Risk Group Turnover in Epidemic Models

1. Introduction

Heterogeneity in transmission risk is a ~~critical~~ consistent characteristic~~feature~~ of epidemics of sexually transmitted infections (STI) (Anderson and May, 1991). This heterogeneity is often demarcated by identifying specific ~~sub-~~populations whose risks of acquisition and onward transmission of STI are the highest, such that their specific unmet ~~STI~~ prevention and treatment needs can sustain local epidemics of STI (Yorke et al., 1978). An important indicator in the appraisal of STI epidemics is the contribution of high-risk groups to the overall epidemic (Case et al., 2012).

Indicators of "contribution" are used to help prioritize or target interventions to reach groups at highest risk (Case et al., 2012; Shubber et al., 2014). Transmission models are increasingly being used to quantify "contribution" by simulating counterfactuals where transmission ~~from and/or to~~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., 2016; Mukandavire et al., 2018). This is often referred to as the transmission population attributable fraction (TPAF) and counterfactuals have included setting susceptibility to zero and/or setting infectiousness to zero to calculate the TPAF (Mishra et al., 2012). 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 a failure to provide effective treatment in a particular risk group (Mishra et al., 2016).

An epidemiologic phenomenon that is often overlooked in transmission models, but is well-described in the context of sexual behaviour, is the movement between risk groups (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 and retirement from formal sex work – a period of time in an individual's sexual life-course that represents a greater risk of STI susceptibility and onward transmission (Watts et al., 2010). Similarly, individuals may have a larger number of sexual partnerships or experience greater STI-associated vulnerabilities during specific times within their sexual life-course (Marston and King, 2006).

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 of HIV (Henry and Koopman, 2015); and the ~~level~~ coverage of ~~universal treatment~~antiretroviral therapy~~treatment~~ (ART) required to achieve HIV epidemic control (Henry and Koopman, 2015). Yet how, and the extent to which, turnover influences the TPAF remains unknown.

Implementations of risk group turnover in compartmental transmission models also vary widely. In works by Koopman et al. and Stigum et al., rates of movement between two risk groups were balanced analytically based on the size of the groups; Boily et al. (2015) used a

Commented [SS1]: Is this the final title? I think something a little more descriptive may catch a bit more attention perhaps – The impact/importance of risk group turnover in epidemic models for sexually transmitted infections or perhaps something that highlights that this is contributing a new method/guidance for how to parameterize turnover in models?

Commented [SB2R1]: Describing the Impact of Turnover in Epidemic Models of Sexually Transmitted Infections

Maybe?

Commented [SB3]: Rather than saying important, would explain. Ie, try to avoid editorializing in the introduction and just provide empiric evidence as to why this is the case.

Commented [SB4]: Why quotations? I see that you are using this as the link to the last sentence of the last paragraph but would avoid this. Suspense is not really ideal here... 😊

Commented [SB5]: Also don't think quotations are ideal. Could just explain contribution as the traditional PAF and that also sets up nicely for why the tPAF is an advance.

Commented [SB6]: Really think the PAF as an old concept needs to be introduced in a sentence and then can include the TPAF as to why it is an advance for infectious diseases.

Commented [SB7]: Again, editorializing and is setting up potentially adversarial relationships with reviewers. I just don't think you need the sentence and can still highlight the need for turnover in models.

Commented [SB8]: Can cite our recent Lancet HIV piece...

Commented [SB9]: I would really frame the specific element of what you aim to study towards the last paragraph of intro.

Formatted: Normal

100-year burn-in period to equilibrate a complex system of demographic transitions before introduction of HIV into the modelled system; others restricted the system to unidirectional turnover – e.g. from high to low risk (Eaton and Hallett, 2014).

Challenges in implementing turnover include incorporation of data-driven epidemiologic constraints. For example, data may suggest that the relative sizes of ~~subgroups~~ specific populations in the model ~~(~~such as the population of sex workers~~)~~ have remained constant over time. Data may also suggest that heterogeneity in risk behaviour of individuals entering into the model (reflecting ~~sexual~~ coital debut) may be different from the risk behaviour of individuals already in the model (reflecting sexually active adults). Data may indicate the average duration of a high-risk period of one's sexual life-course (Watts et al., 2010), or how their behaviour changes following that period. Such data should be reflected in implementations of turnover, but it is not always clear how to do so. Moreover, without an exact analysis of the turnover implementation, a burn-in period may be required to equilibrate the system, resulting in potentially unwanted changes to the initial group sizes.

In this paper, we aim to examine the influence of risk group turnover on the TPAF of an illustrative STI without STI-attributable mortality. First, we propose a solution to the challenges outlined above of including epidemiologic data while avoiding the need for a burn-in period. That is, we introduce a unified framework for parameterizing risk group turnover, based on available data. We then examine the mechanisms by which turnover influences group-specific STI incidence and prevalence (Experiment 1). We then examine how inclusion/exclusion of turnover to produce the epidemic features of the same setting influences the values parameters related to heterogeneity during model fitting (Experiment 2). Finally, we compare the TPAF of the highest-risk groups estimated from two different settings: one with and one without turnover, and from two models that reproduce the same setting: one model with and one model without turnover (Experiment 3).

2. System

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

2.1 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 $x = \{x_1, \ldots, x_G\}$. 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 assume that individuals entering into the population originate from another exogenous population $e = \{e_1, \ldots, e_G\}$. We make this assumption so that the exogenous population e

Commented [LW10]: Can u explain a bit more about this approach?

Not clear what exactly burn-in period mean here.

Given lots of argument below was based on avoiding the need for a burn-in period, it is necessary to let readers know what it is and what is the limitation in this approach.

Commented [SB11]: Would break this sentence up and likely not have it be its own paragraph

Commented [LW12]:

Find this sentence a bit too abstract.

Wondering if can provide a bit more detail here. Like the one sentence summary of other approaches mentioned above:

e.g., In works by Koopman et al. and Stigum et al., rates of movement between two risk groups were balanced analytically based on the size of the groups;

Then in ur approach, what was done? What was the available data?

Commented [LW13]: Should experiment be split into 2?

To me, the first part is a continuation of Experiment 1 and the latter is a continuation of experiment 2.

Also, in the write-up, might be clearer for the readers if we rearrange the experiments (1, 3(first half); 2, 3 (second half)

Commented [LW14]: What do u mean by unified system of equations

Formatted: Normal

may have different relative risk group sizes from the system population x.[1] That is, $\hat{e}_i$ may not necessarily equal $\hat{x}_i$. However, since the entry rate ν is relative to the size N of population x, it will be convenient and otherwise inconsequential to assume that population e also has size N. That way, the total number of individuals entering into population x per year is given by νN, and the number of individuals entering into group x_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 × G matrix φ. The element φ ij corresponds to the proportion of individuals in group x_i who move from group x_i to group x_j each year. An example matrix is given in Eq. (1), where we write the diagonal elements as $*$ since they represent transitions from a group to itself.

$$\phi = \begin{bmatrix} * & x_1 \to x_2 & \cdots & x_1 \to x_G \\ x_2 \to x_1 & * & \cdots & x_2 \to x_G \\ \vdots & \vdots & \ddots & \vdots \\ x_G \to x_1 & x_G \to x_2 & \cdots & * \end{bmatrix} \quad (1)$$

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

1 We could equivalently stratify the rate of entry ν by risk group while keeping the exogenous population e equal to the system population x. However, we find this formulation more difficult to work with in the subsequent sections.

Commented [LW15]:
I think the logic of this paragraph should be modified a bit.

When we say entry rate, we usually define it from the perspective of the current population size (N).

So with an entry rate of v, and current population size of N, the total number of individuals entering into population x per year is vN.

As u mentioned, we vN comes from e, which may have different risk group distribution, and the number of individuals entering into group x_i can thus be given by $\hat{e}_i \nu N$.

