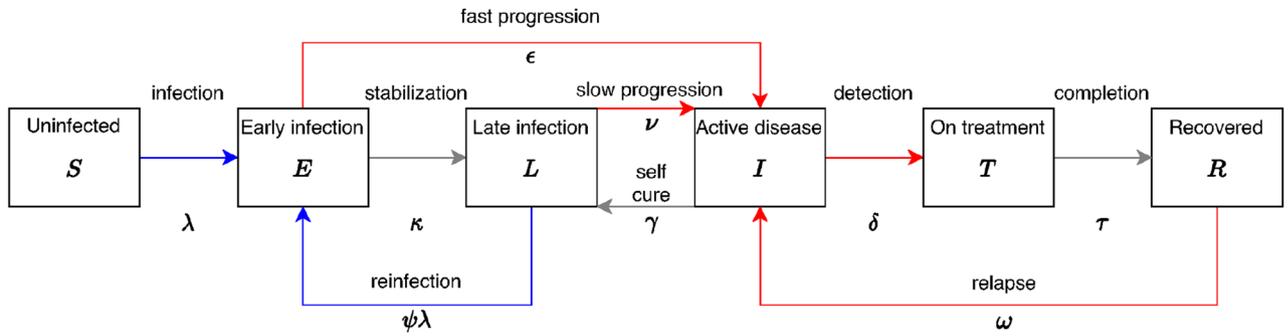


Figure 1. Model structure, duplicated for each sex. Red transitions are sex-dependent to represent different risks in disease progression and care access. Blue transitions represent infection, and are also sex-dependent due to assortative mixing. Not shown: death from all states at rate μ (and at an additional rate μ_t from active TB disease); birth into Uninfected, at rate to keep population fixed.

TB prevalence. Bayesian calibration of our model to two exemplar settings allows us to characterize parameter regions compatible with observations, and assess whether the residual uncertainties affect predicted intervention impact.

Methods

Conceptual framework, model structure, and priors. We use a conceptual framework similar to that developed for considering other social determinants of TB²⁰, and implement these mechanisms within a standard compartmental transmission model structure for adults over 15 years of age (see Fig. 1). We consider three types of mechanism by which men may come to have a higher TB prevalence than women: (1) risk factors directly affecting the natural history of infection, progression and relapse; (2) risk factors that lead to worse care access and therefore longer duration of TB disease among men; (3) risk factors relating to transmission patterns. Data are lacking to support sex-specific mortality and self-cure from untreated TB²¹, and so we did not include this possibility.

A key interest for us is the role of assortative mixing between men and women. We model this using a single parameter that interpolates between random mixing between the sexes (contacts equally shared across each sex), and no mixing between different sexes (only same-sex contacts), without changing the total contact rates (assumed the same for men and women). See Eq. (2) for mathematical specification. Natural history risk aspects affected may include increased risk of infection, increased risk of progression following (re)infection, as well as increased risk of relapse. Causes may include immunological differences between sexes as well as different levels of modifiable proximal TB risk factors such as smoking. We model these as a hazard ratio for progression or relapse, i.e. we modelled differential fast progression for men as $\epsilon_m = \alpha \times \epsilon / (1 + \alpha)$ and for women as $\epsilon_f = \epsilon / (1 + \alpha)$, so that the average progression parameter, ϵ , can be informed by literature that does not stratify by sex. We modelled the sex-specific rates of reactivation (ν) and relapse (ω) analogously, using the same parameter α . The prior for factor α is chosen based on genotypic studies that compared reactivation rates between men and women²², but we increased the reported standard deviation to 1 to allow for greater differences in other settings: see Table 1. Worse access to care may reflect differences in recognition and response to illness, barriers in accessing care, and health systems features; these are modelled in the same way as differential progression (with δ playing the role of ϵ and τ the role of α), i.e. as a hazard ratio in moving from symptomatic disease to treatment (see “Online Appendix”).

We modelled a fixed population size with births equal to deaths, and half of the population being male at all times. With $i = m, f$ for men and women, respectively, the ordinary differential equations for this model are given in Eqs. (1) and (2).

$$\begin{aligned}
 \frac{dS_i}{dt} &= \mu_t I_i + \mu N_i + (1 - \theta)\tau T_i - (\mu + \lambda_i)S_i & \frac{dE_i}{dt} &= \lambda_i S_i + \psi \lambda_i L_i - (\mu + \kappa + \epsilon_i)E_i \\
 \frac{dL_i}{dt} &= \kappa E_i + \gamma I_i - (\mu + \nu_i + \psi \lambda_i)L_i & \frac{dI_i}{dt} &= \epsilon_i E_i + \nu_i L_i + \omega R_i - (\mu + \mu_t + \gamma + \delta_i)I_i \\
 \frac{dT_i}{dt} &= \delta_i I_i - (\tau + \mu)T_i & \frac{dR_i}{dt} &= \theta \tau T_i - (\omega + \mu)R_i,
 \end{aligned} \tag{1}$$

Here, the states and rates are defined in Fig. 1. The total populations are $N_i = S_i + E_i + L_i + I_i + T_i + R_i$. A fraction $(1 - \theta)$ of those on anti-TB treatment are assumed to die. The forces-of-infection in men and women are given by

$$\begin{pmatrix} \lambda_m \\ \lambda_f \end{pmatrix} = M \begin{pmatrix} \beta_m d_m \\ \beta_f d_f \end{pmatrix} = \begin{pmatrix} 1 - (1 - \rho)/2 & (1 - \rho)/2 \\ (1 - \rho)/2 & 1 - (1 - \rho)/2 \end{pmatrix} \begin{pmatrix} \beta_m d_m \\ \beta_f d_f \end{pmatrix}, \tag{2}$$

where $d_i = I_i/N_i$. The state I represents bacteriologically-positive TB. We also consider differential infectiousness by sex as $\beta_m = f_m \times \beta$ and $\beta_f = f_f \times \beta$ where f_m and f_f are the fractions of prevalent TB cases that are sputum smear positive among men and women, respectively. This allows β to be identified with the Styblo ratio

