# Online Appendix: Supplementary Digital Content (SDC)

Supplement to:

Beck E, Birkett M, Armbruster B, Mustanski B. A data-driven simulation of HIV spread among young men who have sex with men: the role of age and race mixing, and STIs.

Submitted to *JAIDS*, April, 2015.

Supplementary Digital Content contains

SDC 1: Crew 450 study

SDC 2: Sexual partnership formation and dissolution

SDC 3: Disease Transmission

SDC 4: Population size, age range, and simulated time horizon

SDC 5: Implementation and Validation

SDC 6: Addendum to results

SDC 7: Counterfactual scenarios

For a detailed table of contents see page ii.

# Table of Contents

I. Supplemental Digital Content 1: Crew 450 study ..... 1
1. Overview SDC 1 ..... 1
II. Supplemental Digital Content 2: Sexual partnership formation and dissolution ..... 5
1. Overview SDC 2 ..... 5
2. Sexual partnership formation ..... 8
2.1 Individual and network based representation and formation of partnerships ..... 8
2.2 Estimated individual attributes of YMSM ..... 9
2.3 One-night-partnerships ..... 13
2.4 Partnerships outside the study population ..... 21
2.5 Partnerships within the study population ..... 29
3. Sexual partnership dissolution ..... 39
4. Sexual orientation and sex-role: tendency and behavior ..... 41
4.1 Notation ..... 42
4.2 Sexual orientation ..... 42
4.3 Desired sex-role, sex-frequency and actual behavior in within-partnerships ..... 44
III. Supplemental Digital Content 3: Disease Transmission ..... 52
1. Overview ..... 52
2. HIV ..... 54
2.1 Course of the HIV infection ..... 54
2.2 HIV Transmission ..... 59
2.3 Interactions with other diseases ..... 70
3. Neisseria Gonorrhea (NG) and Chlamydia trachomatis (CT) ..... 71
3.1 Course of NG and CT infection ..... 72
3.2 NG and CT transmission ..... 78
3.3 Interactions with other diseases ..... 82
4. Testing and treatment coverage ..... 88
4.1 Testing ..... 88
4.2 HIV treatment coverage ..... 89
IV. Supplemental Digital Content 4: Population size, age range, and simulated time horizon ..... 90
1. Age range ..... 90
2. Simulated YMSM population size ..... 90

2.1 YMSM population size at t=0 ..... 90
3.2 Aging out, death, and birth processes ..... 92
3. Simulated time horizon ..... 97
V. Supplemental Digital Content 5: Implementation and Validation ..... 98
1. Implementation ..... 98
2. Validation ..... 102
2.1 Partnership formation and dissolution ..... 102
2.2 Disease transmission ..... 106
2.3 Validation summary ..... 126
VI. Supplemental Digital Content 6: Counterfactual Scenarios ..... 127
1. No race-assortative mixing ..... 127
2. No difference in HIV prevalence ..... 128
3. No racial assortativity and no difference in HIV prevalence ..... 128
4. No increased HIV transmissibility and susceptibility due to NG and CT ..... 128
5. No HIV transmission in outside partnerships and 50% reduction in HIV transmission risk in outside partnerships ..... 129
VII. Supplemental Digital Content 7: Addendum to results ..... 130
1. HIV, NG, and CT prevalence among outside (older) MSM over time ..... 130
2. HIV infections per 100 partnership years, male-male partnerships only ..... 133
References ..... 135

# I. Supplemental Digital Content 1: Crew 450 study

# 1. Overview SDC 1

Data were collected as part of the ongoing longitudinal study of young men who have sex with men (YMSM) conducted in Chicago starting in December 2009 and ending in February 2013. An individual was eligible for participation if they were between the ages of 16 and 20 years, assigned a male sex at birth, spoke English, reported a sexual encounter with a male or an identity of gay/bisexual, and was available for at least 2 years of follow-up.

Data were collected every six months starting at baseline (T1) and data collection continued over 3.5 years until the final data collection wave (T7). The time between data collection waves T5 and T6 was 12 months instead of 6 such that data at T6 and T7 were collected at 36 and 42 months after enrollment respectively. At each data collection wave, participants were administered a computer-assisted self-interview that lasted approximately 1 hour, and were compensated $45 for their time.

Questions about demographic characteristics such as age, race/ethnicity, and sexual identity were assessed at baseline. Participants were asked a series of questions about their sexual behaviors in the past 6 months. Items included overall number of male and female sex partners as well as overall number of partners with whom they had oral, vaginal, or anal sex. In in-depth interviews about their three most recent sex-partners they were asked about the number of vaginal or anal sex acts and number of condomless vaginal or anal sex acts for each of these partnerships. They were also asked about characteristics of their three most recent partners, including the type of partner (serious, casual, one-night stand, etc.), gender, race/ethnicity, age, and HIV status.

Individual (ego) data of YMSM collected at baseline (T1) and partnership data of YMSM collected at T1 and T2 of n=421 YMSM were used to design and parameterize our simulation model. Data collection for waves T1 and T2 was completed at time of parameterization of the model. 29 YMSM (6.4%) were excluded from the sample of the Crew 450 study due to incomplete data. Partnership data and HIV/STI testing data of waves T1 through T7 of the Crew 450 study were used to validate our simulation model.

Table 1 provides an overview over demographic and biomedical variables of the n=421 YMSM. Table 2 provides an overview over partnership variables relevant to design and parameterize our simulation model.

**Table 1:** Demographic and biomedical variables of n=421 YMSM.

| **Individual (ego) variables of n=421 YMSM** | **Percentage** |
|---|---|
| **Demographics at baseline (T1)** | |
| *Race/ethnicity* | |
| Non-Hispanic Black | 53.4% |
| Hispanic or Latinos | 19.7% |
| Non-Hispanic White | 18.1% |
| Other[a] | 8.8% |
| *Age (years)* | |
| 16 | 9.5% |
| 17 | 14.7% |
| 18 | 20.9% |
| 19 | 29.7% |
| 20 | 25.2% |
| *Sexual orientation (self-identified)* | |
| only gay/homosexual | 49.4% |
| mostly gay/homosexual | 23.3% |
| bisexual | 21.9% |
| mostly heterosexual | 2.6% |
| only heterosexual | 0.7% |

| Individual (ego) variables of n=421 YMSM | Percentage |
|---|---|
| other[b] | 2.1% |
| *Socioeconomic status (SES) (low/high)* | |
| SES 1[c] low | 37.1 |
| SES 2[d] low | 58.0 |
| **Biomedical at baseline (T1) (tested in study)** | |
| HIV positive | 7.1 |
| NG[e] positive (urethral) (T2) | 4.0 |
| CT[f] positive (urethral) (T2) | 5.3 |
| circumcised | 59.9 |

[a] Other races: Asian or Pacific Islander, Native American, Multiracial and Other
[b] Other sexual orientation (self-identified): pansexual, trysexual, queer and other
[c] SES 1: SES background category 1 – parental higher education, i.e. high if one or both parents attended college
[d] SES 2: SES background category 2 – parental higher education degree, i.e. high if one or both parents obtained a BA
[e] NG: Neisseria gonorrheae; only tested at T2
[f] CT: Chlamydia Trachomatis; only tested at T2

Table 2: Partnership variables of partnerships mentioned by n=421 YMSM.

| Partnership variables | Explanation |
|---|---|
| **Overall partnerships (n=3678)[a]** | |
| tpfm | total number of sex partners per YMSM in last 6 months stratified by sex of partner and sex-type (i.e., anal, oral or vaginal) |
| **Partnerships named in-depth interviews (n=1651)[b]** | |
| length | length of partnership in days |
| page | age of the partner |
| hivstatusk | Assumed/known HIV status of YMSM partner; assumed by participant. |
| hivstatusk_type | Describes whether YMSM assumed HIV status of partner or knows through disclosure, tests, or other information |
| hivstatusp | HIV status of partnership (i.e., ++ (p), +- (d),--(n)). Combines knowledge of YMSM (ego) about his own HIV status and knowledge of ego about partner HIV status (hivstatusk). |
| racemix | racemix of partnership where partner was categorized as either Black, Latino, White or Other (i.e., race-mixes BB, |

| Partnership variables | Explanation |
|---|---|
| | BL, BW, BO, LL, LW, LO, WW, WO and OO possible) |
| serious | relationship considered serious |
| alcuse | alcohol consumption in relationship |
| druguse | drug (i.e., poppers, marihuana, etc., no injection drug use ) consumption in relationship |
| subuse | substance consumption in relationship (either alcohol, drugs or both) |
| partcon | partner concurrency in relationships where either none, one or both partners had sexual relationships outside the partnership |
| visex | number of sex-acts with a female in last 6 months |
| arsex | number of anal sex-acts being receptive (position of ego) in last six months in partnership |
| aisex | number of anal sex-acts being insertive (position of ego) in last six months in partnership |
| oral-only | relationship was oral-sex only |
| uai[c] | unprotected anal/vaginal intercourse in partnership in last six months |

[a] n=421 named in T1, T2 or T1 and T2 in total 3678 sex-partners.

[b] n=421 named in T1, T2 or T1 and T2 in total 1651 sex-partners in the in-depth interviews where detailed data about the sex-partners were collected. Partnership variables such as hivstatusp or racemix were determined based upon partner characteristics. Partnership variables length, page, hivstatusp, racemix and partcon are categorical variables; partnership variables disclosed, serious, alcuse, druguse, subuse, oral-only and unprotected anal intercourse (uai) are binary variables.

[c] The binary variable uai was determined by classifying a categorical variable which asks about the condom usage in a partnership stratified by sex of partner and sex-position respectively. Categories of condom usage where ‘never’, ‘less than half the time’, ‘about half the time’, ‘more than half the time’ and ‘always’. Except for ‘always’ we classified the variable uai as positive.

# II. Supplemental Digital Content 2: Sexual partnership formation and dissolution

# 1. Overview SDC 2

In this discrete-time stochastic agent-based network simulation model of HIV spread among YMSM we assume HIV to be transmitted only through sexual partnerships. The design of our partnership formation and dissolution model simulating the formation and dissolution of partnerships of YMSM was informed by data of the ongoing longitudinal Crew 450 study (see SDC 1). We incorporated partnership characteristics which are known to impact HIV spread such as the fact that HIV transmission risk is dependent on the type of sexual intercourse (i.e. vaginal, anal, or oral), partnership seriousness (e.g., casual sex[1]), frequency of sex, sex-position or sexual risk behaviors in a partnership such as condom usage[2]), and characteristics of the surrounding sexual network (e.g. the level of partnership concurrency among the target population[3]). Data of the Crew 450 study was used to parameterize the model.

Next we introduce our partnership formation and dissolution model describing how YMSM form different types of partnerships. We model three main types of partnerships: *one-night-partnerships*, *outside-partnerships* and *within-partnerships*. Table 3 gives an overview over the different partnership-types and the corresponding partnership attributes. The categorization of sexual partnerships among YMSM into these three main partnership-types was based upon the survey design and available data of the Crew 450 study. In particular, the significant amount of female sex partners of YMSM (11.3% of all named sex-contacts were female); the fact that YMSM have sexual partnerships with other YMSM enrolled in the study population; and differences in the sexual behavior of YMSM with an older male partner compared to a younger partner (see sections SDC 2.2.4 and SDC 2.2.5) led us to classify

partnerships into these three main groups. This is differs from the classification of partnerships into 'main partnerships' and 'casual partnerships' in other models of HIV spread among MSM[2,4]. The age range of the study population was defined to be 16 to 21.8 years which represents the age range of study participants across waves T1 and T2 since partnership data of T1 and T2 were used to parameterize the model. Our simulation first determines in each time step the partnerships that are formed before determining their characteristics (see SDC 2.2.3, 2.2.4., and 2.2.5). Finally, YMSM are assumed to dissolve their partnerships each time step with a certain probability specified by the partnership attributes (see section SDC 2.3) in this discrete-time simulation.

Table 3: Partnership-types and attributes in model of HIV spread among YMSM.

| Partnership-type | Duration | Sex of partner | Age of partner | Other attributes |
|---|---|---|---|---|
| **One-night-partnerships** | 1 day | female, male | all ages | racemix<br>HIV, NG, CT status of partnership<br>drug, alcohol, substance use (ego)<br>oral-only<br>sex-role position (ego)<br>sex-role versatility<br>HIV status disclosure (partner)<br>UAI |
| **Outside-partnerships** | > 1day | female, male | female: all ages<br>male:<br>older (i.e., > 21.8 years) or younger (i.e., < 16 years) | racemix<br>HIV, NG, CT status of partnership<br>drug, alcohol, substance use (ego)<br>oral-only<br>sex-role position (ego)<br>sex-role versatility<br>sex-frequency<br>HIV status disclosure (partner)<br>seriousness<br>mean length<br>UAI |
| **Within-partnerships**<br>(partnerships among YMSM) | > 1 day | male | YMSM (i.e., 16 – 21.8 years) | racemix<br>HIV, NG, CT status of partnership<br>drug, alcohol, substance use (partnership)<br>oral-only<br>sex-role position (partnership)<br>sex-role versatility<br>sex-frequency<br>HIV status disclosure (partnership)<br>seriousness<br>mean length<br>UAI |

# 2. Sexual partnership formation

## 2.1 Individual and network based representation and formation of partnerships

In each simulated time step we assume YMSM to form one-night-partnerships and outside-partnerships independent of other YMSM in the study population. Thus, the formation of an one-night-partnership or outside-partnership only depends on the individual characteristics of the YMSM including his current partnership status but not on the behavior of other YMSM forming one-night-partnerships or outside-partnerships. We model the formation of one-night-partnerships and outside-partnerships as ego individual events, i.e. YMSM form these with a certain rate and thus these partnerships can be seen rather as an additional attribute of each YMSM but not as a link between two YMSM in the study population. We reason that we do not know in detail about the individual characteristics and sexual networks of female and older male sex-partners of YMSM in the Crew 450 study which therefore does not allow for an accurate representation of one-night-partnerships and outside-partnerships as ties in an actual network. In particular, we do not know about the partnership dynamics and sexual interactions in these networks and between these networks. Also, for one-night-partnerships we only have limited data about the age of the partner and thus we do not stratify one-night-partnerships by age (see also discussion in section 2.3.1.1 in SDC 2).

A sub-study of the longitudinal Crew 450 study with n=175 participants (PI: Birkett, R03DA033906) showed the existence of partnerships among study participants. Using the detailed data of the in-depth interviews in the Crew 450 study where participants where asked about their individual characteristics and their sexual partnership behavior, we're able to model within-partnerships (i.e., partnerships among YMSM) as a network where each tie in the network represents a partnership between two YMSM. In our model, two YMSM form a within-partnership (i.e., a tie between two YMSM of duration longer than a day) dependent on the specific characteristics of the potential partnership which includes the combination

of individual attributes such as age and race mixing as well as the partnership status (i.e., number of currently ongoing partnerships) of both YMSM.

## 2.2 Estimated individual attributes of YMSM

We use multivariate regression analysis to determine YMSM specific partnership formation rates and partnership attributes once a partnership has been formed. Independent variables of these regression models are individual variables shown in Table 1, partnership variables shown in Table 2, and estimated individual (ego) variables shown in Table 4. The estimated individual variables are either modified individual variables of the Crew 450 study such as *age** or represent individual characteristics which were inferred from observed behavior in partnerships such as *druguse** or the number of concurrent partnerships *momentary degree*$_{one}$. Especially for one-night-partnerships and outside-partnerships these estimated variables allow for more detail in modelling the partnership formation process and actual partnership attributes. This allows us to capture a broad range of hypothesized mechanisms contributing to HIV transmission among YMSM[5] using our model of partnership formation and dissolution. The procedures to estimate the individual variables *age**, *druguse**, *alcuse**, *substanceuse** are described in Table 4 whereas the estimation of the desired sex-role *R* and the desired sex-frequency *F* of a YMSM is described in section SDC 2.4. Following we describe the estimation of the individual attributes *momentary degree*$_{one}$ and *momentary degree*$_{part}$ in detail.

### 2.2.1 Number of concurrent partnerships

We assume the formation of partnerships to depend on the number of already ongoing partnerships at time of the formation of a new partnership (i.e., number of concurrent partnerships)[2,4]. We do not consider one-night-partnerships to impact the partnership formation, i.e. we only consider to impact the current number of outside-partnerships and within-partnerships to impact the formation of one-night-partnerships, outside-partnerships, and within-partnerships. In our model, we denote the number of concurrent partnerships (i.e., outside-partnerships and within-partnerships) at time of partnership formation as

momentary degree. We assume outside- and within-partnerships which are non-oral sex only partnerships and serious oral-sex only to impact the formation of one-night-partnerships, outside-partnerships, and within-partnerships.

Table 4: Estimated attributes and number of concurrent partnerships of YMSM at T1 and T2.

| **Estimated variables of individuals** | **Corresponding variable Crew 450 study** | **Explanation[a]** |
|---|---|---|
| **Individual attributes** | | |
| age* | age[b] | Difference of age of YMSM to overall average age of n=421 YMSM at T1/T2 being 19.19 years |
| alcuse* | alcuse[c] | YMSM has used alcohol in a sexual partnership at T1/T2 |
| druguse* | druguse[c] | YMSM has used drugs in a sexual partnership at T1/T2 |
| subuse* | subuse[c] | YMSM has either used alcohol, drugs or both alcohol and drugs in a partnership at T1/T2 |
| R | aisex[c], arsex[c] | Desired sex-role of YMSM determined based upon observed sex-role behavior in his male-male partnerships[d] |
| F | aisex[c], arsex[c], visex[c] | Desired sex-frequency of YMSM determined based upon observed sex-role behavior in his partnerships[d] |
| **Number of ongoing partnerships** | | |
| Momentary degree$_{one}$ | number mentioned partnerships, partcon[c] | Number of concurrent partnerships[e] (outside and within) when forming a one-night-partnership at T1/T2. |
| Momentary degree$_{part}$ | number mentioned partnerships, lenght[c], partcon[c] | Number of concurrent partnerships[e] (outide and within) when forming a new outside-partnership or within-partnership at T1/T2 |

[a] For details see discussion in section SDC 2.2.4. Variables age*, desired sex-role and desired sex-frequency are continuous variables; alcuse*, druguse* and subuse* are binary variables.
[b] Corresponding Crew 450 variable is individual (ego) variable, see Table 2.
[c] Corresponding Crew 450 variable is partnership variable, see Table 3.
[d] For details see discussion in section SDC 2.4.
[e] A concurrent partnership is an outside-partnership or within-partnership which is either non-oral sex only or serious oral-sex only.

#### 2.2.1.1 Number of concurrent partnerships when forming a one-night-partnership: Momentary degree$_{one}$

Using the data of the in-depth interviews about the last three sex-partners at T1 and T2 we estimated the momentary degree for each YMSM. The momentary degree$_{one}$ denotes the number of partnerships (see definition of such a partnership in section SDC 2.2.2.1) of a YMSM at the time of the interview. Since in-depth interviews were limited to the last three sex partners we assume YMSM in our model to have either 0, 1, 2, or 3 or more simultaneous ongoing partnerships at the time of the interview.

We assume the formation of one-night-partnerships to depend on the momentary degree$_{one}$. Since we do not know about the actual time when a one-night-partnership happened throughout the last six months and the corresponding number of currently ongoing partnerships but only the total number of one-night-partnerships during the six months before the interview, we assume the momentary degree$_{one}$ of each YMSM at the time of the interview (i.e., the total number of ongoing partnerships at the time of the interview) to approximate the average degree of each YMSM over the last six months before the interview. Based on this assumption, the momentary degree$_{one}$ approximates the degree of a YMSM when forming a one-night-partnership.

Figure 1 shows the approximate momentary degree distribution of n=421 YMSM at T1 and T2. In comparison, the momentary degree distribution of the Explore study used in Beyrer et al.[2] and Goodreau et al.[4] shows 60.0% percent of men having no main partnership, 38.3% having 1 main partnership and 1.7% having two main partnerships. Thus, our findings of the n=421 YMSM are comparable for the percentage of YMSM having no ongoing partnership. However, our findings show a higher percentage for degree 2 and 3 which are likely to be attributable to the relatively young age of our study population (i.e., baseline age of Crew 450 study was 16-20 years) compared to the older age of the study population of the Explore study (i.e., baseline age was 16 to more than 40 years[6]), attributable to differences in the design of survey questions and the different definition of a partnership in our model compared to a main partnership in the Explore study[6].

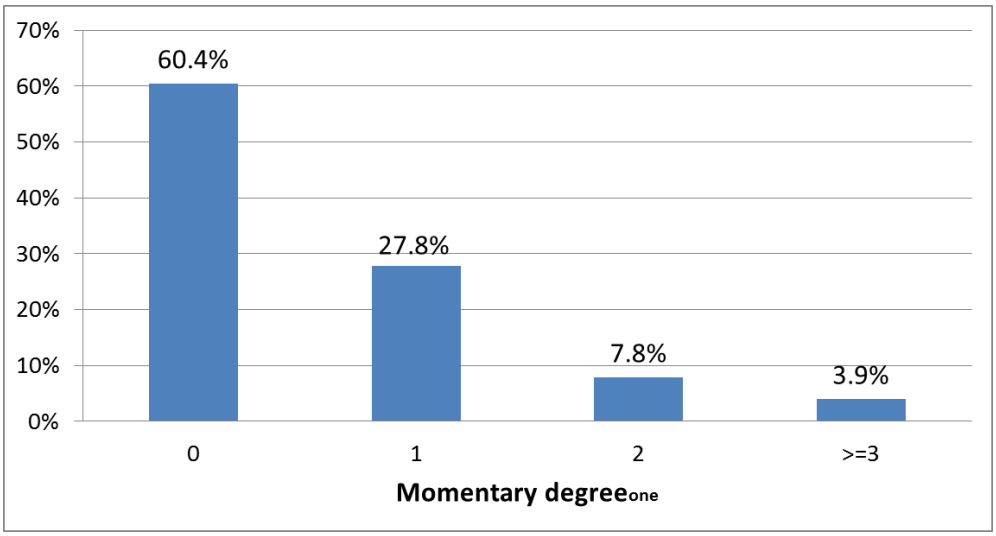


Figure 1: Momentary degree distribution of momentary degree$_{one}$ of n=421YMSM at T1 and T2. For a definition of partnership see discussion in SDC 2.2.2.1.

### 2.2.1.2 Number of concurrent partnerships when forming an outside or within-partnership:

### Momentary degree$_{part}$

Using data about the length of partnerships and partnership concurrency behavior discussed in the in-depth interviews we estimated the number of ongoing partnerships for each YMSM at the time of partnership formation (i.e., formation of an outside-partnership or within-partnership during the six months before the interview). For a definition of an ongoing partnership see section SDC 2.2.2.1. In the in-depth interviews YMSM could name at most up to three of their last partnerships. Thus, we assume a YMSM can only have either 0, 1 or 2 or more simultaneously ongoing partnerships when forming a new partnership and therefore the estimated momentary degree$_{part}$ differs from the estimated momentary degree$_{one}$ at the interview. Figure 2 shows the momentary degree$_{part}$, i.e. the number of ongoing concurrent partnerships of a YMSM when forming a new outside- or within-partnership.

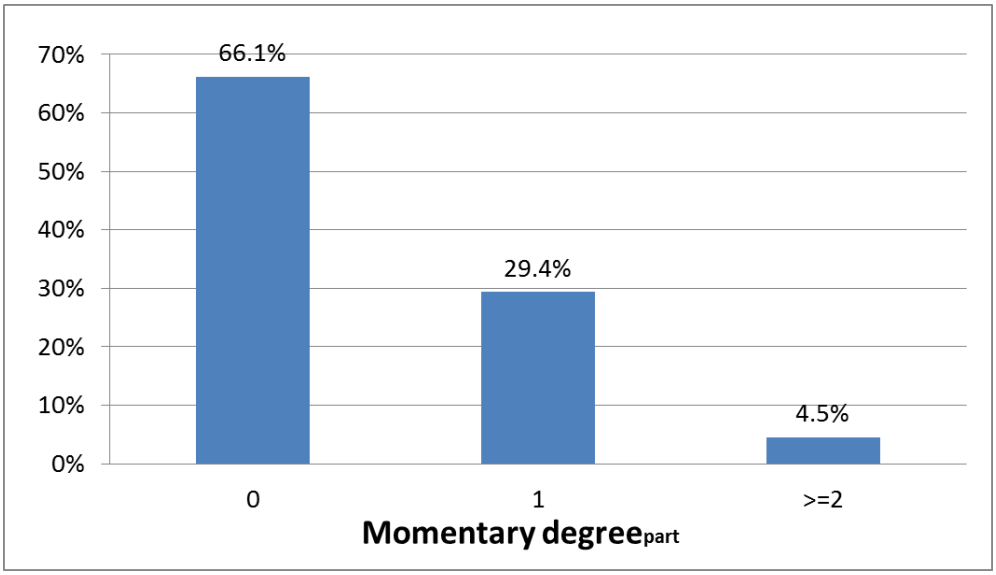


Figure 2: Number of concurrent partnerships momentart degree$_{part}$ when forming a new outside- or within-partnership during the last six months before the interview at T1 and T2 (n=889 newly formed partnerships). For a definition of an ongoing partnership see section SDC 2.2.2.1.