Up till here, it is all independent of our assumption that e has size of N.

We assume that population e also has size N given the rational u described here.

That way, the number of individuals entering into group x_i specifically is given by $e_i \nu$.

Formatted: English (United States)

Formatted: Normal

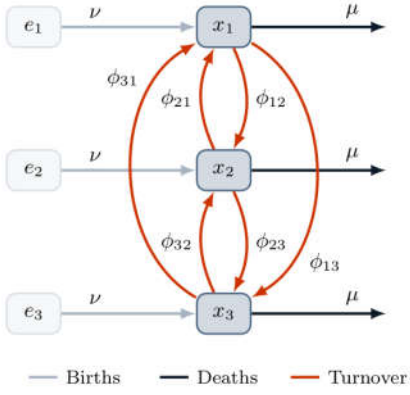


Figure 1: System of risk groups and flows between them for $G = 3$

[skipped the rest of this section for Word][2][3][4]

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

	Name	Eq.	E.g.	Data requirements
1.	Constant group size	(10)	(11)	all values of $\hat{x}_i$ and ν
2.	Specified elements	(12)	(13)	any value of $\hat{e}_i$ or ϕ_{ij}
3.	Group duration	(14)	(15)	any value of δ_i
4.	Relative rates of turnover	(16)	(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.

3. Experiment

We aimed to determine and understand the influence of risk group turnover on the contribution of the highest risk group to the overall epidemic, as measured by the transmission population attributable fraction (TPAF). However, to understand the underlying mechanism, we first examined the influence of turnover on the equilibrium incidence and prevalence projected among risk groups, as well as overall. Since the influence may depend on the duration of infectiousness, we also explored the sensitivity of these

2 dummy so footnote numbers line up
3 dummy
4 dummy

Formatted: English (United States)

Commented [LW16]: I would suggest putting the system and previous approach section into appendix as we discussed at the meeting.

Formatted: English (United States)

Commented [SB17]: Is this STI incidence and prevalence? I would keep specifying so all are clear here.

Formatted: Normal

results to different treatment rates. Next, we examined how turnover may influence inferred model parameters through model fitting. Finally, we compared the TPAF of the highest risk group with and without turnover, before and after model fitting. Our experiments can be summarized as in Table 2.

Table 2: Summary of experiments

Experiment	Outputs	Stratifications	Turnover	Treatment	Fitting
1.1.	prevalence	overall	any vs none	fixed, 2 rates	none
1.2.	prevalence, incidence	overall, by group	varied	fixed, 1 rate	none
1.3.	prevalence, incidence	overall, by group	varied	varied	none
2.	inferred contact rate	by group	any vs none	fixed, 1 rate	prevalence by group
3.1.	TPAF	highest risk	any vs none	fixed, 1 rate	none
3.2.	TPAF	highest risk	any vs none	fixed, 1 rate	prevalence by group

TPAF: transmission population attributable fraction

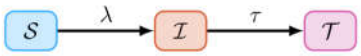


Figure 2: Modelled health states. $\mathcal{S}$: susceptible; $\mathcal{I}$: infected; $\mathcal{T}$: treated; λ: force of infection; τ: treatment.

3.1. Model & Simulations

To run our experiments, we developed a deterministic single-sex SIT model which simulates transmission in a population with heterogeneity in risk. The model is not representative of a specific infection but includes balancing contacts as per sexually transmitted infections (Garnett and Anderson, 1994). The model includes three health states: susceptible S, infectious I, and treated T (Figure 2), and G = 3 levels of risk: high H, medium M, and low L. Risk strata are defined by different number of contacts per year so that individuals in risk group i are assumed to form contacts at a rate C_i per year. The probability of contact 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 contacts within each group:

$$\rho_{ik} = \frac{C_k x_k}{\sum_k C_k x_k} \tag{22}$$

The biological probability of transmission is defined as β per contact. Individuals transition from the infectious I to susceptible S health-state via a force of infection λ per year, per susceptible in risk group i:

$$\lambda_i = C_i \sum_k \rho_{ik} \beta \frac{\mathcal{I}_k}{x_k} \tag{23}$$

Individuals are assumed to transition from the infectious I to treated T health-state at a rate

Commented [LW18]: This paragraph is essentially ur objective statement, which can be merged with the last paragraph of the Intro.

Formatted: English (United States)

Commented [SB19]: I get saying any STI—but some are curative and some aren't. ie, you can treat some folks but the infection may be sustained and in others it is totally gone. So just ask the question as whether we want to specify more about which type of STI might be included in a model like this.

Formatted: English (United States)

Commented [SB20]: I think first time need to say susceptible, infected, treated given we are also talking about STI here.

Commented [SB21]: Not sure what this means.

Formatted: English (United States)

Commented [SS22]: Check: Not from the susceptible to the infectious?

Formatted: English (United States)

Formatted: Normal

τ per year, reflecting diagnosis and treatment. The treatment rate is not stratified by risk group. Individuals in the treated *T* health-state are not infectious nor susceptible, and individuals cannot become re-infected.

As described in Section 2, 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}. The turnover rates φ and distribution of individuals entering the model by risk group ê were computed using the methods outlined in Section 2.2.2 and the following assumptions. First, we assumed that the proportion of individuals entering each risk group ê was equal to the proportion of individuals across risk groups in the model $\hat{x}$. Second, we assumed that the average duration spent in each risk group δ is known. Third, we assumed that the absolute number of individuals moving between two risk groups in either direction is balanced. The system of equations which results from these assumptions is given in Appendix A.2. To meet all three conditions, there is only one possible value for each element in φ and ê – i.e. A is full rank. In other words, by specifying these three conditions, we ensure that a unique set of φ and ê is computed.

Using the above three assumptions, we need to specify the values of $\hat{x}$, δ, ν, and μ. Such parameters could be derived from data as described in Section 2.2.2; however, in this experiment, we use the illustrative values summarized in Table 3. After resolving the system of equations, ê is equal to $\hat{x}$ (assumed), and φ is:

$$\phi = \begin{bmatrix} * & 0.0833 & 0.0867 \\ 0.0208 & * & 0.0158 \\ 0.0058 & 0.0042 & * \end{bmatrix} \quad (24)$$

We then simulated epidemics using these parameters. The model was initialized with *No = 1000* individuals who are distributed across risk groups according to $\hat{x}$. We seeded the epidemic with one infectious individual in each risk group at *t = 0*. There were no treated individuals at the start of the epidemic, and so all individuals except the 3 infectious individuals were susceptible. We numerically solved the system of ordinary differential equations in Python[5] using Euler's method with a time step of *dt = 0.1* years. The full system of model equations is given in Appendix A.1. All comparative analyses are then conducted at equilibrium, defined as a steady state with <1% difference in incidence per year.

5 Code for all aspects of the project is available at: https://github.com/c-uhs/turnover

Commented [LW23]: For this paragraph and the next, have a sub-heading such as:

implement turn-over in in SIT model.

Commented [LW24]: And another strong assumption which was implicit here is that u assume the rate of turn over to be the same irrespective of disease status. And I think it is critical to make it explicit.

So S, I, T all have the same turn over rate,

Commented [SB25]: Do we need to explain why we made these assumptions? Or provide refs?

Formatted: Highlight

Formatted: Highlight

Formatted: Highlight

Commented [LW26]: I would think this is a very strong assumption.

It is essentially saying the turn-over system consistently **swap** individuals in two risk groups.

Formatted: Highlight

Commented [LW27]: After read the whole results section, I think experiment 2 and 3.2 followed values specified in Table 3.

But experiment 1 and 3.1 explored a wide range of turn-over and treatment rates.

it is unclear what do u refer to when u said "this experiment".

Given the organization and length – by the time I reached results of Experiemnt 2, I almost forgot under which values they were done – as the section preceed it explored a range of values of turn-over rate and treatment rate.