Parameters	Description	Distribution	Source
μ_t	TB-related mortality rate (year ⁻¹)	lognormal(-1.58, 0.088)	Ragonnet ²³
γ	TB self-cure rate (year ⁻¹)	lognormal(-1.68, 0.18)	Ragonnet ²³
ϵ	Fast progression rate (year ⁻¹)	lognormal(-2.37, 0.32)	Ragonnet ²⁴
κ	Stabilization rate (year ⁻¹)	lognormal(0.62, 0.068)	Ragonnet ²⁴
ν	Reactivation rate (year ⁻¹)	lognormal(-6.89, 0.58)	Ragonnet ²⁴
ω	Relapse rate (year ⁻¹)	lognormal(-3.95, 0.27)	Crampin ²⁵
α	Relative disease progression	lognormal(-0.298, 0.80)	Shea ²²
β	Effective contact rate (year ⁻¹)	lognormal(1.68, 0.37)	Dodd ²⁶
π	Relative detection rate	lognormal(-0.298, 0.20)	Horton ²
$\rho = 2c - 1$	Assortative mixing	$c \sim \text{lognormal}(-0.57, 0.085)$	Horton ¹⁹
$\delta = \frac{p}{1-p}(\gamma + \mu + \mu_t)$	Case detection rate	$p \sim \text{beta}(5.6, 3.5)$	WHO ¹
ψ	Partial protection	beta(20.7, 77.9)	Andrews ²⁷

Table 1. Priors and sources for model parameters. NB the prior for δ is defined by comparison to competing hazard parameters to give an uninformative prior on the probability of a TB case being notified (p). The inference code implemented priors for the first ten parameters as truncated log-normal to ensure only positive values were considered. We chose the α prior so that the mid value is centered at 1.0²² and increased the reported standard deviation to 1 to allow for greater differences in other settings.

in defining a prior. The assortativity parameter ρ in the mixing matrix M ranges from -1 (mixing only with the opposite sex), through 0 (representing random mixing), to 1 (completely within-sex mixing). The proportion, c , of effective contacts each sex makes with the same sex is given by $c = (\rho + 1)/2$. Below, we will refer to the one-sex model, which is defined by considering one of $i = m, f$ separately with $\rho = 1$ (so that $\lambda_i = \beta_i d_i$). Table 1 defines the priors we use for our model parameters, and the sources we base them on.

Reproduction number and invasion dynamics. In order to explore the impact of mixing on the early dynamics of TB within a susceptible population (invasion) for our model, we analytically calculate R_0 and the M:F ratio during the associated exponential growth phase from the disease-free equilibrium. We calculate how these metrics vary with assortativity at different levels of other parameters.

Equilibria and sensitivity analysis. Similarly, we numerically compute the equilibrium TB prevalence and M:F ratio by running the model for 400 years, and explore how these vary with assortativity. To robustly characterise the importance of different parameters given our priors, including their interactions, we calculated the Sobol' variance-based sensitivity indices²⁸ for the outcomes of total TB prevalence and M:F ratio. We used the SALib Python library²⁹, generating 2000 parameter sets using the priors detailed in Table 1. Code for all analyses is available at <https://github.com/Debebe/tbsex>.

Approach to calibration. We used Mathematica 12.1³⁰ to obtain quadratic equation for the one-sex model equilibrium prevalence by setting the left hand side of Eq. (1) to zero. This was of the form

$$Q(d; \beta, \Psi) = 0,$$

where $d = D/N$, and Ψ represents the parameters other than β . The two-sex model has identical form for men and women (with different parameter values); the dynamics are coupled only through the force of infection in Eq. (2), with $\lambda_i, i = m, f$, playing the role of βd in the one-sex model. This means the condition for the two-sex model to be at equilibrium can be written

$$\begin{aligned} Q(d_m; \lambda_m/d_m, \Psi_m) &= 0 \\ Q(d_f; \lambda_f/d_f, \Psi_f) &= 0. \end{aligned}$$

Approximate inference can then be performed using the log-likelihood

$$LL_{\text{data}} + LL_{\text{constraint}} + \log(\text{prior}(\beta, \Psi)),$$

where

$$LL_{\text{constraint}} = -\left(\frac{Q(d_m; \lambda_m/d_m, \Psi_m)}{Q_{\text{tol}}}\right)^2 - \left(\frac{Q(d_f; \lambda_f/d_f, \Psi_f)}{Q_{\text{tol}}}\right)^2$$

ensures that the prevalences are close to that generated at equilibrium under the input parameters (to within a tolerance Q_{tol}). Smaller values of Q_{tol} result in sampled prevalences that are more tightly peaked around equilibrium values implied by parameters. We used a value of $Q_{\text{tol}} = 10^{-6}$.

The log-prior is the sum over the individual log-priors in Table 1. The data likelihood is given by

	Ethiopia		Uganda	
	Data	Posterior (95% CrI)	Data	Posterior (95% CrI)
Prevalence	277 (208, 347) ³²	283 (232, 336)	401 (292, 509) ³¹	397 (321, 474)
Prevalence M:F ratio	1.2 ³²	1.24 (1.03, 1.45)	4.1 ³¹	3.76 (3.0, 4.5)
Notifications	188	181 (155,208)	119 ³¹	120 (103, 136)
Notifications M:F ratio	1.2	1.11(0.87, 1.40)	2.03 ³¹	2.23 (1.75, 2.86)
Detection M:F ratio (π)	–	0.90 (0.69, 1.16)	–	0.60 (0.46, 0.78)
Progression M:F ratio (α)	–	1.12 (0.95, 1.32)	–	2.75 (2.17, 3.45)
Assortativity (ρ)	–	0.14 (–0.04, 0.34)	–	0.16 (–0.02, 0.36)

Table 2. Posterior estimates for exemplar settings: calibration targets compared to data, and parameters determining sex difference for TB.