# 2.3 One-night-partnerships

## 2.3.1 Formation rate of one-night-partnerships: Regression analysis

Using a multivariate negative-binomial regression model we determined the formation rate of one-night-partnerships for each YMSM per time-step. Independent variables of the regression model were individual attributes (see Table 1) and estimated individual attributes (see Table 4). The outcome variable was the number of one-night-partnerships for each YMSM throughout the last six months before the interview.

### 2.3.1.1 Number of one-night-partnerships

We estimated the number of one-night-partnerships for each YMSM during the 6 months before the interview using the total number of sex-partners and the number of one-night-partnerships mentioned in the in-depth interviews. If a YMSM reported having more than 3 sex-partners (in-depth interviews were only performed on the last three) we assumed these ‘excessive’ sex-partners to be sex-partners of one-night-partnerships and thus calculated the number of one-night-partnerships for each YMSM during the

last 6 months before the interview as the sum of excessive sex-partners plus one-night-partnerships mentioned in the in-depth interviews. The following facts support this assumption. 39% of all study participants reported more than 3 sex-partners during the last six months and data about the partnerships with 'excessive' sex-partners show a high similarity to the one-night-partnerships of the in-depth survey: 32.1% of these were oral-sex only compared to 25.6% of the one-night-partnerships in the in-depth interviews whereas only 9.2% and 14.2% of partnerships mentioned in the in-depth interviews were classified as oral-only. Additionally, only 7.6% of all YMSM reporting more than three sex-partners did exclusively report one-night-partnerships in the in-depth interviews. Further, we assume sexual intercourse in partnerships to happen more frequently on average than one-night-partnerships which implies that study participants named most of their partnerships in the in-depth survey. Figure 3 shows the distribution of the number of one-night-partnerships during the last six months for n=421 YMSM at T1 and T2.

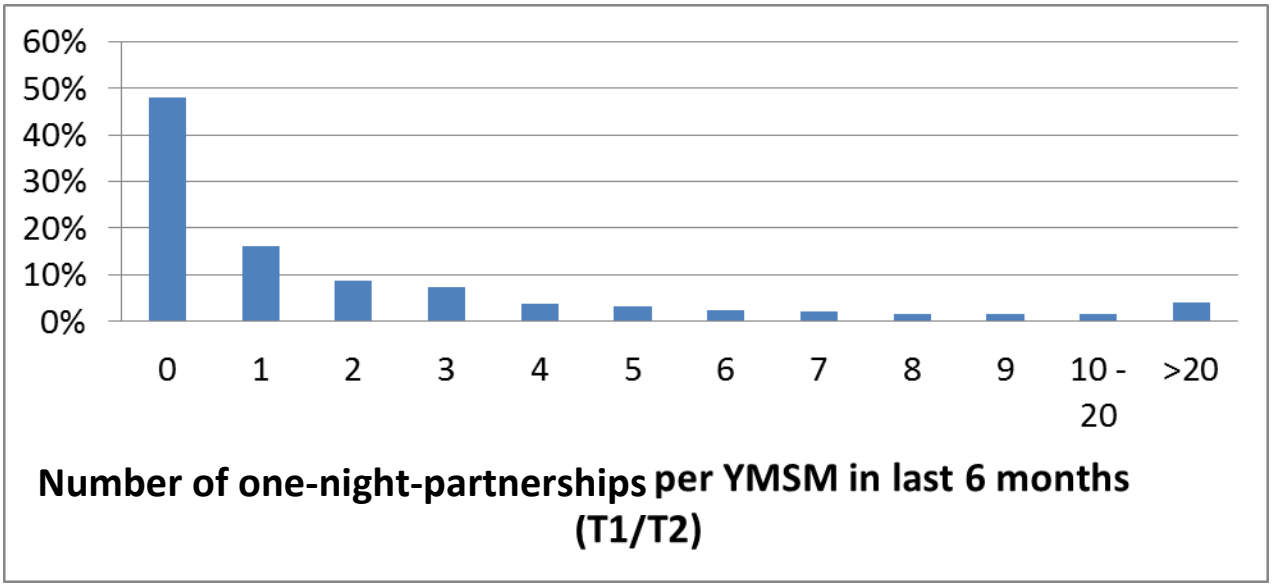


Figure 3: Distribution of number of one-night-partnerships during last six months for n=421 YMSM at in six months before interview at T1 and T2.

### 2.3.1.2 Results

We fit a multivariate negative-binomial regression model to determine the formation rate of one-night-partnerships for each YMSM per time-step using a stepwise elimination of variables approach. All

analysis were conducted using R version 3.0.3[7]. Table 5 shows the average multivariate coefficient estimates for significant variables and the corresponding p-values for the final negative-binomial regression model.

| **Covariate** | **Estimate** | **p-value** |
|---|---|---|
| **(Intercept)** | 0.2865 | p<0.01 |
| **Momentary degree$_{one}$_2** | 0.6895 | p<0.01 |
| **Momentary degree$_{one}$_3** | 1.4854 | p<0.001 |
| **druguse*** | 0.4074 | p<0.01 |
| **alcuse*** | 0.4281 | p<0.01 |
| **age*** | -0.1562 | p<0.01 |
| **raceOther** | 0.7558 | p<0.001 |
| **F** | 0.6072 | p<0.02 |

Table 5: Average multivariate coefficient estimates and p-values of the negative-binomial regression model on number of one-night-partnerships during six months before the interview at T1 and T2 averaged over ten realizations for desired sex-frequency F (n=766). All parameters are significant, i.e. p<0.05 for all ten realizations of the desired sex-frequency. Intercept: Black, Latino and White race, age 19.19, momentary degree$_{one}$ 0, no alcohol and drug usage. Interpretation of variable age: the younger the study participant, the more one-night-partnerships he has. Average value of variance adjustment parameter theta over ten realizations of desired sex-frequency F is 0.3559.

If the variables "desired sex-role R" or the "desired sex-frequency F" are significant variables in a regression model we always show the results of these regression models as average coefficients and average p-values over ten realizations of R and F. For a detailed discussion about the estimation of R and F see section SDC 2.4.

Using the coefficient estimates of the negative-binomial regression model we determine the number of one-night-partnerships $n_{out,i}$ of YMSM i, i.e.

$$n_{out,i} \sim NB(f(\mu_i), \theta) \qquad (1)$$

where

$$f(\mu_i) = \frac{\frac{\mu_i}{\theta}}{1+\frac{\mu_i}{\theta}} \quad (2)$$

$\mu_i$: mean of YMSM i forming one-night-partnerships.

$$\mu_i = \exp(\sum_j c_{int} + c_{one,j}\, x_{one,ij})\, \frac{d_I}{d_T} \quad (3)$$

$c_{int}$: intercept of the negative-binomial regression model for one-night-partnerships, see Table 5

$c_{one,j}$: coefficient of negative-binomial regression model for one-night-partnerships of ego-attribute j, see Table 2

$x_{one,ij} = \begin{cases} 1, & if\ YMSM\ i\ has\ ego-attribute\ j \\ 0, & otherwise \end{cases}$.

$d_T$: duration of simulation time-step in days

$d_I$: average time between interviews in days, $d_I = 177.4\ days$.

$\theta$: overdispersion parameter of negative-binomial regression model for one-night-partnerships, see Table 5.

## 2.3.2 One-night-partnership attributes

Once a one-night-partnership has been formed the one-night-partnership attributes sex of the partner, race of the partner, oral-sex only, sex-role and sex-role versatility, HIV status disclosure as well as the usage of condoms will be determined in sequential order (i.e., first we determine the sex of the one-night partner, then the race of the one-night partner using the sex of the one-night partner, etc.). We did not stratify one-night-partnerships by age of the male partner because data about the age of the excessive one-night-partnership partners was not available (see also discussion in SDC 2.2.3.1.1). This order was determined based upon expert opinion and the analysis of the correlation of these attributes. We used data of n=411 one-night-partnerships mentioned in the in-depth interviews at T1 and T2 for our analyses and the parameterization of the one-night-partnerships in our model.

### 2.3.2.1 Sex of the one-night partner

Dependent on the sexual orientation YMSM i will form a one-night-partnership with a woman with probability $p_{i,one}$. See section SDC 2.4 for details on sex-role and sexual orientation.

### 2.3.2.2 Race of the one-night partner

Given the sex of the one-night partner we determine the race of the one-night partner by use of the race mixing matrices shown in Table 6 and Table 7.

| Race mixing | Black (partner) | Latino (partner) | White (partner) | Other (partner) | n |
|---|---|---|---|---|---|
| Black | 77.91% | 6.75% | 27.98% | 6.75% | 162 |
| Latino | 10.00% | 58.75% | 27.50% | 3.75% | 80 |
| White | 8.42% | 18.95% | 64.21% | 8.42% | 95 |
| Other | 29.63% | 14.81% | 44.44% | 11.11% | 54 |

Table 6: Racial mixing in male-male one-night-partnerships (n=391).

| Race mixing | Black (partner) | Latino (partner) | White (partner) | Other (partner) | n |
|---|---|---|---|---|---|
| Black | 64.55% | 9.09% | 14.55 | 11.82% | 110 |
| Latino | 13.33% | 46.67% | 13.33% | 26.67% | 15 |
| White | 0% | 18.18% | 63.64% | 18.18% | 11 |
| Other | 50.00% | 8.33% | 8.33% | 33.33% | 12 |

Table 7: Racial mixing in female-male one-night-partnerships and female-male partnerships (n=148). Because of the low total number of female-male one-night-partnerships and partnerships compared to male-male one-night-partnerships and male-male partnerships named in the in-depth interviews as well as similar racial mixing patterns in female-male one-night-partnerships and female-male outside-partnerships we assume the same racial mixing patterns for female-male one-night-partnerships and female-male outside-partnerships.

### 2.3.2.3 Oral-sex only one-night-partnerships

Next, we determine whether the one-night-partnership is oral-sex only using the conditional probabilities P(oral-only|female-male)=0.15 (n=20) and P(oral-only|male-male)=0.2737 (n=391).

### 2.3.2.4 Sex-role behavior in one-night-partnerships

#### *Male-male one-night-partnerships*

For a discussion about the estimation of the sex-role $R_{i,one}$ of a YMSM in an one-night-partnership see section SDC 2.4. One-night-partnerships are assigned the one-night-partnership attribute intra-event sex-role versatility, i.e. whether a YMSM is both insertive and receptive during an one-night-partnership or just insertive or receptive. (Intra-event) sex-role versatility is considered important to HIV transmission due to the different transmission risks associated with the insertive and receptive sex-role [8]. We estimated the likelihood that a study participant will be intra-event sex-role versatile in a male-male one-night-partnership to be $P_{one}$(sex-role versatile|male-male)= 0.2254 (n=284). In comparison, Goodreau et al.[4] discuss a parameter estimate of the NHBS-08-SF study where in case of casual sex 32% of study participants where estimated to be intra-event sex-role versatile. In case a YMSM is predicted to be sex-role exclusive in the one-night-partnership, i.e., he is exclusively either insertive or receptive ($R_{i,one} \in \{1,0\}$), he can't be sex-role versatile.

#### *Female-male one-night-partnerships*

Our analysis of sex-role versatility in one-night-partnerships showed that YMSM also change sex position during a female-male one-night-partnership, i.e. YMSM can have exclusively penile-vaginal or both anal-insertive and penile-vaginal sex with a woman during an one-night-partnership. We estimated the probability of a YMSM in a female-male one-night-partnership to be sex-role versatile to be P(versatile|female-male)=0.7059 (n=17). Friedman et al.[9] estimates MSM to be 1.6 times more likely to have unprotected anal intercourse with women compared to men which have exclusively sex with women and Herbenick at al.[10] estimates the fraction of 20 to 24 year old women who had penile-anal intercourse in the past year to be 23.4%. Of these, 79.9% had penile-vaginal intercourse in the past year.

### 2.3.2.5 Knowledge about HIV status of sex-partner and HIV status disclosure

In the in-depth interviews, YMSM were asked about the HIV-status of their sex-partner and how they obtained the information about the HIV status of their sex-partner, i.e. partnership variables hivstatusk and hivstatusk_type (for a definition of these variables see Table 2). We used this information to estimate for each one-night-partnership and each partnership whether the HIV status was disclosed through the partner. We stratified disclosure status by race of the ego.

Table 8 shows the percentages of one-night-partnerships stratified by ego-race where the partner disclosed the HIV status to the ego. Overall, HIV status was disclosed in 62.4% of all sex-partnerships (n=1651). This percentage is comparable to the findings of the Chicago Department of Public Health HIV 2011 MSM surveillance report[11] where 63.8% of all Black MSM reported knowing the HIV status of their most recent sex-partner as well as 63.2% of all White MSM and 56.3% of all Latino MSM. For casual sex partners, Serovich and Mosak[12] observed 37% of HIV positive MSM disclosing their HIV status to all of their sex partners and Perry et al.[13] reported HIV positive individuals to be less likely to disclose HIV status when having casual sex.

| Race of study participant | Percentage of one-night-partnerships disclosed |
|---|---|
| Black | 34.1% (n=176) |
| Latino | 41.46% (n=82) |
| White | 50.0% (n=98) |
| Other | 47.28% (n=55) |

Table 8: Percentage of one-night-partnerships where sex-partner was estimated to have disclosed the HIV status to the study participant (n=411).

### 2.3.2.6 Unprotected anal intercourse (UAI)/use of condoms in one-night-partnerships

To determine the likelihood of a YMSM to have UAI in an one-night-partnership, i.e. no usage of condoms, we used a multivariate logistic regression model with a stepwise elimination of variables approach and the outcome no condom usage. Table 9 shows the results of the final multivariate logistic regression model.

| | **Estimate** | **P value** |
|---|---|---|
| **(Intercept)** | -1.0109 | p<0.001 |
| **SES 1 Low** | 0.7108 | p<0.01 |
| **CT positive** | 1.0039 | p<0.05 |

Table 9: Multivariate coefficient estimates of the logistic regression model on the outcome unprotected sex in one-night-partnerships (n=301). Oral-sex only one-night-partnerships were excluded from analysis. Intercept: high SES 1, no Chlamydia infection. Interpretation: If low SES 1and/or infected with Chlamydia likelihood higher to engage in unprotected sexual intercourse.

Thus, the probability of a YMSM to have unprotected sex in a one-night-partnership is

$$P(UAI|one-night) = \frac{\exp(intercept + c_{ses1,one-night} x_{i,ses1} + c_{CT,one-night} x_{CT,i})}{1 + \exp(intercept + c_{ses1,one-night} x_{i,ses1} + c_{CT,one-night} x_{CT,i})} \quad (4)$$

where

intercept: intercept of the logistic-regression model for unprotected sex (UAI) in a one-night-partnership

$c_{ses1,one-night}$: regression coefficient of ego-attribute socio economic status 1 for above regression model, see Table 4

$c_{CT,one-night}$: regression coefficient of ego-attribute infection with CT

$$x_{i,ses1} = \begin{cases} 1, if\ YMSM\ i\ has\ low\ ses1 \\ 0, otherwise \end{cases}$$

$$x_{CT,i} = \begin{cases} 1, if\ YMSM\ i\ has\ CT \\ 0,\ otherwise \end{cases}.$$

# 2.4 Partnerships outside the study population

## 2.4.1 Formation rate of outside-partnerships: Regression analysis

Using a multivariate Poisson regression model we determined the formation rate of outside-partnerships for each YMSM per time-step. Independent variables of the regression model were individual attributes (see Table 1) and estimated individual attributes (see Table 4). Outcome variable was the number of newly formed outside-partnerships for each YMSM throughout the last six months before the interview. We defined an outside-partnership as a partnership of a YMSM with either a female partner or an older MSM partner lasting longer than one day (see Table 3).

### 2.4.1.1 Number of new outside-partnerships

Using data about the length of named partnerships in the in-depth interview we determined whether an outside-partnership was newly formed within six months before the interviews at T1 and T2. Figure 4 shows the distribution of the number of newly formed outside-partnerships of n=421 YMSM within six months before the interview at T1 and T2.

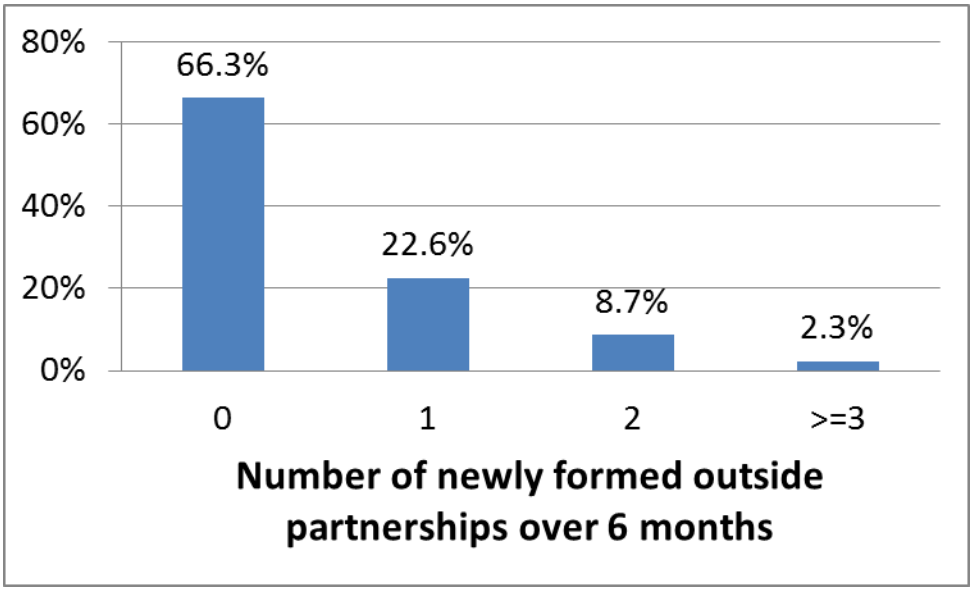

Figure 4: Distribution of newly formed outside-partnerships of n=421 YMSM in six months before the interviews at T1 and T2. YMSM were only asked about the last three sex-partners in the in-depth interviews. Thus the number of newly formed partnerships in the survey is limited to 3.

### 2.4.1.2 Results

We fit a Poisson regression model with stepwise elimination of variables to determine the number of newly formed outside-partnerships during the six months before the interview using data of T1 and T2. Table 10 shows the multivariate coefficient estimates for significant variables and the corresponding p-values of the final Poisson regression model.

| Covariate | Estimate | p-value |
|---|---|---|
| **(Intercept)** | -1.05129 | p<0.001 |
| **alcuse*** | 0.28373 | p<0.01 |
| **age*** | 0.19826 | p<0.001 |
| **bisexual** | 0.33944 | p<0.01 |
| **mostly hetero** | 0.60102 | p<0.05 |

Table 10: Multivariate coefficient estimates of the Poisson regression model on the outcome number of newly formed outside-partnerships within six months before an interview (n=421 YMSM). Intercept: no alcohol usage, age 19.19, sexual orientation gay, mostly gay, hetero or other.

Using the regression coefficients of the Poisson model we determine the formation rate of new outside-partnerships $\lambda_{out,i}$ for each YMSM i per time-step as follows

$$\lambda_{out,i} = \exp\left(intercept + \sum_j c_{out,j} x_{out,ij}\right) \frac{d_T}{d_I} \qquad (5)$$

where

intercept: intercept of Poisson regression model for outside-partnerships

$c_{out,j}$: coefficient of Poisson regression model for outside-partnerships of ego-attribute j, see Table 10

$$x_{out,ij} = \begin{cases} 1, & if\ YMSM\ i\ has\ ego-attribute\ j \\ 0, & otherwise \end{cases}$$

$d_T$: duration ofc simulation time-step in days

$d_I$: average time between interviews in days, $d_I = 177.4\ days$.

## 2.4.2 Outside-partnerships attributes

Similar to one-night-partnerships we determine the outside-partnership attributes after the outside-partnership has formed. First, the outside-partnership attributes sex of the partner, race of the partner, oral-sex only, sex-role and sex-role versatility, sex-frequency and HIV status disclosure will be determined in sequential order using previously determined attributes (e.g., the likelihood of the partnership to be oral-sex only depends on the race and the sex of the partner). Second, the seriousness of the partnership, mean length and the usage of condoms will be determined in sequential order using regression analysis. As in case of one-night-partnerships this order was determined based upon expert opinion and the statistical analysis of the correlation of these attributes. We used data of n=512 outside-partnerships named in the in-depth interviews at T1 and T2 for our analyses and the parameterization of our model.

### 2.4.2.1 Sex of the outside partner

Dependent on the sexual orientation YMSM i will form an outside relationship with a woman with probability $p_{i,out}$. See section SDC 2.4 on sex-role and sexual orientation for details.

### 2.4.2.2 Race of the outside partner

Given the sex of the partner in an outside relationship we determine the race of the partner using the following race mixing matrices shown in Table 11 and Table 12

| Race mixing | Black (partner) | Latino (partner) | White (partner) | Other (partner) | n |
|---|---|---|---|---|---|
| **Black** | 81.36% | 7.91% | 4.52% | 6.21% | 179 |
| **Latino** | 8.86% | 46.84% | 32.91% | 11.39% | 79 |
| **White** | 12.22% | 4.44% | 75.56% | 7.75% | 91 |

| | | | | | |
|---|---|---|---|---|---|
| **Other** | 51.43% | 22.86% | 22.86% | 2.86% | 35 |

Table 11: Racial mixing in male-male outside-partnerships (n=384).

| Race mixing | Black (partner) | Latino (partner) | White (partner) | Other (partner) | n |
|---|---|---|---|---|---|
| **Black** | 64.55% | 9.09% | 14.55% | 11.82% | 110 |
| **Latino** | 13.33% | 46.67% | 13.33% | 26.67% | 15 |
| **White** | 0% | 18.18% | 63.64% | 18.18% | 11 |
| **Other** | 50.00% | 8.33% | 8.33% | 33.33% | 12 |

Table 12: Racial mixing in female-male outside-partnerships and female-male one-night-partnerships (n=148). Because of the low total number of female-male one-night-partnerships and female-male partnerships compared to male-male one-night-partnerships and male-male partnerships as well as similar racial mixing patterns in female-male one-night-partnerships and female-male outside-partnerships we assume the same racial mixing patterns for female-male one-night-partnerships and female-male partnerships.

### 2.4.2.3 Oral-sex only outside-partnerships

In female-male outside-partnerships the probability of an oral-only partnership is $P_{out}$(oral-only|female-male)=0.0078 (race mixing covariates were not significant in a regression analysis, n=128). For male-male outside-partnerships we determine the likelihood of a partnership to be oral-sex only using the regression coefficients of a multivariate Logistic regression model on the outcome oral-sex only and the race mix as independent variables. Table 13 shows the results of the final multivariate Logistic regression model after applying stepwise elimination of variables approach.

| Covariate | Estimate | p-value |
|---|---|---|
| **(Intercept)** | -2.1383 | p<0.001 |
| **White-white** | 0.6959 | p<0.05 |

Table 13: Multivariate coefficient estimates of the Logistic regression model on the outcome oral-sex only male-male outside-partnership (n=384). Intercept– no white-white outside-partnership.

We determine the probability of a newly formed partnership to be oral-sex only using the coefficients shown in Table 13 and the approach discussed in section 2.3.2.6 and equation (4).