Formatted: English (United States)

Formatted: Normal

Table 3: Model parameters. All rates have units year^{-1}; durations are in years; parameters stratified by risk group are written [high, medium, low] risk.

Symbol	Description	Value
β	transmission probability per contact	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	rate of contact formation among individuals in each risk grouop	[25 5 1]
ν	rate of population entry	0.05
μ	rate of population exit	0.03

Commented [SS28]: Note typo in C: group (currently spelled grouop)

Commented [SB29R28]: It's the Canadian spelling!

Formatted: English (United States)

3.2. Experiment 1: Influence of turnover on equilibrium incidence and prevalence

Experiment 1 examined the influence of turnover on equilibrium incidence and prevalence, where incidence is defined as λ i from Eq. (23), and prevalence is defined as $\hat{I}i = Ii/Xi$.

Experiment 1.1: Overall prevalence with vs without turnover. First, we compare the overall prevalence predicted by the model with and without turnover. The model with turnover is as described above. The model without turnover has all rates φ = 0. Following Eq. (14), this means that in the model without turnover, the time spent within each risk group is equal to the average duration in the modelled population μ^{-1} . The comparison of prevalence is repeated for two different treatment rates, in order to illustrate variability in the relative difference.

Experiment 1.2: Influence of turnover rates on incidence and prevalence. Second, we determined the influence of different turnover rates on equilibrium incidence and prevalence at a fixed duration of infectiousness (treatment rate). As in similar experiments (Zhang et al., 2012; Henry and Koopman, 2015), the rates of turnover were scaled by a single parameter. However, because the model used here has G = 3 risk groups, multiplying by a set of base rates φ by a scalar factor would have resulted in changes to the relative population size of risk groups $\hat{x}$. Thus, we controlled the rates of turnover using the duration of individuals in the high risk group δ_H, such that a shorter period of time in the high risk group corresponded to higher rates of turnover among all groups. The duration of individuals in the medium risk group δ M was then defined as a value between δ H and the maximum duration μ^{-1} which scales with δ_H following the equation: $\delta_M = \delta_H + \kappa\,\mu^{-1} - \delta_H$, with κ = 0.3. The duration of individuals in the low risk group δ_L similarly scaled with δ_H, but the value was not required to calculate φ; it can be determined from φ afterwards using Eq. (14). In this way, each value of δ_H was used to define a unique set of turnover rates φ whose elements

Formatted: Normal

all scaled inversely with the duration in the high risk group δ_H . The value of δ_H was then varied from 33 to 3 years to examine the influence of different turnover rates.

Experiment 1.3: Influence of turnover rates at various treatment rates. Next, we conducted a 2-way sensitivity analysis to examine how the influence of turnover might vary at different treatment rates. The treatment rate controls the duration of infectiousness δ_I as in $\delta_I = \tau^{-1}$. Treatment rate τ was varied from 1 to 0.05, implying a duration of infectiousness of 1 to 20 years. The duration of time spent in the high risk group δ_H was varied from 33 to 3 years as in Experiment 1.2. We examined the influence of the rates of turnover on equilibrium prevalence and incidence across the range of treatment rates using multiple 1D plots and 2D surface plots.

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

In Experiment 2, we examined the influence of turnover on the parameter values inferred via model fitting. Specifically, we fit the model to 20% infection prevalence among the ~~high risk~~high-risk group, 3% among the low risk group, and 5% overall, with and without turnover. We fit the contact rates C of all risk groups by minimizing the negative log-likelihood of each predicted prevalence versus the target.[6] We then compared the inferred contact rates C in the model with versus without turnover. The ratio of fitted (or posterior) contact rates C_H/C_L represents the degree of risk heterogeneity in the population, after fixing all other parameters, which produces the given infection prevalence.

3.4. Experiment 3: Influence of turnover on the TPAF of the high risk group

Finally, Experiment 3 sought to examine how the estimated contribution of highest risk group to overall transmission, as measured by the transmission population attributable fraction (TPAF), varies with versus without turnover. The TPAF of a risk group i is defined as:

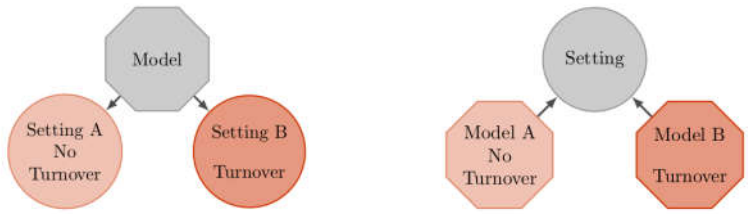
$$\mathrm{TPAF}_i(t) = \frac{I_0(t) - I_i(t)}{I_0(t)} \tag{25}$$

where I_o (t) is the cumulative number of new infections by time t under usual conditions, and I_i (t) is the cumulative number of new infections assuming no transmission from risk group *i*. Both I_o (t) and I_i (t) are calculated starting from a system at equilibrium. There are two ways to consider the comparison of TPAF with versus without turnover; these are illustrated in Figure 3, and explained in the following sections.

6 Sample sizes of 50, 750, and 1000 were assumed to generate binomial distributions for the high, low, and overall prevalence targets respectively, and the minimization was performed using the SLSQP method (Kraft, 1988) from the SciPy Python package (scipy.optimize.minimize).

Formatted: English (United States)

Formatted: Normal



(a) Experiment 3.1: same parameters, different prevalence (no model fitting)

(b) Experiment 3.2: same prevalence, different parameters (fitted model)

Figure 3: Two approaches to comparing TPAF with and without turnover

Experiment 3.1: Two settings. First, we compared two hypothetical settings, which had identical populations (including behaviour), except that setting A had no turnover: $\varphi = 0$, while setting B had turnover: φ from Eq. (24). As a result, model parameters were identical (except turnover), but the infection prevalence predicted for each setting was different. Following equilibration of the model in both settings, the TPAF of the high risk group was then estimated.

Experiment 3.2: Two models. Second, we compared two models, which were identical in structure except that Model A had no turnover and Model B had turnover. In this case, both models were fitted to the same "setting", as defined by overall and group-specific equilibrium infection prevalence (from Experiment 2). As a result, prevalence was the same in both models, but the group-specific contact rates inferred via fitting were different. As before, the TPAF of the high risk group was then estimated in each model after equilibration.

4. Results

4.1. Experiment 1: Influence of turnover on equilibrium incidence and prevalence

First, we present general trends in equilibrium prevalence and incidence with respect to turnover.

Experiment 1.1: Overall prevalence with vs without turnover. Figure 4 shows the influence of turnover on overall prevalence at two different treatment rates. In both scenarios, turnover appeared to slow the initial epidemic growth as indicated by overall prevalence. However, at equilibrium, the influence of turnover on overall prevalence depended on the treatment rate: prevalence was higher with turnover for $\tau = 0.1$ (Figure 4a), while it was higher without turnover for $\tau = 0.2$ (Figure 4b). Experiments 1.2 and 1.3 aimed to clarify and explain this influence through exploration of group-specific incidence and prevalence at equilibrium under different rates of turnover φ and treatment τ.

Experiment 1.2: Influence of turnover rates on incidence and prevalence. Figure 5 illustrates trends in equilibrium prevalence versus turnover among the high and low risk

Formatted: English (United States)

Commented [LW30]: I believe all the experiments were done with 3 risk groups- -but all results were shown just for high and low --- what happened to the middle group?

It is mentioned so less at the end – the reader may almost interpret as if there were just two risk groups in ur whole experiment.

Commented [LW31]: I think there is a lot of contents packed here for the readers to consume. To increase readability,

I may suggest trying to separate out

1) face "Results" u want to show the readers (e.g., patterns/messages u want them to take home even if they are not 100% clear about the underlying reasons);

2)underling "explanations" u want to provide

For example,

1) can include 1 figure, which has 6 panels:

Panel 1-3 shows under low treatment rate, what is the relationship between turnover rate with overall prevalence, high risk prevalence and low risk prevalence.

