Contribution of high risk groups to overall transmission is underestimated without turnover: a mechanistic modelling analysis ☆

Jesse Knight[a], Stefan D. Baral[b], Sheree Schwartz[b], Linwei Wang[a], Huiting Ma[a], Katherine Young[c], Harry Hausler[c], Sharmistha Mishra[a,d,e,f,*]

[a] *MAP Centre for Urban Health Solutions, Unity Health Toronto*
[b] *Deptartment of Epidemiology, Johns Hopkins Bloomberg School of Public Health*
[c] *TB HIV Care, South Africa*
[d] *Department of Medicine, Division of Infectious Disease, University of Toronto*
[e] *Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto*
[f] *Instituof Medical Sciences, University of Toronto*

Abstract

Background. Epidemic models are often used to estimate the contribution of high risk groups to the overall epidemic, helping to inform intervention priorities. While many epidemic models stratify populations by heterogeneity in risk, few consider movement of individuals between risk groups, which we call turnover. Its not clear how turnover can be modelled based on epidemiologic data, or how inclusion of turnover in epidemic models will influence the projected importance of interventions reaching high risk groups. Methods. We developed a framework for modelling risk group turnover in deterministic compartmental transmission models which incorporates available data and assumptions as constraints. We applied this framework to an illustrative STI/HIV model to examine the influence of turnover on: the equilibrium infection prevalence and incidence across risk groups; the inferred level of risk heterogeneity; and the importance of interventions reaching the highest risk group. Results. The equilibrium prevalence ratio between the highest and lowest risk groups decreased with increasing rates of turnover. The level of risk heterogeneity inferred via model fitting to the same prevalence targets was then higher with turnover than without. After model fitting, the estimated importance of interventions reaching the highest risk group was also higher with turnover than without. Implications. Equilibrium prevalence predicted by epidemic models can be strongly influenced by turnover. The projected importance of interventions prioritizing on high risk groups could be underestimated if fitted models do not simulate turnover which is present in reality. Collection of data which can be used to parameterize risk group turnover in epidemic models should be prioritized.

Keywords: mathematical modelling, transmission, risk heterogeneity, turnover, demographic movement, STI, HIV, population attributable fraction

☆ On behalf of the Siyaphambili study team

* Corresponding author (sharmistha.mishra@utoronto.ca)

Abbreviations: HIV: human immunodeficiency virus, tPAF: transmission population attributable fraction

Contents

1. Introduction

Heterogeneity in transmission risk is a consistent characteristic of epidemics of sexually transmitted infections (STI) (Anderson and May, 1991). This heterogeneity is often demarcated by identifying specific populations whose risks of acquisition and onward transmission of STI are the highest, such that their specific unmet prevention and treatment needs can sustain local epidemics of STI (Yorke et al., 1978). Disproportionate risk can be conferred in several ways at the individual-level (higher number of sexual partners), partnership-level (reduced condom use within specific partnership types), or structural-level (stigma as a barrier to accessing prevention and treatment services) (Baral et al., 2013). The contribution of high risk groups to the overall epidemic can then be used as an indicator in the appraisal of STI epidemics, helping to guide intervention priorities (Shubber et al., 2014; Mishra et al., 2016).

Traditionally, contribution to an epidemic was quantified using either: the classic *population attributable fraction* (PAF) via the relative risk of incident infections within a risk group versus the rest of the population and the relative size of the risk group (Hanley, 2001); or the distribution of new infections across subsets of a population (Case et al., 2012; Mishra et al., 2014). So when small risk groups experience disproportionately higher rate of incident infections – e.g. 5 percent of a population acquire 30 percent of STI infections – contribution is interpreted as 5 percent of the population contributing to 30 percent of all infections (Prüss-Ustün et al., 2013). However, the classic PAF does not account for chains of (indirect) transmission, and has been shown to underestimate the contribution of some higher-risk groups to cumulative STI infections, especially over time (Mishra et al., 2014). Thus, transmission models are increasingly being used to quantify contribution by accounting for indirect transmission and estimating the *transmission population attributable fraction* (tPAF). The tPAF is estimated by simulating counterfactual scenarios where transmission between specific subgroups is stopped, and the relative difference in cumulative infections in the total population over various time-periods is measured (Mishra et al., 2014; Mukandavire et al., 2018). Transmission can be stopped by setting susceptibility and/or infectiousness to zero in the model (Mishra et al., 2014). The tPAF is then interpreted as the fraction of all new infections that stem, directly and indirectly, from a failure to prevent acquisition and/or to provide effective treatment in a particular risk group (Mishra et al., 2016; Mukandavire et al., 2018; Maheu-Giroux et al., 2017).

There is limited evidence on how model structure might influence the tPAF of higher risk groups (Mishra et al., 2016; Mukandavire et al., 2018; Maheu-Giroux et al., 2017), especially movement of individuals between risk groups, an epidemiologic phenomenon that is well-described in the context of sexual behaviour (Watts et al., 2010). Such movement is often referred to in the STI epidemiology literature as *turnover* (Watts et al., 2010). For example, turnover may reflect entry into or retirement from formal sex work, or other periods associated with higher STI susceptibility and onward transmission due to more partners and/or vulnerabilities (Marston and King, 2006; Watts et al., 2010). Risk group turnover has been shown

to influence the predicted equilibrium prevalence of an STI (Stigum et al., 1994; Zhang et al., 2012); the fraction of transmissions occurring during acute HIV infection (Zhang et al., 2012); the basic reproductive number R_0 (Henry and Koopman, 2015); and the coverage of antiretroviral therapy required to achieve HIV epidemic control (Henry and Koopman, 2015). Yet how, and the extent to which, turnover influences tPAF has yet to be examined.

There is variability in how turnover has been previously implemented (Stigum et al., 1994; Koopman et al., 1997; Eaton and Hallett, 2014; Boily et al., 2015), in large part because of four main assumptions or epidemiologic constraints surrounding movement between risk groups. For example, in the context of turnover, the relative size of specific populations in the model may be constrained to remain constant over time (Stigum et al., 1994; Koopman et al., 1997; Eaton and Hallett, 2014), such as the proportion of individuals who sell sex. Second, some individuals may enter into high risk groups at an early age, and subsequently settle into lower risk groups; thus the distribution of risks among individuals entering into the transmission model may be assumed to be different from the distribution of risks among individuals already in the transmission model (Eaton and Hallett, 2014). Third, turnover may be constrained to reflect the average duration of time spent within a given risk group (Boily et al., 2015), such as duration engaged in formal sex work (Watts et al., 2010). Finally, turnover could reflect data on how sexual behaviour changes following exit from a given risk group (Boily et al., 2015). Most prior models used some combination of these constraints, based on their specific data or research question, but to date there is no unified approach to modelling turnover.

In this study, we explored the mechanisms by which turnover may influence the tPAF of a high risk group using an illustrative STI model with treatment-induced immunity and without STI-attributable mortality. First, we developed a unified approach to implementing turnover based on epidemiologic constraints. We then sought the following objectives: 1) understand the mechanisms by which turnover influences group-specific STI prevalence and ratios of prevalence between risk groups; 2) examine how inclusion/exclusion of turnover in a model influences the level of risk heterogeneity inferred during model fitting; and 3) examine how inclusion/exclusion of turnover in a model influences the estimated tPAF of the highest risk group after model fitting to a particular setting.

2. Methods

We developed a new, unified framework for implementing turnover. We then simulated a deterministic compartmental model of an illustrative STI, with turnover as per the framework, to conduct out experiments.

2.1. A unified framework for implementing turnover

We developed a framework for implementing turnover, as depicted in Figure 1 and detailed in Appendix A. In the framework, the simulated population is divided into G risk groups. The number of

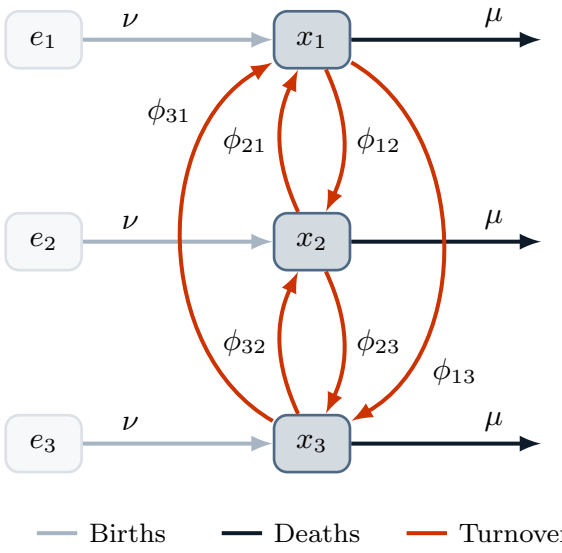


Figure 1: System of $G = 3$ risk groups and turnover between them.
x_i: number of individuals in risk group i; e_i: number of individuals available to enter risk group i; ν: rate of population entry; μ: rate of population exit; ϕ_{ij}: rate of turnover from group i to group j.

individuals in group $i \in [1, \ldots, G]$ is denoted x_i, and the relative size of each group is denoted $\hat{x}_i = x_i/N$, where N is the total population size. Individuals enter the population at a rate ν and exit at a rate μ per year. The distribution of risk groups at entry into the model is denoted $\hat{e}_i$, which may be different from $\hat{x}_i$. The total number of individuals entering group i per year is therefore given by $\nu\hat{e}_i N$. Turnover rates are collected in a $G \times G$ matrix ϕ, where ϕ_{ij} is the proportion of individuals in group i who move from group i into group j each year. The framework is independent of the disease model, and thus transition rates ϕ do not depend on health states.

The framework assumes that: 1. the relative sizes of risk groups $\hat{\boldsymbol{x}} = [\hat{x}_1, \ldots, \hat{x}_G]$ are known and should remain constant over time; and 2. the rates of population entry ν and exit μ are known, but that they may vary over time. An approach to estimate ν and μ is detailed in Appendix A.2.1. The framework then provides a method to estimate the values of the parameters $\hat{\boldsymbol{e}}$ and ϕ, representing G and $G(G-1) = G^2$ total unknowns. In the framework, $\hat{\boldsymbol{e}}$ and ϕ are collected in the vector $\boldsymbol{\theta} = [\hat{\boldsymbol{e}}, \boldsymbol{y}]$, where $\boldsymbol{y} = \text{vec}_{i \neq j}(\phi)$. To uniquely determine the elements of $\boldsymbol{\theta}$, a set of linear constraints are constructed. Each constraint k takes the form $b_k = A_k\boldsymbol{\theta}$, where b_k is a constant and A_k is a vector with the same length as $\boldsymbol{\theta}$. The values of $\boldsymbol{\theta}$ are then obtained by solving:

$$\boldsymbol{b} = A\,\boldsymbol{\theta} \tag{1}$$

using existing algorithms for solving linear systems (LAPACK, 1992).

The framework defines four types of constraints, which are based on assumptions, that can used to solve for the values of $\hat{\boldsymbol{e}}$ and ϕ via $\boldsymbol{\theta}$. The frameworks is flexible with respect to selecting and combining these

Table 1: Summary of constraint types for defining risk group turnover

Constraint	Assumption	Parameters	Types of data sources for parameterization
1. Constant group size	the relative population sizes of groups are known or assumed, and assumed to not change over time	$\hat{x}_i$	demographic health surveys (The DHS Program, 2019), key population mapping and enumeration (Abdul-Quader et al., 2014)
2. Specified elements	the relative numbers of people entering into each group upon entry into the model or after leaving another group are known or assumed	$\hat{e}_i$, ϕ_{ij}	demographic health surveys (The DHS Program, 2019), key population surveys (Baral et al., 2014)
3. Group duration	the average durations of individuals in each group are known or assumed	δ_i	cohort studies of sexual behaviour over time (Fergus et al., 2007), key population surveys (Watts et al., 2010; Baral et al., 2014)
4. Turnover rate ratios	ratios between different rates of turnover are known or assumed	ϕ_{ij}	demographic health surveys (The DHS Program, 2019), key population surveys (Baral et al., 2014)

ϕ_{ij}: rate of turnover from group i to group j; $\hat{x}_i$: proportion of individuals in risk group i; $\hat{e}_i$: proportion of individuals entering into risk group i; δ_i: average duration spent in risk group i.

Figure 2: Modelled health states: $\mathcal{S}$: susceptible; $\mathcal{I}$: infected; $\mathcal{T}$: treated; and transitions: λ: force of infection; τ: treatment.

constraints, guided by the availability of data. However, exactly G^2 non-redundant constraints must be specified to produce a unique solution, such that exactly one value of $\boldsymbol{\theta}$ satisfies all constraints. Table 1 summarizes the four types of constraints, with their underlying assumptions, and the types of data that can be used in each case. Additional details, including constraint equations, examples, and considerations for combining constraints, are in Appendix A.2.2.

2.2. Transmission model

We developed a deterministic, compartmental model of an illustrative sexually transmitted infection with 3 risk groups. We did not simulate a specific pathogen, but rather constructed a biological system that included susceptible, infectious, and treated (or recovered/immune) health states. The transmission model therefore was mechanistically representative of sexually transmitted infections like HIV, where effective antiretroviral treatment represents a health state where individuals are no longer susceptible nor infectious (Maartens et al., 2014), or hepatitis B virus, where a large proportion of individuals who clear their acute infection develop life-long protective immunity (Ganem and Prince, 2004).

The model is represented by a set of coupled ordinary differential equations (Appendix B.1) and includes three health states: susceptible $\mathcal{S}$, infectious $\mathcal{I}$, and treated $\mathcal{T}$ (Figure 2), and $G = 3$ levels of risk: high H, medium M, and low L. Risk strata are defined by different number of partners per year, so that individuals

in risk group i are assumed to form partnerships at a rate C_i per year. The probability of partnership formation ρ_{ik} between individuals in group i and individuals in risk group k is assumed to be proportionate to the total number of available partnerships within each group:

$$\rho_{ik} = \frac{C_k x_k}{\sum_{\mathrm{k}} C_{\mathrm{k}} x_{\mathrm{k}}} \tag{2}$$

The biological probability of transmission is defined as β per partnership. Individuals transition from the susceptible $\mathcal{S}$ to infectious $\mathcal{I}$ health state via a force of infection λ_i per year, per susceptible in risk group i:

$$\lambda_i = C_i \sum_k \rho_{ik}\, \beta\, \frac{\mathcal{I}_k}{x_k} \tag{3}$$

Individuals are assumed to transition from the infectious $\mathcal{I}$ to treated $\mathcal{T}$ health state at a rate τ per year, reflecting diagnosis and treatment. The treatment rate does not vary by risk group. Individuals in the treated $\mathcal{T}$ health state are neither infectious nor susceptible, and individuals cannot become re-infected.