### 2.4.2.4 Sex-role behavior and sex-frequency in outside-partnerships

#### *Sex-role behavior in outside-partnerships*

For a discussion about how to determine the sex-role $R_{i,out}$ of a YMSM in a male-male outside-partnership see section SDC 2.4. Similar to one-night-partnerships, the YMSM can engage in intra-event sex-role versatility if the YMSM is not sex-role exclusive, i.e. $R_{i,out} \in (0,1)$. The probability to engage in intra-event sex-role versatility in each sex-act is $P_{out}$(sex-role versatile|male-male)=$P_{one}$(sex-role versatile|male-male)=0.2254. We use the intra-event sex-role versatility estimate of one-night-partnerships because we could not estimate the likelihood of sex-role versatility specifically for outside-partnerships. Our estimate of $P_{out}$(sex-role versatile|male-male) is comparable to the estimates of the NHBS-08-SF study where the fraction of intra-event sex-role versatile sex in main-partnerships is estimated to be 0.22 [2]. Similar to the intra-event sex-role versatility in male-male outside-partnerships, we assume the likelihood of intra-event anal/vaginal versatility for outside female-male partnerships to be the same as in case of female-male one-night-partnerships, i.e. $P_{out}$(versatile|female-male)= $P_{one}$(versatile|female-male)=0.7059.

#### *Sex-frequency in outside-partnerships*

The sex-frequency $F_{i,out}$ of YMSM i denotes the number of sex-acts per time step in an outside-partnership. For details about the sex-frequency $F_{i,out}$ see section SDC 2.4.

### 2.4.2.5 HIV status disclosure in outside-partnerships

Similar to one-night-partnerships, we determined whether the partner of the YMSM disclosed their HIV status in an outside-partnership using the partnership attributes hivstatusk and hivstatusk_type (see Table 2). Table 14 shows coefficients and p-values of significant variables of the final multivariate Logistics regression model on the outcome disclosure of HIV status after applying stepwise elimination of variables

approach. We determined independent variables of this regression model to be the racemix combinations of outside-partnerships.

| Covariate | Estimate | P value |
|---|---|---|
| **Intercept** | 0.5306 | p<0.001 |
| **BL** | 1.0788 | p<0.05 |
| **BW** | 0.9734 | p<0.05 |
| **LO** | 1.2611 | p<0.05 |
| **WW** | 0.6792 | p<0.05 |

Table 14: Multivariate logistic regression analysis on the outcome disclosure of HIV status in an outside-partnership, n=509. Intercept: racial mixing combinations BB, BO, LL, LW and WO.

We determine the probability that partners of YMSM disclose their HIV status in newly formed outside-partnerships using the coefficients shown in Table 14 and the approach discussed in section SDC 2.2.3.2.6 and equation (4). In absence of data about the timing of HIV disclosure in YMSM partnerships we assume the disclosure of the HIV status to happen instantaneously after the formation of the partnership and in case the partners get to know about their positive HIV status throughout the partnership.

### 2.4.2.6 Seriousness of outside-partnership

We use multivariate Logistic regression modeling to determine if an outside-partnership is serious. In the in-depth interviews, YMSM named whether they considered a partnership to be serious or not. Independent variables of the regression model also include the partnership attributes sex of partner, race of partner, oral-sex only, sex-role behavior and sex-frequency as well as HIV disclosure status. Table 15 shows coefficients and p-values of significant variables of the final multivariate Logistics regression model after applying a stepwise elimination of variables approach.

| Covariate | Estimate | P value |
|---|---|---|
| **(Intercept)** | -0.5871 | p<0.001 |
| **Momentary degree$_{part}$>=1** | -1.0638 | p<0.001 |

| | | |
|---|---|---|
| **LW** | -1.3001 | p<0.02 |
| **BO** | 0.8784 | p<0.02 |
| **SES 1 low** | 0.4918 | p<0.02 |
| **female** | 1.0258 | p<0.001 |
| **disclosed(non)/hivstatusp_d/p** | -0.4539 | p<0.05 |

Table 15: Multivariate coefficient estimates of the Logistic regression model on the outcome seriousness of an outside-partnership (n=509). Variable momentary degree$_{part}$>=1denotes the aggregated variable of momentary degree$_{part}$ 1 and momentary degree$_{part}$ >=2. Variables where aggregated because momentary degree$_{part}$ >=2 was not significant but had similar coefficient compared to the significant covariate momentary degree$_{part}$ 1. Variable disclosed(non)/hivstatusp_d/p aggregates both effects of disclosure status and sero-sorting, i.e. either not knowing about the partners HIV status and ego assumes he is HIV negative or knowledge of at least one HIV infection in partnership (i.e., ego knows that at least he is HIV positive) and thus to be less likely to form a new partnership. Variables were grouped because of similar effects on outcome. Intercept: no ongoing relationship at the time of outside-partnership formation, all racemixes except of Latino-White and Black-Other, high SES 1, male-male outside-partnership, and ego knows HIV status of partner and assumes he is HIV negative.

We determine the probability that a newly formed outside-partnership is serious using the coefficients shown in Table 15 and the approach discussed in section SDC 2.2.3.2.6 and equation (4). Due to lack of data we assume the seriousness of an outside-partnership to be determined right after partnership formation although we know that this partnership characteristic might be established throughout the partnership. Seriousness of a partnership can change in case one or both partners get to know and/or disclose their eventual HIV-positive status.

### 2.4.2.7 Mean length $\mu_{out,i}$ of outside-partnerships

We assume the length of a partnership to follow an exponential distribution [14] (i.e., a geometric distribution in case of discrete-time; for details see discussion in section SDC 2.3). Thus, the probability of a partnership to dissolve in the next time step is constant over time (it can of course vary from partnership to partnership but not on how long the partnership has been ongoing). Based upon this assumption the actual length of an outside-partnership is

$$T_{i,k} \sim \mu_{out,i,k}\exp(1) \qquad (6)$$

where

$T_{i,k}$: actual length of outside-partnership k of individual i

$\mu_{out,i}$: mean length of outside-partnership k of individual i (see discussion below)

$\exp(1)$: exponential random variable with parameter $\lambda = 1$

We determine the mean length of an outside-partnership $\mu_{out,i}$ using a multivariate linear regression model with a stepwise elimination of variables approach. Outcome variable is the length of an outside-partnership minus 15 days (length is categorized variable in Crew 450 study with shortest length being 15 days of a partnership) and independent variables also include the seriousness of an outside-partnership in addition to the independent variables used in the regression model to determine the seriousness of the partnership. Table 16 shows the results of the final multivariate linear regression model.

| **Covariates** | **Estimate** | **p-value** |
|---|---|---|
| **Intercept** | 137.4 | p<0.001 |
| **Momentarydegree$_{part}$>=1** | -76.18 | p<0.01 |
| **age*** | 30.34 | p<0.01 |
| **female** | 115 | p<0.001 |
| **seriousYes** | 154.27 | p<0.001 |
| **BO** | -86.02 | p<0.05 |
| **LL** | 89.8 | p<0.05 |

Table 16: Multivariate coefficient estimates of the linear regression model on the outcome length of outside-partnerships (n=509). Intercept attributes: no ongoing partnership at time of partnership formation, age 19.19, male-male partnership, not serious, not Black-Other or Latino-Latino. Residual standard error is 272.4 with 504 degrees of freedom.

Using the coefficient estimates of the linear regression model shown in Table 16 we determine the mean length of a newly formed outside-partnership as follows

$$mean\ length = 15 days + \exp(intercept\ + \sum_{j \in \{pship\ and\ ego\ attributes\}} c_j\ x_{ij}) \qquad (7)$$

where

intercept: intercept of the linear regression model for length of an outside-partnership

$c_j$: regression coefficient of covariates specified in Table 16

$$x_{ij} = \begin{cases} 1, if\ YMSM\ i\ has\ attribute\ j\ or\ partnership\ YMSM\ i\ is\ in\ has\ attribute\ j \\ 0, otherwise \end{cases}.$$

#### 2.4.2.8 Unprotected anal intercourse/use of condoms in outside-partnerships

To determine the likelihood of UAI per sex-act in an outside-partnership we use a multivariate Logistic regression model with a stepwise elimination of variables approach. Outcome variable is UAI (i.e., no condom usage) in an outside-partnership and independent variables include all ego-attributes and partnership attributes discussed above. Table 17 shows the results of the Logistic regression model.

| Covariates | Estimate | P value |
|---|---|---|
| **Intercept** | -1.0197 | p<0.001 |
| **LL** | 0.7031 | p<0.05 |
| **LW** | 0.9911 | p<0.01 |
| **LO** | 1.3565 | p<0.01 |
| **druguse*** | 0.4006 | p<0.05 |
| **female** | -1.2955 | p<0.001 |
| **seriousYes** | 0.9626 | p<0.001 |

Table 17: Multivariate coefficient estimates of the Logistic regression model on the outcome UAI (no use of condoms) of outside-partnerships (n=509). Intercept-attributes: no drugusage, male-male partnership, not serious, not Latino-Latino, no Latino-White or Latino-Other racemix.

We calculate the probability of UAI per sex-act in a newly formed outside-partnership using the Logistic regression coefficient estimates shown in Table 14 and the approach discussed in section SDC 2.2.3.2.6 and equation (4).

# 2.5 Partnerships within the study population

## 2.5.1 Formation of within-partnerships

In our simulation model we represent within-partnerships, i.e. partnerships among YMSM, using a network model. In this network model, two YMSM form a tie $e_{ij}$ with probability $p_{ij,t}$ in time step *t*. Once the tie is formed, the partnership attributes such as the length and seriousness of the tie will be determined. A tie in the network is assumed to exist until the partnership dissolves.

We model the formation of a tie $e_{ij}$ in time step *t* as a Bernoulli random variable, i.e. the probability that a tie $e_{ij}$ will form in time step t is

$$P_t(e_{ij}=1) = 1 - P(e_{ij}=0) = \begin{cases} p_{ij,t}\text{, if } e_{ij}=0 \text{ at t}-1 \\ 0, \ otherwise \end{cases} \quad (8)$$

where

$p_{ij,t}$: probability of tie formation between *i* and *j* in time step *t*.

## 2.5.2 Formation probability $p_{ij,t}$ of within-partnerships

We determine the tie formation probability $p_{ij,t}$ using a multivariate Logistic regression model. The outcome variable of this regression model is $e_{ij}$, i.e. whether a within-partnership (i.e., tie) was newly formed ($e_{ij}=1$) or not ($e_{ij}=0$) during the six months before the interview. Using the length of named within-partnerships in the in-depth interview we determined that 533 within-partnerships were newly formed in the six months before T1 and T2. Next, we created a list of all potential ties among the n=421 YMSM, i.e. all combinations of age-difference, HIV status of the partnership, racial mixing, STI mixing, drug and alcohol usage of both partners, partnership concurrency status of both partners and desired sex-roles of both partners. Because the formation probability $p_{ij,t}$ depends on the number of potential ties and thus the number of YMSM in the population (here n=421 for the parameterization of the within partnership formation model) we describe in the following section SDC 2.2.5.3 a procedure to scale the formation probability $p_{ij,t}$ in case of population sizes different from n=421 YMSM. A potential tie is a tie where two YMSM do not have the same desired exclusive sex-role, e.g. a tie cannot exist if both YMSM have the desired sex-role to be exclusively receptive. We then matched the empirical 533 newly formed within-partnerships of the Crew 450 study with the created list of potential ties based on the above discussed partnership attributes. Thus, $e_{ij}=1$ if a tie could be matched otherwise $e_{ij}=0$. Since it is possible for one of the 533 ties to match multiple potential ties, we used stochastic matching. This

involved randomizing the order of the potential ties allocating each empirical tie to the first available potential match with a 0.25 probability. We created ten different realizations, i.e. ten different tie lists and matches, because of the stochastic matching procedure as well as using ten different realizations of the desired sex-role of each YMSM (for details on the realizations of the desired sex-role see section SDC 2.4). Independent variables of the Logistic regression model are partnership attributes $x_{ij}$ of all potential ties $e_{ij}$ among the 421 YMSM and combinations of individual attributes of the two YMSM i and j.

Table 18 shows the average multivariate coefficient estimates and average p-values of significant variables of the final Logistic regression model over ten realizations of the stochastic tie matching after applying stepwise elimination of variables.

| **Covariates** | **Average Estimate*** | **P value** |
|---|---|---|
| **Intercept** | -4.20434 | p<0.001 |
| **page(abs)** | -0.1535 | p<0.001 |
| **avgage** | -0.14863 | p<0.01 |
| **hivstatusp_p+hivstatusp_d** | -0.9887 | p<0.001 |
| **BL** | -2.21078 | p<0.001 |
| **BW** | -1.6542 | p<0.001 |
| **BO** | -0.79469 | p<0.001 |
| **LL** | 0.758832 | p<0.001 |
| **LW** | -0.48766 | p<0.01 |
| **LO** | -0.56515 | p<0.05 |
| **WW** | 0.789174 | p<0.001 |
| **WO** | -0.60813 | p<0.05 |
| **partcon_p+partcon_d** | -0.87654 | p<0.05 |

Table 18: Multivariate coefficient estimates of the Logistic regression model on the outcome newly formed partnership in T1 and T2 over all feasible partnerships among 421 YMSM (n>70,000). Variable page(abs) denotes the absolute age-difference between the two YMSM; variable avgage denotes the average age of both YMSM; hivstatusp_p (seropositive couple) and hivstatusp_d (serodiscordant couple) were combined since hivstatusp_d was not statistically significant but coefficient was also negative such as the coefficient of hivstatusp_p and thus combined; the combined variable (partcon_p+partcon_d), i.e. at least one of the partners has one ongoing/concurrent partnership, was included because it was statistically significant (i.e., p<0.05) in 4 out of 10 realizations. Intercept attributes: age difference 0, average age of both partners 19.29, HIV seronegative partnership, Black-Black or Other-Other partnership, both partners do not have concurrent partnerships at time of partnership formation.

Using these coefficient estimates we can calculate the probability $\boldsymbol{p_{ij,t}}$

$$\boldsymbol{p_{ij,t}} = \left(\frac{\exp(\boldsymbol{intercept} + \sum_{j\in\{\boldsymbol{pship\ and\ ego\ attributes}\}} \boldsymbol{c_{ij,t}x_{ij,t}})}{1+\exp(\boldsymbol{intercept} + \sum_{j\in\{\boldsymbol{pship\ and\ ego\ attributes}\}} \boldsymbol{c_{ij,t}x_{ij,t}})}\right) \frac{\boldsymbol{d_T}}{\boldsymbol{d_{Overall}}} \qquad (9)$$

where

intercept: intercept of the Logistic regression model for the existence of a newly formed tie among all feasible ties of 421 YMSM specified in Table 18

$c_{ij,t,k}$: regression coefficient of partnership attribute k specified in Table 18

$$x_{ij,t,k} = \begin{cases} 1, if\ partnership\ of\ YMSM\ i\ and\ j\ at\ time\ t\ has\ attribute\ j \\ 0, otherwise \end{cases}$$

$d_{Overall}$: average time interval of partnership formation for 533 new partnerships over 421 YMSM, $d_{Overall}$=332 days. $d_{Overall}$ > 177.4 days because we regressed over partnerships newly formed in six months before T1 and in six months before T2; $d_{Overall}$< 354.8 days since not all 421 YMSM completed both T1 and T2.

## 2.5.3 Scaling of $p_{ij,t}$ for different population sizes

In section SDC 2.2.5.2 the within partnership formation probability per time-step $p_{ij,t}$ was determined by matching newly formed ties among the $n_e$=421 YMSM to a list of all potential ties among this empirical population size. Thus, the within partnership formation probabilities determined in section SDC 2.2.25.2 depend on the size of the empirical population and need to be scaled in case of the actual simulation because the size of the simulated YMSM population differs from the empirical population size (i.e., $n_s$=4484 YMSM at t=0 (see also SDC 4.1.1) vs. $n_e$=421 YMSM). The scaling factors $s_{x,y,t}$ will adjust the within partnership formation probability $p_{ij,t}$ between two YMSM i and j at time-step t such that the number of ties a YMSM stratified by race forms on average per time-step in the empirical population will be maintained in the simulated YMSM population. $s_{x,y,t}$ adjusts the changes of the population sizes stratified by race in the empirical vs. the simulated YMSM population and thus the scaled probabilities are

$$p_{ij,t,scaled} = \sqrt{s_{x,y,t}\ s_{y,x,t}}\, p_{ij,t} \qquad (10)$$

where

i,j: YMSM i and YMSM j engaging in a within partnership

t: time-step t in the simulation

$p_{ij,t}$: within tie formation probability between YMSM i and j at time-step t. Determined in SDC 2.2.5.2 using the empirical population of n=421 YMSM

$s_{x,y,t}$: scaling factor corresponding to the difference in the population sizes stratified by race and $p_{ij,t}$ at time-step t where x denotes the race of YMSM i and y denotes the race of YMSM j. $x, y \in \{Black, Latino, White, Other\}$. Note, that per definition $p_{ij,t}$ is symmetric, i.e. $p_{ij,t} = p_{ji,t}$ but not the scaling factors $s_{x,y}$, i.e. $s_{x,y} \neq s_{y,x}$ because of different sizes of the populations stratified by race.

The scaling factor $s_{x,y,t}$ is defined as

$$s_{x,y,t} = \left(\frac{n_{e,x}}{n_{s,x,t}}\right) \frac{n_{e,x} n_{e,y}}{n_{s,x,t} n_{s,y,t}} \qquad (11)$$

where

$n_{e,x}$: size of population of race x in empirical population. x denotes the race of the YMSM i. The population sizes of the empirical YMSM population of total size n=421 are: Blacks $n_{e,B}$=225, Latinos $n_{e,L}$=83, Whites $n_{e,W}$=76, Others $n_{e,O}$=37.

$n_{s,x}$: size of population of race x in simulated population at time-step t. x denotes the race of the YMSM i. The population sizes of the simulated YMSM population of total size n=4484 at time-step t=0 are: Blacks $n_{s,B}$=1575, Latinos $n_{s,L}$=1245, Whites $n_{s,W}$=1368, Others $n_{s,O}$=296.

Given the above population sizes of the empirical YMSM population and the simulated YMSM population stratified by race and applying (10) and (11), Table 19 shows the scaling factors $s_{x,y,t}$ and $s_{y,x,t}$ of the within partnership formation probabilities at time-step t=0.

| race-mix | $s_{x,y,0}$ | $s_{y,x,0}$ |
|---|---|---|
| **Black, Black** | 0.143 | 0.143 |
| **Black, Latino** | 0.067 | 0.143 |
| **Black, White** | 0.056 | 0.143 |
| **Black, Other** | 0.125 | 0.143 |
| **Latino, Latino** | 0.067 | 0.067 |
| **Latino, White** | 0.056 | 0.067 |
| **Latino, Other** | 0.125 | 0.067 |
| **White, White** | 0.056 | 0.056 |
| **White, Other** | 0.125 | 0.056 |

| | | |
|---|---|---|
| **Other, Other** | 0.125 | 0.125 |

Table 19: Scaling factors $s_{x,y,t}$ and $s_{y,x,t}$ of the within partnership formation probabilities at time-step t=0.

## 2.5.4 Within-partnership attributes

Similar to one-night-partnerships and outside-partnerships, we determine the within-partnership attributes after the formation of a within-partnership. First, the partnership attributes oral-sex only, sex-role, sex-frequency, and intra-event sex-role versatility and disclosure status are determined in sequential order using point estimates. Second, the within-partnerships attributes seriousness, mean length and UAI are determined using regression analysis. Similar to one-night-partnerships and outside-partnerships this order was determined based upon expert opinion and the statistical analysis of the correlation of these attributes. We used data of n=728 within-partnerships (i.e., ongoing and newly formed) named in the in-depth interviews at T1 and T2 for our analysis and the parameterization of our model.

### 2.5.4.1 Oral-sex only within-partnerships

Based on outcomes of a statistical analysis we assumed the attribute oral-sex only of a within-partnership to be dependent on the race-mix of the partnership. Table 20 shows the probability of a partnership to be oral-sex only dependent on the race-mix of the within-partnership.

| Race-mix within-partnership | Probability of partnership to be oral-sex only |
|---|---|
| **BB** | 0.1057 ($p<0.02$) |
| **BL** | 0.1429 |
| **BW** | 0.1429 |
| **BO** | 0.1429 |
| **LL** | 0.1429 |
| **LW** | 0.1429 |
| **LO** | 0.1429 |
| **WW** | 0.2840 ($p<0.001$) |
| **WO** | 0.1429 |
| **OO** | 0.1429 |

Table 20: Probabilities of partnerships within the study population to be oral-sex only in dependence of race-mix of the partnership. B-Black, L-Latino (n=728). W-White, O-Other. The overall probability of a partnership within the study population to be oral-sex only is 0.1429. BB and WW race-mixes had statistically significant different likelihoods of oral-sex only partnershsips compared to other race-mixes.

### 2.5.4.2 Sex-role behavior and sex-frequency in within-partnerships

#### *Sex-role behavior in within-partnerships*

In a partnership between YMSM i and YMSM j, i is assumed to take on the receptive sex-role with probability $R_{ij}$ and j with probability $1 - R_{ij}$. These probabilities are defined in section SDC 2.4. If $R_{ij} \in (0,1)$, i.e. both YMSM are not strictly sex-role exclusive in the partnership, both YMSM can be also intra-event sex-role versatile (i.e., changing sex-role within a sex-act). In absence of data about the likelihood of intra-event sex-role versatility in within-partnerships we assume the probability of sex-role versatility in a within-partnership to be the same as in a one-night-partnership, i.e. $P_{within}$(sex-role versatile|male-male)=$P_{one}$(sex-role versatile|male-male)=0.2254.

#### *Sex-frequency in within-partnerships*

The sex-frequency $F_{ij}$ denotes the number of sex-acts per time step in a within-partnership between YMSM i and j. For details about the sex-frequency $F_{ij}$ see section SDC 2.4.

### 2.5.4.3 HIV status disclosure in within-partnerships

We assume HIV status disclosure to happen in a within-partnership only if both YMSM disclose their HIV status to each other. Further, we assume that YMSM can only disclose their HIV status if they know about their HIV status and assume that YMSM do disclose their HIV status in a partnership independent of their YMSM partner. The CDPH HIV MSM surveillance report[11] shows that in the Chicago area in 2011, 78% of all HIV-positives were aware of their HIV infection. Applying this finding to our Crew 450 cohort, we take a conservative approach and assume that 78% of all YMSM knew at T1 and T2 about their HIV status regardless of being HIV positive or negative.

Statistical analysis showed that the within-partnership attributes oral-sex only and HIV disclosure status are not independent. Based upon the variables hivstatusk and hivstatusk_type we estimated that HIV status was disclosed in 48.08% of all oral-sex only and 71.6% of all non oral-sex only within-partnerships. Using these estimates and the percentage of YMSM who know about their HIV status we determine the probability of a YMSM to disclose his HIV status in a partnership using Bayes' rule. The probabilities of HIV status disclosure of a YMSM in a within-partnership are $P_{within}$(disclose|know HIV status, non oral-sex only)=0.9180 and $P_{within}$(disclose|know HIV status, oral-sex only)=0.6164.

### 2.5.4.4 Seriousness of within-partnerships

We use multivariate Logistic regression modeling to determine if a within-partnership is serious. Independent variables of the regression model also include the partnership attributes race of partner, oral-sex only, sex-role behavior and sex-frequency as well as HIV disclosure status. Table 21 shows coefficients and p-values of significant variables of the final multivariate Logistics regression model after applying stepwise elimination of variables approach.

| Covariates | Estimate | p-value |
|---|---|---|
| **(Intercept)** | -0.3005 | 0.1 |
| **alcuse** | 0.5090 | p<0.01 |
| **oral-only** | -0.9902 | p<0.001 |
| **hivstatusp_d** | 0.9726 | p<0.02 |
| **disclosed** | 1.0280 | p<0.001 |
| **BL** | -1.5562 | p<0.01 |
| **BW** | -0.9546 | p<0.02 |
| **LW** | -1.1911 | p<0.01 |
| **partcon_p** | -1.7529 | p<0.001 |
| **partcon_d** | -0.8876 | p<0.001 |

Table 21: Multivariate coefficient estimates of the Logistic regression model on the outcome seriousness of a within-partnership (n=728). hivstatusp_d: both partners know about their HIV status and assume/know the status of the other; partcon_p: both partners have sex outside the relationship, partcon_d: one partner has sex outside the relationship. Intercept attributes: no alcohol usage, non oral-sex only partnership, both partners are HIV negative or HIV positive (know themselves, assume/know the other), status not disclosed, not Black-Latino, Black-White or Latino-White partnership, both partners have no sex outside the partnership.

We determine the probability that a newly formed within-partnership is serious using the coefficients shown in Table 21 and the approach discussed in section SDC 2.2.3.2.6 and equation (4). Seriousness of a partnership can change in case one or both partners get to know and disclose an eventual HIV infection.