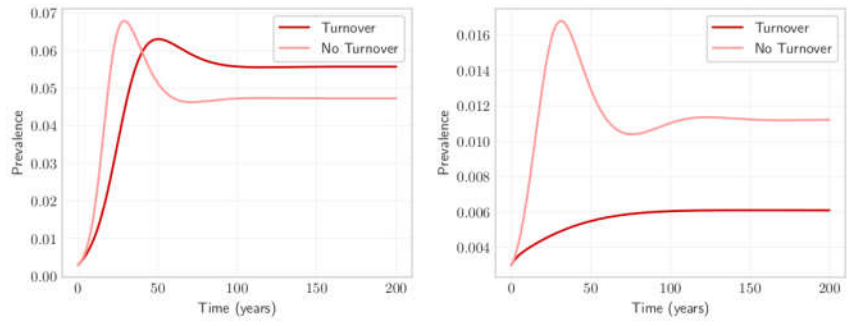
Panel 4-6 shows under low treatment rate, what is the relationship between turnover rate with overall prevalence, high risk prevalence and low risk prevalence.

2)can include things u investigated to explain the patterns u find (such as influence on incidence, etc.) some figures of 2) can be included in the Appendix only.

Commented [LW32]: Language like this sounds like method rather than results.

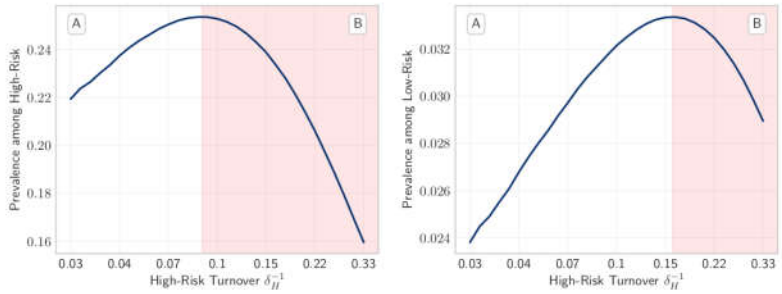
Formatted: Normal

groups for fixed treatment rate (τ = 0.1). For both groups, the same profile was observed: at low turnover, increasing turnover increased prevalence, up to a maximum value (region A), and increasing turnover beyond this point (region B) then decreased prevalence. In the ~~high risk~~high-risk group (Figure 5a), this transition occurred at a lower rate of turnover, while in the low risk group (Figure 5b), the transition occurred at a higher rate of turnover. This peaked prevalence profile can be explained by the interaction between two factors: the movement of individuals between risk groups and incidence.



(a) Treatment rate $\tau = 0.1$ (b) Treatment rate $\tau = 0.2$

Figure 4: Overall projected prevalence with and without risk group turnover under two different treatment rates τ.



(a) High risk (b) Low risk

Figure 5: Equilibrium prevalence among high and low risk groups versus turnover, as controlled by the duration in the high risk group δ_H. Turnover shown in log scale.

Commented [LW33]: U should also notethe treatment rate in Figure 5 legend.

Commented [SDB34]: I guess a style issue and happy for it to go either way.

Commented [LW35]: This figure can be dropped – or specify turnover rate here.

I think figure 5 by including x axis=0 u will have similar information.

The time-trajectory here is of less importance.

Formatted: English (United States)

Formatted: English (United States)

Formatted: Normal

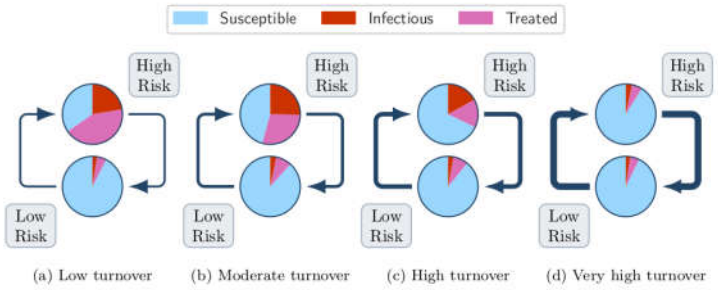


Figure 6: Average health states of individuals moving between high and low risk groups due to different rates of turnover.

The movement of individuals between risk groups is depicted in Figure 6 for four rates of turnover and a simplified system (two risk groups). Recall that, by our assumptions, the distribution of health states among individuals leaving a risk group was equal to the distribution of health states within the group. Therefore, a higher proportion of individuals leaving the high risk group were infectious as compared to individuals entering the high risk group, who were mostly susceptible. Thus, in the high risk group, turnover yielded a net replacement of infectious individuals with susceptible individuals. However, turnover similarly replaced treated individuals with susceptible individuals in this group. If incidence is sufficiently high, infection of these susceptible individuals can outpace the loss of infectious individuals via turnover. As a result, prevalence among the high risk group can actually increase with increasing turnover. In fact, this is what we observed in this model for moderate rates of turnover (Figure B.2b), yielding the increase in prevalence among high risk individuals shown in region A of Figure 5a.[7] Among the low risk group, turnover yielded a net exchange susceptible individuals for infectious and treated individuals. As a result, moderate rates of turnover also increased prevalence among the low risk group, as shown in region A of Figure 5b.

Next, and in order to explain the reversal of these trends at higher rates of turnover (region B), we examined the second factor in the influence of turnover on infection prevalence: incidence. Consider the force of infection equation, Eq (23). As shown in Appendix A.4, the dynamic component in this expression is the proportion of available partnerships which are offered by infectious individuals, denoted as C_I. This component can be further broken down into the following two sub-factors: $\hat{C}_I$ the average contact rate among infectious individuals, and $\hat{I}$ overall prevalence. Thus, the influence of turnover on incidence can be understood

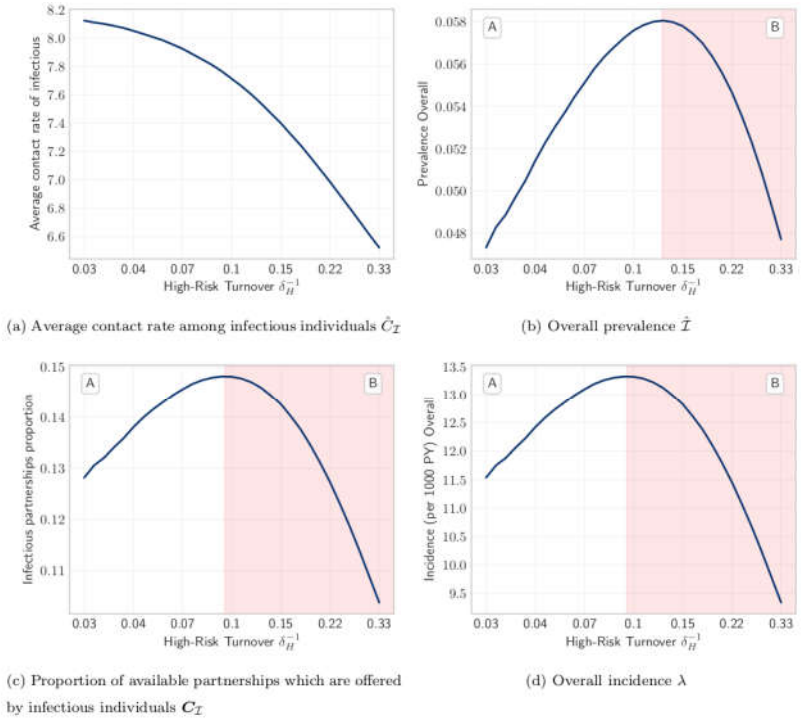
7 If prevalence is defined as to include both infectious individuals and individuals on treatment, such as in the context of HIV, then infection prevalence among the high risk group monotonically decreased with increasing turnover (Knight et al., 2019).