2.2.1. Implementing turnover within the transmission model

As described in Section 2.1, individuals enter the model at a rate ν, exit the model at a rate μ, and transition from risk group i to group j at a rate ϕ_{ij}, health state. The turnover rates ϕ and distribution of individuals entering the model by risk group $\hat{\boldsymbol{e}}$ were computed using the methods outlined in Appendix A.2.2, based on the following three assumptions. First, we assumed that the proportion of individuals entering each risk group $\hat{\boldsymbol{e}}$ was equal to the proportion of individuals across risk groups in the model $\hat{\boldsymbol{x}}$. Second, we assumed that the average duration of time spent in each risk group $\boldsymbol{\delta}$ was known. Third, we assumed that the absolute number of individuals moving between two risk groups in either direction was balanced, meaning that if 10 individuals moved from group i to group j, then another 10 individuals moved from group j to group i. These three assumptions were selected because they reflect the common assumptions underlying turnover in prior models (Zhang et al., 2012; Henry and Koopman, 2015) and also to avoid any dominant direction of turnover. That is, we wanted to study the influence of movement between risk groups in general, as compared to no movement, and at various rates of movement, rather than movement predominantly from some groups to some other groups. The system of equations formulated from the above assumptions and constraints is given in Appendix B.2. To satisfy all three assumptions, there was only one possible value for each element in ϕ and $\hat{\boldsymbol{e}}$. That is, by specifying these three assumptions, we generated a unique set of ϕ and $\hat{\boldsymbol{e}}$.

Under the above three assumptions, we still needed to specify the particular values of the parameters $\hat{\boldsymbol{x}}$, $\boldsymbol{\delta}$, ν, and μ. Such parameter values could be derived from data as described in Appendix A.2.2. However, in all our experiments, we used the illustrative values summarized in Table 2. After resolving the system of

Table 2: Default model parameters for experiments

Symbol	Description	Default value
$\boldsymbol{\beta}$	transmission probability per partnership	0.03
τ	rate of treatment initiation among infected	0.1
N_0	initial population size	1000
$\hat{\boldsymbol{x}}$	proportion of system individuals by risk group	[0.05 0.20 0.75]
$\hat{\boldsymbol{e}}$	proportion of entering individuals risk by risk group	[0.05 0.20 0.75]
$\boldsymbol{\delta}$	average duration spent in each risk group	[5 15 25]
C	number of partners per year by individuals in each risk group	[25 5 1]
ν	rate of population entry	0.05
μ	rate of population exit	0.03

All rates have units year^{-1}; durations are in years; parameters stratified by risk group are written [high, medium, low] risk.

equations Eq. (1) using these values, $\hat{\boldsymbol{e}}$ was equal to $\hat{\boldsymbol{x}}$ (assumed), and ϕ was:

$$\phi = \begin{bmatrix} * & 0.0833 & 0.0867 \\ 0.0208 & * & 0.0158 \\ 0.0058 & 0.0042 & * \end{bmatrix} \tag{4}$$

We then simulated epidemics using ϕ above and the parameters shown in Table 2. The transmission model was initialized with $N_0 = 1000$ individuals who were distributed across risk groups according to $\hat{\boldsymbol{x}}$. We seeded the epidemic with one infectious individual in each risk group at $t = 0$ in an otherwise fully susceptible populatuon. We numerically solved the system of ordinary differential equations (Appendix B.1) in Python using Euler's method with a time step of $dt = 0.1$ years. Code for all aspects of the project is available at: `https://github.com/c-uhs/turnover`.

2.3. Experiments

We designed three experiments to examine the influence of turnover on simulated epidemics. We analyzed all outcomes at equilibrium, defined as steady state at $t = 500$ years with $< 1\%$ change in incidence per year.

2.3.1. Experiment 1: Mechanisms by which turnover influences equilibrium prevalence

We designed Experiment 1 to explore the mechanisms by which turnover influences the equilibrium STI prevalence of infection, and the ratio of prevalence between risk groups (prevalence ratios). We defined prevalence as $\hat{\mathcal{I}}_i = \dfrac{\mathcal{I}_i}{x_i}$. Similar to previous studies (Zhang et al., 2012; Henry and Koopman, 2015), we

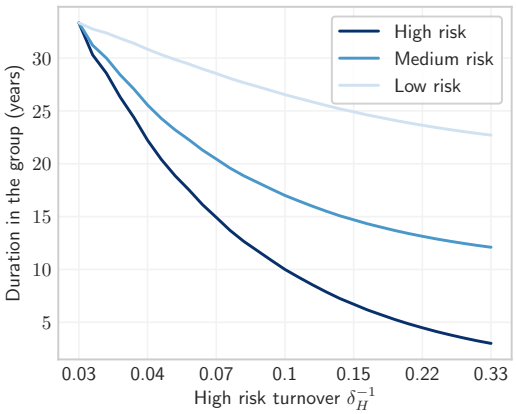


Figure 3: Average duration of time spent in each risk group versus turnover. Turnover rate (log scale) is a function of the duration of time spent in the high risk group δ_H, where shorter time spent in the high risk group yields faster turnover.

varied the rates of turnover using a single parameter. However, because our model had $G = 3$ risk groups, multiplying a set of base rates ϕ by a scalar factor would change the relative population sizes of risk groups $\hat{\boldsymbol{x}}$. Instead of a scalar factor, we controlled the rates of turnover using the duration of time spent in the high risk group δ_H, because of the practical interpretation of δ_H in the context of STI transmission, such as the duration in formal sex work (Watts et al., 2010). A shorter δ_H yielded higher rates of turnover among all groups. The duration of time spent in the medium risk group δ_M was then defined as a value between δ_H and the maximum duration μ^{-1} which scaled with δ_H following: $\delta_M = \delta_H + \kappa\left(\mu^{-1} - \delta_H\right)$, with $\kappa = 0.3$. The duration of time in the low risk group δ_L similarly scaled with δ_H, but due to existing constraints, specification of δ_H and δ_M ensured only one possible value of δ_L. Thus, each value of δ_H yielded a unique set of turnover rates ϕ whose elements all scaled inversely with the duration in the high risk group δ_H.

We varied δ_H across a range of 33 to 3 years, reflecting a range from the full duration of simulated sexual activity $\mu^{-1} \approx 33$ years, through an average reported duration in sex work as low as 3 years (Watts et al., 2010). The resulting duration of time spent in each group versus turnover in the high risk group δ_H^{-1} is shown in Figure 3. For each set of turnover rates, we plotted the equilibrium prevalence in each risk group, and the prevalence ratios between high/low, high/medium, and medium/low risk groups. In order to understand the mechanisms by which turnover influenced prevalence and prevalence ratios (Objective 1), we additionally plotted the four components which contributed to gain/loss of infectious individuals in each risk group, based on Eq. (B.2): 1) net gain/loss via turnover of infectious individuals, 2) gain via incident infections, 3) loss via treatment, and 4) loss via death. The influence of turnover on prevalence was only mediated by components 1 and 2, since components 3 and 4 were defined as constant rates which did not change with turnover; as such, our analysis focused on components 1 and 2. Finally, to further understand trends in incident infections versus turnover (component 1), we factored equation Eq. (3) for incidence λ_i

into constant and non-constant factors, and plotted the non-constant factors versus turnover.

2.3.2. Experiment 2: Inferred risk heterogeneity with vs without turnover

We designed Experiment 2 to examine how the inclusion versus exclusion of turnover influences the inference of transmission model parameters related to risk heterogeneity, specifically the numbers of partners per year C_i across risk groups. The ratio of partner numbers C_H / C_L is one way to measure of how different the two risk groups are with respect to acquisition and transmission risks. Indeed, ratios of partner numbers are often used when parameterizing risk heterogeneity in STI transmission models (Mishra et al., 2012).

First, we fit the transmission model with turnover and without turnover, to equilibrium infection prevalence across risk groups. Specifically, we held all other parameters at their default values and fit the numbers of partners per year in each risk group C_i to reproduce the following: 20% infection prevalence among the high risk group, 8.75% among the medium risk group, 3% among the low risk group, and 5% overall. To identify the set of parameters (i.e. partner numbers C in each risk group) that best reproduced the fitting targets, we minimized the negative log-likelihood of group-specific and overall prevalence. Sample sizes of 500, 2000, 7500, and 10,000 were assumed to generate binomial distributions for the high, medium, low, and overall prevalence targets respectively, reflecting typical sample sizes in nationally representative demographic and health surveys (The DHS Program, 2019), multiplied by the relative sizes of risk groups in the model $\hat{\boldsymbol{x}}$. The minimization was performed using the SLSQP method (Kraft, 1988) from the SciPy Python `minimize` package. To address our second objective, we compared the fitted (posterior) ratio of partners per year C_H / C_L in the model with turnover versus the model without turnover.

2.3.3. Experiment 3: Influence of turnover on the tPAF of the highest risk group

We designed Experiment 3 to examine how the tPAF of the highest risk group varies when estimated by a model with versus without turnover (Objective 3). We calculated the tPAF of risk group i by comparing the relative difference in cumulative incidence between a base scenario, and a counterfactual where transmission from group i is turned off, starting at the fitted equilibrium. That is, in the counterfactual scenario, infected individuals in the high risk group could not transmit the infection. The tPAF was calculated over different time-horizons (1 to 50 years) as (Mishra et al., 2014):

$$\text{tPAF}_i(t) = \frac{\int_{t_{eq}}^{t} I_b(\tau)\, d\tau - \int_{t_{eq}}^{t} I_c(\tau)\, d\tau}{\int_{t_{eq}}^{t} I_b(\tau)\, d\tau} \tag{5}$$

where t_{eq} is the time corresponding to equilibrium, $I_b(t)$ is the rate of new infections at time t in the base scenario, and $I_c(t)$ is the rate of new infections at time t in the counterfactual scenario. We then compared the tPAF generated from the fitted model with turnover to the tPAF generated from the fitted model without turnover.

3. Results

3.1. Experiment 1: Mechanisms by which turnover influences equilibrium prevalence

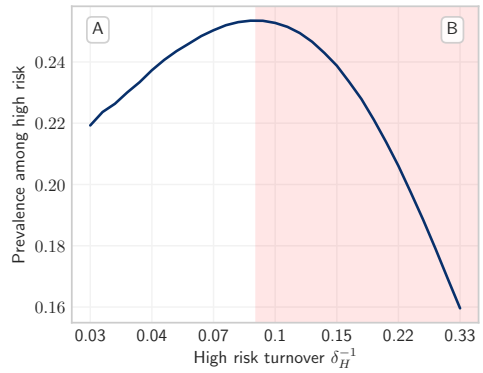
Figure 4 illustrates trends in equilibrium STI prevalence among the high, medium, low, risk groups, at different rates of turnover, controlled by duration of time spent in the highest risk group. We make two observations in Figure 4. First, there was an inverted U-shaped relationship between STI prevalence and turnover in all three risk groups. That is, equilibrium STI prevalence increased with slow versus no turnover (region A), reached a maximum, and then declined under faster turnover (region B). Second, the threshold turnover rate at which group-specific prevalence peaked varied by risk group. The high risk group had the lowest turnover threshold (Figure 4a), while the low risk group had the highest (Figure 4c). To explain the inverted U-shape and different turnover thresholds by group, we examined the components contributing to prevalence, first in the high risk group, and then in the low risk group.

Figure 5 shows the four components which contributed to gain/loss of infectious individuals in each risk group: 1) net gain/loss via turnover, 2) gain via incident infections, 3) loss via treatment, and 4) loss via death, for each risk group in the model, at equilibrium under different rates of turnover. Figure 6 also illustrates the distribution of health states within each risk group and among individuals moving between risk groups under four different rates of turnover, for reference.

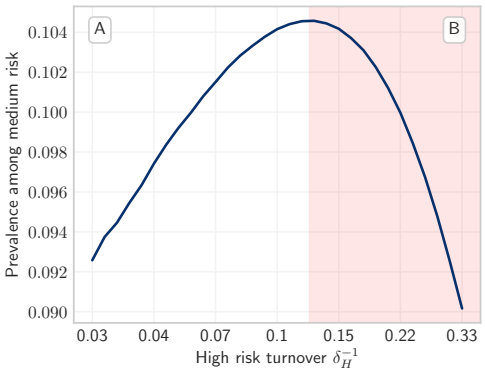
3.1.1. Influence of turnover on equilibrium prevalence in the high risk group

From Figure 6, we see that the high risk group had a larger proportion of infectious individuals compared to other groups, and so infectious individuals leaving the high risk group via turnover were replaced by mainly susceptible individuals from lower risk groups. Turnover therefore yielded a net loss of infectious individuals in the high risk group via turnover (Figure 5a, yellow), which acted to decrease prevalence in the group. However, turnover similarly caused a net replacement of treated individuals with susceptible individuals in the high risk group (Figure 6). Loss of treated individuals acted to reduce herd immunity, thereby increasing the rate of incident infections, which we observed for slow rates of turnover versus no turnover (Figure 5a, red). Under slow rates of turnover, increasing turnover added more incident infections to the high risk group via incidence than were lost via turnover (Figure 5a), which explains why equilibrium prevalence in the high risk group increased with turnover in region A (Figure 4a).