### 2.5.4.5 Mean length $\mu_{with,ij}$ of within-partnerships

We assume that the duration of partnerships within the study population $T_{with,ij}$ has a geometric distribution (see also discussion in SDC 2.2.4.2.7). We determine the mean length of a partnership within the study population using a multivariate linear regression model with a stepwise elimination of variables approach. Outcome variable is the length of a within-partnership minus 15 days and independent variables also include the seriousness of a within-partnership. Table 22 shows the results of the multivariate linear regression model.

| Covariates | Estimate | p-value |
|---|---|---|
| **(Intercept)** | 131.61 | p<0.001 |
| **page(abs)** | -24.95 | p<0.02 |
| **druguse** | 91.33 | p<0.001 |
| **BW** | -106.09 | p<0.05 |
| **WW** | -65.42 | p<0.05 |
| **LL** | 66.82 | p<0.05 |
| **LW** | -138.33 | p<0.01 |
| **OO** | 292.49 | p<0.01 |
| **partcon_p** | 145.29 | p<0.001 |
| **seriousYes** | 127.04 | p<0.001 |

Table 22: Multivariate coefficient estimates of the multivariate linear regression model on the outcome length of within-partnerships (n=726). Intercept attributes: absolute age difference 0, no drugusage, not Black-White, not White-White, not Latino-Latino, not Latino-White or Other-Other race-mixing, either none or only one partner has sex outside the relationship, non-serious relationship. Residual standard error is 269.5 with 716 df.

We determine the mean length of a newly formed outside-partnership using the coefficient estimates of the linear regression model shown in Table 22 and the approach discussed in section SDC 2.2.4.2.7.

### 2.5.4.6 Unprotected anal intercourse/use of condoms in within-partnerships

To determine the likelihood of UAI per sex-act in a within-partnership we use a multivariate Logistic regression model with a stepwise elimination of variables approach. Again, outcome variable is UAI (i.e., no condom usage) in a within-partnership and independent variables include all ego-attributes and partnership attributes discussed above. Table 23 shows the results of the final multivariate Logistic regression model.

| Covariates | Estimate | p-value |
|---|---|---|
| **(Intercept)** | -1.68203 | p<0.001 |
| **NG** | 0.983616 | p<0.02 |
| **SES 2 low** | 0.527127 | p<0.01 |
| **page(abs)** | 0.178967 | p<0.05 |
| **hivstatusp_d** | -0.88445 | p<0.05 |
| **alcuse** | 0.3781 | p<0.05 |
| **LL** | 0.597545 | p<0.02 |
| **OO** | 2.291779 | p<0.05 |
| **length** | 0.000679 | p<0.02 |
| **seriousYes** | 0.907858 | p<0.001 |

Table 23: Multivariate coefficient estimates of the multivariate Logistic regression model on the outcome UAI (no use of condoms) of within-partnerships (n=726). Intercept attributes: none of the partners is infected with NG, both partners have high SES 2, absolute age-difference 0, both partners are HIV negative or HIV positive (know themselves, assume/know the other), no alcohol usage, no Latino-Latino or Other-Other race mix, partnership length 0 days and non-serious partnership.

# 3. Sexual partnership dissolution

We assume the duration of both sexual partnerships within and outside the study population to follow an exponential distribution [14]. Thus, we assume the partnership duration to follow a geometric distribution in our discrete-time simulation model. The probability of a partnership dissolving in the next time step remains constant over time (it can of course vary from partnership to partnership but not on how long the partnership has been ongoing) and thus the probability that a partnership has not ended at time-step k after the partnership began is denoted as $P(X = k) = (1 - p)^{k-1}p$, where the mean length of the partnership defines the dissolution parameter p, i.e. $E[X] = \mu = \frac{1}{p}$.

Figure 5 and Figure 6 show the distribution of the survival function of the normalized partnership lengths of within partnerhips $T_{with,ij}$ and outside-partnerships $T_{out,i,k}$ respectively, i.e., the actual length $T_{with,ij}$ and $T_{out,i,k}$ divided by the mean length $\mu_{ij}$ and $\mu_{out,i,k}$ respectively. The survival function is the probability that the partnership is longer than x. Thus, the point (2,0.3) would correspond to 30% of the partnerships have normalized length ≥2 (i.e., the actual length is ≥2x the predicted length).

We used a log scale since the survival function would be a straight line for an exponential distribution. The black line shows the model assuming an exponential distribution of rate 1. We also see that for the longest 10-20% of the partnerships (y-values of 0.1-0.2 and below --- note the logarithmic scale) that the data better fits an exponential distribution with rate 0.4 or 0.6. This means that for the longest normalized partnerships, the actual probability of them dissolving is about half of what we predicted, i.e., that if a partnership seems to last unexpectedly long (if the partnership length T is bigger than the prediction $\mu$), then it is about half as likely to dissolve each time step as we assume. Since this is only for a small fraction of the relationships, we suggest a good fit of our model (assuming exponential distribution rate 1).

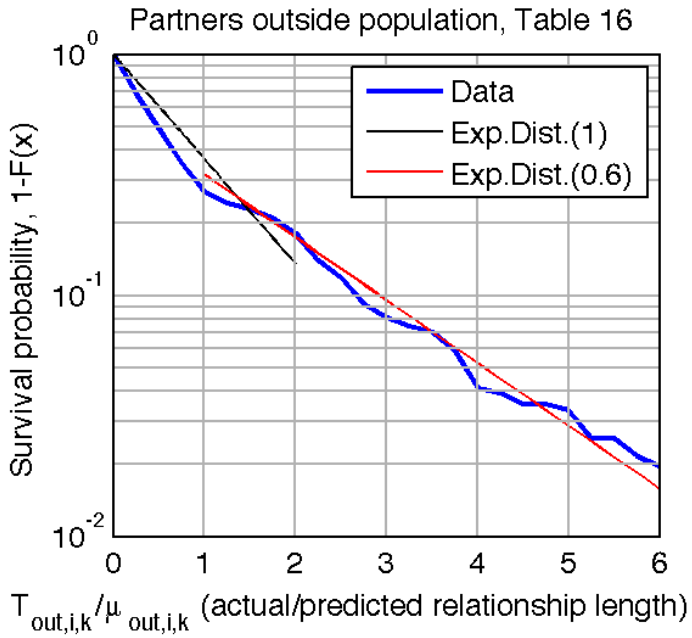


Figure 5: Distribution of the survival function of the normalized outside-partnership duration $T_{out,i,k}$ based on linear regression model shown in Table 16 to predict the mean length of an outside-partnership $\boldsymbol{\mu_{out,i,k}}$.

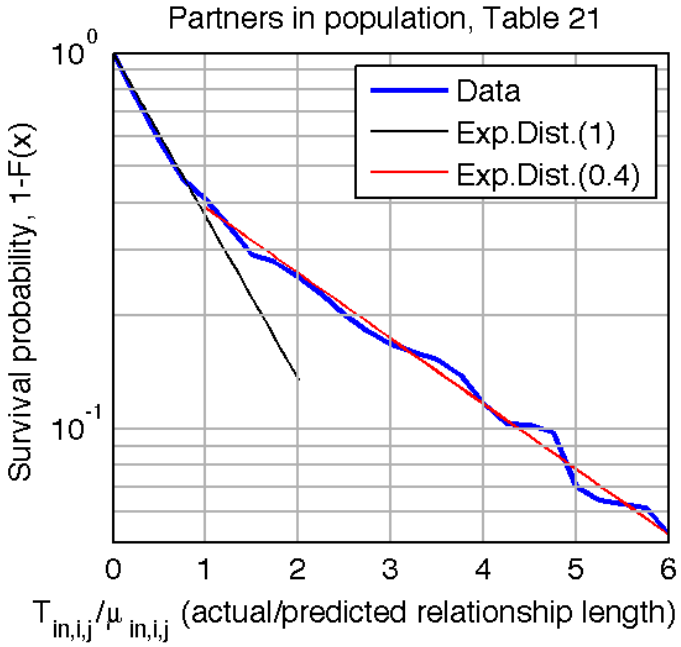


Figure 6: Distribution of the survival function of the normalized within-partnerships duration $T_{with,ij}$ based on the linear regression model shown in Table 22to predict the mean length of a partnership within the study population $\boldsymbol{\mu_{ij}}$.

# 4. Sexual orientation and sex-role: tendency and behavior

In our model of partnership formation and dissolution each YMSM identifies himself with a sexual tendency. The sexual tendency of a YMSM consists of his sexual orientation, his desired sex-role, and his desired sex-frequency. The (reported) sexual tendency of a YMSM may differ from the actual sexual behavior observed in his partnerships [15]. Thus, we developed a probabilistic model where the actual sexual behavior of a YMSM in a partnership is determined based on his sexual tendency, the sexual tendency of his partner, and the overall sexual behavior in the cohort. The sexual orientation of a YMSM will determine his likelihood to choose either a female or a male partner. We use latent variables to model the influence of the sexual tendency components desired sex-role and desired sex-frequency on actual sexual partnership behavior because these variables are not directly observable in a partnership but assumed to influence the sex-role and sex-frequency behavior in partnerships. Figure 7 shows a schematic illustration of how sexual tendency influences sexual behavior in our model of partnership formation and dissolution and in the following sections we introduce the probabilistic model to determine actual sexual behavior of a YMSM in a partnership.

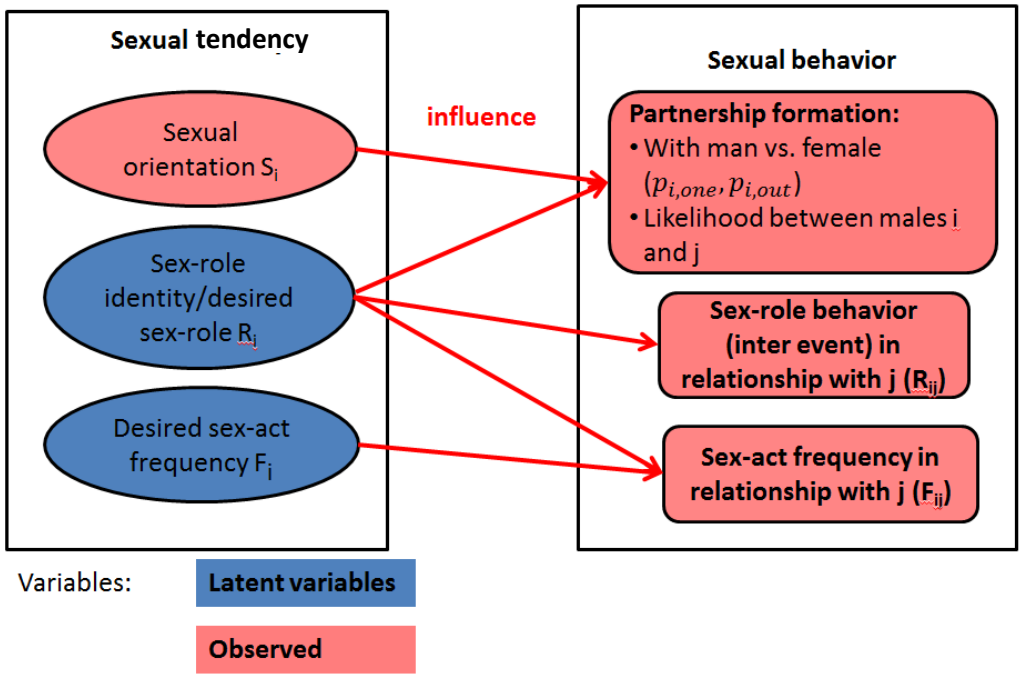


Figure 7: Schematic illustration showing influence of sexual orientation and sex-role tendency on partnership formation, sex-role behavior, and sex-act frequency in the model of partnership formation and dissolution among YMSM.

## 4.1 Notation

i: indexes individuals (i.e., YMSM of age 16-21.8 years)

e=(i,j): within-partnership between YMSM i and j

$S_x$ set of possible *x*-values

$S_o$: set of sexual orientations; $S_o$={only gay/homosexual, mostly gay/homosexual, bisexual, mostly heterosexual, only heterosexual, other}, see also discussion in SDC 1.

$O_i$: sexual orientation individual i identifies himself with, $O_i \in S_o, for\ all\ i$

$S_r$: set of sexual roles; $S_r$ ={insertive, receptive, versatile}

## 4.2 Sexual orientation

When forming a new one-night-partnership or partnership the decision of a YMSM to choose his sex partner to be either male or female will depend upon the sexual orientation of the YMSM. We model this decisqion of the YMSM using the probability $p_i$ which is conditional on the sexual orientation $O_i$ of

YMSM i. Given the sexual orientation $O_i$, $p_i$ denotes the conditional probability that the newly formed partnership of YMSM i will be a female-male partnership stratified by partnership-types one-night-partnership and outside-partnership. Thus,

$\boldsymbol{p_{i,one} = P(X_i = 1 | O_i,\ Y_i = 1)}$: probability that one-night-partnership of YMSM i will be a female-male one-night-partnership given his sexual orientation $O_i$

$\boldsymbol{p_{i,out} = P(X_i = 1 | O_i,\ Y_i = 0)}$: probability that outside-partnership of YMSM i will be female-male partnership given his sexual orientation $O_i$

where

$Y_i$ : partnership-type of newly formed partnership, i.e. one-night or outside-partnership

$$Y_i = \begin{cases} 1, & if\ next\ partnership\ of\ i\ is\ one-night \\ 0, & otherwise \end{cases}$$

$X_i$ : sex of partner in newly formed partnership of i, i.e. female-male or male-male partnership.

$$X_i = \begin{cases} 1, & if\ next\ partnership\ of\ i\ is\ female-male \\ 0, & otherwise \end{cases}$$

We estimated $p_{i,one}$ and $p_{i,out}$ for each sexual orientation using data about the sex of the partner for one-night-partnerships and outside-partnerships named in the in-depth interviews. Table 24 shows the empirical estimates for $p_{i,one}$ and $p_{i,out}$.

| Sexual orientation | $p_{i,one}$ | $p_{i,out}$ |
|---|---|---|
| **Only gay/homosexual** | 0.005 (n=207) | 0.000 (n=237) |
| **Mostly gay/homosexual** | 0.020 (n=102) | 0.011 (n=95) |
| **Bisexual** | 0.152 (n=79) | 0.123 (n=143) |
| **Mostly heterosexual** | 0.111 (n=9) | 0.857 (n=26) |
| **Only heterosexual** | 1.0 (n=1) | 0.500 (n=8) |
| **Other** | 0.231 (n=13) | 0.000 (n=13) |

Table 24: Empirical estimates of conditional probabilities $\boldsymbol{p_{i,one}}$ and $\boldsymbol{p_{i,out}}$ of YMSM that sexual partner in a one-night-partnership or outside-partnership will be female. N=411 for one-night-partnerships and n=512 for outside-partnerships.

## 4.3 Desired sex-role, sex-frequency and actual behavior in within-partnerships

In our model, YMSM identify themselves with one of the three desired sex-roles receptive, versatile or insertive [2,16]. YMSM are considered sex-role versatile within a partnership if they take on both the insertive and receptive sex-role in male-male anal sex-acts.

We model the desired sex-role of a YMSM using the latent variable $R_i$. $R_i$ denotes the fraction of anal sex-acts in a male-male partnership where YMSM i desires to be receptive. For example, $R_i = 1$ implies that YMSM i desires to be exclusively receptive in all of his male-male sex-acts. The actual sex-role behavior in a partnership might differ from the desired sex-role identity $R_i$. $R_{ij}$ denotes the actual fraction of sex-acts in a within-partnership between YMSM i and j where YMSM i will be receptive. Thus, $R_{ij}$ describes the inter-event sex-role behavior of YMSM i and j, i.e. if $R_{ij} \in (0,1)$ then YMSM i will take on the receptive sex-role with probability $R_{ij}$ in all sex-acts in the within-partnership between YMSM i and j. However, given the inter-event sex-role YMSM i and j can also change the sex-role within a sex-act (i.e., intra-event sex-role versatile) if both YMSM are not sex-role exclusive in their partnership (i.e., $R_{ij} \neq \{0,1\}$).

Similar to the desired sex-role $R_i$ we assume the YMSM to also have a desired sex-frequency $F_i$, i.e. the desired number of sex-acts per time-step in a partnership. Since the desired sex-frequency $F_i$ may also differ from the observed sex-frequency we model the influence of $F_i$ on the actual sex-frequency $F_{ij}$ in a within-partnership between YMSM i and j.

Following, we summarize the above introduced notation as well as we introduce $r_{ij}$ and $f_{ij}$, the empirical fraction of receptive sex-acts and the empirical sex-frequency in partnerships between two individuals:

$R_i$ : desired fraction of YMSM i to be receptive in a male-male sex-act (latent variable)

$R_{ij}$ : fraction of YMSM i being receptive in sex-acts within a male-male partnership with YMSM j. $R_{ij} = 1 - R_{ji}$.

$F_i$ : desired sex-act frequency of YMSM i in a partnership (latent variable)

$F_{ij}$ : sex-act frequency of YMSM i and j in a within-partnership. $F_{ij} = F_{ji}$

$r_{ij}$: empirical fraction of receptive sex-acts in male-male partnership of individual i and j

$f_{ij}$: empirical frequency of sex-acts in female-male and male-male partnership of individual i and j

## 4.3.1 Desired fractions $R_i$ and $F_i$ of YMSM i

We determined $R_i$ and $F_i$ using the number of insertive and receptive sex-acts of each partnership named in the in-depth interviews at T1 and T2, i.e. we used the fraction of receptive sex-acts $avg(r_{ij}: j)$ observed in all male-male partnerships of YMSM i and the average sex-act frequency $avg(f_{ij}: j)$ observed in all partnerships of YMSM i respectively to determine $R_i$ and $F_i$. We assumed the desired sex-role and sex-frequency to be independent of other individual attributes and we assumed no difference in the desired sex-frequency of female-male to male-male partnerships. For YMSM who named less than two (male-male) partnerships we bootstrap sampled $avg(r_{ij}: j)$ and $avg(f_{ij}: j)$ from the empirical distributions of $avg(r_{ij}: j)$ (n=265) and $avg(f_{ij}: j)$ (n=265). We created 10 samples of $avg(r_{ij}: j)$ and $avg(f_{ij}: j)$ for all n=421 YMSM which are used to determine 10 different sets of $R_i$ and $F_i$ for n=421 YMSM.

### 4.3.1.1 Calculation of $R_i$ and $F_i$

***Notation***

$avg(r_{ij}: j)$ : observed fraction of YMSM i being receptive in male-male sex-acts in partnerships (no one-night-partnership data included; if no empirical estimate was available, value was sampled from empirical distribution). $avg(r_{ij}: j) \in [0,1]$

$avg(avg(r_{ij}: j))$: overall fraction of YMSM being receptive in male-male sex-acts in partnerships. $avg(avg(r_{ij}: j)) \in [0,1]$

$E[R_{ij}|R_i]$ : true (expected) fraction of YMSM i in partnership with YMSM j to be receptive given his desired sex-role $R_i$. $E[R_{ij}|R_i] \in [0,1]$

$E[R_j]$: expected desired fraction of the partner YMSM j to be receptive in sex-acts in male-male partnerships. $E[R_j] \in [0,1]$

$avg(f_{ij}{:}\, j)$ : observed average sex-act frequency of YMSM i in partnerships (if no empirical estimate was available, value was sampled from empirical distribution). $avg\big(f_{ij}{:}\, j\big) \in [0, \infty)$

$avg(avg\big(f_{ij}{:}\, j\big))$: observed average sex-act frequency among all YMSM. $avg(avg\big(f_{ij}{:}\, j\big)) \in [0, \infty)$

$E[F_{ij}|F_i]$ : true (expected) sex-act frequency in partnership between YMSM i and j given YMSM i's desired sex-frequency $F_i$. $E[F_{ij}|F_i] \in [0, \infty)$

$E[F_j]$: expected sex-act frequency of YMSM j in a partnership. $E[F_j] \in [0, \infty)$

***Calculation of $R_i$***

Following we describe a probability model to determine the desired sex-role of $R_i$. We assume the true (expected) fraction of YMSM i to be receptive in a partnership with YMSM j given his desired sex-role $R_i$ to be

$$E[R_{ij}|R_i] = E[Rj] + \frac{R_i}{2} - \frac{\delta(E[R_j])}{2}. \qquad (12)$$

Further, we assume

$$E[R_{ij}|R_i] \approx avg(r_{ij}{:}\, j) \qquad (13),$$

i.e. that the observed fraction represent the true fraction and

$$E[R_j] \approx avg(avg\big(r_{ij}{:}\, j\big)) \qquad (14).$$

Defining the threshold functions

$$\delta(x) = \begin{cases} avg\left(avg\big(r_{ij}{:}\, j, r_{ij} \neq 1\big)\right), if\ avg\big(r_{ij}{:}\, j\big) > avg(avg\big(r_{ij}{:}\, j\big)) \\ avg\left(avg\big(r_{ij}{:}\, j, r_{ij} \neq 0\big)\right), otherwise \end{cases} \qquad (15)$$

and

$$\vartheta_R(x) = \begin{cases} 1, \ if\ x \geq 1 \\ x, \ if\ 0 < x < 1 \\ 0, otherwise \end{cases} \qquad (16)$$

as well as $R_i \in [0,1]$ we can rearrange (10) such that

$$R_i = \vartheta_R \left(2\, avg(r_{ij}{:}\, j) - 2\, avg\left(avg(r_{ij}{:}\, j)\right) + \delta(avg\left(avg(r_{ij}{:}\, j)\right))\right). \quad (17)$$

Figure 8a shows the distribution of $R_i$ for the total population n=421 given one realization of $avg(r_{ij}{:}\, j)$. We observe for this realization that approximately 57% of the total population has a desired sex-role exclusivity. Figure 8b shows the comparison of the observed sex-role behavior $E[R_{ij}|R_i]$ to the desired sex-role $R_i$.

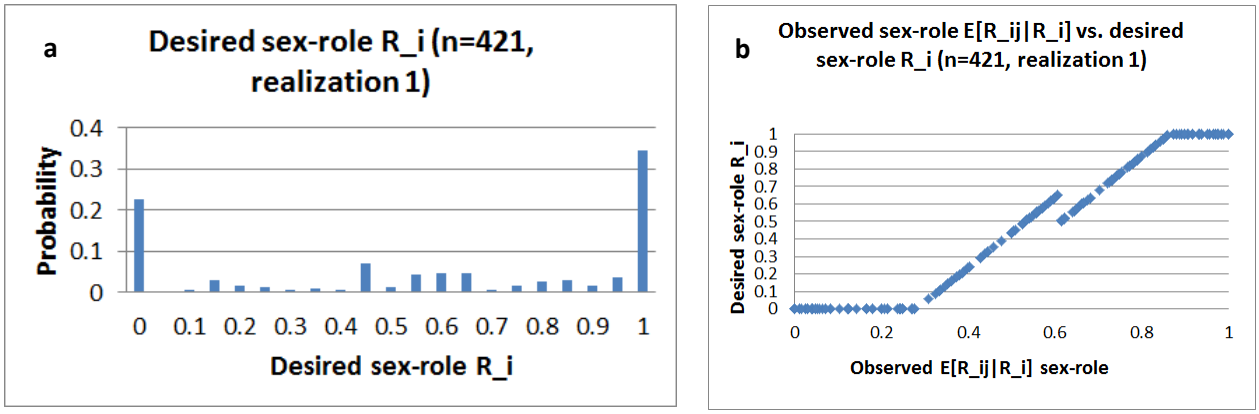


Figure 8a: Distribution of the desired sex-role $R_i$ of n=421 YMSM for one realization of $\boldsymbol{avg(r_{ij}{:}\, j)}$. Figure 8b: Comparison of the observed sex-role behavior $\boldsymbol{E[R_{ij}|R_i]}$ to the desired sex-role $\boldsymbol{R_i}$ for the entire population of n=421 YMSM.