Formatted: English (United States)

Commented [SDB36]: Isnt it always just two states—ie, is this more of a simplified system than the others?

Formatted: Normal

through the influence of turnover on these two sub-factors.[8]



(a) Average contact rate among infectious individuals $\hat{C}_{\mathcal{I}}$

(b) Overall prevalence $\hat{\mathcal{I}}$

(c) Proportion of available partnerships which are offered by infectious individuals $\boldsymbol{C}_{\mathcal{I}}$

(d) Overall incidence λ

Figure 7: Incidence and the dynamic factors of incidence versus turnover. The product of components (a) and (b) is proportional to (c) the proportion of total available contacts which are with infectious individuals and (d) overall incidence.

As shown in Figure 7a, turnover decreased the first sub-factor: $\hat{C}_I$ the average contact rate among infectious individuals. This is because, as noted above, turnover causes in a net movement of infectious individuals from high to low risk. However, at low to moderate rates of turnover, turnover increased the second sub-factor: $\hat{I}$ overall prevalence (Figure 7b, region A). Now, under the conditions in region A, overall prevalence $\hat{I}$ increased faster with turnover than the average contact rate of infectious people $\hat{C}_I$ decreased. Thus, as a product of these two sub-factors, C_I increased with turnover in region A (Figure 7c) and incidence increased proportionally (Figure 7d).

It therefore follows that the peak in incidence, and the transition between regions A and B

8 Since we assumed proportionate mixing among risk groups, and only consider heterogeneity in contact rate C_i , incidence in each risk group is the same as overall except for a scale factor proportional to C_i .

Formatted: English (United States)

Commented [SDB37]: It may be common to write like this in the modeling world, but less so in traditional scientific writing. Ie, normally results would be laid out in an agnostic way and then discussions really provides the interpretation of why this was or was not expected. And the editorializing "therefore" would not be there as much…

Formatted: Normal

in Figure 7c occurred when the dominating sub-factor between $\hat{I}$ and $\hat{C}_I$ reversed. That is, the transition occurred when the average contact rate of infectious people $\hat{C}_I$ decreased with turnover faster than prevalence $\hat{I}$ increased with turnover. Then, as rates of turnover continued to increase, declining incidence also reversed the upward trend in prevalence, and incidence and prevalence decreased across all groups in a snowball effect. This mechanism then explains the observations shown in region B throughout.

Finally, we note that the rates of turnover which maximized incidence (Figure 7d) were lower than those which maximized overall prevalence (Figure 7b). This is because incidence represents the number of new infections per susceptible, not per individual; so a system at slightly higher prevalence can be maintained by a slightly lower incidence, provided the proportion of susceptibles has decreased more than prevalence has increased. We also note that infection prevalence among the high risk group peaked at lower rates of turnover than prevalence among the low risk group. This is because turnover yields a net movement of infectious individuals from high to low risk; so prevalence among the high risk group can decrease with turnover even as incidence increases, whereas prevalence among the low risk group can increase with turnover even as incidence decreases.

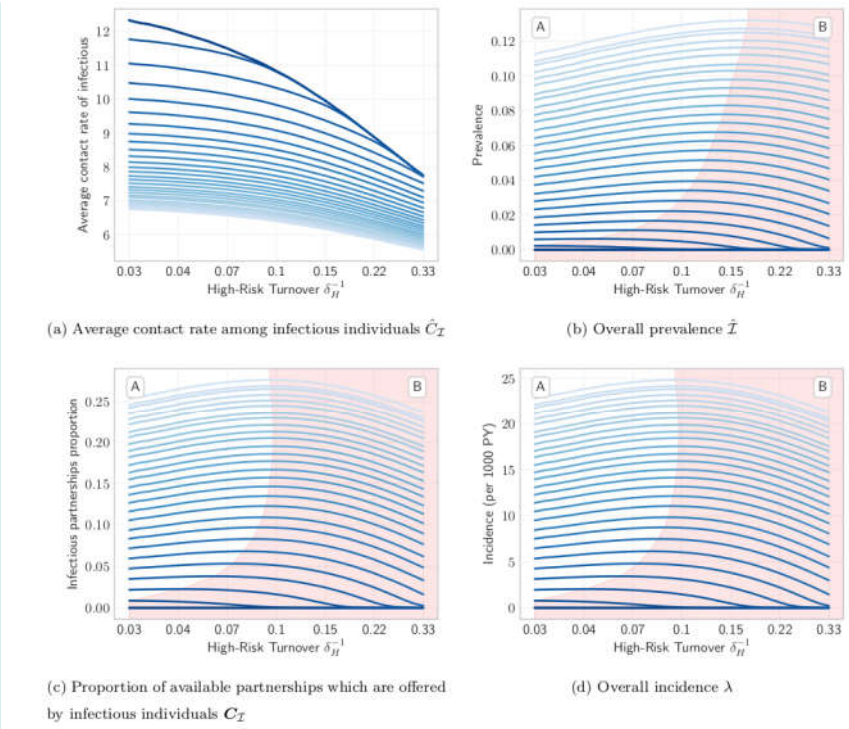
Experiment 1.3: Influence of turnover rates at various treatment rates. So far, the influence of turnover on equilibrium incidence and prevalence was explored for a single treatment rate. Experiment 1.3 explored additional treatment rates τ. First, the sub-factors of incidence are again shown in Figure 8. Increasing the treatment rate τ increased the average contact rate of infectious individuals $\hat{C}_I$ (Figure 8a). This is because increasing treatment concentrated infections in the highest risk group (Figure B.7), so that $\hat{C}_I$, on average, increased. However, increasing treatment reduced overall prevalence $\hat{I}$ (Figure 8b), due to the herd immunity effect of treated individuals. The dominant effect was that of $\hat{I}$, which tended towards zero faster than $\hat{C}_I$ tended towards infinity, and so incidence declined with treatment (Figure 8d).

Commented [SDB38]: Why is this a snowball effect?

Commented [SDB39]: Does that mean this implies a curative STI? That can have repeat episodes?

Commented [LW40]: If u were to explain everything under one treatment rate- then don't bring up two scenairos (high vs low treatment rate) at the beginning.

Formatted: Normal



(a) Average contact rate among infectious individuals $\hat{C}_{\mathcal{I}}$

(b) Overall prevalence $\hat{\mathcal{I}}$

(c) Proportion of available partnerships which are offered by infectious individuals $\boldsymbol{C}_{\mathcal{I}}$

(d) Overall incidence λ

Figure 8: Incidence and the dynamic factors of incidence versus turnover, for a range of treatment rates. Darker blue indicates higher treatment rate. The product of components (a) and (b) is proportional to (c) the proportion of total available contacts which are with infectious individuals and (d) overall incidence.

Figure 8d also shows that the rate of turnover which maximized incidence decreased with treatment. That is, as treatment rates increased, turnover was more likely to decrease incidence than it was to increase incidence (region B grew). This effect can be explained as follows. Recall that the mechanism by which turnover decreased incidence was through reduction of the average contact rate of infectious individuals $\hat{C}_I$, due to net movement of infectious individuals from high to low risk (Figure 6). If treatment increased the concentration of infections in the high risk group, then the average contact rate of infectious individuals $\hat{C}$ I would be more sensitive to redistribution of those infectious individuals to lower risk groups via turnover. In Figure 8a, this is shown as the larger downward slope of $\hat{C}_I$ versus turnover at higher treatment rates (darker blue). Therefore, at higher treatment rates, turnover was more likely to decrease incidence because movement of infectious individuals from high to low risk had a larger impact on the average contact rate of infectious individuals.