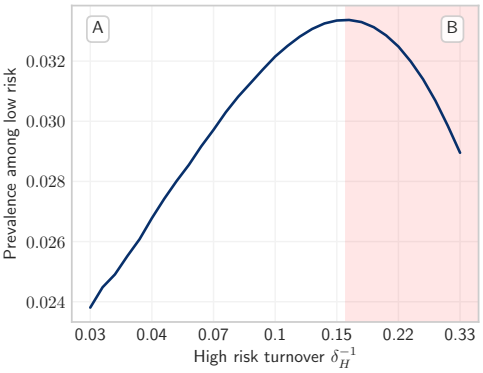
Under faster rates of turnover, the rate of incident infections in the high risk group did not keep pace with the loss of infectious individuals via turnover (Figure 5a); thus prevalence declined with increasing turnover (Figure 4a, region B). Two phenomena contributed to decreasing rate of incident infections in the high risk group under fast versus slow rates of turnover. First, as herd immunity in the high risk group was reduced via turnover, the additional infections attributable to further reducing this barrier to transmission decreased. Second, the net movement of infectious individuals from high to low risk (Figure 6) reduced the



(a) High risk

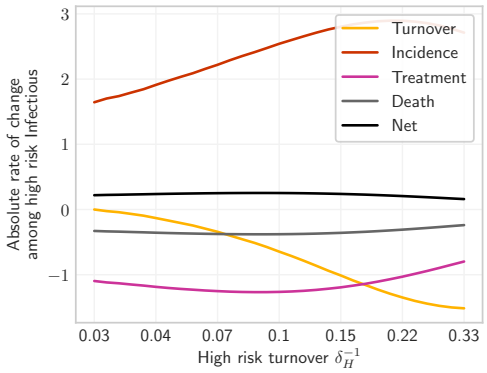


(b) Medium risk

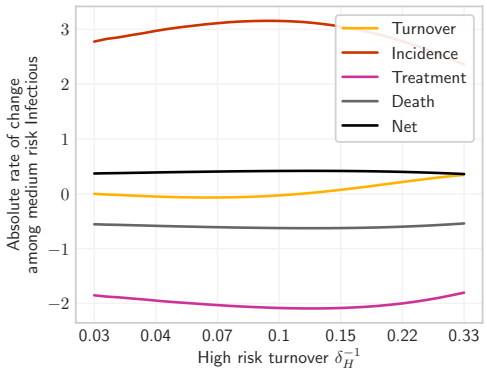


(c) Low risk

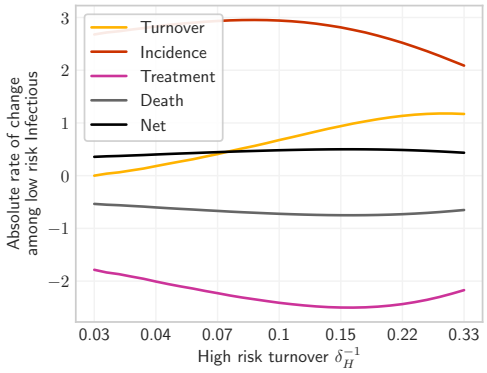
Figure 4: Relationship between equilibrium STI prevalence in high, medium, and low risk groups versus turnover rate. Regions A and B denote where prevalence increases and decreases with turnover, respectively. See Figure 3 for x-axis definition.



(a) High risk



(b) Medium risk



(c) Low risk

Figure 5: Absolute rates of change at equilibrium (number of individuals gained/lost per year) among infectious individuals in each risk group, broken down by type of change: gain via incident infections, loss via treatment, loss/gain via turnover, loss via death, and net change. Based on Eq. (B.2). Rates of change do not sum to zero due to population growth. See Figure 3 for x-axis definition.

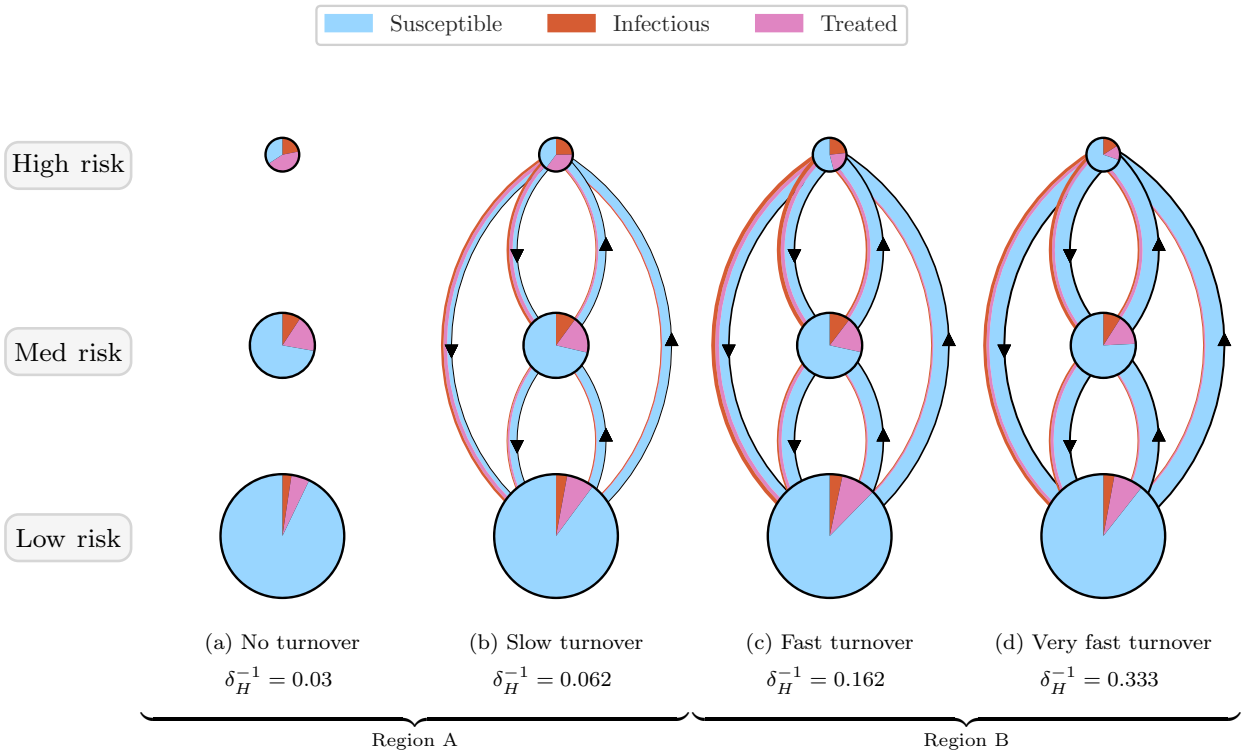
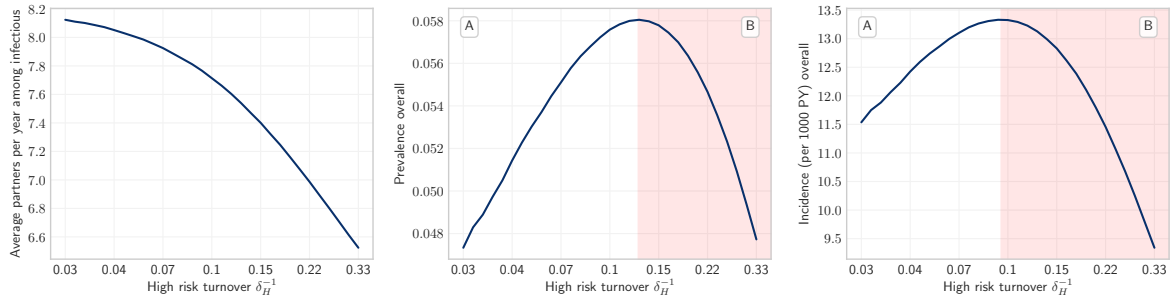


Figure 6: Depiction of health states of individuals in each risk group and of individuals moving between risk groups, obtained from models at equilibrium under four overall rates of turnover. Circle sizes are proportional to risk group sizes. Circle slices and arrow widths are also proportional to the proportion of health states within risk groups and among individuals moving between risk groups, respectively. However, circle sizes and arrow widths do not have comparable scales.

average number of partners per year among infectious individuals (Figure 7a). As shown in Appendix B.4, under our modelling assumptions, incidence in all risk groups was proportional to the average number of partners per year among infectious individuals (Figure 7a), and overall prevalence (Figure 7b). Thus, as the average number of partners per year among infectious individuals decreased with turnover, incidence also decreased (Figure 7c). Moreover, as equilibrium incidence decreased across all groups with faster turnover for the same reason, equilibrium prevalence then decreased, which in turn reduced incidence, and so on in a mutually reinforcing exponential decline (Figure 4, region B).

3.1.2. Influence of turnover on equilibrium prevalence in the low risk group

As shown in Figure 6, the low risk group was composed mainly of susceptible individuals at equilibrium, so increasing turnover yielded a net replacement of susceptible individuals with infectious and treated individuals. The net influx of infectious individuals acted to increase equilibrium prevalence in the low risk group (Figure 5c, yellow). The net influx of treated individuals contributed to a small increase in herd immunity, but the herd immunity effect in the low risk group remained negligible due to the small proportion of treated



(a) Average number of partners per year among infectious individuals

(b) Overall prevalence

(c) Overall incidence

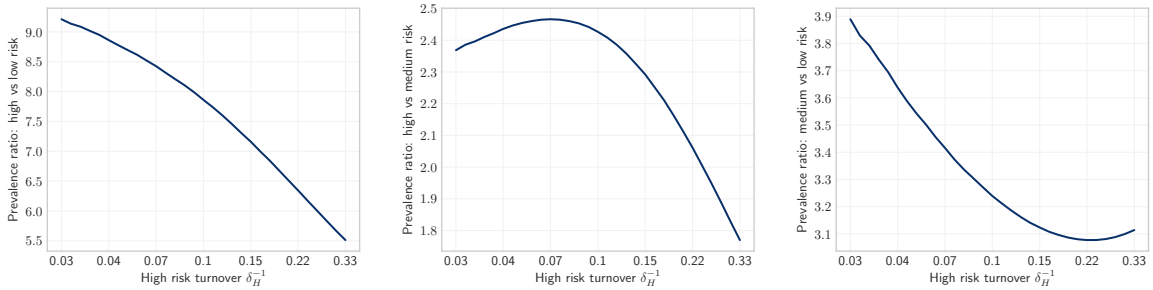
Figure 7: Overall incidence and the non-constant factors of incidence versus turnover. The product of factors (a) and (b) is proportional to (c) overall incidence. See Appendix B.4 for proof. See Figure 3 for x-axis definition.

relative to susceptible individuals. Incident infections also increased with slow versus no turnover in the low risk group (Figure 5c, red), driven by increasing overall incidence (Figure 7c), ultimately attributable to reduced herd immunity in the high risk group (see Section 3.1.1). The net gain of infectious individuals via turnover and increasing incident infections then both increased equilibrium prevalence in the low risk group under slow versus no turnover (Figure 4c, region A).

Faster turnover still contributed a net gain of infectious individuals via movement of already infected individuals. However, as in the high risk group, the rate of incident infections decreased in the low risk group under faster rates of turnover (Figure 5c, red), due to decreasing overall incidence (Figure 7c) via decreasing number of partners per year among infectious individuals (Figure 7a). As a result, equilibrium prevalence decreased in the low risk group as turnover increased (Figure 4c, region B), following the same exponential decline as in the high risk group under faster rates of turnover.

3.1.3. Prevalence ratios and different turnover rates for peak prevalence among risk groups

The movement of individuals between risk groups via turnover yielded a net loss of infectious individuals from the high risk group, but a net gain of infectious individuals in the low risk group (Figure 5, yellow). By comparison, the influence of turnover on the rate of incident infections was relatively similar among high versus low risk groups (Figure 5, red). Thus, the net effect of increasing turnover in the high risk group *tended towards* decreasing prevalence, whereas in the low risk group, the net effect *tended towards* increasing prevalence. Since these two tendencies act to bring prevalence between the two groups closer together, we summarize these two tendencies as the "homogenizing effect" of turnover on group-specific prevalence. The homogenizing effect is demonstrated in Figure 8, which shows the equilibrium prevalence ratios in pair-wise comparisons of all three risk groups. The prevalence ratio between highest and lowest risk groups monotonically decreased with turnover (Figure 8a). The homogenizing effect is also why, in the high risk group, lower rates of turnover were required to have a net decreasing effect on prevalence, as compared



(a) High vs Low risk (b) High vs Medium Risk (c) Medium vs Low Risk

Figure 8: Equilibrium prevalence ratios between risk groups under different rates of turnover. See Figure 3 for x-axis definition.

to in the low risk group. That is, the threshold rate of turnover at which prevalence in the high risk group peaked and then declined was lower than in the low risk group (Figure 4).

3.2. Experiment 2: Inferred risk heterogeneity with vs without turnover

Next, we compared the fitted parameters in models with versus without turnover ($\delta_H = 5$ vs $\delta_H = 33$ years). Before model fitting, the predicted prevalence ratio between high and low risk groups was lower with turnover versus without: 6.7 vs 9.2. This reflects the homogenizing effect of turnover on group-specific prevalence. As shown in Figure 8a, the high-to-low prevalence ratio consistently declined with turnover. Thus, when fitting the model to target prevalence values, the fitted numbers of partners per year C would have to compensate for this difference in prevalence ratio with versus without turnover.

After fitting the number of partners per year, both models predicted the target equilibrium infection prevalence values of 20%, 8.75%, 3%, and 5% among the high, medium, low risk groups, and overall (Figure C.3). However, in order to do so, the ratio of fitted partners between high and low risk groups (C_H / C_L) was higher with turnover than without: 23.9 vs 15.2 (Table 3). That is, the inferred level of risk heterogeneity was higher in the model with turnover than in the model without turnover. In order to observe the same prevalence ratio in a system with turnover, the homogenizing effect of turnover must be overcome by greater heterogeneity in risk, as compared to a system without turnover.

3.3. Experiment 3: Influence of turnover on the tPAF of the highest risk group

Finally, we compared the predicted tPAF of the highest risk group with and without turnover, after fitting to the same prevalence data (Figure 9). The tPAF approaches 1 for both models over the 50 year period, indicating that unmet treatment needs of the highest risk group are central to epidemic persistence in both scenarios. SB: Crucial. The estimated tPAF of the highest risk group is also higher in the model with turnover versus in the model without turnover over all time horizons. This increase in tPAF of the highest risk group can be attributed to a higher ratio of fitted partners C_H / C_L in the model with turnover

Table 3: Equilibrium partnership formation rates and prevalence among the high and low risk groups predicted by the models with and without turnover, before and after model fitting.