### *Calculation of $F_i$*

We determine $F_i$ using a similar approach. We assume the expected sex-act frequency of YMSM i in a partnership with j given the desired sex-act frequency $F_i$ of YMSM i to be

$$E[F_{ij}|F_i] = \frac{F_i}{2} + \frac{E[F_j]}{2} \quad (18).$$

Rearranging (16) and taking expectations on both sides yields

$$E[F_i] = 2E[E[F_{ij}|F_i]] - E[E[F_j]] \quad (19).$$

Assuming

$$E\left[E[F_j]\right] \approx [E[E[F_{ij}|F_i]] \approx avg\left(avg(f_{ij}{:}\, j)\right) \quad (20),$$

yields

$$E[F_i] = E[E[F_{ij}|F_i]] \quad (21).$$

Further, assuming

$$E[F_{ij}|F_i] \approx avg(f_{ij}{:}j) \quad (22)$$

and introducing the threshold function

$$\vartheta_F(x) = \begin{cases} x, & if\ x > 0 \\ 0, & otherwise \end{cases} \quad (23)$$

we obtain

$$F_i = \vartheta_F\left(2avg(f_{ij}{:}j) - avg\left(avg(f_{ij}{:}j)\right)\right) \quad (24)$$

Figure 9a shows the distribution of $F_i$ for the total population of n=421 YMSM given one realization of $avg(f_{ij}{:}j)$. Figure 9b compares the observed sex-frequency behavior $E[F_{ij}|F_i]$ to the desired sex-frequency $F_i$ in male-male partnerships.

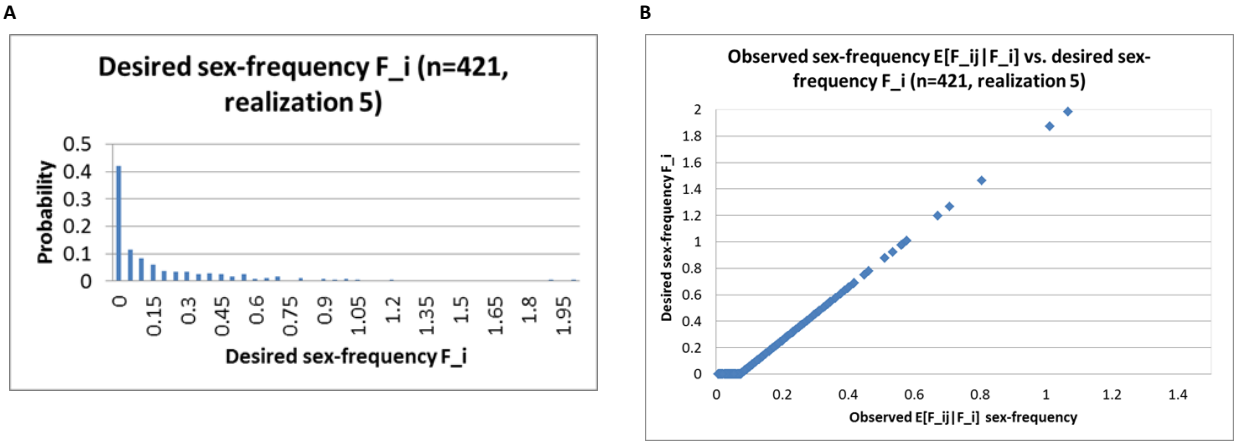


Figure 9a: Distribution of the desired sex-frequency $F_i$ in male-male partnerships of n=421 YMSM for one realization of $\boldsymbol{avg(f_{ij}{:}j)}$. Figure 5b: Comparison of the observed sex-frequency behavior $\boldsymbol{E[F_{ij}|F_i]}$ to the desired sex-frequency behavior $\boldsymbol{F_i}$ for the entire population of n=421 YMSM.

### 4.3.1.2 Desired sex-role $R_i$ and partnership formation within study population

If YMSM i and j identify themselves as being sex-role exclusive, i.e. $R_i, R_j \in \{0,1\}$, a partnership can only be formed if both identify themselves with different exclusive sex-roles, i.e. a partnership can only be formed if

$$0 < R_i + R_j < 2 \qquad (25).$$

### 4.3.2 Sex-role behavior $R_{ij}$ and sex-frequency $F_{ij}$ in within-partnerships

Following we describe the probability model to determine $R_{ij}$, the actual fraction YMSM i will be receptive in sex-acts of the within-partnership with YMSM j as well as the probability model to determine the sex-act frequency $F_{ij}$ in this partnership.

#### 4.3.2.1 Probability model for $R_{ij}$

*Notation:*

$R_{ij}$ : fraction of YMSM i to be receptive in sex-acts in within-partnership with YMSM j. $R_{ij} = 1 - R_{ji}$

$Z_{ij}^r$ : (stochastic) error-term accounting for statistical errors in $R_{ij}$

$\sigma_r$: standard deviation of error-term $Z_{ij}^r$

The actual fraction of YMSM i to be receptive in sex-acts within the partnership with YMSM j (given that the partnership can be formed) is

$$R_{ij} = \vartheta_R \left( \frac{1}{2} + \frac{R_i + R_j}{2} + \sigma_r Z_{ij}^r \right) \qquad (26)$$

where

$$Z_{ij}^r \sim N(0,1), for\ all\ partnerships\ (i,j) \qquad (27)$$

and $\sigma_r = 1.13$. $\sigma_r$ is determined using a coordinate-line search algorithm such that the first two moments of the generated distribution of $R_{ij}$ matches the first two moments of the empirical distribution of the observed fractions $r_{ij}$.

### 4.3.2.2 Probability model for $F_{ij}$

*Notation:*

$F_{ij}$ : actual sex-act frequency in within-partnership of YMSM i and j. $F_{ij} = F_{ji}$

$Z_{ij}^{f}$ : (stochastic) error-term accounting for statistical errors in $F_{ij}$

$\sigma_f$: standard deviation of error-term $Z_{ij}^{f}$

$\rho$ : correlation coefficient measuring correlation between $R_{ij}$ and $F_{ij}$ (i.e., corr(|R_ij-1/2|,F_ij)), $\rho_e = -0.223$ (p<0.01) (i.e., the more partners tend in a relationship to be sex-role exclusive the less sex they have)

The sex-act frequency in a within-partnership of YMSM i and j (given that the partnership can be formed) is

$$F_{ij} = \vartheta_F\left(\frac{F_i + F_j}{2} + \rho\left(\left|R_{ij} - \frac{1}{2}\right| - E\left|R_{ij} - \frac{1}{2}\right|\right) + \sigma_f Z_{ij}^{f}\right) \tag{28}$$

where

$$E\left|R_{ij} - \tfrac{1}{2}\right| \approx \left|avg\left(r_{ij}: j\right) - \tfrac{1}{2}\right|, for\ all\ YMSMi \tag{29}$$

and

$$Z_{ij}^{f} \sim N(0,1),\ for\ all\ partnerships\ (i,j) \tag{30}$$

with $\sigma_f = 0.01$ and $\rho = -0.12$ . $\sigma_f$ and $\rho$ were determined using a coordinate-line search algorithm such that the first two moments of the generated distribution of $F_{ij}$ match the moments of the empirical distribution of the observed sex-frequencies $f_{ij}$.

## 4.3.3 Sex-role behavior and sex-frequency in one-night-partnerships and outside-partnerships

### 4.3.3.1 Sex-role behavior in one-night-partnerships and outside-partnerships

We assume that $R_i$ accurately reflects the sex-role behavior in one-night-partnerships and outside-partnerships. Thus, the probability that YMSM i will be receptive in a one-night-partnership is

$$R_{i,one} = R_i + \sigma_r Z_{ij}^r . \quad (31)$$

For outside-partnerships, $R_{i,out}$ denotes the fraction of YMSM i to be receptive in sex-acts where

$$R_{i,out} = R_i + \sigma_r Z_{ij}^r. \quad (32)$$

### 4.3.3.2 Sex-act frequency in outside-partnerships

We assume the same desired sex-frequency $F_i$ for outside-partnerships and within-partnerships. Thus, the sex-act frequency of YMSM i in an outside-partnership $F_{i,out}$ can be determined as follows:

$$F_{i,out} = F_i + \rho\left(\left|R_i - \tfrac{1}{2}\right| - E\left|R_i - \tfrac{1}{2}\right|\right) + \sigma_f Z_{ij}^f, for\ all\ YMSM\ i. \quad (33)$$

# III. Supplemental Digital Content 3: Disease Transmission

# 1. Overview

We model the simultaneous spread of HIV, Neisseria gonorrhea (NG) and Chlamydia trachomatis (CT) among YMSM and assume HIV, NG and CT to be transmitted through sexual intercourse only. Table 25 provides an overview over the infectious diseases, the interactions among these diseases, and the variables describing the course of the disease and the disease transmission in our model. The design of our model was informed using available data about HIV, NG and CT infections in the Crew 450 study (i.e., biomedical data in the Crew 450 study were only collected for the diseases HIV, NG, and CT. Thus, we limited our model of disease spread among YMSM to HIV, NG, and CT). The model was parameterized using data of local and national surveillance reports[11,17-21] and estimates of published studies. Following, we discuss for each disease the modelled course of infection, the disease transmission, and interactions with other diseases. We conclude this section with a discussion of testing and treatment coverage.

Table 25: Overview of modelled infectious diseases, interactions among the diseases and model variables describing the course of infection and the disease transmission.

| Infectious disease | Interaction with other diseases | Course of infection | Transmission[a] |
|---|---|---|---|
| **HIV** | - | HIV 5-stage model stratified by<br>• testing<br>• ART[b] treatment<br>• full/partial suppression<br>• race | sex-role position<br>sex-frequency<br>UAI<br>male-male anal intercourse/ female-male anal intercourse/ penile-vaginal intercourse<br>HIV level of infectiousness<br>• stage of infection<br>• ART treatment<br>• full/partial suppression<br>• race<br>NG, CT co-infections |
| **Neisseria gonorrheae (NG)/ Chlamydia Trachomatis (CT)** | increases HIV susceptibility and HIV transmissibility | Duration of infection stratified by<br>• site of infection (urethral/rectal)<br>• symptomatic/asymptomatic<br>• treatment<br>ceasing of sexual activity | sex-role position<br>sex-frequency<br>UAI<br>male-male anal intercourse/ female-male anal intercourse/ penile-vaginal intercourse |

[a] Transmission of HIV, NG, and CT through sexual intercourse only.
[b] ART: antiretroviral therapy.

# 2. HIV

## 2.1 Course of the HIV infection

### 2.1.1 Disease progression model

Upon the actual infection a HIV positive progresses through various stages of the HIV infection to full-blown AIDS and finally AIDS related death. We model the progression of the HIV infection over time using a modified version of the HIV-stage model of Hollingsworth et al.[22]. Dependent on treatment coverage HIV infected go through up to five stages of the HIV infection[23]: primary infection (P), asymptomatic period without antiretroviral therapy (A-No ART) and with ART (A-ART), acquired immune-deficiency syndrome (AIDS) period while being sexual active (AIDS) and AIDS period while not being sexual active (0). We consider an individual to only be part of the study population as long as he is sexually active, i.e. the individual will leave the study-population once he reaches the infection stage '0'.

Hollingsworth et al.[22] estimates the duration of stage 'A' ($d_P$) to be 3 months, the duration of stage 'AIDS' ($d_{AIDS}$) to be 9 months and the duration of stage '0' ($d_0$) to be 10 months. The duration of stages 'A-No ART' ($d_{A\text{-}No\text{-}ART}$) and 'A-ART' ($d_{A\text{-}ART}$) vary due to status of treatment coverage, time of treatment initiation, and whether the individual is fully or partially suppressed (see discussion in section 3.2.3 in SDC 3). Figure 10 shows the progression of the HIV infection throughout different stages for sexually active HIV positive Black YMSM stratified by treatment coverage and viral suppression levels. Table 26 shows the estimated durations of the sexually active stages throughout the HIV infection stratified by race, treatment and viral suppression levels.

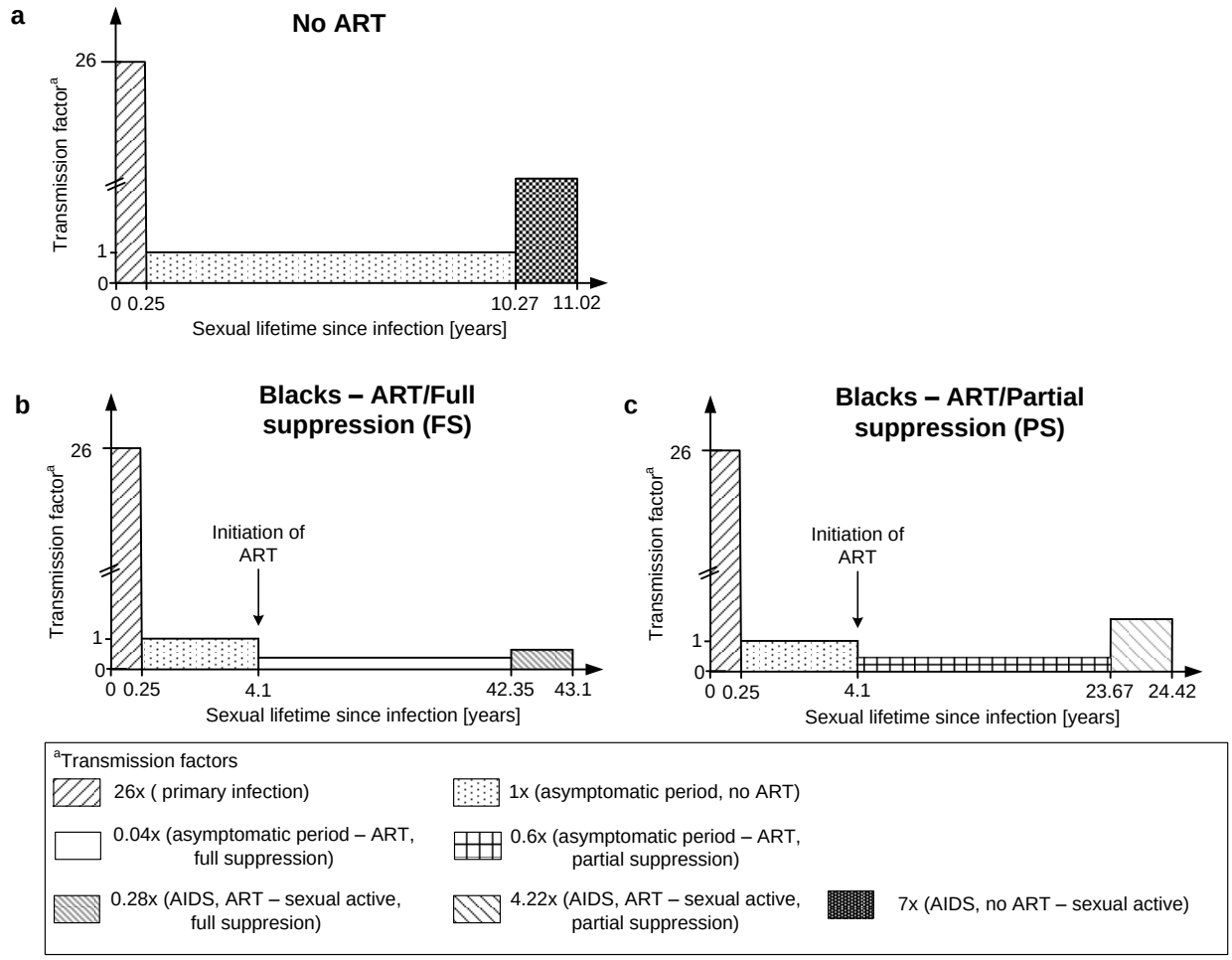


Figure 10: Stages of HIV infection for sexually active Black HIV positive YMSM stratified by treatment and full or partial suppression. In the style of [23]. Durations for other races are listed inTable 26.

Table 26: Estimated duration of sexually active stages of HIV infection in years for all races stratified by treatment status (ART) and level of suppression.

| Type HIV+ | Sexual lifetime after infection (y) | Duration stages of HIV infection (y)[a] | | | | | Source |
|---|---|---|---|---|---|---|---|
| | | $d_P$ | $d_{A\text{- No Art}}$ | $d_{A\text{-ART-FS}}$ | $d_{A\text{-ART-PS}}$ | $d_{AIDS}$ | |
| All races- No ART[b] | 11.02 | 0.25 | 10.02 | - | - | 0.75 | [22,24] |
| Blacks – ART | | | | | | | |
| FS[c] | 43.11 | 0.25 | 3.85 | 38.25 | - | 0.75 | [2,22,25,26] |
| PS[c] | 32.1 | 0.25 | 3.85 | - | 27.25 | 0.75 | [2,22,25-27] |
| Whites[d] – ART | | | | | | | |
| FS[c] | 52.07 | 0.25 | 3.35 | 47.72 | - | 0.75 | [2,22,25,26] |
| PS[c] | 41.07 | 0.25 | 3.35 | - | 36.72 | 0.75 | [2,22,25-27] |
| Latinos – ART | | | | | | | |
| FS[c] | 26.8 | 0.25 | 4.75 | 21.05 | - | 0.75 | [2,22,25,26] |
| PS[c] | 15.8 | 0.25 | 4.75 | - | 10.05 | 0.75 | [2,22,25-27] |

[a] $d_P$-duration primary infection period; $d_{A\text{- No Art}}$-duration asymptomatic period without ART treatment; $d_{A\text{-ART-FS}}$ - duration asymptomatic period with ART treatment and full suppression; $d_{A\text{-ART-FS}}$- duration asymptomatic period with ART treatment and partial suppression; $d_{AIDS}$-duration AIDS stage; in years (y)

[c] We estimate the duration of the sexually active stage for HIV positives not receiving ART (i.e., dA-No ART+dAIDS+d0) to be 11.6 years using Todd's et al. [24] estimate of the age adjusted median survival time after seroconversion.

[c] FS: full suppression of HIV viral load; PS: partial suppression of HIV viral load.

[d] Due to lack of data we assume that expected durations of whites apply to other races[2].

## 2.1.2 Treatment (ART)

### 2.1.2.1 Life-expectation with ART treatment

Mills et al.[25] estimated the life expectancy of HIV infected after the initiation of ART, i.e. $d_{A\text{-}ART+}$ $d_{AIDS}+d_0$, to be 27.9 years. Assuming full suppression, the initiation of ART at the corresponding eligibility threshold of 250 CD4 cells/µl in the above study, and ART to only extend the asymptomatic period (i.e., the duration of the primary infection period ($d_P$) and AIDS period ($d_{AIDS}$) will remain the same[23]), we assume the HIV infected to receive 26.32 years of treatment throughout the asymptomatic

period ($d_{A\text{-}ART}$). During these 26.32 years the CD 4 cell count will decline from 250 CD4 cells/µl at the initiation of ART treatment to 100 CD 4 cells/µl at the beginning of AIDS. Assuming a linear decline of the CD 4 cell count during this time[23], the expected yearly decline in CD 4 cells/µl is $r_{FS}$= 5.7 CD 4 cells/µl y.

### 2.1.2.2 Treatment initiation

We assume that treatment will be only initiated if the HIV infected YMSM was tested positive.

Treatment initiation is dependent on CD4 cell count: Swindells et al.[28] estimated the median CD4 cell count at the initiation of treatment to be 318 CD 4 cells/µl for Blacks, 372 CD 4 cells/µl for Whites, and 220 CD 4 cells/µl for Latinos. Using these CD 4 cell counts Goodreau et al.[4] estimated the time of treatment initiation after HIV seroconversion/infection to be 4.1 years for Blacks, 3.6 years for Whites and 5.0 years for Latinos. Applying the estimates of Swindell et al.[28] and Goodreau et al.[4] to the YMSM population in Chicago, assuming full suppression, and the rate of decline in the CD4 cell count to be $r_{FS}$ we estimate the duration of the asymptomatic period without ($d_{A\text{- No Art}}$) and with treatment ($d_{A\text{- Art-FS}}$) to be

$$d_{A\text{- No Art}} = \text{time to treatment} - d_P \quad (34)$$

$$d_{A\text{- Art-FS}} = (\text{CD 4 cells/µl at treatment initiaton} - 100 \text{ CD 4 cells/µl})/ r_{FS}. \quad (35)$$

### 2.1.2.3 Partial and full suppression

The immune response of individuals receiving treatment varies because of differences in the resistance to treatment, the fact that men go on and off for treatment, switching of regimens and adherence problems[2,4]. Similar to Goodreau et al.[4] we account for these differences in the immune response to treatment by assuming that HIV positives who receive ART can be either fully suppressed (FS) or partially suppressed (PS). Thus, PS accounts in our model for all possible differences in the immune response to ART ranging from low adherence to high resistance to treatment. We apply the estimates of Weintrob et al.[29] and

Goodreau et al.[4] and estimate the percentage of HIV positive men who are fully suppressed while being on treatment to be 62.7% for Blacks, 75% for Whites, and 68.9% for Latinos.

Further, we assume that HIV positives who are partially suppressed will progress faster to AIDS ($d_{AIDS}$) than fully suppressed HIV positives[2,4,27]. May et al.[27] estimated the life expectancy of partially suppressed HIV positives after initiation of ART to be 11 years shorter than the life expectancy of fully suppressed HIV positives after initiation of ART. Assuming the duration of $d_{AIDS}$ and $d_0$ not to be affected by PS we estimate the duration of the asymptomatic period with treatment and partial suppression to be

$$d_{A\text{- }Art\text{-}PS} = d_{A\text{- }Art\text{-}FS} - 11\ y. \qquad (36)$$

In absence of data we assume the estimates for the time of treatment initiation and the percentages of FS for Others to be the same as for Whites[4].

### 2.1.3 Death rates

We estimated the durations of the different stages of the HIV infection shown in Table 26 using estimates of the median survival time after HIV infection[24] and the median survival time after initiation of ART[25]. Thus, the estimated durations in our model already account for the HIV related increase in mortality.

# 2.2 HIV Transmission

## 2.2.1 Mode of HIV transmission

In our model of HIV spread among YMSM we assume HIV to be transmitted through anal or vaginal sexual intercourse only. We do not model HIV transmission through oral sex because of the low transmission risk for unprotected receptive oral intercourse being 0.04% per sex-act[8] and oral sex being not considered as efficient mode of HIV transmission[30].

## 2.2.2 HIV Transmission risk per sex-act

### 2.2.2.1 Baseline HIV transmission risk and circumcision status

We assume the HIV transmission risk per sex-act to depend on the sex-role behavior, the type of sexual intercourse (i.e., male-male anal, female-male anal or penile-vaginal), the stage of the infection and the treatment status of the infected, condom usage, circumcision status of the HIV negative and NG and CT (co-)infections of HIV positive and negative. We define the baseline HIV transmission risk as the HIV transmission risk per sex-act of a HIV negative and a HIV positive where the HIV positive is in the asymptomatic stage of the infection and does not receive treatment, the HIV positive and negative do not have NG and CT infections and both have unprotected sex (i.e., no condom usage).

The baseline HIV transmission risk per sex-act for male-male anal intercourse depends on the sex-role behavior of the HIV negative. Vittinghoff et al.[8] estimated the baseline HIV transmission risk to be 0.82% per sex-act for HIV negative being receptive in male-male anal intercourse. For insertive male-male anal intercourse we stratify the baseline HIV transmission risk per sex-act by circumcision status. We assume the HIV transmission risk of circumcised men having insertive male-male anal intercourse to be lower compared to uncircumcised men because of the hypothesized association of circumcision status and HIV transmission risk in case of insertive anal male-male intercourse[31,32]. Using the estimates of Goodreau et al.[4] we assume the baseline HIV transmission risk for insertive male-male anal intercourse with a HIV

positive to be 0.126% and 0.314% per sex-act for circumcised and non-circumcised men respectively. We assume circumcision status to have no impact on HIV transmission risk per sex-act for receptive anal intercourse[32]. In case of intra-event sex-role versatility for male-male anal intercourse we assume the HIV transmission risk per sex-act to equal the sum of the receptive and insertive HIV transmission risk per sex-act stratified by circumcision status.

We assume the baseline HIV transmission risk for penile-vaginal intercourse, i.e. the HIV transmission risk for a YMSM having insertive penile-vaginal sex, to be 0.05% per sex-act[33]. Assuming a circumcision prevalence of 16% among males[34] in the cohort of Leynaert et al. and a risk reduction of 60% in HIV transmission risk for insertive penile-vaginal sex for circumcised men compared to non-circumcised men[35] we estimate the HIV transmission risk to be 0.022% and 0.055% per insertive penile-vaginal sex-act for circumcised and non-circumcised men having respectively. Data of the Crew 450 study show that most of the YMSM (see section 2 in SDC 2) are sex-role versatile while having sex with a woman, i.e. YMSM have both anal-insertive and vaginal-insertive sex with probability P(versatile|female-male)=0.7059 during a sex-act. We assume the HIV transmission risk per episode of female-male anal insertive intercourse to be the same as for male-male anal insertive intercourse stratified by circumcision status.