$$LL_{\text{data}} = -\frac{(d - d_{\text{target}})^2}{2\sigma_d^2} - \frac{(r - r_{\text{target}})^2}{2\sigma_r^2} - \frac{(n_m - n_{\text{target}}^m)^2}{2s_m^2} - \frac{(n_f - n_{\text{target}}^f)^2}{2s_f^2},$$

where d and r are the overall prevalence and M:F ratios, respectively, and n_i are the sex-specific notification rates. For Uganda, the observed values of d and r are 401 per 100,000 and 4.1 respectively³¹, while for Ethiopia the corresponding values were 277 per 100,000 and 1.2³². We used the n_m and n_f values of respectively 161 and 79 per 100,000 for Uganda³¹, and 208 and 169 per 100,000 for Ethiopia¹ (see Table 2). We assumed the standard deviations were 10% of their corresponding targets, i.e. $\sigma_d = d_{\text{target}}/10$, $\sigma_r = r_{\text{target}}/10$, $s_m = n_{\text{target}}^m/10$, and $s_f = n_{\text{target}}^f/10$.

Calibration was performed using no-U-turn Hamiltonian Markov chain Monte Carlo using Stan³³. Three chains were run for 10,000 iterations and the first 5000 runs were excluded as burn-in.

Exemplar settings and data. We chose one setting with high M:F ratio (Uganda) and one with low M:F ratio (Ethiopia). For these exemplar settings, we extracted data on sex-stratified bacteriologically-confirmed TB prevalence from national prevalence surveys of Uganda (2014–2015)³⁴ and Ethiopia (2010–2011)³². In both settings, the proportions prevalent with smear-positive disease (which we used to parametrize relative infectiousness) were higher in men: $f_m = 0.45$ vs $f_f = 0.35$ in Ethiopia³², and $f_m = 0.43$ vs $f_m = 0.38$ in Uganda³¹. We also compiled notification and case detection rates from WHO reports on the same year of prevalence surveys in those settings¹.

Sex-specific prevalence and notification rates for each setting were used as calibration targets (see previous section). All data used in this study represent only individuals older than 15 years. For Ethiopia, sex-stratified notifications for bacteriologically confirmed TB were not available in the year of prevalence survey (2011). We therefore generated male-to-female ratios for new and relapse TB notifications from the closest year (i.e., 2013), and applied these values to generate sex-specific bacteriologically confirmed notification calibration targets for 2011. Models for the two exemplar settings were identical and used the same prior distributions (Table 1), but used setting-specific values for background mortality (from United Nations World Population Prospects estimates) and anti-TB treatment success (from World Health Organization collated treatment outcome data).

Relevance to interventions; contribution of assortativity to prevalence M:F ratios. We simulated an intervention (eliminating the sex gap in access to care) and evaluated the relative impact on TB M:F ratio and prevalence after 10 years. Equalizing male case detection was implemented by setting the detection rate, δ_m , for men equal to its estimated value for women. We reported posterior mean and 95% quantiles of percentage reductions for each setting. Anticipating non-identifiability and potential posterior correlations between assortativity and other parameters influencing the M:F ratio, we assessed whether different combinations of parameters with similar posterior probability resulted in different dynamics under interventions. We selected three pairs of parameters that generate sex differences for TB: α , determining the relative risk of progression; π , determining the relative detection rate; and ρ , determining the assortativity of social mixing. For each parameter pair we drew an elliptical region with major and minor radii of respectively 0.05 and 0.02 on their joint posterior distributions. The ellipses were centered at the 25, 50 and 75th quantiles for each variable, and were labelled respectively ‘L’ for low, ‘M’ for middle, and ‘U’ for upper. We used all samples within ellipses for parameter pairs of interest and random samples of all other parameters to simulate the impact of intervention. The number of parameters within each ellipse varied between 90 and 422 depending on sampled region, parameter combination and the exemplar setting.

To assess the contribution of assortativity to M:F ratios in these settings, we evaluated the mean percentage reduction in equilibrium M:F ratio (and 95% quantiles) from moving to a counterfactual with random mixing, across 200 samples from the full posterior.

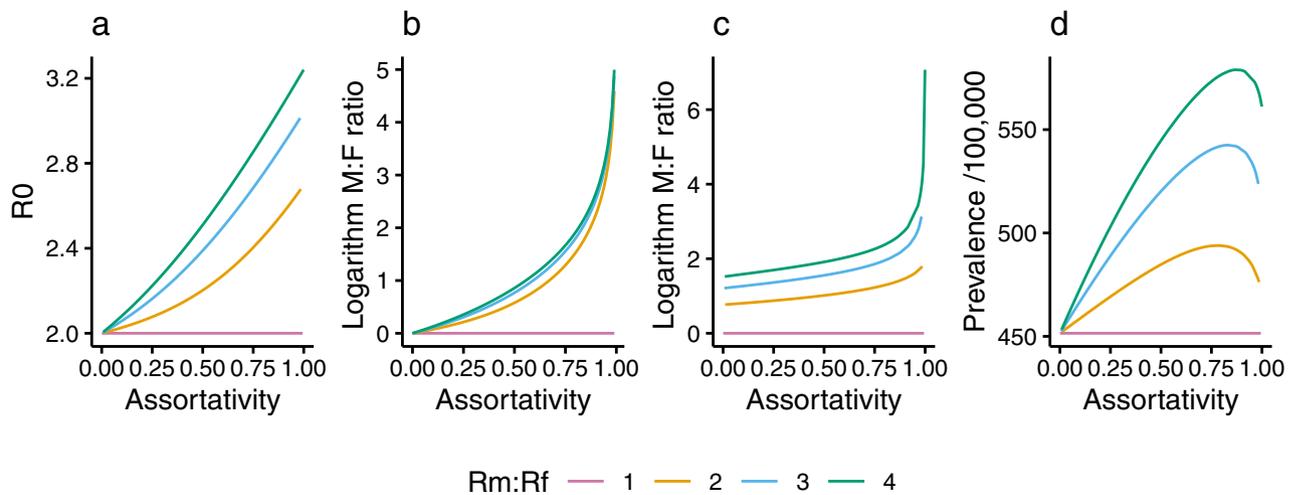


Figure 2. Influence of assortativity (ρ) at invasion and equilibria for different R_m/R_f . The effect of ρ on: (a) R_0 ; (b) M:F ratio during early exponential growth from the disease-free equilibrium; (c) M:F ratio in TB prevalence at endemic equilibrium; (d) TB prevalence at endemic equilibrium. In (c, d), β and α were solved to give $R_0 = 2$ and fix R_m/R_f , ρ was varied, and $\pi = 1$. Other parameters were the means of priors in Table 1.