	Number of Partners			Prevalence		
Context	High	Low	High/Low	High	Low	High/Low
No Turnover	25.0	1.0	25.0	21.9%	2.4%	**9.2**
Turnover	25.0	1.0	25.0	21.6%	3.2%	**6.7**
No Turnover [fit]	23.5	1.5	**15.2**	20.0%	3.0%	6.7
Turnover [fit]	24.3	1.0	**23.9**	20.0%	3.0%	6.7

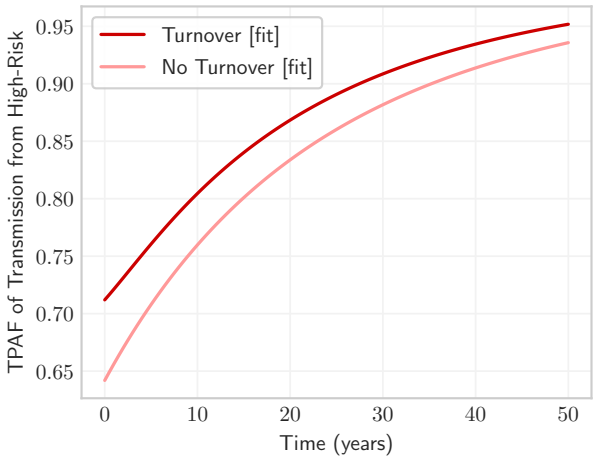


Figure 9: Transmission population attributable fraction (tPAF) of the high risk group in models with and without turnover, after fitting the number of partners per year to group-specific prevalence.

is higher than in the model without (Experiment 2). The increased partner ratio affords a higher risk of onward transmission to the highest risk group in the model with turnover, and thus an increase in tPAF. This result then implies that models which fail to capture turnover dynamics which are present in reality may underestimate the tPAF of high risk groups.

4. Discussion

We developed a new framework for modelling turnover of individuals among risk groups, based on a flexible combination of data-driven constraints. We then used this framework to explore the influence of turnover on the contribution of the highest risk group to onward transmission in an illustrative STI epidemic without STI-attributable mortality. We found that failure to model turnover when simulating settings with turnover could result in underestimation of the contribution of the highest risk group to onward transmission.

Influence of turnover on incidence & prevalence. We found that turnover influences the overall equilibrium STI prevalence and incidence in a pattern consistent with previous works (Stigum et al., 1994; Zhang et al., 2012; Henry and Koopman, 2015). However, unlike previous works, we also illustrated the influence of turnover on group-specific prevalence, and demonstrated mechanistically how this occurs. For example, we showed that some infections among the low risk groups were acquired by individuals during a previous period of higher risk (see Figure C.8 for additional results). We can summarize this influence of turnover, including reduced ratio of STI prevalence between the highest and lowest risk groups, as "reducing heterogeneity in risk" via movement of individuals between risk groups. Henry and Koopman (2015) demonstrated that such reduction in risk heterogeneity through turnover decreases the basic reproductive number and thus, means epidemic control could be easier to achieve. Our findings on the mechanisms by which increasing turnover reduces the STI treatment rate to achieve zero prevalence further supports these insights from Henry and Koopman (2015).

Implications for interventions. Our comparison of models with and without turnover, calibrated to the same epidemic, showed that if turnover exists in a given setting but it is ignored in a model, the tPAF of high risk groups will be underestimated by the model. This is because heterogeneity in risk must be higher in the presence of turnover than in a model without turnover in order to produce the same epidemic features. Although we examined a single parameter to capture risk (frequency of partner change), these insights would be generalizable to any other component of susceptibility or infectivity, because the risk per susceptible individual (force of infection) includes both biological transmission probabilities and frequency of partner change (Anderson and May, 1991). In the context of models with assortative mixing (individuals are more likely to form partnerships with other individuals in their risk group), the difference between the tPAFs estimated by the model with versus without turnover is therefore expected to be even larger. The public health implication of models ignoring turnover which is present in reality is that the tPAF of high risk groups will be systematically underestimated, potentially misguiding resources away from high risk groups. Follow on research should quantify the size of this potential bias in tPAFs generated from models without turnover, and characterize the epidemiologic conditions under which the bias would be largest.

Turnover framework. The proposed framework provides a flexible way to parameterize risk group turnover based on available epidemiologic data and/or assumptions. The approach has five advantages. First, the framework defines how specific epidemiologic data and assumptions can be used as constraints to help define rates of turnover (Table 1). These data and assumptions are further discussed in the next paragraph. Second, the framework allows such constraints, which take the form $b_k = A_k\boldsymbol{\theta}$ to be chosen and combined in a flexible way, depending on which data are available, or which assumptions are most plausible. While it is necessary that constraints do not conflict one another, it is not necessary that a complete set of G^2 constraints be defined (where G is the number of risk groups), since optimization techniques can be used to calculate, for

example, the smallest possible values of ϕ which satisfy the given constraints. Third, this flexible approach also allows the framework to scale to any number of risk groups, G. Fourth, we have shown how several previous implementations of turnover (Stigum et al., 1994; Eaton and Hallett, 2014; Henry and Koopman, 2015) can be recreated exactly using the proposed framework (Appendix A.3). In so doing, we highlight which specific assumptions are the same and which are different across the different implementations. Fifth and finally, the framework avoids the need for a burn-in period to establish a demographic steady-state before introducing infection, which was required in some previous models (Boily et al., 2015).

The types of data which can be used by this approach include the following four categories (Table 1). First, the proportion of total individuals in each risk group $\hat{x}_i$ is required for constraints of type 1. In the context of STIs, risk group size estimates may be obtained from demographic health surveys (The DHS Program, 2019), and from mapping and enumeration of marginalized populations, such as sex workers (Abdul-Quader et al., 2014). For example, one risk group may be defined by any engagement in casual sex within the past year, and the corresponding proportion of the population could be estimated from self-reported behaviours in demographic health surveys (The DHS Program, 2019). Second, the proportion of individuals who enter into each risk group $\hat{e}_i$ upon entry into the model can be used to define additional constraints of type 2. Such proportions could be obtained the same way as $\hat{x}_i$ above, except using only data from individuals who recently became sexually active. For example, among women who became sexually active in the past year, what proportion also engaged in casual sex within the past year. If data on sexual debut is not available, then recent entry into sexual activity could be approximated using a suitable age range. Third, the average duration of time spent within each risk group δ_i can similarly be used to define type 3 constraints. Cross-sectional survey questions asked of female sex workers such as "for how many years have you been a sex worker?" may be used to obtain estimates of duration in sex work, with the recognition that such data are censured (Watts et al., 2010). Longitudinal, or cohort studies that track the self-reported sexual behaviour over time can also provide estimates of duration of time spent in other risk groups (Fergus et al., 2007). Fourth, similar studies may provide data on trajectories in risk behaviour of individuals over time, which can be used to estimates specific transition rates ϕ (type 2 constraints) or ratios of transition rates (type 4 constraints). For example, upon retirement from sex work, twice as many former sex workers may enter into monogamous relationships, as compared to those who continue to form multiple partnerships.

Limitations. There are six limitations of the study that are important when considering the proposed turnover framework and the implications of our results. First, our approach does not account for infection-attributable mortality, such as HIV-attributable mortality. It is well-established that HIV-attributable mortality will reduce the relative size of higher risk groups, and that alone can cause an epidemic to decline (Boily and Mâsse, 1997). As such, many models of HIV transmission that include very small ($< 3\%$ of the

population) high risk groups, such as female sex workers, often do not constrain the relative size of the sub-group populations to be stable over time (Pickles et al., 2013). Second, we assumed a single-sex population and did not stratify by age. In the context of real-world STI epidemics, the relative size of risk groups may differ by both sex and age, such as the number of females who sell sex, versus the number of males who sell sex, versus the number of males who pay for sex (clients). Future work on the proposed framework could incorporate both infection-attributable mortality, and age-sex stratifications into the model. Third, we assumed that turnover rates were not affected by the health status of individuals, which could change the observed influence of turnover on equilibrium STI prevalence and incidence. While changes to health state may affect the sexual behaviour of individuals, its not clear what overall patterns emerge at the population level. Fourth, we did not include individuals becoming re-susceptible – an important feature of many STIs such as syphilis and gonorrhoea (Fenton et al., 2008). As shown by Fenton et al. (2008) and Pourbohloul et al. (2003), the re-supply of susceptible individuals following STI treatment will fuel an epidemic, and so the influence of turnover on STI incidence, prevalence, and tPAF may be different, and warrants future study. Fifth, our analyses were restricted to equilibrium STI prevalence and incidence. The influence of turnover at different phases of an epidemic – growth, mature, declining – are expected to vary, and thus represents an important topic for future investigation. Finally, our analyses reflected an illustrative STI epidemic in a population with illustrative risk strata. Future work should explore more realistic systems for specific STIs, such as the work by Johnson and Geffen (2016).

Conclusion. In conclusion, turnover in risk will influence epidemic model outputs, including projected incidence, prevalence, and measures of the contribution of high risk groups to overall STI transmission. Turnover should therefore be considered in transmission models with heterogeneity in risk. The methods presented here illustrate how epidemiologic data can be used to parametrize turnover in epidemic models. We hope that this will support accurate estimation of the importance of addressing the unmet needs of high risk populations – including gay men and other men who have sex with men, transgender women, people who use drugs, and people of all genders who sell sex – to achieve population-level transmission reduction.

Acknowledgements

We would like to thank Kristy Yiu (MAP Centre for Urban Health Solutions, Unity Health Toronto) for logistical support, the Siyaphambili research team for helpful discussions, and Carly Comins (Johns Hopkins University) for facilitating the modelling discussions with the wider study team. SM is supported by an Ontario HIV Treatment Network and Canadian Institutes of Health Research New Investigator Award.

Contributions

JK and SM conceptualized the study and drafted the manuscript; JK designed the experiments with input from LW, HM, and SM. JK developed the unified framework and conducted the modelling, experiments, and analyses; conducted the literature review, and drafted the first version of the manuscript. LW, HM, SB, and SS led substantial structural revisions to the manuscript, including assessment of epidemiological constraints and assumptions; and provided critical discussion surrounding implications of findings. All authors contributed to interpretation of the results and manuscript revision.

Funding

The study was supported by the National Institutes of Health, Grant number: NR016650; the Center for AIDS Research, Johns Hopkins University through the National Institutes of Health, Grant number: P30AI094189.

Conflicts of Interest

Declarations of interest: none.

References

Abdul-Quader, A.S., Baughman, A.L., Hladik, W., 2014. Estimating the size of key populations: Current status and future possibilities 9, 107–114. doi:DOI 10.1097/COH.0000000000000041.

Anderson, R.M., May, R.M., 1991. Infectious diseases of humans: dynamics and control. Infectious diseases of humans: dynamics and control. .

Baral, S., Ketende, S., Green, J.L., Chen, P.A.A., Grosso, A., Sithole, B., Ntshangase, C., Yam, E., Kerrigan, D., Kennedy, C.E., Adams, D., 2014. Reconceptualizing the HIV epidemiology and prevention needs of female sex workers (FSW) in Swaziland. PLoS ONE 9, e115465. doi:DOI 10.1371/journal.pone.0115465.

Baral, S., Logie, C.H., Grosso, A., Wirtz, A.L., Beyrer, C., 2013. Modified social ecological model: A tool to guide the assessment of the risks and risk contexts of HIV epidemics. BMC Public Health 13, 482. doi:DOI 10.1186/1471-2458-13-482.

Boily, M.C., Mâsse, B., 1997. Mathematical models of disease transmission: A precious tool for the study of sexually transmitted diseases. Canadian Journal of Public Health 88, 255–265. doi:DOI 10.1007/bf03404793.

Boily, M.C., Pickles, M., Alary, M., Baral, S., Blanchard, J., Moses, S., Vickerman, P., Mishra, S., 2015. What really is a concentrated HIV epidemic and what does it mean for West and Central Africa? Insights from mathematical modeling. Journal of Acquired Immune Deficiency Syndromes 68, S74–S82. doi:DOI 10.1097/QAI.0000000000000437.

Case, K., Ghys, P., Gouws, E., Eaton, J., Borquez, A., Stover, J., Cuchi, P., Abu-Raddad, L., Garnett, G., Hallett, T., 2012. Understanding the modes of tranmission model of new HIV infection and its use in prevention planning. Bulletin of the World Health Organization 90, 831–838. doi:DOI 10.2471/blt.12.102574.

DataBank, 2019. Population estimates and projections. URL: https://databank.worldbank.org/source/population-estimates-and-projections.

Eaton, J.W., Hallett, T.B., 2014. Why the proportion of transmission during early-stage HIV infection does not predict the long-term impact of treatment on HIV incidence. Proceedings of the National Academy of Sciences 111, 16202–16207. doi:DOI 10.1073/pnas.1323007111.

Fenton, K.A., Breban, R., Vardavas, R., Okano, J.T., Martin, T., Aral, S., Blower, S., 2008. Infectious syphilis in high-income settings in the 21st century. The Lancet Infectious Diseases 8, 244–253. doi:DOI 10.1016/S1473-3099(08)70065-3.

Fergus, S., Zimmerman, M.A., Caldwell, C.H., 2007. Growth trajectories of sexual risk behavior in adolescence and young adulthood. American Journal of Public Health 97, 1096–1101. doi:DOI 10.2105/AJPH.2005.074609.

Ganem, D., Prince, A.M., 2004. Hepatitis B Virus Infection — Natural History and Clinical Consequences. New England Journal of Medicine 350, 1118–1129. doi:DOI 10.1056/NEJMra031087.

Hanley, J.A., 2001. A heuristic approach to the formulas for population attributable fraction. Journal of Epidemiology and Community Health 55, 508–514. doi:DOI 10.1136/jech.55.7.508.

Henry, C.J., Koopman, J.S., 2015. Strong influence of behavioral dynamics on the ability of testing and treating HIV to stop transmission. Scientific Reports 5, 9467. doi:DOI 10.1038/srep09467.

ICAP, 2019. PHIA Project. URL: https://phia.icap.columbia.edu.

Johnson, L.F., Geffen, N., 2016. A Comparison of two mathematical modeling frameworks for evaluating sexually transmitted infection epidemiology. Sexually Transmitted Diseases 43, 139–146. doi:DOI 10.1097/OLQ.0000000000000412.

Koopman, J.S., Jacquez, J.A., Welch, G.W., Simon, C.P., Foxman, B., Pollock, S.M., Barth-Jones, D., Adams, A.L., Lange, K., 1997. The role of early HIV infection in the spread of HIV through populations. Journal of Acquired Immune Deficiency Syndromes 14, 249–58. URL: http://www.ncbi.nlm.nih.gov/pubmed/9117458.

Kraft, D., 1988. A software package for sequential quadratic programming. Technical Report DFVLR-FB 88-28. DLR German Aerospace Center — Institute for Flight Mechanics. Koln, Germany.

LAPACK, 1992. LAPACK: Linear Algebra PACKage. URL: http://www.netlib.org/lapack.

Lawson, C.L., Hanson, R.J., 1995. Solving least squares problems. volume 15. SIAM.

Maartens, G., Celum, C., Lewin, S.R., 2014. HIV infection: Epidemiology, pathogenesis, treatment, and prevention. The Lancet 384, 258–271. doi:DOI 10.1016/S0140-6736(14)60164-1.