### 2.2.2.2 HIV transmission risk and stages of HIV-infection

Transmissibility of the HIV infected varies for each stage of the infection[22]. Figure 10a shows transmission factors for a HIV infected without ART treatment throughout the stages of the HIV infection relative to the transmission risk $c_{A-NoART}$ in the asymptomatic period. We set $c_{A-NoART}$ =1 because the baseline HIV transmission risk assumes the HIV positive to be in the asymptomatic period without treatment. We use the estimates of Hollingsworth et al.[22] and assume the transmission risk during each period to be constant[22].

### 2.2.2.3 HIV transmission risk and ART treatment

Cohen et al.[26] estimated the reduction in HIV transmissibility to be 96% for HIV positive receiving treatment in the asymptomatic period. Thus, we assume the transmissibility of fully suppressed HIV positives in the asymptomatic period receiving ART to be $c_{ART,FS}$ = 0.04.

We determine the HIV transmissibility of partially suppressed HIV positives in the asymptomatic period receiving ART by relating the HIV viral load with the HIV transmission risk[36]. Based upon a mathematical model of Wilson et al.[36] we calculate the relative reduction of the HIV transmission risk per sex-act using the median HIV RNA viral load of partially suppressed HIV infected in the asymptomatic period. Given the HIV RNA viral load of HIV positives without treatment in the asymptomatic period to be $v_0$=4.5 log10 copies/mL[36], the HIV RNA viral load of a partially suppressed HIV positives in the asymptomatic period $v_{0,PS}$=3.1 log10 copies/mL[37], and the HIV transmissibility of HIV positives without treatment in the asymptomatic period $\beta_0$ =1, we determine the HIV transmissibility of partially suppressed HIV positive in the asymptomatic phase $c_{ART,PS}$ by use of the following formula

$$c_{ART,PS} = 2.45^{\log10\,(v0,PS\,/\,v0)} * \beta_{0,FS} = 0.865. \qquad (37)$$

Granich et al.[38] and Armbruster et al.[23] assume the effect of ART on HIV transmission risk of HIV positives in the AIDS stage to be the same as the effect of ART on the HIV transmission risk in the asymptomatic period. Further, we assume this effect to account for both fully suppressed and partially suppressed HIV positives. Thus, $c_{ART-AIDS,\cdot} = c_{ART,\cdot} c_{AIDS}$ and $c_{ART-AIDS,FS}$=0.28 as well as $c_{ART-AIDS,PS}$= 6.055 for fully suppressed HIV positive and partially suppressed HIV positive in the AIDS stage respectively.

Figure 10a and Figure 10b show the HIV transmission factors for Black HIV positive YMSM with full and partial suppression. We assume the same HIV transmission factors across all races.

#### 2.2.2.4 HIV transmission risk and condom usage

Weller and Davis[39] estimate condom usage to reduce HIV transmission risk by 80% for penile-vaginal intercourse. We assume condom usage to reduce HIV transmission risk by 80% for penile-vaginal, female-male anal intercourse and male-male anal intercourse.

#### 2.2.2.6 HIV transmission risk and NG and CT (co-)infections

For a discussion of the impact of NG and CT infections on HIV transmissibility and HIV susceptibility see section 3 in SDC 3.

## 2.2.3 HIV transmission risk for different types of partnerships

We determine the HIV transmission risk per sex-act of a HIV negative with a HIV positive by multiplying the baseline HIV transmission probability per sex-act with the corresponding risk factor.

In comparison to YMSM, the detailed HIV status (i.e., infection status, stage, suppression level and treatment) and thus the HIV transmissibility of the sex-partners in one-night-partnerships and outside-partnerships is unknown. Table 27 shows the estimates for the HIV prevalence and the HIV transmissibility of female and male partners in one-night-partnerships and outside-partnerships of YMSM stratified by race. HIV prevalence estimates and HIV transmissibility estimates were derived using local surveillance and census data as well as data shown in Table 26 and Figure 10.

Table 27: Parameter estimates of HIV prevalence at time-step t=0 and HIV transmissibility of female and male partners in one-night-partnerships and outside-partnerships of YMSM stratified by race.

| **Parameter** | **Value** | | **Source** |
|---|---|---|---|
| | **Male** | **Female** | |
| Outside HIV prevalence[a] at time t=0[c]: $p_{out,race,0.}$ | | | |

| Parameter | Value | | Source |
|---|---|---|---|
| | Male | Female | |
| Blacks, $p_{out,Blacks,0}$ | 29.14% | 0.40% | 11,17-19,40 |
| Latinos, $p_{out,Latinos,0}$ | 8.6% | 0.064% | |
| Whites, $p_{out,Whites,0}$ | 10.57% | 0.0158% | |
| Others, $p_{out,Others,0}$ | 17.21% | 0.0489% | |
| HIV transmissibility[b] of HIV+, $c_{out,.}$ | | | |
| Blacks, $c_{out,Blacks}$ | 2.249 | 1.590 | 11,17-19 |
| Latinos, $c_{out,Latinos}$ | 2.153 | 1.456 | |
| Whites, $c_{out,Whites}$ | 1.083 | 1.663 | |
| Others, $c_{out,Others}$ | 1.744 | 1.611 | |

[a]The analysis of the age distribution of YMSM partners in the Crew 450 study shows that 6.37% of all partners of YMSM were of age 30-39. Since only few partners of YMSM are 40 years and older we assume 93.63% of all YMSM partners to be younger than 30. Thus, the outside HIV prevalence and other age-related estimates are weighted averages of estimates for 18-29 year old and 30-39 year old MSM or women in Chicago. For women, estimates within each age-group stratified by race are a weighted average of estimates of IDU and heterosexual women where weights were determined using the fraction of HIV+ women who got infected through heterosexual contact and IDU respectively.
[b] Transmissibility of HIV+ sex-partners of YMSM in one-night-partnerships or outside-partnerships. For comparison the transmissibility of HIV+ in the asymptomatic period $d_A$ without treatment is $c_{A\text{-No Art}}=1$. For details about calculation of transmission factors see section 2.2.3 in SDC 3.
[c]For update of outside HIV prevalence for t>0 see section SDC 3.2.2.4

We calculate the HIV transmissibility $c_{out,i,j}$ of a HIV positive sex-partner of race i and sex j of a YMSM in an one-night-partnership or outside-partnership using the following formula

$$c_{out,i,j} = a_{i,j}(c_P * d_P * f_{T,i,j} + c_{A-No\,Art}(1 - d_P * f_{T,i,j})) + (1 - a_{i,j})(t_{i,j} * c_{T,i} + (1 - t_{i,j})c_{NT,i}). \quad (38)$$

The HIV transmissibility $c_{T,i}$ of a HIV positive of race i receiving treatment is

$$c_{T,i} = f_{FS,i}c_{FS,i} + (1 - f_{FS,i})\, c_{PS,i} \quad (39)$$

with the HIV transmissibility $c_{FS,i}$ of a HIV positive of race i receiving treatment who is fully suppressed is

$$c_{FS,i} = (c_{A-No\,ART} * (d_{A-No\,ART,i} - 0.5/f_{T,i,j}) + c_{A-ART,FS} * d_{A-ART,FS,i} + c_{AIDS-ART,FS} * d_{AIDS})/(d_{A-No\,ART,i} - 0.5/f_{T,i,j} + d_{A-ART,FS,i} + d_{AIDS}) \quad (40)$$

and the HIV transmissibility $c_{PS,i}$ of a HIV positive of race i receiving treatment who is partially suppressed is

$$c_{PS,i} = (c_{A-No\ ART} * (d_{A-No\ ART,i} - 0.5/f_{T,i,j}) + c_{A-ART,PS} * d_{A-ART,PS,i} + c_{AIDS-ART,PS} * d_{AIDS})/(d_{A-No\ ART,i} - 0.5/f_{T,i,j} + d_{A-ART,PS,i} + d_{AIDS}). \quad (41)$$

The HIV transmissibility $c_{NT,i}$ of a HIV positive not receiving treatment is

$$c_{NT,i} = (c_{A-No\ ART} * (d_{A-No\ ART} - 0.5/f_{T,i,j}) + c_{AIDS-No\ ART} * d_{AIDS})/(d_{A-No\ ART} - 0.5/f_{T,i,j} + d_{AIDS}), \quad (46)$$

where

i: race of partner in one-night-partnership or outside-partnership, $i \in \{Black, Latino, White, Other\}$

j: sex of partner in one-night-partnership or outside-partnership, $j \in \{female, male\}$

$a_{i,j}$: fraction of HIV positive with race i and sex j who are unaware of their infection. In 2011, 34% of Black MSM, 19.4% of White MSM, 39% of Latino MSM and 31.9% of Other MSM (weighted average of other 3 races) in Chicago [11] were not aware of their HIV infection (fractions are weighted averages of estimates of 18-29 and 30-39 year olds). For women, 27.8% of Blacks, 31.1% of Whites, 23.4% of Latinos, and 29.7% of Others in Chicago were unaware of their HIV infection [17,19].

$f_{T,i,j}$: Fraction of individuals at risk of race i and sex j who have been tested at least once throughout the last year. Thus, we assume that all individuals at risk get tested on average every $1/f_{T,i,j}$ years. We also assume that the duration of a HIV infection of a newly infected is uniformly distributed within the time interval $[0,1/f_{T,i,j}]$. In 2011, 71.1% of Black YMSM, 52.8% of White MSM, 55.6% of Latino MSM and 57.3% of Other MSM (weighted average of other 3 races) in Chicago were tested at least once throughout the last year [11]. For women, 38.2% of Blacks, 40.8% of Whites, 34.8% of Latinos and 39.7% of Others in Chicago were tested at least once throughout the last year [17-19].

$t_{i,j}$: fraction of HIV positives of race i and sex j receiving ART. In 2011, 83.9% of HIV positive Black MSM, 100% of HIV positive White MSM, 81.8% of HIV positive Latino MSM and 90.6% of HIV positive Other MSM (weighted average of other 3 races) received ART in Chicago [11]. For women, 55.7% of HIV positive Blacks, 55.9% of HIV positive Whites, 55.4% of HIV positive Latinos and 55.8% of HIV positive Others received ART in Chicago [17-19].

$f_{FS,i}$: fraction of HIV positive individuals of race i who receive ART who are fully suppressed. For details see section 2.1.2.3 in SDC 3.

$c_P$, $c_{\mathbf{A\text{-}ART,FS}}$, $c_{A\text{-}ART,PS}$, $c_{\mathbf{AIDS\text{-}ART,FS}}$, $c_{\mathbf{AIDS}\text{-}ART,PS}$, $c_{AIDS\text{-}No\ ART}$, $c_{A\text{-}No\ Art}$: HIV transmissibility coefficients of HIV positive for different stages of HIV infection, with/without treatment and partially or fully suppressed. For details see sections 2.1.1 and 2.1.2 in SDC 3.

$d_P$, $d_{AIDS}$, $d_{A\text{-}No\ Art,i}$, $d_{A\text{-}ART,PS,i}$, $d_{A\text{-}ART,FS,i}$: duration of HIV infection stages stratified by race, treatment and full or partial suppression. For details see section 2.1.1 in SDC 3.

## 2.2.4 Update of outside HIV prevalence over time

YMSM can form one-night partnerships and outside partnerships with partners who are not YMSM. We determine the likelihood that these partners are infected with HIV using the HIV prevalence among older MSM in Chicago (i.e., outside MSM) in case of male partners and the HIV prevalence among women in Chicago in case of female partners. Because of the large number of YMSM aging out of the YMSM population age 16 to 21.8 years and aging into the outside MSM population over the simulated time horizon of 15 years we assume that changes in the HIV prevalence in this target population will impact the outside MSM HIV prevalence, i.e. the HIV prevalence among older MSM age 21.8 to 39 years in Chicago (see section SDC 3.2.2.3 for a discussion of the outside MSM HIV prevalence). Because we do not know about the size of the female population engaging in sexual activity with YMSM as well as the size of the corresponding heterosexual network we assume that changes in the HIV prevalence of the YMSM population do not significantly impact the HIV prevalence among females in Chicago over time.

In the following, we apply the simple principle that the number of HIV infected within the population of MSM age 21.8 to 39 years per time-step equals the number of HIV infected YMSM entering this population per time-step (i.e., inflow) minus the number of HIV infected MSM age 21.8 to 39 years leaving this population per time-step (i.e., outflow) plus the number of secondary infections caused by already infected MSM age 21.8 to 39 in this age-group per time-step (i.e., additional infections), in order to update the outside HIV prevalence $p_{out,r,t}$ among MSM age 21.8 to 39 years in Chicago stratified by race r at time-step t:

$$p_{out,race,t} = 0.9363\, p_{out,r,t,21.8-29} + 0.0637\, p_{out,r,t,30-39} \qquad (42)$$

where

0.96363 and 0.0637: constants denoting the fractions of YMSM outside partners who are 21.8 to 29 years old and 30 to 39 years old respectively, see also Table 27

$p_{out,r,t,21.8-29}$: HIV prevalence among outside MSM age 21.8 to 29 years stratified by race r with $r \in \{Black, Latino, White, Other\}$ at time-step t.

At time-step t=0[11] : $p_{out,Black,0,21.8-29}$=0.279, $p_{out,Latino,0,21.8-29}$=0.081, $p_{out,White,0,21.8-29}$=0.10, $p_{out,Other,0,21.8-29}$=0.1718, and

$$p_{out,r,t,21.8-29} = \frac{i_{r,t,21.8-29}}{n_{MSM,r,21.8-29}}, \text{ t>0} \qquad (43)$$

$i_{r,t,21.8-29}$: number of HIV infected age 21.8 to 29 years of race r at time-step t

$n_{MSM,r,21.8-29}$: population size of MSM age 21.8 to 29 years of race r. For details see SDC 4.1

$p_{out,r,t,30-39}$: HIV prevalence among outside MSM age 30 to 39 years at time t stratified by race r.

At time-step t=0[11]: $p_{out,Black,0,30-39}$=0.474, $p_{out,Latino,0,30-39}$=0.16, $p_{out,White,0,30-39}$=0.189, $p_{out,Other,0,30-39}$=0.2413, and

$$p_{out,r,t,30-39} = \frac{i_{r,t,30-39}}{n_{MSMr_{,,}30-39}}, \text{ t>0} \qquad (44)$$

$i_{r,t,30-39}$: number of HIV infected age 30 to 39 years of race r at time-step t

$n_{MSM,r,30-39}$: population size of MSM age 30 to 39 years of race r. For details see SDC 4.1

***Number of HIV infected outside MSM age 21.8 to 29 years of race r at time-step t: $i_{r,t,21.8-29}$***

We determine the number of HIV infected age 21.8 to 29 years of race r at time-step t by

$$i_{r,t,21.8-29} = p_{out,r,t-1,21.8-29}\ n_{MSM,r,21.8-29}\left(1 + R_{0,HIV,r,21.8-29}\right) + n_{r,21.8,t}\ p_{r,t}\ -\ n_{r,21.8,t}\, p_{out,r,t-1,21.8-29} \qquad (45)$$

In (45), $n_{r,21.8,t}\ p_{r,t}$ denotes the number of infected YMSM age 21.8 years and older of race r at time-step t who age out of the YMSM population and age into the outside MSM population age 21.8 to 29 years. $n_{r,21.8,t}\ p_{out,r,t,21.8-29}$ denotes the number of infected MSM of race r within the group of 21.8 to 29 year old MSM who age-out of this MSM group and age into the outside MSM group age 30 to 39 years at time-step t. We assume the HIV prevalence of those aging out of the MSM group age 21.8 to 29 years to be the same as the overall HIV prevalence in this age group. Further, we assume the population size of the MSM group age 21.8 to 29 years to be constant over time. Thus, the number of MSM aging out of this MSM group at time-step t always equals the number of YMSM aging into this MSM group at time-step t being $n_{r,21.8,t}$.

$p_{out,r,t-1,21.8-29}\ n_{MSM,r,21.8-29}\left(1 + R_{0,HIV,r,21.8-29}\right)$ denotes the number of HIV infected in this age-group at time t-1 plus the number of secondary infections infected by the already infected (1st generation) in time-step t-1 in this age-group which are determined using the HIV basic-reproduction number $R_{0,HIV,r,21.8-29}$ per time-step t. We determine the HIV rate of secondary infections $R_{0,HIV,r,21.8-29}$ per time-step t for this age-group using the average number of YMSM aging into the MSM group age 21.8 to 29 years $n_{r,21.8}$, the race stratified HIV prevalence estimates of MSM ages 18 to 29 years in 2011 in Chicago[11] $p_{out,r,18-29}$ and the Crew 450 HIV prevalence estimates adjusted to the simulated population size $p_{YMSM,r,t=0}$. Assuming the system in steady state we obtain the HIVrate of secondary infections per time-step t by

$$R_{0,HIV,r,21.8-29} = \frac{n_{r,21.8}\left(p_{out,r,21.8-29} - p_{YMSM,r,t=0}\right)}{p_{out,r,18-29}\ n_{MSM,r,21.8-29}}. \qquad (46)$$

**Table 28** shows the estimates for $R_{0,HIV,r,21.8-29}$.

***Number of HIV infected outside MSM age 30 to 39 years of race r at time-step t: $i_{r,t,30-39}$***

We use the same approach as in case of $\boldsymbol{i_{r,t,21.8-29}}$ in order to determine the number of HIV infected outside MSM age 30 to 39 years of race r at time-step t, i.e.

$$i_{r,t,30-39} = p_{out,r,t-1,30-39}\ n_{MSM,r,30-39}\left(1 + R_{0,HIV,r,30-39}\right) + n_{r,21.8,t}\ p_{out,r,t,21.8-29} - n_{r,21.8,t}\ p_{out,r,t-1,30-39} \quad (47)$$

Again, we assume that $n_{MSM,r,30-39}$, the size of the total population of the outside MSM group age 30 to 39, remains constant over time and thus the number of MSM age 21.8 to 29 years aging out of the age group 21.8 to 29 years and aging into the outside MSM group age 30 to 39 years per time-step equals the number of outside MSM age 30 to 39 who age out of this age group in each time step. Assuming the system to be in the steady state, we determine the HIV rate of secondary infections $R_{0,HIV,r,30-39}$ per time-step t for this age-group using the average number of YMSM aging into the MSM group age 21.8 to 29 years $n_{r,21.8}$, the race stratified HIV prevalence estimates of MSM age 18 to 29 years in 2011 in Chicago[11] $p_{out,r,18-29}$ and the estimates of the race stratified HIV prevalence of MSM age 30 to 39 years in Chicago $p_{out,r,30-39}$:

$$R_{0,HIV,r,3039} = \frac{n_{r,21.8}\left(p_{out,r,30-39} - p_{out,r,18-29}\right)}{p_{out,r,30-39}\ n_{MSM,r,30-39}}. \quad (48)$$

**Table 28** shows the estimates for $R_{0,HIV,r,30-39}$ per time-step t.

| Race | $R_{0,HIV,r,21.8-29}$ | $R_{0,HIV,r,30-39}$ |
|---|---|---|
| **Black** | 0.0310 | 0.0017 |
| **Latino** | 0.0026 | 0.0019 |
| **White** | 0.0039 | 0.0018 |
| **Other** | 0.0032 | 0.0012 |

Table 28: HIV rate of secondary infections per time step t (0.5 months) for MSM age groups 21.8 to 29 years and 30 to 39 years.

The simulated trajectory of the outside MSM HIV prevalence for both age groups and the overall case is shown in Figure 24 section SDC 7.1.

Given the per time-step HIV rate of secondary infections shown in Table 28, the corresponding HIV basic reproduction numbers over a ten year horizon for the outside MSM age group 21.8 to 29 years would be 0.74 for Black MSM, 0.61 for Latino MSM, 0.93 for White MSM, and 0.77 for Other MSM. The corresponding HIV basic reproduction numbers (i.e., number of secondary infections caused by already infected) over a ten year horizon for MSM age group 30 to 39 years would be 0.42 for Black MSM, 0.45 for Latino MSM, 0.34 for White MSM, and 0.30 for Other MSM. We observe that with increasing age the HIV basic reproduction number decreases which aligns with the findings of McCormick et al.[41] who showed using a mathematical model that the HIV basic reproduction number declines over the duration of the infection. McCormick et al.[41] estimated the HIV basic reproduction number within the first ten years of the infection to be 1.9 for HIV infected without treatment and 1.4 for HIV infected with treatment as well as the HIV basic reproduction number within years 10 to 20 of the infection to be 0.6 in case of HIV infected without treatment and 0.4 for HIV infected with treatment. Considering the case of a YMSM who got infected at age 20 and not taking into account the HIV basic reproduction number attributable to the primary infection period which is estimated to be 0.36[41], our estimate of 0.93 of the HIV basic reproduction number for White MSM age 21.8 to 29 years who all receive treatment in our model is comparable to the HIV basic reproduction number estimate of McCormick et al. for HIV infected receiving treatment in the first ten years. Additionally, our HIV basic reproduction number estimates for MSM age 30 to 39 years aligns with the estimates of McCormick et al. for HIV infected with treatment within year 10 to 20 of their HIV infection. Further, the study of McCormick et al. was published in 2007. Thus, advancements in antiretroviral therapy and the scaling of treatment coverage and prevention efforts which took place until 2011 in Chicago[11] are likely to decrease the HIV basic reproduction number in comparison to estimate of 2007.

## 2.3 Interactions with other diseases

Based upon the opinion of our experts we assume that neither the HIV infection nor ART treatment impacts transmissibility and susceptibility of NG and CT.

# 3. Neisseria Gonorrhea (NG) and Chlamydia trachomatis (CT)

There is biological evidence that ulcerative and non-ulcerative sexually transmitted infections (STI) increase the risk of HIV transmission[42-45]. Among the few modelling studies which quantify the impact of STIs on the HIV epidemic, Chesson and Pinkerton[46] estimated the percentage of heterosexual HIV infections attributable to the STIs Syphilis, CT, NG and genital herpes to be within the range of 6.3% to 12.6% of all HIV infections in the US in 1996. 73% of the HIV infections attributable to the above mentioned STIs were attributable to NG and CT[46].

We model the simultaneous spread of HIV and the STIs NG and CT because NG and CT are hypothesized to significantly contribute to the increased risk and racial disparities in HIV among YMSM[5]. Further, YMSM in the Crew 450 study were tested for urethral NG and CT and the urethral NG and CT prevalence levels in the beginning of the study (T2) were 4.0% and 5.3% respectively. It is our goal to cover the most important aspects of the transmission of CT and NG as well as their influence on HIV susceptibility and HIV transmissibility in order to quantify the impact of the STIs NG and CT on HIV transmission among YMSM.

To the best of our knowledge we do not know of any published mathematical models which study the simultaneous spread of HIV, CT and NG among (Y)MSM as well as include urethral and rectal infections as mode of transmission. We informed and parameterized our model of simultaneous HIV, NG and CT spread using available data of the Crew 450 study, expert opinion, findings of epidemiological studies[42-45], and modelling approaches of the spread of HIV and CT among MSM[47], of HIV and NG among a heterosexual population[48] and models studying solely the transmission of either CT[49-52] or NG[49].

# 3.1 Course of NG and CT infection

## 3.1.1 Disease progression model

Both NG and CT are 'SIS'- type diseases, i.e. susceptible individuals get infected, recover from the infection and become susceptible again. We consider the course of NG and CT infections to be similar: individuals get infected through sexual contact; the infection is either symptomatic or asymptomatic; a fraction of individuals with symptoms will seek treatment and will cease sex. Asymptomatic infected and symptomatic infected who do not cease sex can transmit the disease to other individuals; the duration of treatment for symptomatic NG and CT infections are similar; individuals recover from symptomatic and asymptomatic infections and become susceptible again. Assuming NG and CT to have a similar course of the infection we follow the modeling approach of Chen et al.[49], i.e. we assume the same disease progression model for NG and CT but parameterize the model differently for each disease. Figure 11 shows a schematic illustration of the disease progression model and Table 29 shows the parameter values of the disease progression model for both NG and CT.