Finally, Figure 9 summarizes trends in overall equilibrium incidence and group-specific

Commented [LW41]: I found this figure hard to interpret/

Formatted: English (United States)

Formatted: Normal

prevalence with respect to both turnover φ and treatment rate τ.[9] Treatment consistently decreased equilibrium incidence (as noted above), as well as prevalence, at all rates of turnover. As suggested in Experiment 1.2, infection prevalence increased with turnover for moderate rates of turnover and treatment, among all groups, and overall. However, for higher rates of turnover, prevalence among each risk group peaked, and then declined. As shown in Figure 8, the rate of turnover at which the peak occurs decreased with treatment rate. Finally, for high rates of treatment and/or turnover, the product of the average contact rate of infectious individuals $\hat{C}_I$ and prevalence $\hat{I}$ was too low to sustain the epidemic. That is, the basic reproductive number R_o declined to less than one, and no epidemic was observed.[10]

Commented [LW42]: As mentioned in my comments above:

Choose one as the main result – for example, prevalence.

And incidence as a tool to assist in interpretation/explanation.

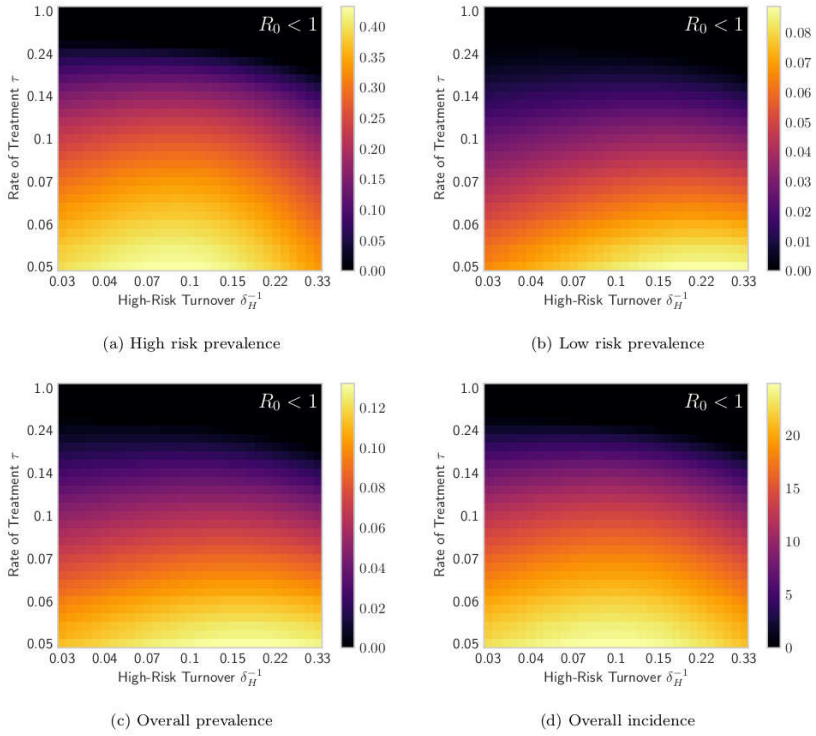
U have a lot of information – and have to strategically simply things for the reader…

Otherwise -- -way too much information to grasp and readers will be lost.

9 Figures 9c and 9d are the surface projections of the profiles shown in Figures 8b and 8d, respectively.

10 In fact, it can be shown that for extreme rates of turnover, a heterogeneous system (e.g. Full model) will converge on a homogeneous system (e.g. model V1). This result is shown in Figure B.9.

Formatted: Normal



(a) High risk prevalence

(b) Low risk prevalence

(c) Overall prevalence

(d) Overall incidence

Figure 9: Equilibrium prevalence and incidence for different rates of turnover ϕ (log scale) and treatment τ (log scale).

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

Before model fitting, the predicted prevalence ratio between high and low risk groups was lower with turnover than without: 6.7 vs 9.2. This reflects the "homogenizing" effect of turnover on the average risk experienced by individuals in the model. As shown in Figure B.7b, the high-to-low prevalence ratio consistently declined with turnover for all turnover and treatment rates explored. Thus, when fitting the model to target prevalence values, the fitted contact rates C would have to compensate for this difference. After fitting the contact rates, both models predicted the desired equilibrium infection prevalence values of 20%, 3%, and 5% among the high risk group, low risk group, and overall. However, in order to do so, the ratio of fitted contact rates between high and low risk groups (C_H/C_L) was higher with turnover than without: 23.8 vs 15.1. That is, the inferred level of risk heterogeneity was higher in the model with turnover than in the model without turnover. This is because, in order to observe the same prevalence ratio in a system with turnover, the "risk homogenizing" effects of turnover must be overcome by greater heterogeneity in risk, as

Formatted: English (United States)

Commented [SDB43]: I really feel like a separate paper… just a lot of info here!!

Commented [LW44]: Refresh the readers on under what values of treatment rate and turnover rate did u conduct these experiments.

Commented [SS45]: ???

Formatted: Normal

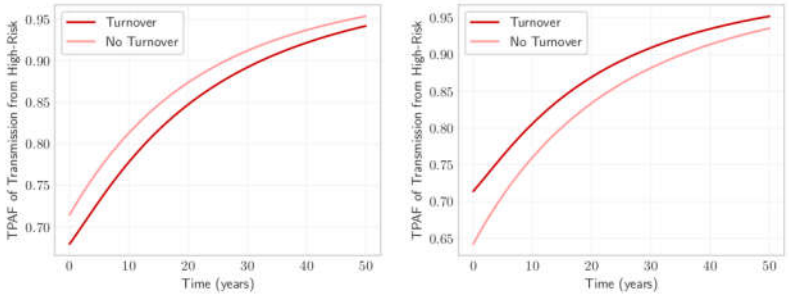
compared to a system without turnover. These results are also summarized in Table 4.

Table 4: Equilibrium contact rates C and prevalence P among the high H and low L risk groups predicted by the models with and without turnover, before and after model fitting.

Context	C_H	C_L	C_H / C_L	P_H	P_L	P_H / P_L
No Turnover (Setting A)	25.0	1.0	25.0	22%	2%	**9.2**
Turnover (Setting B)	25.0	1.0	25.0	22%	3%	**6.7**
No Turnover [fit] (Model A)	23.5	1.6	**15.1**	20%	3%	6.6
Turnover [fit] (Model B)	24.3	1.0	**23.8**	20%	3%	6.7

4.3. Experiment 3: Influence of turnover on the ~~TPAF~~tTPAF of the high risk group

Finally, we compared the predicted TPAF of the high risk group with and without turnover in: two settings (same parameters, different prevalence); and two models (same prevalence, different parameters). These results are shown in Figure 10. The TPAF approaches 1 for all models over the 50 year period, indicating that unmet treatment needs of the high risk group are central to epidemic persistence in all scenarios. Additionally, no TPAFs intersect during this period, so relative differences between TPAFs can be described irrespective of time horizon.



(a) No fitting (b) Fitted contact rates

Figure 10: Transmission population attributable fraction (TPAF-from) of the high risk group with and without turnover, and with and without fitted contact rates to group-specific prevalence.

Experiment 3.1: Two settings. Figure 10a shows the estimated TPAF of the high risk group in two different settings – with and without turnover. In this case, the estimated TPAF is lower in Setting B with turnover versus in Setting A without turnover. This can be attributed to the larger equilibrium prevalence ratio without model fitting (Experiment 2, Table 4), which results in more onward transmission from the high prevalence high risk group. In other words, the importance of reaching the high risk group in a context without turnover is

Formatted: English (United States)

Commented [SDB46]: Isnt it normally noted like this?

Commented [LW47]: I would suggest dropping the TPAF under two settings section as whole.

Not a very relevant question on its own.

Realistically speaking,

1) How can u identify two settings with exact same behavioral parameters, but one with and another without turn-over?

2)Even if there is, comparing the tPAF of high risk group in setting A (no turnover) to high risk group in setting B(turnover) is not a really fair comparison, as high risk group in setting A has a prevalence ratio of 9.2 wheares high risk group in setting B has a prevalence ratio of 6.7.

Commented [SDB48]: Crucial.