Maheu-Giroux, M., Vesga, J.F., Diabaté, S., Alary, M., Baral, S., Diouf, D., Abo, K., Boily, M.C., 2017. Changing Dynamics of HIV Transmission in Côte d'Ivoire: Modeling Who Acquired and Transmitted Infections and Estimating the Impact of Past HIV Interventions (1976-2015). Journal of Acquired Immune Deficiency Syndromes 75, 517–527. doi:DOI 10.1097/QAI.0000000000001434.

Malthus, T.R., 1798. An Essay on the Principle of Population.

Marston, C., King, E., 2006. Factors that shape young people's sexual behaviour: a systematic review. Lancet 368, 1581–1586. doi:DOI 10.1016/S0140-6736(06)69662-1.

Mishra, S., Boily, M.C., Schwartz, S., Beyrer, C., Blanchard, J.F., Moses, S., Castor, D., Phaswana-Mafuya, N., Vickerman, P., Drame, F., Alary, M., Baral, S.D., 2016. Data and methods to characterize the role of sex work and to inform sex work programs in generalized HIV epidemics: evidence to challenge assumptions. Annals of Epidemiology 26, 557–569. doi:DOI 10.1016/j.annepidem.2016.06.004.

Mishra, S., Pickles, M., Blanchard, J.F., Moses, S., Boily, M.C., 2014. Distinguishing sources of HIV transmission from the distribution of newly acquired HIV infections: Why is it important for HIV prevention planning? Sexually Transmitted Infections 90, 19–25. doi:DOI 10.1136/sextrans-2013-051250.

Mishra, S., Steen, R., Gerbase, A., Lo, Y.R., Boily, M.C., 2012. Impact of High-Risk Sex and Focused Interventions in Heterosexual HIV Epidemics: A Systematic Review of Mathematical Models. PLoS ONE 7, e50691. doi:DOI 10.1371/journal.pone.0050691.

Mukandavire, C., Walker, J., Schwartz, S., Boily, M.C., Danon, L., Lyons, C., Diouf, D., Liestman, B., Diouf, N.L., Drame, F., Coly, K., Muhire, R.S.M., Thiam, S., Diallo, P.A.N., Kane, C.T., Ndour, C., Volz, E., Mishra, S., Baral, S., Vickerman, P., 2018. Estimating the contribution of key populations towards the spread of HIV in Dakar, Senegal. Journal of the International AIDS Society 21, e25126. doi:DOI 10.1002/jia2.25126.

Pickles, M., Boily, M.C., Vickerman, P., Lowndes, C.M., Moses, S., Blanchard, J.F., Deering, K.N., Bradley, J., Ramesh, B.M., Washington, R., Adhikary, R., Mainkar, M., Paranjape, R.S., Alary, M., 2013. Assessment of the population-level effectiveness of the Avahan HIV-prevention programme in South India: A preplanned, causal-pathway-based modelling analysis. The Lancet Global Health 1, e289–e299. doi:DOI 10.1016/S2214-109X(13)70083-4.

Pourbohloul, B., Rekart, M.L., Brunham, R.C., 2003. Impact of mass treatment on syphilis transmission: A mathematical modeling approach. Sexually Transmitted Diseases 30, 297–305. doi:DOI 10.1097/00007435-200304000-00005.

Prüss-Ustün, A., Wolf, J., Driscoll, T., Degenhardt, L., Neira, M., Calleja, J.M.G., 2013. HIV Due to Female Sex Work: Regional and Global Estimates. PLoS ONE 8, e63476. doi:DOI 10.1371/journal.pone.0063476.

Shubber, Z., Mishra, S., Vesga, J.F., Boily, M.C., 2014. The HIV modes of transmission model: A systematic review of its findings and adherence to guidelines. Journal of the International AIDS Society 17, 18928. doi:DOI 10.7448/IAS.17.1.18928.

Stigum, H., Falck, W., Magnus, P., 1994. The core group revisited: The effect of partner mixing and migration on the spread of gonorrhea, chlamydia, and HIV. Mathematical Biosciences 120, 1–23. doi:DOI 10.1016/0025-5564(94)90036-1.

The DHS Program, 2019. Data. URL: https://www.dhsprogram.com.

Watts, C., Zimmerman, C., Foss, A.M., Hossain, M., Cox, A., Vickerman, P., 2010. Remodelling core group theory: the role of sustaining populations in HIV transmission. Sexually Transmitted Infections 86, iii85–iii92. doi:DOI 10.1136/sti.2010.044602.

Yorke, J.A., Hethcote, H.W., Nold, A., 1978. Dynamics and control of the transmission of gonorrhea. Sexually Transmitted Diseases 5, 51–56. doi:DOI 10.1097/00007435-197804000-00003.

Zhang, X., Zhong, L., Romero-Severson, E., Alam, S.J., Henry, C.J., Volz, E.M., Koopman, J.S., 2012. Episodic HIV Risk Behavior Can Greatly Amplify HIV Prevalence and the Fraction of Transmissions from Acute HIV Infection. Statistical

Communications in Infectious Diseases 4. doi:DOI 10.1515/1948-4690.1041.

A. Turnover Framework

We introduce a system of parameters and constraints to describe risk group turnover in deterministic epidemic models with heterogeneity in risk. We then describe how the system can be used in practical terms, based on different assumptions and data available for parameterizing turnover in risk. We conclude by framing previous approaches to this task using the proposed system.

A.1. Notation

Consider a population divided into G risk groups. We denote the number of individuals in risk group $i \in [1, \ldots, G]$ as x_i and the set of all risk groups as $\boldsymbol{x} = \{x_1, \ldots, x_G\}$. The total population size is $N = \sum_i x_i$, and the relative population size of each group is denoted as $\hat{x}_i = x_i/N$. Individuals enter the population at a rate ν per year, and exit at a rate μ per year. We model the distribution of risk groups among individuals entering into the population as $\hat{\boldsymbol{e}}$, which may be different from individuals already in the population $\hat{\boldsymbol{x}}$.[1] Thus, the total number of individuals entering into population $\boldsymbol{x}$ per year is given by νN, and the number of individuals entering into group i specifically is given by $\hat{e}_i \nu N$.

Turnover transitions may then occur between any two groups, in either direction. Therefore we denote the turnover rates as a $G \times G$ matrix ϕ. The element ϕ_{ij} corresponds to the proportion of individuals in group i who move from group i to group j each year. An example matrix is given in Eq. (A.1), where we write the diagonal elements as $*$ since they represent transitions from a group to itself.

$$\phi = \begin{bmatrix} * & x_1 \to x_2 & \cdots & x_1 \to x_G \\ x_2 \to x_1 & * & \cdots & x_2 \to x_G \\ \vdots & \vdots & \ddots & \vdots \\ x_G \to x_1 & x_G \to x_2 & \cdots & * \end{bmatrix} \tag{A.1}$$

Risk groups, transitions, and the associated rates are also shown for $G = 3$ in Figure A.1.

A.2. Parameterization

Next, we construct a system like the one above which reflects the risk group dynamics observed in a specific context. We assume that the relative sizes of the risk groups in the model ($\hat{\boldsymbol{x}}$) are already known, and should remain constant over time. Thus, what remains is to estimate the values of the parameters: ν, μ, $\hat{\boldsymbol{e}}$, and ϕ, using commonly available sources of data.

[1] We could equivalently stratify the rate of entry ν by risk group; however, we find that the mathematics in subsequent sections are more straightforward using $\hat{\boldsymbol{e}}$.

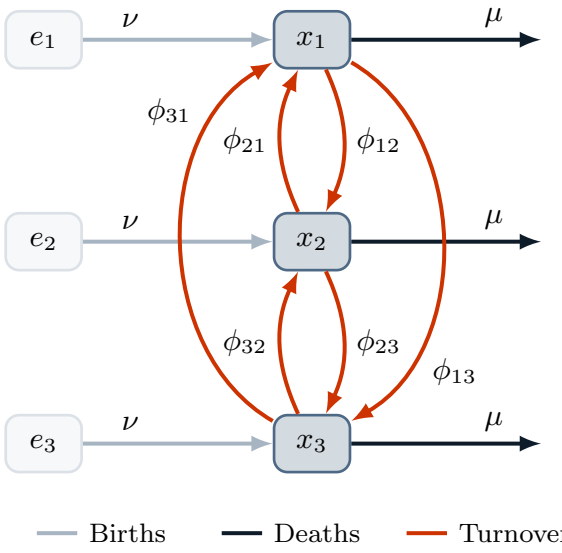


Figure A.1: System of $G = 3$ risk groups and turnover between them.
x_i: number of individuals in risk group i; e_i: number of individuals available to enter risk group i; ν: rate of population entry; μ: rate of population exit; ϕ_{ij}: rate of turnover from group i to group j.

A.2.1. Total Population Size

The total population size $N(t)$ is a function of the rates of population entry $\nu(t)$ and exit $\mu(t)$, given an initial size N_0. We allow the proportion entering the system to vary by risk group via $\hat{\boldsymbol{e}}$, while the exit rate has the same value for each group. We assume that there is no disease-attributable death. Because the values of ν and μ are the same for each risk group, they can be estimated independent of $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{e}}$, and ϕ.

The difference between entry and exit rates defines the rate of population growth:

$$\mathcal{G}(t) = \nu(t) - \mu(t) \tag{A.2}$$

The total population may then be defined using an initial population size N_0 as:

$$N(t) = N_0 \exp\left(\int_0^t \log\left(1 + \mathcal{G}(\tau)\right) d\tau\right) \tag{A.3}$$

which, for constant growth, simplifies to the familiar expression (Malthus, 1798):

$$N(t) = N_0(1 + \mathcal{G})^t \tag{A.4}$$

Census data, such as (DataBank, 2019), can be used to source the total population size in a given geographic setting over time $N(t)$, thus allowing Eqs. (A.3) and (A.4) to be used to estimate $\mathcal{G}(t)$.

If the population size is assumed to be constant, then $\mathcal{G}(t) = 0$ and $\nu(t) = \mu(t)$. If population growth occurs at a stable rate, then $\mathcal{G}$ is fixed at a constant value which can be estimated via Eq. (A.4) using any

two values of $N(t)$, separated by a time interval τ:

$$\mathcal{G}_\tau = \frac{N(t+\tau)^{\frac{1}{\tau}}}{N(t)} - 1 \tag{A.5}$$

If the rate of population growth $\mathcal{G}$ varies over time, then Eq. (A.5) can be reused for consecutive time intervals, and the complete function $\mathcal{G}(t)$ approximated piecewise by constant values. The piecewise approximation can be more feasible than exact solutions using Eq. (A.3), and can reproduce $N(t)$ accurately for small enough intervals τ, such as one year.

Now, given a value of $\mathcal{G}(t)$, either $\nu(t)$ must be chosen and $\mu(t)$ calculated using Eq. (A.2), or $\mu(t)$ must be chosen, and $\nu(t)$ calculated. Most modelled systems assume a constant duration of time that individuals spend in the model $\delta(t)$ (Anderson and May, 1991) which is related to the rate of exit μ by:

$$\delta(t) = \mu^{-1}(t) \tag{A.6}$$

In the context of sexually transmitted infections, the duration of time usually reflects the average sexual life-course of individuals from age 15 to 50 years, such that $\delta = 35$ years. The duration δ may also vary with time to reflect changes in life expectancy. The exit rate $\mu(t)$ can then be defined as $\delta^{-t}(t)$ following Eq. (A.6), and the entry rate $\nu(t)$ defined as $\mathcal{G}(t) - \mu(t)$ following Eq. (A.2).

A.2.2. Turnover

Next, we present methods for resolving the distribution of individuals entering the risk model $\hat{\boldsymbol{e}}(t)$ and the rates of turnover $\phi(t)$, assuming that entry and exit rates $\nu(t)$ and $\mu(t)$ are known. Similar to above, we first formulate the problem as a system of equations. Then, we explore the data and assumptions required to solve for the values of parameters in the system. The (t) notation is omitted throughout this section for clarity, though time-varying parameters can be estimated by repeating the necessary calculations for each t.

The number of risk groups G dictates the number of unknown elements in $\hat{\boldsymbol{e}}$ and ϕ: G and $G(G-1)$, respectively. We collect these unknowns in the vector $\boldsymbol{\theta} = [\hat{\boldsymbol{e}}, \boldsymbol{y}]$, where $\boldsymbol{y} = \text{vec}_{i \neq j}(\phi)$. For example, for $G = 3$, the vector $\boldsymbol{\theta}$ is defined as:

$$\boldsymbol{\theta} = \begin{bmatrix} \hat{e}_1 & \hat{e}_2 & \hat{e}_3 & \phi_{12} & \phi_{13} & \phi_{21} & \phi_{23} & \phi_{31} & \phi_{32} \end{bmatrix} \tag{A.7}$$

We then define a linear system of equations which uniquely determine the elements of $\boldsymbol{\theta}$:

$$\boldsymbol{b} = A\,\boldsymbol{\theta} \tag{A.8}$$

where A is a $M \times G^2$ matrix and $\boldsymbol{b}$ is a M-length vector. Specifically, each row in A and $\boldsymbol{b}$ defines a constraint: an assumed mathematical relationship involving one or more elements of $\hat{\boldsymbol{e}}$ and ϕ. For example, a simple constraint could be to assume the value $\hat{e}_2 = 0.20$. Each of the following sections introduces a type of constraint, including: assuming a constant group size, specifying elements of $\boldsymbol{\theta}$ directly, assuming

an average duration in a group, and assuming relative rates of turnover. Constraints may be selected and combined together based on availability of data and plausibility of assumptions. However, a total of $M = G^2$ constraints must be defined in order to obtain a "unique solution": exactly one value of $\boldsymbol{\theta}$ which satisfies all constraints. The values of $\hat{\boldsymbol{e}}$ and ϕ can then be calculated algebraically by solving Eq. (A.8) with $\boldsymbol{\theta} = A^{-1}\boldsymbol{b}$, for which many algorithms exist (LAPACK, 1992).