Results

Influence of assortativity on invasion dynamics and equilibria. *The one-sex model.*

For a compartmental TB transmission model such as this, we found the mean number of secondary cases generated at disease free equilibrium (R_0) as $R_0 = (\text{infectiousness}) \times (\text{mean duration}) \times (\text{progression probability}) = \beta \times T \times P$. For TB models the structure of Eq. (1), there are two routes of progression from infection to active disease—fast and slow—so that:

$$R_0 = \beta \times T \times (P_{\text{fast}} + P_{\text{slow}}).$$

In our model,

$$P_{\text{fast}} = \frac{\epsilon}{\epsilon + \kappa + \mu} \quad P_{\text{slow}} = \frac{\kappa + \mu}{\epsilon + \kappa + \mu} \cdot \frac{\nu}{\nu + \mu} \quad T_1 = \frac{1}{\delta + \mu_t + \gamma + \mu},$$

where T_1 is the mean duration of a single TB episode. In reality (and our model) multiple episodes of TB are possible due to relapse and self-cure followed by reactivation. The probability of each subsequent episode was therefore the sum over these routes, $P_+ = P_{\text{self-cure}} + P_{\text{relapse}}$, and the mean time with TB, T , given as a sum over each possible number of episodes:

$$P_{\text{self-cure}} = \gamma T_1 \frac{\nu}{\nu + \mu} \quad P_{\text{relapse}} = \theta T_1 \frac{\omega}{\omega + \mu} \quad T = T_1 (1 + P_+ + P_+^2 + \dots) = \frac{T_1}{1 - P_+}.$$

These formulae were used to compute R_m and R_f —the values of R_0 for a one-sex model with parameters for men and women, respectively.

The two-sex model. For the two-sex model, R_0 was calculated as the largest eigenvalue of the next generation matrix G , given by

$$G = \begin{pmatrix} P_m & 0 \\ 0 & P_f \end{pmatrix} M \begin{pmatrix} \beta_m T_m & 0 \\ 0 & \beta_f T_f \end{pmatrix},$$

where M is the mixing matrix defined in Eq. (2), and P_m , β_m , and T_m are the values of P , β , and T for men (and similarly for women, with subscript f). This was solved exactly to give

$$R_0 = \frac{1}{2} (1 + \rho) \bar{R} \left[1 + \sqrt{1 - 16\phi\rho / [(1 + \phi)(1 + \rho)]^2} \right],$$

where $\bar{R} = (R_m + R_f)/2$ and $\phi = R_m/R_f$. Solving for the corresponding eigenvector allowed calculation of the M:F incidence ratio during exponential growth as a function of ρ , which was found to satisfy

$$\frac{MF(\rho)}{MF(0)} = 2 \left(\frac{1 - \rho}{1 + \rho} \right) \left[1 - \phi + (1 + \phi) \sqrt{1 - 16\phi\rho / [(1 + \phi)(1 + \rho)]^2} \right]^{-1},$$

where $MF(0) = P_m/P_f$. Thus, for a given \bar{R} and $MF(0)$ (respectively), R_0 and the early M:F incidence ratio only depend on assortativity (ρ) and the ratio of male:female one-sex R_0 values (ϕ). The behaviour of R_0 and $MF(\rho)$ as assortativity varies is shown for various choices of ϕ in Fig. 2a,b.