### 3.1.1.1 Site of infection

Because we focus on the most important aspects of NG and CT transmission and their impact on HIV we assume NG and CT infections in our model to be either urethral or rectal. We do not consider dual-site infections in our model (i.e., both urethral and rectal) because of the low prevalence of dual site-infections among NG or CT infections (5.6% and 6.2%[53]). Further, we do not consider pharyngeal infections in our model. Kent et al.[53] studied NG and CT among MSM in San Francisco and estimated the prevalence of pharyngeal infections among all CT infections to be 6.6%. Similarly, a recent study among 17898 MSM in the US found the percentage of pharyngeal infections among all types of CT infections to be 6.4%[54]. In

case of NG, pharyngeal infections are more prevalent

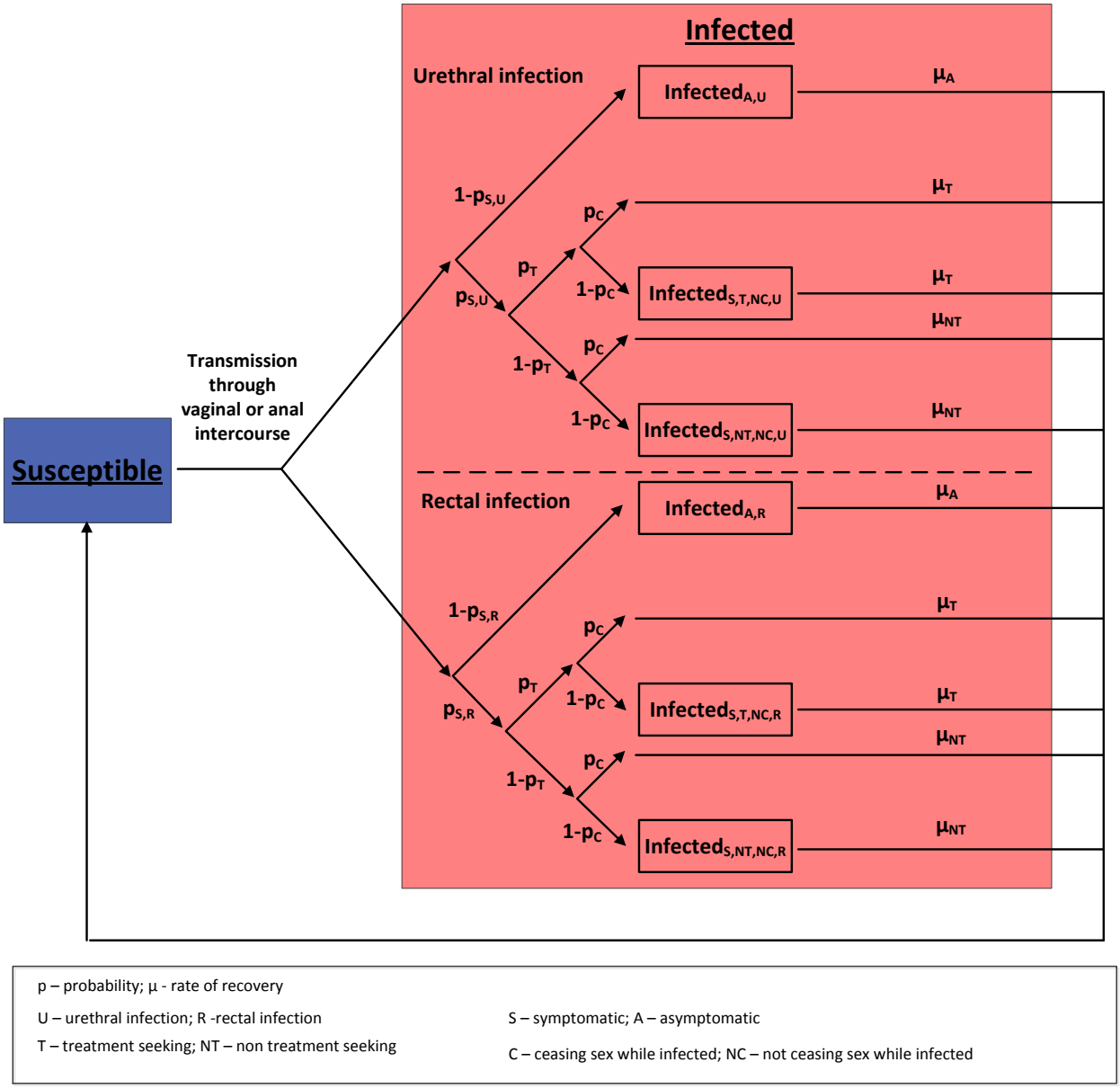


Figure 11: Schematic illustration of the disease progression model of Neisseria Gonorrhea (NG) and Chlamydia (CT). Infected are only highlighted (box) if sexually active, i.e. able to spread the disease. Asymptomatic infected do not receive treatment except when being tested positive (not shown in this diagram, see discussion in section 3.1.2 in SDC 3).

compared to CT (36.4% vs. 6.6% [53], 26.4% vs. 6.4% [54]). However, we do not consider pharyngeal infections in our model because of the low prevalence in case of CT and the fact that NG pharyngeal infections in MSM have a 9 times lower bacterial load compared to rectal infections[55] as well as there have been no studies reported which quantify the transmission risk of a pharyngeal infected to the male

urethra[55]. Additionally, our model does not incorporate oral HIV infections as a mode of HIV transmission (see discussion in section 2 in SDC 3) and less is known about the impact of pharyngeal NG infections on HIV transmission among MSM[55].

We do not model dual-site infections for one disease. However, a YMSM can be infected at two different sites with two different diseases.

### 3.1.1.2 Asymptomatic vs. symptomatic infections

We assume a NG or CT infection to either be symptomatic or asymptomatic, i.e. in our model asymptomatic infections won't become symptomatic over the course of the asymptomatic infection[49,50]. We use the estimates of Kent et al.[53] to determine the fractions of urethral ($p_{S,U}$, $p_{A,U}$=1- $p_{S,U}$) and rectal NG and CT infections ($p_{S,R}$, $p_{A,R}$=1- $p_{S,R}$) to be symptomatic or asymptomatic (see Table 29).

### 3.1.1.3 Ceasing of sexual activity

NG or CT infected who have a symptomatic infection are likely to cease their sexual activity until they receive treatment[46,56]. Kramer et al.[56] estimated the fraction of individuals with symptomatic STI's who cease their sexual activity to be 0.8. In absence of data we assume that 80% of all NG and CT infected individuals who have a symptomatic infection do cease sex throughout their entire infection regardless of the site of infection and treatment status.

### 3.1.2 Treatment and recovery

We assume individuals to recover from their NG and CT infections with a constant rate $\mu$ which is the inverse of the average duration d of a NG or CT infection, i.e. $\mu = \frac{1}{\mathrm{d}}$. The average duration of a NG or CT infection is dependent on the type of infection (i.e., symptomatic or asymptomatic), the site of infection (i.e., rectal or urethral) and the treatment status in case of symptomatic infections. Table 29 shows the average duration for the different types of NG and CT in our model. Estimates of the average treatment durations do incorporate the incubation period.

In their model of HIV and CT transmission Vriend et al.[47] estimated the duration of an urethral asymptomatic CT infection to be 240 days. In absence of data for the duration of urethral asymptomatic NG infections we estimate their duration to be the same as the duration of urethral asymptomatic CT infections. Reviewing modelling studies of CT transmission Davies et al.[52] suggest the duration of an asymptomatic CT infection to be 497 days. Because the underlying studies in the review of Davies et al.[52] focused either on vaginal infections in females or urethral infections in males as well as taking into account that men are hypothesized to have a shorter duration of an asymptomatic infection[57] we assume the duration of a rectal asymptomatic CT infection to be 497 days. In absence of estimates for the duration of rectal asymptomatic NG infections as well as studies showing a lower prevalence of rectal NG compared to rectal CT among MSM[54,58] (see also discussion in SDC 5), we assume the duration of a rectal asymptomatic NG infection to be 300 days using a shorter estimate which is oriented at the estimate of the duration of an asymptomatic CT infection of Geisler et al.[57].

Symptomatic NG or CT infected individuals do not always seek treatment because the infected might perceive the symptoms not to be severe, might have problems with the access to clinical care, or might have the belief that symptoms will vanish quickly without treatment[59]. Table 29 shows the parameter estimates for the fraction $p_T$ of individuals with symptomatic infections who seek for care. Due to lack of

data we assume the same parameter values for the fractions of urethral and rectal symptomatic infected who seek for care.

We assume individuals to only receive treatment if they seek for treatment or if they get tested positive for NG or CT. If a NG or CT positive individual gets tested for STIs we assume that the infection is detected regardless of the type of infection (i.e., asymptomatic and symptomatic). However, a rectal NG or CT infection can only get detected if tested rectally for STIs. A positive test result implies immediate treatment and recovery of the infection after completion of treatment. In general, we assume the duration of treatment and recovery to be 13 days for NG and 14 days for CT respectively[47,49]. Further, we assume the mortality risk associated with NG or CT to be negligible.

Table 29: Parameter values of disease progression model for Neisseria Gonorrhea (NG) and Chlamydia Trachomatis (CT).

| **Parameter** | **Value** | | **Source** |
|---|---|---|---|
| | **NG** | **CT** | |
| Fractions of infections being symptomatic | | | |
| urethral, $p_{S,U}$ | 0.9 | 0.58 | [53] |
| rectal, $p_{S,R}$ | 0.16 | 0.14 | [53] |
| Fraction of individuals with symptomatic infection (urethral and rectal) who | | | |
| seek for treatment, $p_T$ | 0.9 | 0.85 | [59] |
| cease sexual activity, $p_C$ | 0.8 | 0.8 | [46] |
| Average duration of infection[a] | | | |
| asymptomatic[b], $d_A$ | | | |
| urethral | 240 | 240 | [47,49] |
| rectal | 300 | 497 | [52] |
| symptomatic (urethral and rectal) | | | |
| treatment, $d_T$ | 13 | 14 | [47,49] |
| no treatment, $d_{NT}$ | 185 | 180 | [46,60] |

[a] Average duration of infection incorporates incubation period.

[b] Treatment for asymptomatic infections only initiated if tested. We assume the duration of the treatment for asymptomatic infections to be the same as the duration of a symptomatic infection with treatment, $d_T$. In absence of

reliable estimates we assume the duration of a urethral asymptomatic NG infection to be of the same length as a urethral asymptomatic CT infection and estimate the duration of a rectal asymptomatic NG infection to be 300 days.

# 3.2 NG and CT transmission

## 3.2.1 Mode of NG and CT transmission

In our model, NG and CT are transmitted through penile-vaginal, female-male anal intercourse and male-male anal intercourse. As discussed in section 3.1 in SDC 3 oral transmission is not considered as a mode of transmission.

Based upon the discussion with our experts we assume that the infected site of the NG or CT infected has to match with the potential infection site of the NG or CT susceptible. For example, an urethral infected person can't transmit NG or CT to a susceptible while being receptive in anal intercourse and a rectal NG or CT infected can't transmit CT or NG while being insertive in anal intercourse.

## 3.2.2 NG and CT transmission risk per sex-act

### 3.2.2.1 NG and CT baseline transmission risk

Similar to the baseline transmission risk of HIV we define the baseline NG or CT transmission risk per sex-act as the transmission risk per sex-act where the NG or CT infected does not receive treatment (both symptomatic and asymptomatic infections) and both partners having unprotected sex.

#### *NG and CT female-to-male transmission*

In 1978, Hooper et al.[61] estimated for NG the female-to-male penile-vaginal baseline transmission to be 19% per sex-act. For CT, Althaus et al.[62] estimated the female-to-male penile-vaginal baseline transmission risk to be 9.5%.

#### *Male-to-male (anal) transmission*

Estimates of epidemiological studies for the NG and CT baseline transmission risk per male-male anal sex-act were not available. Most of the parameter estimates for the NG and CT baseline transmission risk of male-male anal intercourse used in mathematical models are either calibrated estimates of mathematical models[47,49,60] or estimates of female-to-male and male-to-female transmission risks[47,50,63].

Since models and thus the calibrated estimates vary across study populations we use estimates of female-to-male and male-to-female transmission risks of epidemiological studies to estimate the baseline transmission risk per male-male anal sex-act.

Because HIV transmission risk for male-male insertive anal intercourse is higher than for penile-vaginal intercourse we assume the NG and CT baseline transmission risk of male-male insertive anal intercourse to be higher than the NG and CT baseline transmission risk of penile-vaginal intercourse. Thus, we assume the CT baseline transmission probability per male-male insertive anal sex-act to be 32%[64,65] which equals the highest estimate of the CT baseline transmission risk per penile-vaginal sex-act we found in literature. Further, we estimated the baseline transmission risk for one episode of receptive male-male anal intercourse to be 40% for CT[64,65] and 60% for NG[66] which equals the male-to-female NG and CT transmission risks per penile-vaginal sex-act.

***Circumcision***

Findings of randomized controlled trials show that circumcision status is not associated with a significant protective effect against the acquisition of NG and CT[67,68]. Thus, we assume that circumcision to not decrease the baseline transmission risk of NG and CT.

### 3.2.2.2 NG and CT treatment

We assume treatment always to be fully effective for both NG and CT infections[69-71]. Because treatment of NG and CT is assumed to only be fully effective after 7 days[69] we assume for the matter of simplicity that a NG or CT infected who receive treatment has the same infectivity throughout the whole course of his infection, i.e. the infectivity remains the same before and after receiving treatment. Thus, treatment only shortens the duration of a symptomatic infection of NG or CT to 13 respectively 14 days in our model but does not reduce the level of transmissibility. Once an asymptomatic NG or CT infected individual tests positive, he will receive immediate treatment and be cured within 7 days[69].

#### 3.2.2.3 Condom-usage

We assume the efficacy of condoms to prevent NG and CT transmission to be 73%[72] and 70%[73] per penile-vaginal and female-male and male-male anal sex-act respectively. The lower efficacy of condoms in case of NG and CT compared to HIV is attributable to the higher transmission risk of NG and CT in case of condom errors.

### 3.2.3 Calculation of NG and CT transmission risk for different partnership-types

Given the match of the body sites of the infected and susceptible we calculate the NG and CT transmission risk for each sex-act by multiplying the sex-act specific baseline NG and CT transmission risk with the corresponding risk reduction factor for condom-usage in case of protected sex.

Similar to HIV, the NG and CT status of sex-partners in one-night-partnerships and outside-partnerships are unknown. We estimate the likelihood that a female sex partner is infected with NG or CT using the 2011Illinois CDC estimates for the prevalence of NG and CT among 16-24 year old women ($p_{women,NG}$= 2.7%[21] and $p_{women,CT}$= 9.5%[20]). The 2011 MSM surveillance report of the CDPH estimated the prevalence of NG and CT among Chicagoan MSM to be $p_{MSM,NG,Black}$=12.2% and $p_{MSM,CT,\ Black}$=6.7% for Blacks, $p_{MSM,NG,Latinos}$=15.5% and $p_{MSM,CT,\ Latinos}$=10.3% for Latinos and $p_{MSM,NG,Whites}$=5.7% and $p_{MSM,CT,\ Whites}$=7.3% for Whites[11]. We use the weighted average of Blacks, Latinos and Whites to determine the NG and CT prevalence of Others being $p_{MSM,NG,Others}$=10.05% and $p_{MSM,CT,\ Others}$=7.87%. Because the CDPH estimates are not stratified by urethral and rectal infections. we apply the findings of Kent et al.[53] to estimate the fraction of rectal infected among the NG and CT infected MSM in Chicago to be $p_{R,NG}$=58.5% and $p_{R,CT}$=64.7% respectively[58].

## 3.2.4 Update of the outside NG and CT prevalence over time

Similar to the update of the outside MSM HIV prevalence we update the outside MSM NG and CT prevalence for outside MSM age 21.8 to 39 years over the simulated time horizon to account for NG and CT system dynamics among the YMSM population. Given the initial empirical estimates of the outside MSM NG ($p_{MSM,NG,race}$) and CT ($p_{MSM,CT,race}$) prevalence discussed in section SDC 3.3.2.3 at time-step t=0, we update $p_{MSM,NG,race,t}$ and $p_{MSM,CT,race,t}$ for each race at each time-step t>0 similar to (43) and (45) in case of the update of the outside MSM HIV prevalence, i.e. for $p_{MSM,NG,race,t}$

$$p_{MSM,NG,r,t} = \frac{iNG_{r,t,21.8-39}}{n_{MSM,r,21.8-39}}\text{, t>0} \qquad (49)$$

where

$iNG_{r,t,21.8-39}$: number of NG infected age 21.8 to 39 years of race r at time-step t

$n_{MSM,r,21.8-39}$: population size of MSM age 21.8 to 29 years of race r.

$$n_{MSM,r,21.8-39} = n_{MSM,r,21.8-29} + n_{MSM,r,30-39} \qquad (50)$$

For details see SDC 4.1

***Number of NG infected outside MSM age 21.8 to 39 years of race r at time-step t: $iNG_{r,t,21.8-39}$***

We determine the number of NG infected age 21.8 to 39 years of race r at time-step t by

$$i_{r,t,21.8-39} = p_{MSM,NG,r,t}\ n_{MSM,r,21.8-39}\left(1+\beta_{NG,MSM,r}\right) + n_{r,21.8,t}\ p_{YMSM-21.8,NG,r,t} - n_{r,21.8,t}\ p_{MSM,r,t-1} \quad (51)$$

Where

$p_{YMSM-21.8,NG,r,t}$: NG prevalence among aging-out YMSM age 21.8 or older at time-step t

$\beta_{NG,r}$: NG and CT are SIS-type (i.e., susceptible-infected-susceptible) diseases and the NG rate of secondary infection $R_{0,NG,MSM}$ defines the secondary number of infected per time-step an NG infected MSM causes. $\beta_{NG,MSM,r}$ denotes the additional infectivity of NG infected outside MSM (1st generation) with $\beta_{NG,MSM,r} \geq 0$ and thus $R_{0,NG,MSM} = 1 + \beta_{NG,MSM,r}$. We determine $\beta_{NG,MSM,r}$ using

$$\beta_{NG,MSM,r} = \frac{n_{r,21.8}\left(p_{MSM,r,21.8-39} - p_{YMSM,NG,r,0,}\right)}{p_{MSM,r,21.8-39}\ n_{MSM,r,21.8-39}}. \qquad (52)$$

where

$p_{YMSM,NG,r,0,}$: NG prevalence among YMSM at beginning of simulation stratified by race

Table 30 shows the estimates of the additional infectivity parameters stratified by race for NG and CT. The update of the outside MSM CT prevalence uses CT-specific values in (49)-(52).

| Race | $\beta_{NG,r}$ | $\beta_{CT,r}$ |
|---|---|---|
| Black | 0.00045 | 0 |
| Latino | 0.00060 | 0 |
| White | 0 | 0 |
| Other | 0 | 0 |

Table 30: Additional infectivity parameter $\beta_{NG,r}$ and $\beta_{CT,r}$ per time-step t stratified by race for NG and CT update of outside NG and CT MSM prevalence. $\beta_{NG,r} \geq 0$ and thus the rate of secondary infection $R_{0,NG,MSM} = 1 + \beta_{NG,r}$.

Given the additional infectivity parameter for NG and CT per time-step t shown in Table 30 and the average durations of NG and CT infections shown in Table 29 (section SDC 3.3.1.2), the resulting basic reproduction numbers for NG and CT among outside MSM are close to 1, i.e. the NG and CT infected outside MSM is expected to infected another outside MSM over the course of his infection.

Simulated trajectories of the outside NG and CT prevalence are shown in **Figure 25** in section SDC 7.1.

## 3.3 Interactions with other diseases

In our model of simultaneous HIV, NG and CT spread we assume NG and CT to impact HIV susceptibility and transmissibility. Based upon expert opinion we do not assume HIV and ART to impact susceptibility and transmissibility of NG and CT as well as we assume NG and CT to not impact the susceptibility and transmissibility of each other. It remains unknown whether and to which extend HIV and ART impacts NG and CT and whether a pre-infection with one bacterium impacts susceptibility and

transmissibility of the other. The influence diagram in Figure 12 shows a schematic illustration of the interaction of the three diseases HIV, NG and CT.

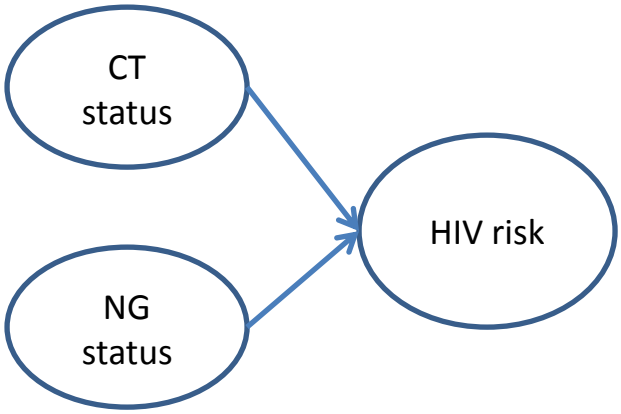


Figure 12: Influence diagram of HIV, NG and CT in the simulation model.

### 3.3.1 Impact of NG and CT on HIV susceptibility and transmissibility

There is biological evidence that NG and CT infections increase HIV susceptibility and transmissibility[43-45]. We model the impact of NG and CT on HIV transmission by multiplying the baseline HIV transmission risk per sex-act with a factor corresponding to either an increase in HIV susceptibility (IS) of the HIV negative or an increase in HIV transmissibility (IT) of the HIV positive. We assume the same IS and IT factors for asymptomatic and symptomatic NG or CT infections[47] as well as we assume NG or CT treatment not to impact IS and IT. Whether the increase in the HIV transmission risk is caused by either IS or IT depends on the HIV, NG and CT infection status of both partners engaging in a sex-act[47]. Based upon the HIV, NG and CT status of both partners we classify sex-partnerships at each time a sex-act is happening as either a

- *Partial discordant –infected partnership [(HIV-,I+); (HIV+,I-)]*: One partner is either infected with NG, CT or both (I+) and HIV negative (HIV-,I+) whereas the other partner is HIV positive and not infected with either NG, CT or both (HIV+,I-);

- *Fully discordant-infected partnership [(HIV-,I-); (HIV+,I+)]*: One partner is neither infected with HIV nor CT or NG (HIV-,I-) whereas the other partner is HIV positive and either infected with NG, CT or both (HIV+,I+).
- *Mixture partnership [(HIV-,I+); (HIV+,I+)]:* The mixture partnership describes a mixture of the partial discordant-infected partnership and the fully discordant-infected partnership. In this case we assume that the effects of IS and IT are not additive, i.e. the effect on HIV transmission in this type of relationship is the larger of IS and IT.

Table 31 shows the IS and IT factors corresponding to each partnership-type.

IS and IT factors depend on the I+ status of the HIV negative or HIV positive, i.e. we stratify by the site of infection and the sex-role position of the individual. Jin et al.[44] estimates the factor of IS for HIV negative having a rectal I+ while being receptive to be 1.754 for NG and 1.638 for CT. In absence of available parameter estimates for the factor of IS in case of HIV negatives having a rectal I+ while being insertive we assume based upon expert opinion that there is no increased susceptibility in this case, i.e. IS=1. We also assume IS=1 for HIV negative having an urethral I+ while being receptive. Koblin et al.[45] estimate the IS for HIV negative having a NG or CT infection to be 1.429 and 1.078 respectively. However, Koblin et al.[45] does not stratify by the site of infection. Assuming a larger surface area inflamed in case of rectal infections compared to urethral infections we assume IS to be lower for urethral I+ having insertive sex compared to rectal I+ having receptive sex. Thus, we use the estimates of Koblin et al. for IS where HIV negative having an urethral I+ while being insertive.

Parameter estimates describing the increased transmissibility (IT) of HIV due to rectal and urethral I+ of HIV positive were not available. Thus, we relied on the estimates of Koblin et al.[45] and expert opinion. We assume that there is increased transmissibility of HIV in case of HIV positive having rectal I+ while being receptive. However, we assume IT and IS for rectal I+ to not be symmetric and assume a lower value for IT (rectal I+) compared to the corresponding IS. Further, we assume that IT is non-symmetric in

case of HIV positives having an urethral I+ while being insertive compared to HIV negatives having an urethral I+ while being insertive. To account for the non-symmetry between IT and IS for urethral I+ we assume the ratio of IS to IT among rectal I+ to also apply for urethral I+ which yields the IT factors 1.244 and 1.010 for NG and CT, respectively.

Rotchford et al.[74] observed significant increases in HIV RNA viral load for NG infected HIV positive but not for CT infected HIV positive. Thus, we assume an increased HIV transmissibility for HIV positive having a rectal NG infection while being insertive (IT=1.244). We do not assume an IT for HIV positive having a rectal CT infection while being insertive.

Our model allows for co-infections of HIV, NG and CT. Given a co-infection of NG and CT in either a HIV positive or HIV negative we assume that the effects of increased HIV susceptibility and transmissibility are not additive, i.e. that if an individual is infected with both NG and CT the resulting effect on the increased HIV susceptibility and transmissibility is the larger of the two. We further assume, that there are no differences in IS and IT factors for different sex-types, i.e. IS does not differ for male-male anal intercourse and female-male anal intercourse.