1. Constant group size. One epidemiologic feature that epidemic models consider is whether or not the relative sizes of risk groups are constant over time (Henry and Koopman, 2015; Boily et al., 2015). Assuming constant group size implies a stable level of heterogeneity over time. To enforce this assumption, we define the "conservation of mass" equation for group i, wherein the rate of change of the group is defined as the sum of flows in / out of the group:

$$\frac{d}{dt}x_i = \nu N\hat{e}_i + \sum_j \phi_{ji}\, x_j - \mu\, x_i - \sum_j \phi_{ij}\, x_i \tag{A.9}$$

Eq. (A.9) is written in terms of absolute population sizes $\boldsymbol{x}$, but can be written as proportions $\hat{\boldsymbol{x}}$ by dividing all terms by N. If we assume that the proportion of each group $\hat{x}_i$ is constant over time, then the desired rate of change for risk group i will be equal to the rate of population growth of the risk group, $\mathcal{G}x_i$. Substituting $\frac{d}{dt}x_i = \mathcal{G}x_i$ into Eq. (A.9), and simplifying yields:

$$\nu\, x_i = \nu\, N\hat{e}_i + \sum_j \phi_{ji}\, x_j - \sum_j \phi_{ij}\, x_i \tag{A.10}$$

Factoring the left and right hand sides in terms of $\hat{\boldsymbol{e}}$ and ϕ, we obtain G unique constraints. For $G = 3$, this yields the following 3 rows as the basis of $\boldsymbol{b}$ and A:

$$\boldsymbol{b} = \begin{bmatrix} \nu x_1 \\ \nu x_2 \\ \nu x_3 \end{bmatrix}; \qquad A = \begin{bmatrix} \nu & \cdot & \cdot & -x_1 & -x_1 & x_2 & \cdot & x_3 & \cdot \\ \cdot & \nu & \cdot & x_1 & \cdot & -x_2 & -x_2 & \cdot & x_3 \\ \cdot & \cdot & \nu & \cdot & x_1 & \cdot & x_2 & -x_3 & -x_3 \end{bmatrix} \tag{A.11}$$

These G constraints ensure risk groups do not change size over time. However, a unique solution requires an additional $G(G-1)$ constraints. For $G = 3$, this corresponds to 6 additional constraints.

2. Specified elements. The simplest type of additional constraint is to directly specify the values of individual elements in $\hat{\boldsymbol{e}}$ or ϕ. Such constraints may be appended to $\boldsymbol{b}$ and A as an additional row k using indicator notation.[2] That is, with b_k as the specified value v, and A_k as the indicator vector, with 1 in the same position as the desired element in $\boldsymbol{\theta}$:

$$b_k = v; \qquad A_k = [0, \ldots, 1, \ldots, 0] \tag{A.12}$$

[2] Indicator notation, also known as "one-hot notation" is used to select one element from another vector, based on its position. An indicator vector is 1 in the same location as the element of interest, and 0 everywhere else.

For example, for $G = 3$, if it is known that 20% of individuals enter directly into risk group 2 upon entry into the model ($\hat{e}_2 = 0.20$), then $\boldsymbol{b}$ and A can be augmented with:

$$b_k = \begin{bmatrix} 0.20 \end{bmatrix}; \qquad A_k = \begin{bmatrix} \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \tag{A.13}$$

since $\hat{e}_2$ is the second element in $\boldsymbol{\theta}$. If the data suggest zero turnover from group i to group j, then Eq. (A.13) can also be used to set $\phi_{ij} = 0$.

Note that the elements of $\hat{\boldsymbol{e}}$ must sum to one. Therefore, specifying all elements in $\hat{\boldsymbol{e}}$ will only provide $G - 1$ constraints, as the last element will be either redundant or violate the sum-to-one rule. As shown in Appendix B.3, the sum-to-one rule is actually implicit in Eq. (A.11), so it is not necessary to supply a constraint like $1 = \sum_i \hat{e}_i$.

3. Group duration. Type 1 Constraints assume that the relative population size of each group remains constant. Another epidemiologic feature that epidemic models considered is whether or not the duration of time spent within a given risk group remains constant. For example, in STI transmission models that include formal sex work, it can be assumed that the duration in formal sex work work remains stable over time, such as in (Mishra et al., 2014; Boily et al., 2015). The duration δ_i is defined as the inverse of all rates of exit from the group:

$$\delta_i = \left(\mu + \sum_j \phi_{ij} \right)^{-1} \tag{A.14}$$

Estimates of the duration in a given group can be sourced from cross-sectional survey data where participants are asked about how long they have engaged in a particular practice – such as sex in exchange for money (Watts et al., 2010). Data on duration may also be sourced from longitudinal data, where repeated measures of self-reported sexual behaviour, or proxy measures of sexual risk data, are collected (The DHS Program, 2019; ICAP, 2019). Data on duration in each risk group can then be used to define ϕ by rearranging Eq. (A.14) to yield: ${\delta_i}^{-1} - \mu = \sum_j \phi_{ij}$. For example, if for $G = 3$, the average duration in group 1 is known to be $\delta_1 = 5$ years, then $\boldsymbol{b}$ and A can be augmented with another row k:

$$b_k = \begin{bmatrix} 5^{-1} - \mu \end{bmatrix}; \qquad A_k = \begin{bmatrix} \cdot & \cdot & \cdot & 1 & 1 & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \tag{A.15}$$

Note that, similar to specifying all elements of $\hat{\boldsymbol{e}}$, specifying δ_i may result in conflicts or redundancies with other constraints. A conflict means it will not be possible to resolve values of ϕ which simultaneously satisfy all constraints, while a redundancy means that adding one constraint does not help resolve a unique set of values $\boldsymbol{\theta}$. For example, for $G = 3$, if Type 2 Constraints are used to specify $\phi_{12} = 0.1$ and $\phi_{13} = 0.1$, and $\mu = 0.05$, then by Eq. (A.14), we must have $\delta_1 = 4$. Specifying any other value for δ_1 will result in a conflict, while specifying $\delta_1 = 4$ is redundant, since it is already implied. There are innumerable situations in which this may occur, so we do not attempt to describe them all. Section A.2.2 describes how to identify conflicts and redundancies when they are not obvious.

4. Turnover rate ratios. In many cases, it may be difficult to obtain estimates of a given turnover rate ϕ_{ij} for use in Type 2 Constraints. However, it may be possible to estimate relative relationships between rates of turnover, such as:

$$r\,\phi_{ij} = \phi_{i'j'} \tag{A.16}$$

where r is a ratio relating the values of ϕ_{ij} and $\phi_{i'j'}$. For example, for $G = 3$, let T_1 be the total number of individuals entering group 1 due to turnover. If we know that 70% of T_1 originates from group 2, while 30% of T_1 originates from group 3, then $0.7\,T_1 = \phi_{23}\,x_2$ and $0.3\,T_1 = \phi_{13}\,x_1$, and thus: $\phi_{23}\left(\frac{0.3\,x_2}{0.7\,x_1}\right) = \phi_{13}$. This constraint can then be appended as another row k in $\boldsymbol{b}$ and A like:

$$b_k = \begin{bmatrix} 0 \end{bmatrix}; \qquad A_k = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \left(\frac{0.3\,x_2}{0.7\,x_1}\right) & \cdot & 1 & \cdot & \cdot \end{bmatrix} \tag{A.17}$$

The example in Eq. (A.17) is based on what proportions of individuals entering a risk group j came from which former risk group i, but similar constraints may be defined based on what proportions of individuals exiting a risk group i enter into which new risk group j. It can also be assumed that the absolute number of individuals moving between two risk groups is equal, in which case the relationship is: $\phi_{ij}\left(\frac{x_i}{x_j}\right) = \phi_{ji}$. All constraints of this type will have $b_k = 0$.

Solving the System. Table A.1 summarizes the four types of constraints described above. Given a set of sufficient constraints on $\boldsymbol{\theta}$ to ensure exactly one solution, the system of equations Eq. (A.8) can be solved using $\boldsymbol{\theta} = A^{-1}\boldsymbol{b}$. The resulting values of $\hat{\boldsymbol{e}}$ and ϕ can then be used in the epidemic model.

However, we may find that we have an insufficient number of constraints, implying that there are multiple values of the vector $\boldsymbol{\theta}$ which satisfy the constraints. An insufficient number of constraints may be identified by a "rank deficiency" warning in numerical solvers of Eq. (A.8) (LAPACK, 1992). Even if A has G^2 rows, the system may have an insufficient number of constraints because some constraints are redundant. In this situation, we can pose the problem as a minimization problem, namely:

$$\boldsymbol{\theta}^* = \arg\min f(\boldsymbol{\theta}), \quad \text{subject to: } \boldsymbol{b} = A\,\boldsymbol{\theta};\ \ \boldsymbol{\theta} \geq 0 \tag{A.18}$$

where f is a function which penalizes certain values of $\boldsymbol{\theta}$. For example, $f = ||\cdot||_2$ penalizes large values in $\boldsymbol{\theta}$, so that the smallest values of $\hat{\boldsymbol{e}}$ and ϕ which satisfy the constraints will be resolved.[3]

Similarly, we may find that no solution exists for the given constraints, since two or more constraints are in conflict. Conflicting constraints may be identified by a non-zero error in the solution to Eq. (A.8) (LAPACK, 1992). In this case, the conflict should be resolved by changing or removing one of the conflicting constraints.

[3] Numerical solutions to such problems are widely available, such as the Non-Negative Lease Squares solver (Lawson and Hanson, 1995), available in Python: `https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.nnls.html`.

Table A.1: Summary of constraint types for defining risk group turnover

Name	Eq.	E.g.	Data requirements
1. Constant group size	(A.10)	(A.11)	all values of $\hat{x}_i$ and ν
2. Specified elements	(A.12)	(A.13)	any value of $\hat{e}_i$ or ϕ_{ij}
3. Group duration	(A.14)	(A.15)	any value of δ_i
4. Turnover rate ratios	(A.16)	(A.17)	any relationship between two turnover rates ϕ_{ij} and $\phi_{i'j'}$

ν: rate of population entry; ϕ_{ij}: rate of turnover from group i to group j; $\hat{x}_i$: proportion of individuals in risk group i; $\hat{e}_i$: proportion of individuals entering into risk group i; δ_i: average duration spent in risk group i.

A.3. Previous Approaches

Few epidemic models of sexually transmitted infections with heterogeneity in risk have simulated turnover among risk groups, and those models which have simulated turnover have done so in various ways. In this section, we review three prior implementations of turnover and each study's objectives for the implementation – e.g. constant relative group sizes over time. We then highlight how the approach proposed in Section A.2 could be used to achieve the same objectives.

Stigum et al. (1994) simulated turnover among $G = 2$ risk groups in a population with no exogenous entry or exit ($\nu = \mu = 0$ and hence $\hat{\boldsymbol{e}}$ is not applicable). Turnover between the groups was balanced in order to maintain constant risk group sizes (Type 1 Constraint),[4] while the rate of turnover from high to low was specified as κ (Type 2 Constraint). Thus, the turnover system used by Stigum et al. (1994) can be written in the proposed framework as:

$$\begin{bmatrix} 0 \\ \kappa \end{bmatrix} = \begin{bmatrix} \hat{x}_1 & -\hat{x}_2 \\ 1 & \cdot \end{bmatrix} \begin{bmatrix} \phi_{12} \\ \phi_{21} \end{bmatrix}, \qquad \hat{e}_1 = \hat{e}_2 = 0 \tag{A.19}$$

Henry and Koopman (2015) also simulated turnover among $G = 2$ risk groups, but considered exogenous entry and exit, both at a rate μ. The authors used the notation f_i for our $\hat{x}_i$, and assumed that the exogenous population had the same distribution of risk groups as the system population: $\hat{e}_i = f_i$ (Type 2 Constraint). The authors further maintained constant risk group sizes (Type 1 Constraint) by analytically balancing turnover between the two groups using: $\phi_{12} = \omega\hat{x}_2$; $\phi_{21} = \omega\hat{x}_1$. However, it can be shown that this analytical approach is also the solution to the following combination of Type 1 and 2 Constraints:

$$\begin{bmatrix} 0 \\ \omega f_2 \end{bmatrix} = \begin{bmatrix} f_1 & -f_2 \\ 1 & \cdot \end{bmatrix} \begin{bmatrix} \phi_{12} \\ \phi_{21} \end{bmatrix}, \qquad \hat{e}_i = f_i \tag{A.20}$$

[4] Due to its simplicity, this constraint is actually an example of both Type 1 and Type 4 Constraints.

Eaton and Hallett (2014) simulated turnover among $G = 3$ risk groups, considering exogenous entry from a population with a unique distribution of risk groups $\hat{\boldsymbol{e}}$. Turnover was considered from high-to-medium, high-to-low, and medium-to-low risk, all with an equal rate ψ; the reverse transition rates were set to zero (six total Type 2 Constraints). In the absence of turnover in the other direction, risk group sizes were maintained using the values of $\hat{e}_i$, computed using Type 1 Constraints as follows:

$$\begin{bmatrix} \nu x_1 + 2x_1\psi \\ \nu x_2 - x_1\psi + x_2\psi \\ \nu x_3 - x_1\psi - x_2\psi \end{bmatrix} = \begin{bmatrix} \nu & \cdot & \cdot \\ \cdot & \nu & \cdot \\ \cdot & \cdot & \nu \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix}, \qquad \begin{aligned} \phi_{12} = \phi_{13} = \phi_{23} = \psi \\ \phi_{21} = \phi_{31} = \phi_{32} = 0 \end{aligned} \tag{A.21}$$

In sum, the framework for modelling turnover presented in this section aims to generalize all previous implementations. In so doing, we hope to clarify the requisite assumptions, dependencies on epidemiologic data, and relationships between previous approaches.