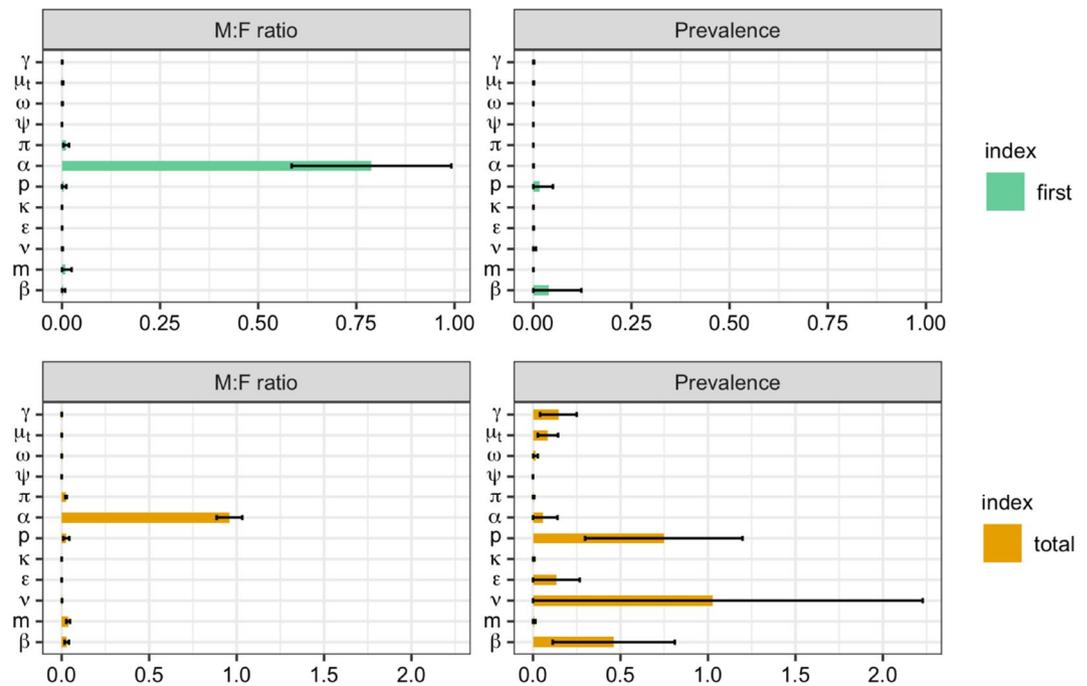


Figure 3. Sobol' first-order (one-way) and total sensitivity indices (including interactions) for equilibrium prevalence M:F ratio and total TB prevalence (shown on x-axes). Distributions of parameters (y-axis) are as in Table 1 (ψ —partial protection, ν —reactivation, ϵ —fast progression, κ —stabilization, ω —relapse, γ —self-cure, μ_t —TB mortality, p —detection, β —effective contact, $\rho = 2m - 1$ —assortativity, π —differential detection, α —differential progression).

Equilibria. The corresponding changes in equilibrium TB prevalence and equilibrium M:F ratio are shown in Fig. 2c,d. Analytical solution was not possible, and (unlike the invasion dynamics) these results may depend on the particular parameter combinations that were used. Parameter values corresponding to the means of priors in Table 1 were used, except for: ρ , which was varied between 0 and 1; β and α , which were solved to fix $R_0 = 2$ and various R_m/R_f ; and π , which was set to 1 to allow R_m/R_f to attain a value of 1.

Sensitivity of all parameters' influence on equilibria. At equilibrium the variance in TB M:F ratio is most influenced by a single model parameter—differential progression α , see Fig. 3. The first-order effect of α on TB M:F ratio is 0.79. On the other hand, the differential detection parameter, π and the mixing parameter ($m = (\rho + 1)/2$ being the proportion of effective contacts that are same sex) have first order effects of 0.01 and 0.009, respectively. All other model parameters have first-order effects less than 0.009. By interacting with other model parameters: α , m , p , β and π attained total-order effects of 0.96, 0.04, 0.03, 0.03 and 0.02 respectively.

For equilibrium TB prevalence, the main influence comes from reactivation rate (ν), detection rate (δ), effective contact rate (β), and fast progression rate (ϵ). Their total effects were attained by their interaction with other parameters (their sum is ≥ 1).

Calibration to exemplar settings. Explicit formulae for the quadratic function $Q(d, \beta; \Psi)$ used in fitting are given in the “Online Appendix”, as are example Markov chain plots of samples. The Gelman–Rubin convergence diagnostic statistic exhibited $\hat{R} < 1.0009$ for all model parameters. For each exemplar setting, the fits to calibration target data and estimates of the parameters modelling sex disparities are shown in Table 2. The smoothed pairs plot for the 8-parameter posterior in Ethiopia is shown in Fig. 4.

Intervention results; contribution of assortativity. Sampling across the whole posterior for each setting, we found the hypothetical intervention reduced TB prevalence after 10 years in comparison to the counterfactual of no intervention by 0.7% (95% CrI -0.3 to 2.7%) for Ethiopia and 8% (95% CrI 4–14%) for Uganda. The M:F ratios reduced over the same time period by 7% (95% CrI -10 to 21%) for Ethiopia and 20% (95% CrI 11–31%) for Uganda. The 3 elliptical regions selected for each parameter pair (π, α), (ρ, α), (ρ, π) in order to assess the impact of remaining uncertainty on projected intervention impacts are shown in Fig. 4 (top right three panels). Comparisons of posterior log-likelihoods for these regions are shown in the “Online Appendix”, alongside time series plots of prevalence under the intervention. The percentage reductions in TB prevalence and M:F ratio for each setting, parameter pair, and L, M, and U sampling regions are given in Table 3.

Sampling across the whole posterior for each setting, we found equilibria with random mixing had a prevalence M:F ratio that was 12% (-0.2 to 30%) lower in Uganda. For Ethiopia, equilibria with random mixing