Formatted: English (United States)

Commented [SS49]: I find it a little confusing that Setting B is with Turnover but in the legend that comes first

Formatted: Normal

higher than in a context with turnover, all other factors being equal.

Experiment 3.2: Two models. Figure 10b shows the estimated TPAF of the high risk group by two different models – with and without turnover – both fitted to the same prevalence. Now the estimated TPAF is higher in Model B with turnover versus in Model A without turnover. This result, opposite to Experiment 3.1, can be explained by two factors. First, the equilibrium prevalence ratios predicted by the models with and without turnover are equalized through model fitting to the same targets. As a result, differences in prevalence in the high risk group no longer contribute to an increased TPAF estimate by Model A without turnover. Second, as shown in Experiment 2, the ratio of fitted contact rates C_H/C_L in model B with turnover are higher than in Model A without. This affords a higher risk of onward transmission to the high risk group in Model B with turnover, and thus an increased 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. Consequently, the importance of prioritizing high risk groups to achieve epidemic control may be similarly underestimated by such models.

Commented [LW50]: I would avoid further comparing experiment 3.1 and 3.2. there is no need to compare them.

Commented [LW51]: I think this is the key message and valid.

Commented [SDB52]: While I agree—this is really interpretation. To me, assumed this would be more in the discussion but could also be more about the style of paper.

5. Discussion

Epidemic models have rarely considered turnover of individuals among risk groups. We developed a unified framework for modelling turnover in risk, and used the 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 this contribution is lower in settings with turnover, but that failure to model turnover when simulating settings with turnover could result in underestimation of the contribution.

Commented [LW53]: I don't think this is a valid stand alone statement as mentioned in my comments above in results

Commented [SDB54]: Of turn over or high risk groups?

Commented [LW55]: I would keep this at the end of discussion before limitation section.

It just flows better with results if we discuss result implications first before moving into method advance.

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 four 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 to be chosen and combined in a flexible way (as rows in the linear system of equations $b = A\theta$), 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 the parameters 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. In so doing, we highlight which specific assumptions are the same and which are different across the different implementations.

The major data needs of this approach include the following four categories. First, the proportion of total individuals in each risk group $\hat{x}_i$ is required. In the context of STIs, risk

Formatted: Normal

group size estimates may be obtained from representative surveys of the modelled population such as demographic health surveys, which collect data on self-reported behaviours or engagement in networks associated with variable risk (USAID, 2019). For example, one risk group may be defined by any engagement in casual sex within the past year. For marginalized populations, such as sex workers, estimates of population size are generated using various mapping and enumeration methods (Abdul-Quader et al., 2014). Second, the proportion of exogenous individuals who enter into each risk group can be used to specify elements ê i as Type 2 Constraints. Such proportions could be obtained the same way as xˆ 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 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 within variable periods of risk (Fergus et al., 2007). Fourth, similar data, if available, can also be used to estimate the rates of transition between specific risk groups (φ).

Influence of turnover on incidence & prevalence. We found that turnover influences the overall equilibrium STI incidence and prevalence, as shown by 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 incidence and prevalence, and demonstrated mechanistically how this occurs. 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. We found that the TPAF of the highest risk group would be lower in settings without turnover. This is important to consider when comparing whom to prioritize in different regions with different epidemiologic features, such as STI prevalence. That is, a setting with turnover may require less of a focused approach on those at highest risk. Such an implication may be counterintuitive, as a shorter period of higher risk among a smaller group means fewer person-years of intervention required. However, as described in the paragraph above, turnover reduces heterogeneity in risk and reduces the STI prevalence ratio between the highest and lowest risk groups, decreasing the contribution of the highest risk group to overall transmission. Public health implications of the above finding are two-fold. First, if epidemiologists and modelers were to generate a geographic

Commented [SDB56]: Well, size estimates are complicated for key populations and DHS likely not ideal. But could reference either this https://www.who.int/hiv/pub/surveillance/final_estimating_populations_en.pdf or one of our pieces like https://www.ncbi.nlm.nih.gov/pubmed/31341935

Commented [SDB57]: aha! Ok!

Commented [LW58]: I would avoid the symbols (e.g., δ. x ….) in discussion section.

Commented [LW59]: Did the previous work demonstrate similar patterns?

Formatted: Highlight

Commented [SS60]: With? These two highlighted sentences seem to be in conflict otherwise

Commented [LW61]: I think u mean higher? Typo?

I don't think it is a fair comparison in two settings as discussed in my comments above on results.

Formatted: Highlight

Formatted: Highlight

Commented [LW62]: I think such conclusion/implication is misleading:

Comparing different high risk groups in two settings and draw conclusion might be unfair.

Formatted: Normal

map of TPAF across different settings with different rates of turnover, then turnover could be an important source of variability in the projected TPAF between settings. Second, if decision-makers were allocating a fixed ceiling of STI prevention and treatment resources for a particular high-risk group (e.g. sex workers) across settings, based on the TPAF, then settings with zero or lower rates of turnover may benefit the most from those resources. The latter implication is a general insight, and does not account for: transmission between sites; how costs of STI interventions vary of the sexual life-course of individuals; and other sources of variability in TPAF not explored in our study.

In contrast to a comparison of settings with and without turnover, our comparison of models with and without turnover, calibrated to the same epidemic, showed that ignoring turnover (if it exists in a given setting) will cause the TPAF to be underestimated. 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 vs. without turnover was even larger (Appendix B.7). The public health implication of models ignoring turnover which is present is that the TPAF of the highest-risk group will be systematically underestimated, potentially misguiding resources away from the highest-risk group. 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.

Limitations. There are five primary limitations of the study that are important when considering the implications of our proposed unified approach to turnover, and the implications of turnover on TPAF. 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 the higher risk groups, and that alone can cause an epidemic to decline (Boily and Mâsse, 1997). As such, many models of HIV transmission, in particular those 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). Additional work on the proposed framework could address these two important limitations to the generalizability of the approach across STI models.

Third, 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

Commented [LW63]: I think this is equivalent as:

if u were to allocate a fixed amount of resources to e.g., the top 20% population with highest risk in a setting with e.g, 20% HIV prevalence vs. in a setting with 15% HIV prevalence, the latter will benefit more. I think this is almost an intuitive conclusion.

Turn-over is not really the cause here – but rather a confounder per-se --- created through the way the experiments were designed.

Commented [SS64]: Over?

Commented [LW65]: I think this is a valid conclusion

Commented [LW66]: I think underltying assumptions used in ur experiments should be highlighted especially the third assumption.

Commented [LW67]: Lacking a "however" statement following this limitation to justify in which case such limitation is of less concern

Commented [LW68]: What do u mean here? U mean future work is required?

Formatted: Normal

treatment will fuel an epidemic, and so the influence of turnover on STI incidence, prevalence, and TPAF may be different, and warrants future study. Fourth, 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 in ~~(~~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. Although the data~~-~~ needs remain potentially challenging, a failure to meet this challenge could lead to misguided information on the importance of addressing the unmet needs of high risk populations – ~~such~~ 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~~as those engaged in sex work~~ – to achieve population-level transmission reduction.

Commented [SS69]: This seems to undercut the paper a bit – if the data needs are allowed to come off as too daunting. I wonder if before highlighting what the failure would result in, if it's worth noting that suggestions were provided of how to estimate / parameterize these data points. Perhaps not necessary, but I was thinking again about the contributions you want to highlight in this paper and the methods that this is advancing, and by ending with this statement it seems to downplay the methodological advance a bit by making it sound difficult to implement. Feel free to ignore.

Commented [SDB70R69]: Yep, this is the balance of the big data grant. Ie, data are weak but not so weak that we cant do things. So while a balanced approach is needed, it is possible!

References

Abu S Abdul-Quader, Andrew L Baughman, and Wolfgang Hladik. Estimating the size of key populations: Current status and future possibilities. 9(2):107–114, mar 2014. DOI 10.1097/COH.0000000000000041.