B. Supplemental Equations

Table B.1: Notation

Symbol	Definition
i	risk group index
j	risk group index for "other" group in turnover
k	risk group index for "other" group in incidence
t	time
$\mathcal{S}_i$	number of susceptible individuals in risk group i
$\mathcal{I}_i$	number of infectious individuals in risk group i
$\mathcal{T}_i$	number of treated individuals in risk group i
N	total population size
ν	rate of population entry
μ	rate of population exit
ϕ_{ij}	rate of turnover from group i to group j
λ_i	force of infection among susceptibles in risk group i
τ	rate of treatment initiation among infected
$\hat{x}_i$	proportion of individuals in risk group i
$\hat{e}_i$	proportion of individuals entering into risk group i
δ_i	average duration spent in risk group i
C_i	number of partners per year among individuals in risk group i
β	probability of transmission per partnership
ρ_{ik}	probability of partnership formation between risk groups i and k

B.1. Model Equations

$$\frac{d}{dt}\mathcal{S}_i(t) = +\sum_j \phi_{ji}\mathcal{S}_j(t) - \sum_j \phi_{ij}\mathcal{S}_i(t) - \mu\mathcal{S}_i(t) + \nu\hat{e}_i N(t) - \lambda_i(t)\mathcal{S}_i(t) \tag{B.1}$$

$$\frac{d}{dt}\mathcal{I}_i(t) = +\sum_j \phi_{ji}\mathcal{I}_j(t) - \sum_j \phi_{ij}\mathcal{I}_i(t) - \mu\mathcal{I}_i(t) + \lambda_i(t)\mathcal{S}_i(t) - \tau\mathcal{I}_i(t) \tag{B.2}$$

$$\frac{d}{dt}\mathcal{T}_i(t) = +\sum_j \phi_{ji}\mathcal{T}_j(t) - \sum_j \phi_{ij}\mathcal{T}_i(t) - \mu\mathcal{T}_i(t) + \tau\mathcal{I}_i(t) \tag{B.3}$$

$$\underbrace{+\sum_j \phi_{ji}\mathcal{T}_j(t)}_{\text{turnover into}} \underbrace{- \sum_j \phi_{ij}\mathcal{T}_i(t)}_{\text{turnover from}} \underbrace{- \mu\mathcal{T}_i(t)}_{\text{death}} \underbrace{+ \nu\hat{e}_i N(t)}_{\text{birth}} \underbrace{+ \lambda_i(t)\mathcal{S}_i(t)}_{\text{incidence}} \underbrace{- \tau\mathcal{I}_i(t)}_{\text{treatment}}$$

B.2. Complete Example Turnover System

$$
\begin{array}{r}
\text{constant group size} \\ \\ \\
\text{specified } e \\ \\ \\ \\ \\
\text{group duration} \\ \\ \\
\text{turnover rate ratios} \\ \\ \\
\end{array}
\left\{ \begin{array}{c}
\left[\begin{array}{c}
\nu x_1 \\ \nu x_2 \\ \nu x_3 \\ e_1^* \\ e_2^* \\ e_3^* \\ {\delta_1}^{-1} - \mu \\ {\delta_2}^{-1} - \mu \\ {\delta_3}^{-1} - \mu \\ 0 \\ 0 \\ 0
\end{array} \right]
\end{array} \right.
=
\left[\begin{array}{ccccccccc}
\nu & \cdot & \cdot & -x_1 & -x_1 & x_2 & \cdot & x_3 & \cdot \\
\cdot & \nu & \cdot & x_1 & \cdot & -x_2 & -x_2 & \cdot & x_3 \\
\cdot & \cdot & \nu & \cdot & x_1 & \cdot & x_2 & -x_3 & -x_3 \\
1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & 1 & 1 & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & 1 & 1 & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 1 \\
\cdot & \cdot & \cdot & x_1 & \cdot & -x_2 & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & x_1 & \cdot & \cdot & -x_3 & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & x_2 & \cdot & -x_3
\end{array} \right]
\left[\begin{array}{c}
e_1 \\ e_2 \\ e_3 \\ \phi_{12} \\ \phi_{13} \\ \phi_{21} \\ \phi_{23} \\ \phi_{31} \\ \phi_{32}
\end{array} \right]
\tag{B.4}
$$

B.3. Redundancy in specifying all elements of $\hat{e}$

Whenever it is assumed that risk groups do not change size, G rows of the form shown in Eq. (A.11) are added to $\boldsymbol{b}$ and A:

$$
\boldsymbol{b} = \left[\begin{array}{c} \nu x_1 \\ \nu x_2 \\ \nu x_3 \end{array} \right]; \qquad A = \left[\begin{array}{ccccccccc}
\nu & \cdot & \cdot & -x_1 & -x_1 & x_2 & \cdot & x_3 & \cdot \\
\cdot & \nu & \cdot & x_1 & \cdot & -x_2 & -x_2 & \cdot & x_3 \\
\cdot & \cdot & \nu & \cdot & x_1 & \cdot & x_2 & -x_3 & -x_3
\end{array} \right] \tag{A.11}
$$

After multiplying by $\boldsymbol{\theta}$, these G rows can be row-reduced by summing to obtain:

$$
\begin{aligned}
[\nu x_1 + \nu x_2 + \nu x_3] &= [\nu e_1 + \nu e_2 + \nu e_3 + 0\,\phi_{12} + 0\,\phi_{13} + 0\,\phi_{21} + 0\,\phi_{23} + 0\,\phi_{31} + 0\,\phi_{32}] \\
\nu\,[x_1 + x_2 + x_3] &= \nu\,[e_1 + e_2 + e_3]
\end{aligned} \tag{B.5}
$$

which therefore implies that $\sum_i x_i = \sum_i e_i$, or equivalently $\sum_i \hat{x}_i = \sum_i \hat{e}_i = 1$. Thus, it is redundant to specify all G elements of $\hat{e}$, as the final element will be dictated by constant group size constraints.

B.4. Factors of Incidence

Substituting the proportional mixing definition of ρ_{ik} into the incidence equation λ_i, we have:

$$
\lambda_i = C_i \sum_k \rho_{ik} \beta \frac{\mathcal{I}_k(t)}{\mathcal{X}_k} \tag{B.6}
$$

$$
= C_i \beta \sum_k \frac{C_k \mathcal{X}_k}{\sum_{\mathrm{k}} C_{\mathrm{k}} \mathcal{X}_{\mathrm{k}}} \frac{\mathcal{I}_k(t)}{\mathcal{X}_k} \tag{B.7}
$$

$$
= C_i \beta \underbrace{\frac{\sum_k C_k \mathcal{I}_k(t)}{\sum_{\mathrm{k}} C_{\mathrm{k}} \mathcal{X}_{\mathrm{k}}}}_{f} \tag{B.8}
$$

We can factor the term f as:

$$f = \frac{\sum_k C_k \mathcal{I}_k}{\sum_k C_k \mathcal{X}_k} \tag{B.9}$$

$$= \frac{\sum_k C_k \mathcal{I}_k}{\sum_k \mathcal{I}_k} \cdot \frac{\sum_k \mathcal{I}_k}{\sum_k \mathcal{X}_k} \cdot \frac{\sum_k \mathcal{X}_k}{\sum_k C_k \mathcal{X}_k} \tag{B.10}$$

which we recognize as the following terms:

$$= \hat{C}_{\mathcal{I}} \cdot \hat{\mathcal{I}} \cdot \hat{C}^{-1} \tag{B.11}$$

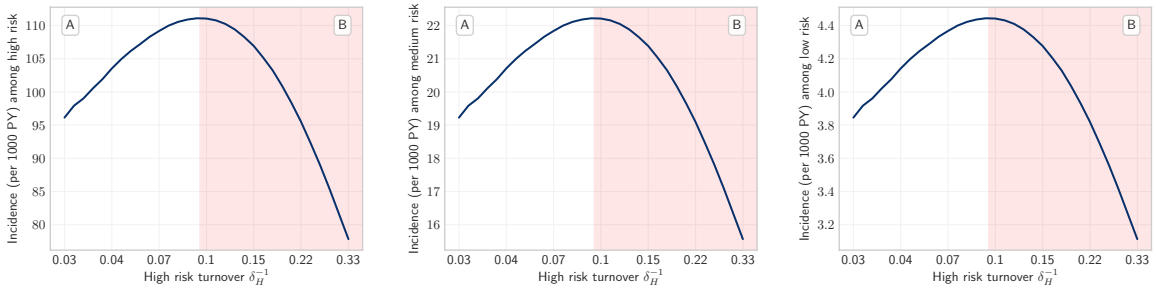
Namely,

1. $\hat{C}_{\mathcal{I}}$ is the average number of partners among infectious individuals
2. $\hat{\mathcal{I}}$ is the proportion of the population who are infectious (overall prevalence)
3. $\hat{C}$ is the average number of partners among all individuals (constant)

Therefore, only two non-constant factors control incidence per susceptible: 1) the average number of partners among infectious individuals $\hat{C}_{\mathcal{I}}$, and 2) overall prevalence $\hat{\mathcal{I}}$. The product of these factors $\hat{C}_{\mathcal{I}}\hat{\mathcal{I}}$, scaled by $\beta C_i/\hat{C}$, then gives λ_i. In fact, the incidence in each group individually is proportional to incidence overall, as C_i is only factor depending on i.

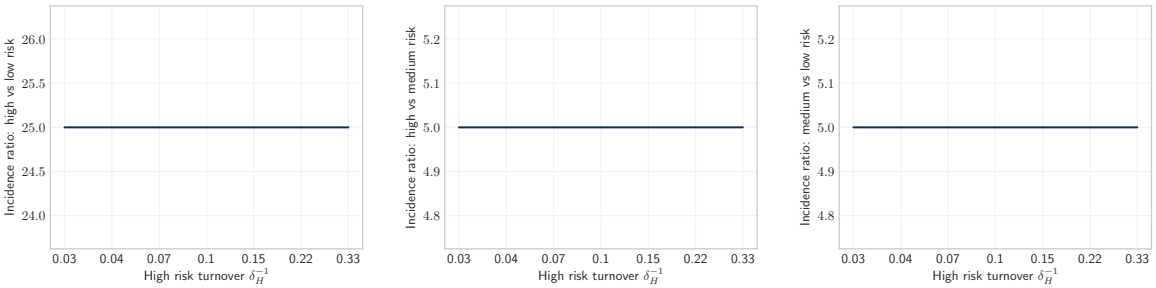
C. Supplemental Results

C.1. Equilibrium Incidence



(a) High risk (b) Medium risk (c) Low risk

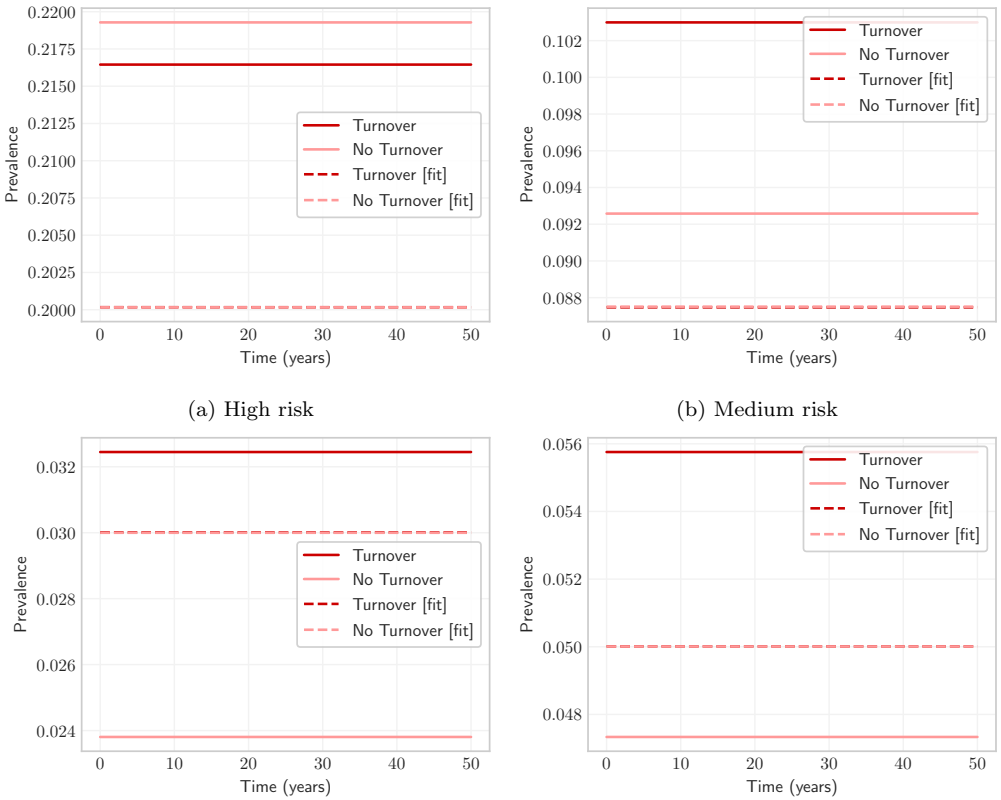
Figure C.1: Equilibrium incidence among high, medium, and low risk groups versus turnover, as controlled by the duration in the high risk group δ_H. Turnover shown in log scale. incidence in each risk group is proportional to overall incidence with C_i as a scale factor.



(a) High vs low risk (b) High vs medium risk (c) Medium vs low risk

Figure C.2: Equilibrium incidence ratios between risk groups under different rates of turnover ϕ. Incidence ratios do not depend on turnover.

C.2. Equilibrium prevalence and number of partners before and after model fitting



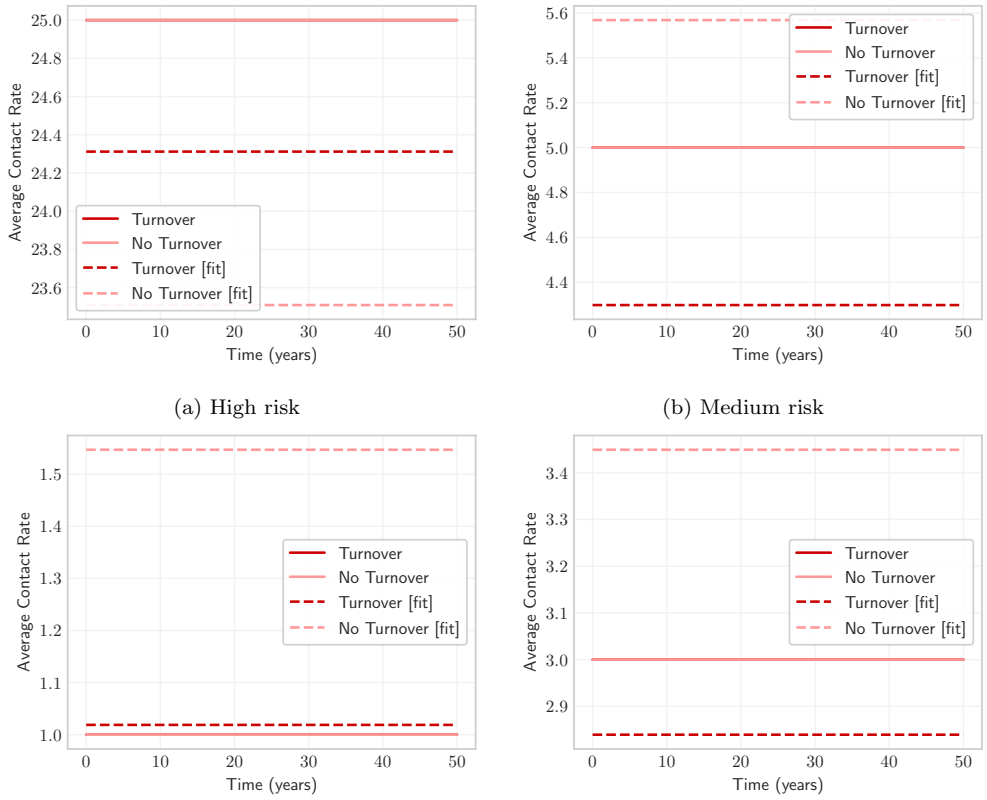
(a) High risk

(b) Medium risk

(c) Low risk

(d) Overall

Figure C.3: Equilibrium STI prevalence among high, medium, and low risk groups as well as overall, with and without turnover, and with and without fitted C_i to group-specific prevalence.