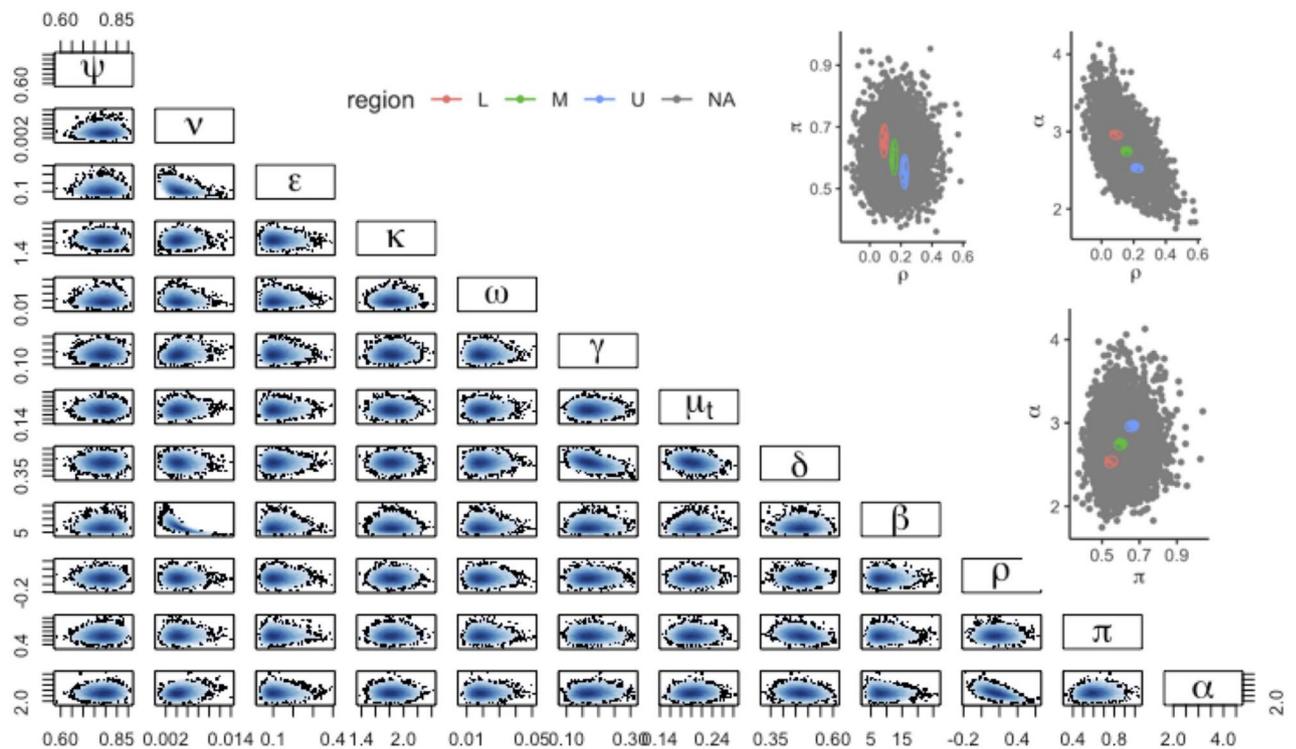


Figure 4. Correlation between model parameters. The lower left corner plot shows the smoothed posterior samples (ψ —partial protection, ν —reactivation, ϵ —fast progression, κ —stabilization, ω —relapse, γ —self-cure, μ_t —TB mortality, p —detection, β —effective contact, ρ —assortativity, π —differential detection, α —differential progression). The upper right plots show the ‘L’, ‘M’, and ‘U’ ellipses used for the sensitivity analysis in Table 3.

Outcome	Parameter pair	Ethiopia			Uganda		
		L	M	U	L	M	U
Prevalence	π, α	0.9 (0.5, 1.4)	0.5 (0.3, 0.8)	0.1 (−0.04, 0.2)	10 (7, 15)	9 (6, 11)	7 (5, 9)
	ρ, α	0.8 (−0.3, 3)	0.6 (−0.3, 3)	0.6 (−0.2, 0.3)	9 (4, 15)	8 (3, 13)	8 (3, 13)
	ρ, π	0.9 (0.3, 1.8)	0.4 (0.04, 1.1)	0.1 (0, 0.3)	10 (7, 13)	9 (6, 12)	7 (5, 9)
TB M:F ratio	π, α	11 (9, 14)	6 (5, 8)	1 (−0.5, 3)	24 (20, 27)	21 (17, 24)	17 (14, 20)
	ρ, α	6 (−9, 21)	5 (−11, 19)	6 (−9, 21)	21 (12, 30)	20 (8, 29)	20 (9, 29)
	ρ, π	11 (8, 14)	6 (4, 9)	1.5 (−1.3, 4)	23 (19, 28)	20 (17, 24)	18 (14, 21)

Table 3. Sensitivity analysis of percentage reduction in TB prevalence and M:F ratios sampling from restricted regions of the posterior. ‘L’, ‘M’, and ‘U’ correspond to different elliptical regions shown top-right in Fig. 4. (ρ —assortativity, π —differential detection, α —differential progression).

has equilibrium prevalence M:F ratio lower by 3% (−0.5 to 10%) compared to equilibrium M:F ratio based on assortativity parameter sampled from the whole region.

Discussion

Using a TB transmission dynamic model and Bayesian inference framework, we determined the extent to which social mixing and other factors contribute to sex disparity in TB prevalence. Our analysis shows that sex disparity in TB burden results from a complex interplay among several factors that determine progression, detection and social mixing, although different factors dominate at different stages of the epidemic. The impact of assortativity is much larger at invasion compared to its role at endemic equilibrium. Our fit to a high M:F prevalence ratio setting (Uganda) suggested 12% (0–30%) of the M:F ratio could be attributed to the effects of assortative mixing.

Tuberculosis invades more rapidly as assortative sex mixing increases because men are more likely to transmit infection to other men due to the more frequent interactions with males¹⁹ as well as due to higher burden of infectious TB among men^{32,35}. Men have higher rates of mixing with other men meaning that they can generate more secondary cases preferentially among men. As a result, early growth sex disparity in TB burden between men and women is extremely wide. The importance of assortativity during invasion dynamics has relevance to TB outbreaks, and potentially to invasion and replacement by new strains (eg. drug-resistant strains) in TB-endemic settings.

At invasion, higher assortative mixing increases overall TB prevalence. At endemic equilibrium, increasing assortative mixing leads to a gradual initial rise in an overall equilibrium prevalence followed by a subsequent fall as assortativity approaches 1 (Fig. 2c). The peak represents a balance between the amplifying effects of increasing assortativity, and the folding in of infections among only the half of the population at higher risk (men) as assortativity approaches 1.

Although, we found lower male case detection rates in Uganda (our high M:F ratio setting), our finding for Ethiopia (our low M:F ratio setting) is different. In Ethiopia, case detection rates between men and women are nearly comparable, with empirical prevalence-to-notification ratios of 1.7 and 1.5, respectively—implying no considerable differences between men and women in accessing TB care. These findings are consistent with some studies³⁶, and explain the limited impacts of our hypothetical intervention in this setting.

Interventions targeting only differential case detection are unlikely to eliminate sex disparities in TB burden alone. We found closing disparities in detection in a high M:F ratio setting (Uganda) would reduce the M:F ratio by around 20% over 10 years, but would also reduce TB prevalence by around 8%. Elimination of sex disparities in TB burden will therefore also require interventions targeting modifiable risk factors that are more common among men.

We found the predicted impact of narrowing case detection gaps was not strongly dependent on posterior parameter region in our high M:F ratio setting. This is reassuring, in that while data does not completely identify the exact contribution of different mechanisms to observed sex disparities in TB, our ignorance need not serve as a barrier to projecting the likely impact of sex-specific interventions. For Ethiopia, with only a small sex difference in baseline detection rates, the exact size of this differential did influence projected impact on M:F ratio, but not appreciably the impact TB prevalence.