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

Marie Claude Boily and Benoı̂t Mâsse. Mathematical models of disease transmission: A precious tool for the study of sexually transmitted diseases. Canadian Journal of Public Health, 88(4):255–265, 1997. DOI 10.1007/bf03404793.

Marie Claude Boily, Michael Pickles, Michel Alary, Stefan Baral, James Blanchard, Stephen Moses, Peter Vickerman, and Sharmistha Mishra. 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, mar 2015. DOI 10.1097/QAI.0000000000000437.

Kelsey Case, Peter Ghys, Eleanor Gouws, Jeffery Eaton, Annick Borquez, John Stover, Paloma Cuchi, Laith Abu-Raddad, Geoffrey Garnett, and Timothy Hallett. Understanding the modes of tranmission model of new HIV infection and its use in prevention planning. Bulletin of the World Health Organization, 90(11):831–838, nov 2012. DOI 10.2471/blt.12.102574.

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

Jeffrey W. Eaton and Timothy B. Hallett. Why the proportion of transmission during early-stage HIV infection does not predict the long-term impact of treatment on HIV incidence.

Formatted: Normal

Proceedings of the National Academy of Sciences, 111(45): 16202–16207, nov 2014. DOI 10.1073/pnas.1323007111.

Kevin A. Fenton, Romulus Breban, Raffaele Vardavas, Justin T. Okano, Tara Martin, Sevgi Aral, and Sally Blower. Infectious syphilis in high-income settings in the 21st century. The Lancet Infectious Diseases, 8(4):244–253, apr 2008. DOI 10.1016/S1473-3099(08)70065-3.

Stevenson Fergus, Marc A Zimmerman, and Cleopatra H Caldwell. Growth trajectories of sexual risk behavior in adolescence and young adulthood. American Journal of Public Health, 97(6):1096–1101, jun 2007. DOI 10.2105/AJPH.2005.074609.

Geoffrey P. Garnett and Roy M. Anderson. Balancing sexual partnership in an age and activity stratified model of HIV transmission in heterosexual populations. Mathematical Medicine and Biology, 11(3):161–192, jan 1994. DOI 10.1093/imammb/11.3.161.

Christopher J. Henry and James S. Koopman. Strong influence of behavioral dynamics on the ability of testing and treating HIV to stop transmission. Scientific Reports, 5(1):9467, aug 2015. DOI 10.1038/srep09467.

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

Leigh F. Johnson and Nathan Geffen. A Comparison of two mathematical modeling frameworks for evaluating sexually transmitted infection epidemiology. Sexually Transmitted Diseases, 43(3):139–146, mar 2016. DOI 10.1097/OLQ.0000000000000412.

Jesse Knight, Linwei Wang, Huiting Ma, Sheree Schwartz, Stefan Baral, and Sharmistha Mishra. The influence of risk group turnover in STI/HIV epidemics: mechanistic insights from transmission modeling. In STI & HIV 2019 World Congress, Vancouver, BC, Canada, 2019. URL https://sti.bmj.com/content/95/Suppl_1/A83.3.

J S Koopman, J A Jacquez, G W Welch, C P Simon, B Foxman, S M Pollock, D Barth-Jones, A L Adams, and K Lange. The role of early HIV infection in the spread of HIV through populations. Journal of Acquired Immune Deficiency Syndromes, 14(3):249–58, mar 1997. URL http://www.ncbi.nlm.nih.gov/pubmed/9117458.

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

LAPACK. LAPACK: Linear Algebra PACKage, 1992. URL http://www.netlib.org/lapack. Charles L Lawson and Richard J Hanson. Solving least squares problems, volume 15. SIAM, 1995.

Thomas Robert Malthus. An Essay on the Principle of Population. 1798.

Cicely Marston and Eleanor King. Factors that shape young people's sexual behaviour: a systematic review. Lancet, 368 (9547):1581–1586, nov 2006. DOI 10.1016/S0140-

Formatted: Normal

6736(06)69662-1.

Sharmistha Mishra, Richard Steen, Antonio Gerbase, Ying Ru Lo, and Marie Claude Boily. Impact of High-Risk Sex and Focused Interventions in Heterosexual HIV Epidemics: A Systematic Review of Mathematical Models. PLoS ONE, 7(11): e50691, nov 2012. DOI 10.1371/journal.pone.0050691.

Sharmistha Mishra, Michael Pickles, James F Blanchard, Stephen Moses, and Marie Claude Boily. 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(1):19–25, feb 2014. DOI 10.1136/sextrans-2013-051250.

Sharmistha Mishra, Marie-Claude Boily, Sheree Schwartz, Chris Beyrer, James F. Blanchard, Stephen Moses, Delivette Castor, Nancy Phaswana-Mafuya, Peter Vickerman, Fatou Drame, Michel Alary, and Stefan D. Baral. 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(8):557–569, aug 2016. DOI 10.1016/j.annepidem.2016.06.004.

Christinah Mukandavire, Josephine Walker, Sheree Schwartz, Marie-Claude Boily, Leon Danon, Carrie Lyons, Daouda Diouf, Ben Liestman, Nafissatou Leye Diouf, Fatou Drame, Karleen Coly, Remy Serge Manzi Muhire, Safiatou Thiam, Papa Amadou Niang Diallo, Coumba Toure Kane, Cheikh Ndour, Erik Volz, Sharmistha Mishra, Stefan Baral, and Peter Vickerman. Estimating the contribution of key populations towards the spread of HIV in Dakar, Senegal. Journal of the International AIDS Society, 21:e25126, jul 2018. DOI 10.1002/jia2.25126.

Annett Nold. Heterogeneity in disease-transmission modeling. Mathematical Biosciences, 52(3-4):227–240, dec 1980. DOI 10.1016/0025-5564(80)90069-3.

Michael Pickles, Marie Claude Boily, Peter Vickerman, Catherine M Lowndes, Stephen Moses, James F Blanchard, Kathleen N Deering, Janet Bradley, Banadakoppa M Ramesh, Reynold Washington, Rajatashuvra Adhikary, Mandar Mainkar, Ramesh S Paranjape, and Michel Alary. 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(5):e289–e299, nov 2013. DOI 10.1016/S2214-109X(13)70083-4.

Babak Pourbohloul, Michael L. Rekart, and Robert C. Brunham. Impact of mass treatment on syphilis transmission: A mathematical modeling approach. Sexually Transmitted Diseases, 30(4):297–305, apr 2003. DOI 10.1097/00007435-200304000-00005.

Zara Shubber, Sharmistha Mishra, Juan F. Vesga, and Marie Claude Boily. The HIV modes of transmission model: A systematic review of its findings and adherence to guidelines. Journal of the International AIDS Society, 17(1):18928, jan 2014. DOI 10.7448/IAS.17.1.18928.

Formatted: Normal

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

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

C. Watts, C. Zimmerman, A. M. Foss, M. Hossain, A. Cox, and P. Vickerman. Remodelling core group theory: the role of sustaining populations in HIV transmission. Sexually Transmitted Infections, 86(Suppl 3):iii85–iii92, dec 2010. DOI 10.1136/sti.2010.044602.

James A Yorke, Herbert W Hethcote, and Annett Nold. Dynamics and control of the transmission of gonorrhea. Sexually Transmitted Diseases, 5(2):51–56, 1978. DOI 10.1097/00007435-197804000-00003.

Xinyu Zhang, Lin Zhong, Ethan Romero-Severson, Shah Jamal Alam, Christopher J Henry, Erik M Volz, and James S Koopman. Episodic HIV Risk Behavior Can Greatly Amplify HIV Prevalence and the Fraction of Transmissions from Acute HIV Infection. Statistical Communications in Infectious Diseases, 4(1), nov 2012. DOI 10.1515/1948-4690.1041.

Formatted: Normal