(a) High risk

(b) Medium risk

(c) Low risk

(d) Overall

Figure C.4: Numbers of partners C_i among high, medium, and low risk groups as well as overall, with and without turnover, and with and without model fitting to group-specific prevalence.

C.3. Influence of turnover on the tPAF of the highest risk group before model fitting

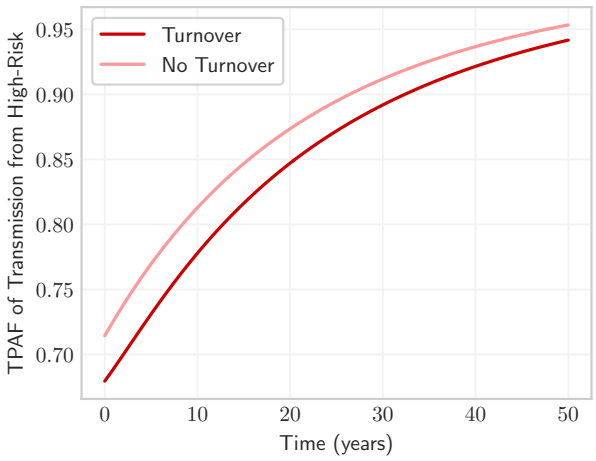
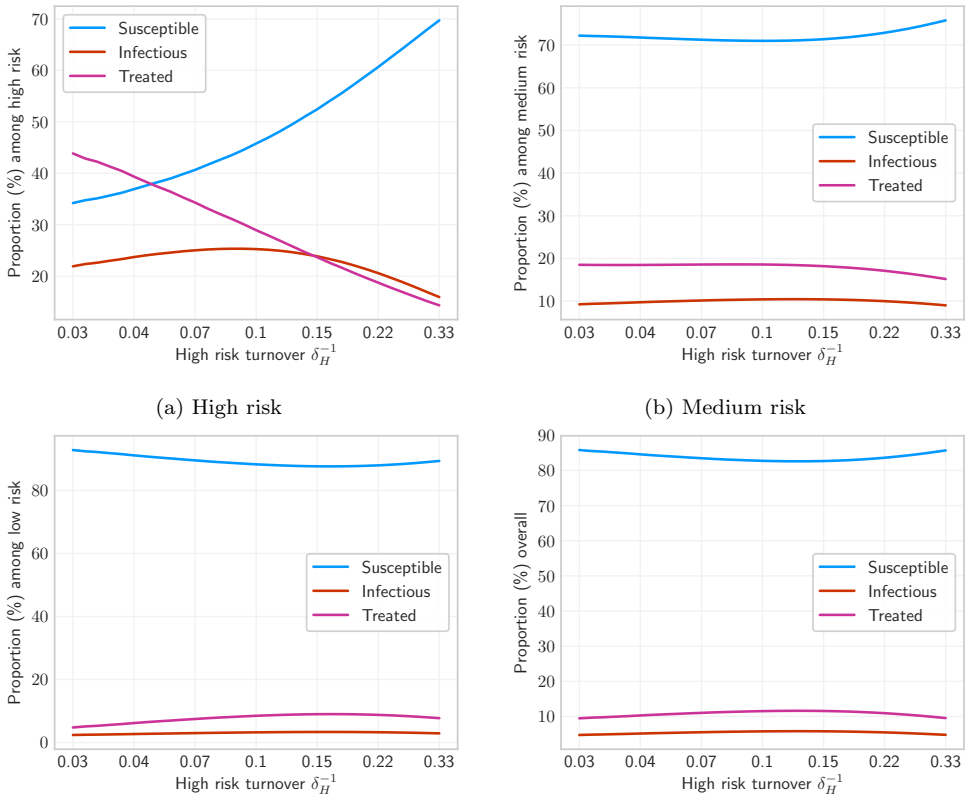


Figure C.5: Transmission population attributable fraction (tPAF) of the high risk group in models with and without turnover, before model fitting.

C.4. Equilibrium health states and rates of transition



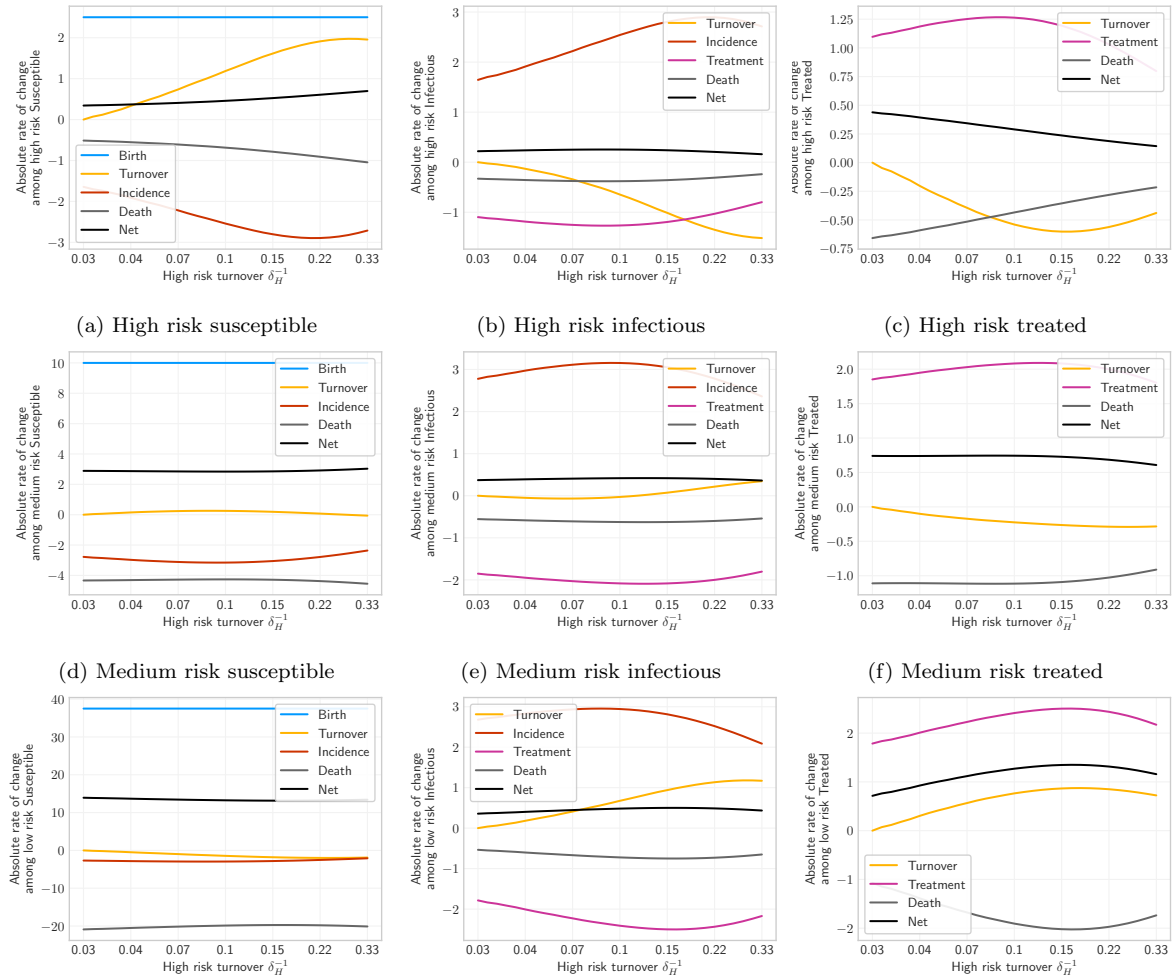
(a) High risk

(b) Medium risk

(c) Low risk

(d) Overall

Figure C.6: Equilibrium health state proportions under different rates of turnover.



(a) High risk susceptible (b) High risk infectious (c) High risk treated

(d) Medium risk susceptible (e) Medium risk infectious (f) Medium risk treated

(g) Low risk susceptible (h) Low risk infectious (i) Low risk treated

Figure C.7: Absolute rates of change at equilibrium (number of individuals gained/lost per year) among individuals in each health state and risk group, broken down by type of change: gain via births, gain via incident infections, loss via treatment, loss/gain via turnover, loss via death, and net change. Based on Appendix. B.1. Rates of change do not sum to zero due to population growth. See Figure 3 for x-axis definition.

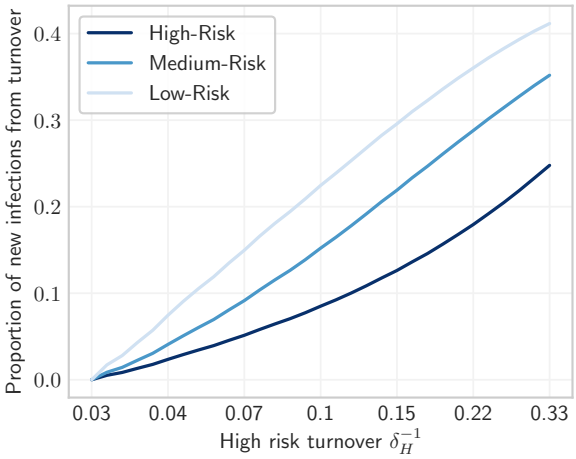
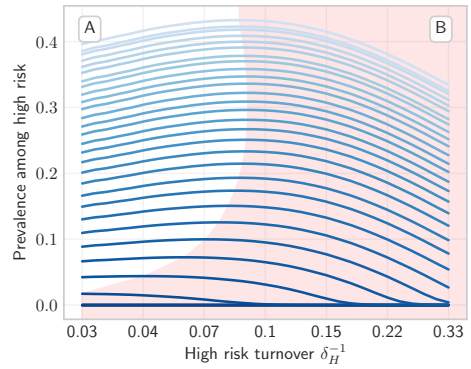


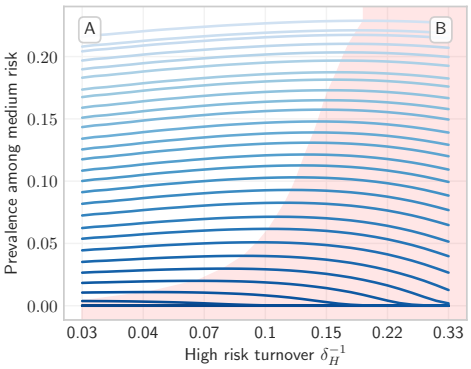
Figure C.8: Proportion of new infectious individuals in each risk group which are from turnover of infectious individuals, as opposed to incident infection of susceptible individuals in the risk group.

C.5. Effect of treatment rate on the influence of turnover on equilibrium prevalence

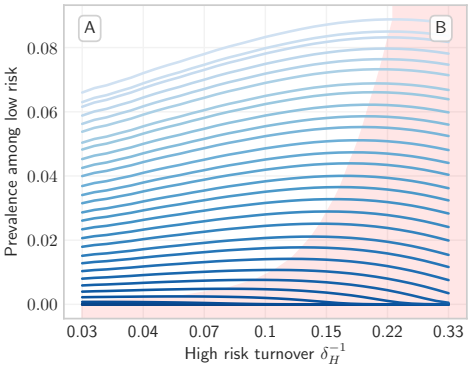
In order to examine the effect of treatment rate τ on the results of Experiment 1 – the influence of turnover on equilibrium prevalence – we recreated Figures 4 and 7 for a range of treatment rates $\tau \in [0.05, 1.0]$. The results are shown in Figure C.9.



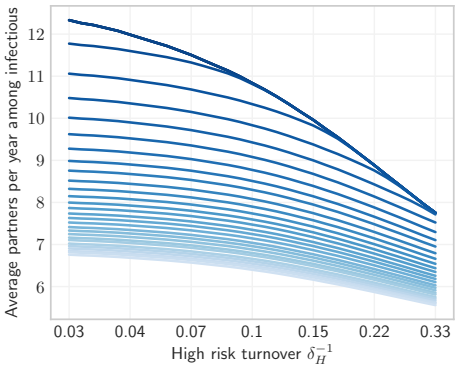
(a) Prevalence among high risk



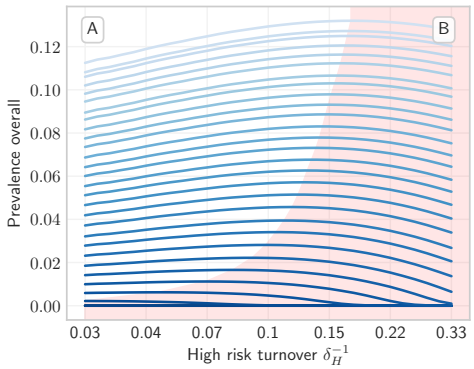
(b) Prevalence among medium risk



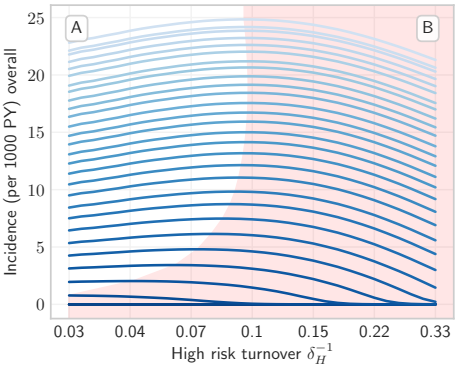
(c) Prevalence among low risk



(d) Average C among infectious individuals $\hat{C}_{\mathcal{I}}$



(e) Prevalence overall



(f) Incidence overall λ

Figure C.9: Relationship between turnover rate and equilibrium STI prevalence in high, medium, and low risk groups, as well as overall STI prevalence and incidence, and average C among infectious individuals, for a range of treatment rates τ. Darker blue indicates higher treatment rate. The threshold turnover rate separating regions A and B decreases with treatment rate, meaning that increasing turnover becomes more likely to decrease equilibrium prevalence as treatment rate increases.