A strength of this study was our inference approach, which treated all model parameters as uncertain, specified via priors, within a Bayesian framework. This was necessary in order to answer our question as to whether residual uncertainty after model fitting allowed for very different model predictions of intervention effect. Bayesian inference was made possible within Stan by introducing auxiliary prevalence variables and using an algebraic condition to fix these to the equilibrium solutions of the model. While Stan is capable of numerically solving ordinary differential equations, our algebraic log-likelihood allowed far more efficient (and in our experience) more reliable inference for a parameter space of this number of dimensions. Our approach may be applicable more widely in cases where a fit to an assumed equilibrium or steady decline through sparse data is sought.

Fitting our model under an assumption of equilibrium is a weakness of our study. Inevitably, TB epidemiology is not at equilibrium in any setting. However, our choice is defensible given that typical TB dynamics exhibit only single-digit annual percentage declines, and given a single prevalence survey and the substantial technical advantages discussed above bought by this assumption. Our model did not consider other subgroups that tend to mix non-randomly in real societies. Assortative mixing by other risk factors for TB, e.g. socioeconomic level, are likely to exhibit analogous behaviour, with disparities in TB burden amplified by mixing. Notably, we did not consider age; varying contact rates by age and simultaneous patterns of non-random mixing by age and sex, together with non-uniform TB burden by age may nuance our findings. In particular, some patterns of sex-dependent age mixing could weaken the effects of sex assortativity we describe: for example, older men preferentially mixing with younger women, as has been documented for sexual partnerships in some settings³⁷. Such mixing patterns could vary between settings and limit generalizability, but have not yet been documented for effective contacts relevant to respiratory infections. Crowded urban locales or accommodation with more men than women in some settings³⁸, for example where young men have moved seeking work, likely intensify the sex-assortative mixing relevant to TB that we model.

We also did not explicitly model HIV, whose prevalence is often unequal by sex. Such differences only appear in our model as averaged effects on progression in each sex. Indeed, the higher prevalence of HIV among women in many sub-Saharan African countries may reduce the sex disparities observed for TB^{2,39}. In our approach, differential progression rates due to sex differences in HIV prevalence or antiretroviral therapy coverage would manifest in the calibrated values for the relative progression parameter.

Assortative mixing has been shown to introduce potential biases in the magnitude of associations between risk factors and outcomes in infectious disease epidemiology⁴⁰. Similarly to this, our work shows that ascribing observed sex differences in TB burden to specific pathways without considering differences in exposure due to assortative mixing will result in measures of association that are biased upwards. These biases may apply to data we have used to parametrize priors for differential progression. However, these data were from a very low transmission setting, and the relatively large prior uncertainty in this parameter (α) allows it to change in fitting to settings with large M:F ratios. The large uncertainty in the prior for α partially explains the dominance of this parameter in the M:F uncertainty analysis (Fig. 3). Finally, we only considered a simple hypothetical intervention around equalizing case detection. Modelled impact may still depend on residual uncertainty for other types of intervention. Future work should consider inclusion of these features we have simplified, and realistic policy options that target disparities in detection, or potentially modifiable risk factors among men.

In conclusion, our analysis showed that assortative mixing does play a role in sex disparities in TB burden, and that this is particularly important under invasion dynamics. We found that equalizing case detection between the sexes alone is insufficient to eliminate disparities in burden. However, in settings with suboptimal case detection and large sex disparities in TB prevalence, this potentially achievable goal could still achieve meaningful reductions in overall TB burden.

Received: 5 November 2020; Accepted: 17 March 2021

Published online: 06 April 2021

References

- World Health Organization. *Global Tuberculosis Report 2019* (Tech. Rep, 2019).
- Horton, K. C., MacPherson, P., Houben, R. M. G. J., White, R. G. & Corbett, E. L. Sex differences in tuberculosis burden and notifications in low- and Middle-Income countries: A systematic review and meta-analysis. *PLoS Med.* **13**, e1002119 (2016).
- Holmes, C. B., Hausler, H. & Nunn, P. A review of sex differences in the epidemiology of tuberculosis. *Int. J. Tuberc. Lung Dis.* **2**, 96–104 (1998).
- Borgdorff, M. W., Nagelkerke, N. J., Dye, C. & Nunn, P. Gender and tuberculosis: a comparison of prevalence surveys with notification data to explore sex differences in case detection. *Int. J. Tuberc. Lung Dis.* **4**, 123–132 (2000).
- Neyrolles, O. & Quintana-Murci, L. Sexual inequality in tuberculosis. *PLoS Med.* **6**, e1000199 (2009).
- Nhamoyebonde, S. & Leslie, A. Biological differences between the sexes and susceptibility to tuberculosis. *J. Infect. Dis.* **209**(Suppl 3), S100–S106 (2014).
- Hudelson, P. Gender differentials in tuberculosis: The role of socio-economic and cultural factors. *Tuberc. Lung Dis.* **77**, 391–400 (1996).
- Bates, M. N. *et al.* Risk of tuberculosis from exposure to tobacco smoke: A systematic review and meta-analysis. *Arch. Intern. Med.* **167**, 335–342 (2007).
- Imtiaz, S., Shield, K.D., Roerecke, M., *et al.* Alcohol consumption as a risk factor for tuberculosis: meta-analyses and burden of disease. *Eur Respir J.* **50**(1), 1700216, 2017.
- Hertz, D. & Schneider, B. Sex differences in tuberculosis. *Semin. Immunopathol.* **41**, 225–237 (2019).
- Bishwakarma, R. *et al.* Epidemiologic link between tuberculosis and cigarette/biomass smoke exposure: Limitations despite the vast literature. *Respirology* **20**, 556–568 (2015).
- Jiménez-Corona, M.-E. *et al.* Gender differentials of pulmonary tuberculosis transmission and reactivation in an endemic area. *Thorax* **61**, 348–353 (2006).
- Dodd, P. J. *et al.* Age- and sex-specific social contact patterns and incidence of mycobacterium tuberculosis infection. *Am. J. Epidemiol.* **183**, 156–166 (2016).
- Glynn, J. R. *et al.* Whole genome sequencing shows a low proportion of tuberculosis disease is attributable to known close contacts in rural Malawi. *PLoS ONE* **10**, e0132840 (2015).
- Verver, S. *et al.* Proportion of tuberculosis transmission that takes place in households in a high-incidence area. *Lancet* **363**, 212–214 (2004).
- Martinez, L. *et al.* Transmission of mycobacterium tuberculosis in households and the community: A systematic review and meta-analysis. *Am. J. Epidemiol.* **185**, 1327–1339 (2017).
- Rehm, J. *et al.* The association between alcohol use, alcohol use disorders and tuberculosis (TB). A systematic review. *BMC Public Heal.* **9**, 450 (2009).
- Mathema, B. *et al.* Drivers of tuberculosis transmission. *J. Infect. Dis.* **216**, S644–S653 (2017).
- Horton, K. C., Hoey, A. L., Béraud, G., Corbett, E. L. & White, R. G. Systematic review and meta-analysis of sex differences in social contact patterns and implications for tuberculosis transmission and control. *Emerg. Infect. Dis.* **26**, 910–919 (2020).
- Hargreaves, J. R. *et al.* The social determinants of tuberculosis: From evidence to action. *Am. J. Public Heal.* **101**, 654–662 (2011).
- Horton, K. C., Sumner, T., Houben, R. M. G. J., Corbett, E. L. & White, R. G. A Bayesian approach to understanding sex differences in tuberculosis disease burden. *Am. J. Epidemiol.* **187**, 2431–2438 (2018).
- Shea, K. M., Kammerer, J. S., Winston, C. A., Navin, T. R. & Horsburgh, C. R. Jr. Estimated rate of reactivation of latent tuberculosis infection in the united states, overall and by population subgroup. *Am. J. Epidemiol.* **179**, 216–225 (2014).
- Ragonnet, R. *et al.* Revisiting the natural history of pulmonary tuberculosis: A Bayesian estimation of natural recovery and mortality rates. *BioRxiv* 729426, <https://doi.org/10.1101/729426> (2019).
- Ragonnet, R. *et al.* Optimally capturing latency dynamics in models of tuberculosis transmission. *Epidemics* **21**, 39–47 (2017).
- Crampin, A. C. *et al.* Recurrent TB: Relapse or reinfection? The effect of HIV in a general population cohort in Malawi. *AIDS* **24**, 417–426 (2010).
- Dodd, P. J., Gardiner, E., Coghlan, R. & Seddon, J. A. Burden of childhood tuberculosis in 22 high-burden countries: A mathematical modelling study. *Lancet Glob. Heal.* **2**, e453–e459 (2014).
- Andrews, J. R. *et al.* Risk of progression to active tuberculosis following reinfection with Mycobacterium tuberculosis. *Clin. Infect. Dis.* **54**, 784–791 (2012).
- Sobol, I. M. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math. Comput. Simul.* **55**, 271–280 (2001).
- Herman, J. & Usher, W. SALib: An open-source python library for sensitivity analysis. *J. Open Source Softw.* <https://doi.org/10.21105/joss.00097> (2017).
- Wolfram Research, Inc. Mathematica, Version 12.1. Champaign, IL (2020).
- The Republic of Uganda, Ministry of Health. The Uganda national tuberculosis prevalence survey (2014–2015).
- Kebede, A. H. *et al.* The first population-based national tuberculosis prevalence survey in Ethiopia, 2010–2011. *Int. J. Tuberc. Lung Dis.* **18**, 635–639 (2014).
- Carpenter, B. *et al.* Stan: A probabilistic programming language. *J. Stat. Softw.* **76**(1), <https://doi.org/10.18637/jss.v076.i01> (2017).
- The Uganda National Tuberculosis Prevalence Survey, 2014–2015 Survey Report—Ministry of Health. <https://www.health.go.ug/cause/the-uganda-national-tuberculosis-prevalence-survey-2014-2015-survey-report-2/>. Accessed 26 Sept 2020.
- Hoa, N. B. *et al.* National survey of tuberculosis prevalence in Viet Nam. *Bull. World Heal. Organ.* **88**, 273–280 (2010).
- Asres, M., Gedefaw, M., Kahsay, A. & Weldu, Y. Patients' delay in seeking health care for tuberculosis diagnosis in East Gojjam Zone, Northwest Ethiopia. *Am. J. Trop. Med. Hyg.* **96**, 1071–1075 (2017).
- Schaefer, R. *et al.* Age-disparate relationships and hiv incidence in adolescent girls and young women: evidence from Zimbabwe. *AIDS (London, England)* **31**, 1461 (2017).
- MacPherson, P. *et al.* Disparities in access to diagnosis and care in Blantyre, Malawi, identified through enhanced tuberculosis surveillance and spatial analysis. *BMC Med.* **17**, 1–11 (2019).
- Aliyu, G., El-Kamary, S. S., Abimiku, A., Blattner, W. & Charurat, M. Demography and the dual epidemics of tuberculosis and hiv: Analysis of cross-sectional data from Sub-Saharan Africa. *PLoS ONE* **13**, e0191387 (2018).
- Lemieux-Mellouki, P. *et al.* Assortative mixing as a source of bias in epidemiological studies of sexually transmitted infections: The case of smoking and human papillomavirus. *Epidemiol. Infect.* **144**, 1490–1499 (2016).

Acknowledgements

DSA is supported under the EDCTP2 programme supported by the European Union (Grant number RIA2016S-1632-TREATS). PJD is supported by a Career Development Award from the UK MRC (MR/P022081/1). This UK funded award is part of the EDCTP2 programme supported by the European Union.

Author contributions

P.J.D. conceived the experiments, D.S.A. conducted the experiments and analysed the results. All authors critically reviewed and edited the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-86869-w>.

Correspondence and requests for materials should be addressed to P.J.D.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021