Table 31: Factors of increased HIV susceptibility (IS) and transmissibility (IT) compared to baseline HIV transmission risk due to infection with Neisseria Gonorrhea (NG) and Chlamydia Trichomatis (CT)[a].

| **Factor of increased HIV susceptibility (IS) and transmissibility (IT)[b]** | **Value** | | **Source** |
|---|---|---|---|
| | **NG** | **CT** | |
| ***Partial discordant-infected partnership [(HIV-,I+); (HIV+,I-)][c]*** | | | |
| I+ has rectal infection | | | |
| IS for receptive sex | 1.754 | 1.638 | [44] |
| IS for insertive sex | 1 | 1 | -[c] |
| I+ has urethral infection | | | |
| IS for receptive sex | 1 | 1 | -[c] |
| IS for insertive sex | 1.429 | 1.078 | [45] |
| ***Fully discordant-infected partnership [(HIV-,I-); (HIV+,I+)][c]*** | | | |
| I+ has rectal infection | | | |
| IT for receptive sex | 1.429 | 1.078 | [45] |
| IT for insertive sex | 1.244 | 1 | [44,45,74,75] |
| I+ has urethral infection | | | |
| IT for receptive sex | 1 | 1 | -[c] |
| IT for insertive sex | 1.244[d] | 1.010[d] | [44,45] |
| ***'Mixture'[(HIV-,I+); (HIV+,I+)][c]*** | | | |
| (HIV-;I+) has rectal infection | | | |
| IS for receptive sex | 1.754 | 1.638 | [44] |
| IS for insertive sex | | | |
| (HIV+;I+) has rectal infection | 1.429 | 1.078 | [45] |
| (HIV+;I+) has urethral infection | 1 | 1 | -[c] |
| (HIV-;I+) has urethral infection | | | |
| IS for receptive sex | | | |
| (HIV+;I+) has rectal infection | 1.244[d] | 1 | [44,45,74,75] |
| (HIV+;I+) has urethral infection | 1.429 | 1.078 | [45] |
| IS for insertive sex | 1.429 | 1.078 | [45] |

[a] For explanation about the derivation of the parameter estimates see section 3.3.1 in SDC 3.

[b] In case of intra-event sex-role versatility the same sex-role specific factors for increased HIV transmission risk apply as in case of no intra-event sex-role versatility.
[c] I: NG or CT infection; I+: NG or CT positive; I-: NG or CT negative
[d] No plausible published estimates available. Based upon expert opinion.
[e] Ratio IS to IT for rectal I+ is 0.569 for NG and 0.122 for CT. Values for IT of urethral I+ are 1+0.569*0.429=1.422 for NG and 1+0.122*0.78=1.010 for CT.

# 4. Testing and treatment coverage

## 4.1 Testing

We use data of the CDPH MSM surveillance report[11] to determine the daily probability of a YMSM to get tested for HIV, NG and CT. The surveillance report provides estimates for the percentage of Chicagoan MSM stratified by race and HIV status who got tested at least once for HIV and the STI's NG, CT and Syphilis during 2011. In absence of more detailed data we derive the daily testing probability of a YMSM using the annual testing estimates while assuming the daily testing probability to be constant[4]. Using the daily testing probabilities we calculate the testing probability per time-step in our discrete-time simulation model.

### 4.1.1 Daily HIV/STI testing probabilities

In 2011, 65.8% of Black MSM, 52.8% of White MSM and 55.6% of Latino MSM were tested for HIV at least once per year in Chicago[11]. Using the weighted average of Blacks, Whites and Latinos we determine the percentage of Other MSM who got tested for HIV at least once throughout 2011 to be 57.33%. The corresponding daily HIV testing probabilities are 0.002935 for Blacks, 0.00222 for Latinos, 0.002055 for Whites and 0.00233 for Others. Among Chicagoan MSM who assumed to be HIV negative 60.4% of Blacks, 53.2% of Whites and 46% of Latinos were tested at least once in 2011 for STI's[11]. Again, using the weighted average of the tested fractions of Black, White and Latino MSM we estimate that 53.4% of Other MSM were tested for STI's in the last year. The CDPH data on STI testing did not specify whether urethral, rectal, and or pharyngeal testing was performed. We assume these STI testing data to represent urethral testing data only since urethral testing is much more common among MSM in Chicago compared to rectal and pharyngeal STI testing[54]. Thus, the corresponding daily testing probabilities for urethral STI's are 0.002535 for Black MSM, 0.001687 for Latino MSM, 0.002078 for White MSM and 0.002096 for Other MSM. We estimate the daily testing probability for rectal STI's using the findings of Patton et

al.[54] who estimate the percentage of Chicagoan MSM who got tested for rectal STI's during the last year to be 7.9%. This corresponds to a daily testing probability for rectal STI's being 0.000225 for all MSM. We assume the daily testing probability for rectal and urethral STI's to be independent.

### 4.1.2 Daily STI testing probability for HIV positive

The 2011 CDPH MSM surveillance report estimates the fraction of HIV positive MSM in Chicago who got tested for STI's at least once during the last year to be 67.7% for Blacks, 81.3% for Whites and 90.9% for Latinos. Again, we use the weighted average of the fractions of Blacks, Latinos and Whites to estimate the percentage of HIV positive Other MSM who got tested for STI's at least once during the last year to be 77.0%. The corresponding daily testing probabilities are 0.00309 for Black MSM, 0.00458 for White MSM, 0.00655 for Latino MSM, and 0.004022 for Other MSM. Again, we assume these STI testing probabilities to be for urethral testing only. Patton et al.[54] did not stratify their estimates for the fraction of MSM being tested for rectal STI at least once during the last year by HIV status. Thus, we assume their estimates to represent the testing behavior of HIV negative MSM and apply the ratio of urethral STI testing rates of HIV positive compared to HIV negative stratified by race[11] to obtain the daily testing probabilities for rectal STI testing being 0.000274 for Black MSM, 0.000873 for Latino MSM, 0.000496 for White MSM and 0.000432 for Other MSM.

## 4.2 HIV treatment coverage

Using the available estimates of the CDPH MSM surveillance report[11] we assume the percentage of tested HIV positive MSM who receive ART to be 84% for Blacks, 100% for Whites and 82% for Latinos. For other races ('Others') we assume the same coverage as for Whites[4].

We assume unlimited coverage for STI treatment, i.e. if a NG or CT infected YMSM decides to seek treatment or gets tested positive for STIs he will always receive treatment and recover.

# IV. Supplemental Digital Content 4: Population size, age range, and simulated time horizon

# 1. Age range

We model the transmission of HIV, NG, and CT among YMSM in Chicago age 16 to 21.8 years over time. This age range represents the age of the study participants enrolled in the underlying Crew 450 study across data collection waves T1 and T2 which were used to parameterize the partnership formation and dissolution model. We simulate the partnership formation and dissolution process as well as the HIV, NG, and CT transmission process within this specific age group over the simulated time horizon and thus do not follow up specific YMSM over the total simulated time. Therefore, YMSM get born into the YMSM population (i.e., age in) and age out of this population when they become 21.8 years or older and did not die beforehand.

# 2. Simulated YMSM population size

## 2.1 YMSM population size at t=0

We chose the size and race mix of our simulated study population such that it is representative of the total YMSM population age 16 to 21.8 years in Chicago. We further assumed that the distribution of YMSM attributes (see SDC 1) other than race and disease status of the n=421 YMSM of the Crew 450 study at T1/T2 is representative of the total YMSM population in Chicago and thus determined the size of the simulated YMSM population based on an estimate of the size of the total YMSM population age 16 to 21.8 years in Chicago and factoring in the empirical data of the n=421 YMSM of the underlying Crew 450 study.

### 2.1.1 Estimate of the size of the total YMSM population in Chicago

Using the 2010 Chicago Census data[40,76,77] and assuming the number of males to be evenly distributed between the ages 21 to 22 we estimate the number of males age 16 to 21.8 years in Chicago to be $n_{16\text{-}21.8}$=110,022. Purcell et al.[78] estimatethe percentage of males in the US who are men who have sex with men (MSM) to be 3.9%. Thus, we estimate the size of the total YMSM population to be $n_{YMSM,Chicago}$=4291.

Similarly, we estimate the size of the outside MSM population age 21.8 to 29 years to be $n_{21.8\text{-}29}$=7968 and the size of the outside MSM population age 30 to 39 years to be $n_{30\text{-}39}$=8599.

### 2.1.2 Estimate of the size of the simulated YMSM population

For individual attributes (see SDC 1) except for race and disease status we assume the n=421 YMSM of the Crew 450 study at T1 and T2 to be a representative sample of the YMSM population age 16 to 21.8 years. Thus, we scale the existing race stratified populations of this n=421 YMSM sample such that the size and race mix of the simulated YMSM population matches the size and race mix of the total YMSM population in Chicago.

Among the n=421 YMSM enrolled in the Crew 450 sample used to parameterize this simulation model, 225 YMSM were Black, 83 YMSM were Latino, 76 YMSM were White, and 37 YMSM were of Other race/ethnicity. To obtain the size of the Black, Latino, White, and Other populations in the simulated YMSM population we chose for each race the integer multiple of the n=421 YMSM race-stratified population such that the overall race mix of the resulting simulated YMSM population was closest to the 2010 Chicago race mix which we applied to the total YMSM population of size $n_{YMSM,Chicago}$=4291. For example, 31.7% of all Chicago males were of White which implies 1360 White YMSM in the total YMSM population. Closest integer multiple of 76 YMSM in the n=421 Crew 450 sample is 18 and thus the number of White YMSM in the simulated YMSM population is 18*76=1368. Table 32 shows the size of the simulated YMSM population with total size $n_{YMSM}$=4484 and the sizes of the subpopulations

stratified by race, the corresponding race mix and the Chicago 2010 census data of the race mix among males in Chicago.

| Race/ethnicity | Number simulated YMSM population at t=0 | Race mix simulated YMSM population | Chicago 2010 census race mix |
|---|---|---|---|
| Blacks | 1575 | 35.1% | 32.9% |
| Latinos | 1245 | 27.8% | 28.9% |
| Whites | 1368 | 30.5% | 31.7% |
| Others | 296 | 6.6% | 6.5% |
| Total | 4484 | 100% | 100% |

Table 32: Total size of simulated YMSM population stratified by race and corresponding race mix percentages. In comparison race mix percentages among males in Chicago in 2010 are shown.

# 3.2 Aging out, death, and birth processes

## 3.2.1 Aging out process

We defined the YMSM population to be YMSM of ages 16 to 21.8 years. Thus, a YMSM ages out of the YMSM population at time-step t if he becomes 21.8 years or older in this time step. If the YMSM has ongoing within partnerships when aging out, i.e. partnerships with other YMSM, these within partnerships are then transformed into outside partnerships for the YMSM partners of the aging out YMSM who remain in the YMSM population. Once the YMSM aged out he becomes part of the outside MSM community and thus is not further tracked by the simulation model.

## 3.2.2 Death process

We only model non-HIV related deaths in our simulation model of HIV, NG, and CT spread among the YMSM population age 16 to 21.8 years and do not consider HIV related death for the following reasons: 1. the approximate mean age of YMSM at time of HIV infection in the simulation model is 19 years, 2. Nakagawa et al.[79] estimate for HIV positives 5 years after the infection that the cumulative death

probability solely caused by HIV to be less than 1% at a very high diagnoses rate which is applicable to the simulated YMSM population, and 3. data of the WISQARS data base[80] showing that in the US HIV causes less than 1.5% of all deaths among male Blacks age 15 to 24 years, less than 0.6% among male Latinos of same age, and less than 0.2% among male Whites and other male races/ethnicities of the same age. In general, we consider HIV related deaths in our simulation model using estimates of the median survival time of HIV infected individuals stratified by treatment status to estimate the average duration of the HIV infection stages in our model (see section SDC 3.3,2.1).

We estimate the non-HIV related mortality rates for YMSM stratified by race using 2009 mortality data of the City of Chicago for males 15 to 24 years[81] and data provided by WISQARS data base about the causes of deaths among males 15 to 24 years in the US[80].

Table 33 shows the daily non-HIV related death probability of YMSM age 16 to 21.8 years. In absence of detailed data, we assume this daily death probability to be constant over the age range of 16 to 21.8 years.

| Race/ethnicity of YMSM age 16 to 21.8 years | Daily non-HIV related death probability |
|---|---|
| Blacks | 0.0000075902 |
| Latinos | 0.0000030301 |
| Whites | 0.0000018320 |
| Others | 0.0000018313 |

Table 33: Daily non-HIV related death probabilities of YMSM age 16 to 21.8 years stratified by race.

## 3.2.3 Birth processes

### 3.2.3.1 Population growth

Based on the 2010 to 2013 census data of the City of Chicago[76] we determined the overall growth rate of the Chicago population to be 0.264% per year which translates into a population growth rate of 0.00011% per time step. We assume the population growth to be constant over 15 years and dependent on the

population size at t=0 (i.e., non-cumulative). Thus, using this data we model the population growth in the simulation model as a Poisson arrival process with mean 0.492, i.e., every time-step the YMSM population increases on average by 0.492 YMSM.

YMSM entering the population due to population growth are duplicates of already existing YMSM in the population (see also SDC 4.2.1) and are chosen such that the overall race mix shown in Table 32 is maintained.

### 3.2.3.2 Birth process: Entering the YMSM population

Because the YMSM population increases over time (see section SDC 4.3.2.3.1), YMSM who age out of the population and YMSM who die are immediately reborn and enter the YMSM population in the following time-step with the same attributes when aging out or dying except for age, HIV and STI status, and having no existing partnerships. Additionally, a YMSM entering the population due to population growth represents a duplicate of an already existing YMSM chosen such that the overall race mix is maintained. However, his age, HIV, and STI status are newly determined at the time he is entering the population as well as he has no existing partnerships when entering.

#### *3.2.3.2.1 Age when entering*

When entering the YMSM population age 16 to 21.8 years each YMSM is assigned a specific age as shown in Table 34. YMSM enter the simulated YMSM population at an average age of 16.69 years and the age distribution of a YMSM entering was determined such that the overall average age of the YMSM population being 19.19 years was maintained throughout the simulation. Figure 13 shows the trajectory of the average age over time with a burn-in period for the aging in process of 2000 time-steps or 83.3 years.

| Probability p | Age |
|---|---|
| 0.6 | 16 + U(0,0.35) |
| 0.25 | 16.9 + U(0,0.35) |
| 0.15 | 17.9 + U(0,0.35) |

Table 34: Age distribution of YMSM entering the YMSM population. Average age when entering is 16.69 years. U(0,0.35) denotes a uniformly random number between 0 and 0.35.

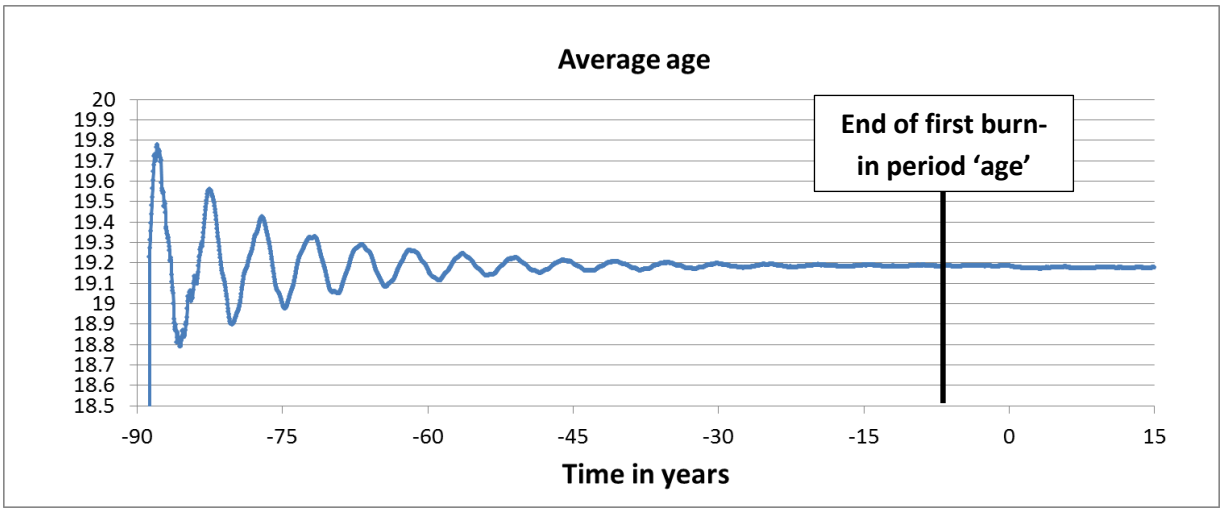


Figure 13: Average age of the YMSM population over time. Y-axis age in years. T=0 start of the actual simulation. T=-5.42 years end of burn-in phase 'age'.

### *3.2.3.2.2 HIV and STI status*

When entering the YMSM population each YMSM gets based on his age and race randomly assigned whether they are HIV, NG, or CT infected. Table 35 shows the disease specific prevalence stratified by race and age used to determine the likelihood of an aging in YMSM to be HIV, NG (urethral), or CT (urethral) positive. These distributions are based on the empirical prevalence of the Crew 450 at T1/T2 (see also SDC 1) except for the case of HIV 16 year olds and urethral CT 16 year olds. These where estimated to be 0.5% and 3.0% respectively based on expert opinion. The race specific disease prevalence was then determined for each age using the disease prevalence ratio of the race specific disease prevalence and the overall disease prevalence in the Crew 450 study at T1/T2, i.e. for HIV among Black YMSM 10.67%/5.49%=1.94 , 3.61%/5.49%=0.66 for Latino YMSM, 1.24%/5.49%=0.23 for White

YMSM, and 5.4%/5.49%=0.99 for Other YMSM. Thus, the aging in HIV prevalence for Black 16 year old YMSM can be calculated as 0.5%*1.94=0.97%.

| Age and disease | Overall | Blacks | Latinos | Whites | Others |
|---|---|---|---|---|---|
| **HIV prevalence** | | | | | |
| 16 | 0.5% | 0.97% | 0.33% | 0.11% | 0.49% |
| 17 | 1.16% | 2.24% | 0.76% | 0.26% | 1.13% |
| 18 | 2.48% | 4.82% | 1.63% | 0.56% | 2.44% |
| **NG prevalence (urethral)** | | | | | |
| 16 | 2.30% | 2.50% | 3.02% | 0.82% | 5.07% |
| 17 | 2.48% | 2.69% | 3.24% | 0.89% | 5.45% |
| 18 | 4.72% | 5.13% | 6.18% | 1.69% | 10.39% |
| **CT prevalence (urethral)** | | | | | |
| 16 | 3.00% | 3.62% | 4.52% | 0.82% | 5.07% |
| 17 | 4.79% | 5.77% | 7.22% | 1.31% | 8.09% |
| 18 | 3.42% | 4.21% | 5.27% | 0.96% | 5.91% |

Table 35: HIV, NG (urethral), and CT (urethral) prevalence stratified by race and age of YMSM when entering the YMSM population.

Data about the rectal NG and CT prevalence among YMSM in the Crew 450 study were not available, thus YMSM were randomly assigned to have a rectal NG or CT infection based on the NG and CT prevalence estimates (i.e., test positivity) of Sullivan et al.[58] who studied a cohort of Black and White YMSM in Atlanta (overall, 8.3% and 11.8% for NG and CT respectively; race stratified - NG: Black, Latino, and Other YMSM 10.8% and 3.0% for White YMSM; race stratified-CT: Black, Latino, and Other YMSM 15.8% and 4.0% for White YMSM; in absence of data for Latinos and Others, the Black YMSM NG and CT prevalence was also assumed for Latino and Other YMSM).

#### *3.2.3.2.2 Partnership status*

When entering the YMSM population at time-step t, we assume YMSM to have no existing within partnerships, i.e. no partnerships with other YMSM as well as no existing outside partnerships (i.e.,

partnerships with either other outside MSM or female partners) at that time. This is rule was chosen independent of the YMSM being already sexual active or not when entering the YMSM population.

# 3. Simulated time horizon

We simulate the transmission of HIV, NG, and CT over 15 years using discrete time-steps of length 0.5 months. This time horizon was chosen such that we are able to study the long range system behavior of HIV, NG, and CT transmission among YMSM and thus be able to determine the impact of initial conditions (i.e., empirical data about disease prevalence etc. at the start of the simulation t=0).

# V. Supplemental Digital Content 5: Implementation and Validation

# 1. Implementation

The stochastic discrete-time agent-based network simulation model was implemented using NetLogo version 5.1.0[82]. The duration of a simulated time-step was set to 0.5 months, or equivalently 15.25 days. Such a time step makes the simulation both computationally feasible and detailed enough to plausibly model partnership formation and disease transmission[83].

In our model of partnership formation and dissolution, we assume that a partnership (i.e., outside- and within-partnerships) cannot be both formed and dissolved within one time step. Further, we assume the formation and dissolution of partnerships to be conditionally independent within a time-step, i.e. the formation and dissolution rates of partnerships only depend on the previous time-step[83]. This implies that the formation of one-night-partnerships, outside-partnerships, and within-partnerships are also conditionally independent within a time-step.

We use discrete-time simulation. Thus, we do not model the exact point in time when an infection or partnership formation or dissolution happens during a time-step. Thus, we assume infections and partnership events to happen on average in the middle of a time-step. For infections we therefore apply a *sequential get and give* logic to model the infection of a YMSM and possible transmissions to other YMSM. Applying the s*equential get and give* logic a YMSM cannot simultaneously get infected and give the infection to other YMSM within one time step. That is, the YMSM gets infected in one time step and can only infect others during the following time-step(s). For example, a YMSM gets infected with NG and has symptomatic NG in one time-step. During this time-step he cannot infect others, only in the next time-step.

Each simulation run was preceded by a three-phase burn-in period of 2000+30+100=2130 time-steps. In the first phase of the burn-in period, we simulated the age in process of the YMSM population (see SDC 4.2) ignoring partnership formation and dissolution processes as well as disease transmission until the average age reached steady-state behavior after 2000 time-steps, see Figure 13. In the second phase of the burn-in period, we simulated the formation of the network of sexual partnerships among YMSM. Starting out with no existing partnerships at the end of the first phase of the burn-in period (i.e., at t=-5.42 years, Figure 13, Figure 14, Figure 15), the second phase of the burn-in period simulated aging in and the formation and dissolution of partnerships among YMSM in each time step, ignoring disease transmission. By the end of the second phase of the burn-in period after 30 time-steps (i.e., t=--4.17), the sexual network among YMSM reached steady-state behavior in addition to the average age, i.e. the partnership formation rates, the average-degree stratified by race and partnership-type, and the global network measurements, closeness centrality and betweenness centrality reached steady-state (see Figure 14). The third phase of the burn-in period, lasting 100 time-steps, continued the aging in process, the partnership formation and dissolution process and simulated the NG and CT disease transmission model, again ignoring HIV transmission. By the end of phase 3, the NG and CT prevalence in addition to the sexual network and average age reached steady-state behavior (see Figure 14 and Figure 15). The third phase of the burn-in period for the NG and CT disease transmission model was required because data about the type of NG or CT infection (i.e., symptomatic or asymptomatic) as well as data about treatment seeking behavior and ceasing of sexual activity among YMSM in the Crew 450 study were not available. Thus, at the beginning of the third phase of the burn-in period, we randomly assigned NG and CT infected YMSM whether they had a symptomatic or asymptomatic NG or CT infection, whether they were treatment seeking, and whether they ceased sexual activity. In addition, data about the prevalence of rectal NG and CT infections among YMSM in the Crew 450 study were not available (i.e., YMSM were randomly assigned to have rectal NG or CT at baseline and when entering the YMSM population in the simulation, see also section SDC 4.3.2). Thus, at the end of the three-phase burn-in period as we begin the actual simulation of HIV, NG and CT spread among $n_{YMSM}$=4484 YMSM over 15 years, the average age is in

steady-state at 19.19 years and we have the sexual network and the NG and CT transmission in steady-state that is representative of the pre-existing sexual partnerships and the NG and CT transmission of YMSM at time of enrollment in the Crew 450 study.

Results for each simulation scenario represent the average over n=10x100=1000 replications. For each scenario we simulated $n_1$=10 different realization of the desired sex-role R and desired sex-frequency F (for a detailed description of the procedure to determine these realizations see section SDC 2.4) for $n_2$=100 replications. Results are expressed as means. The half-width of the 95% confidence intervals (CI) are ≤ 1.5% of the mean unless the 95% CI is explicitly stated.

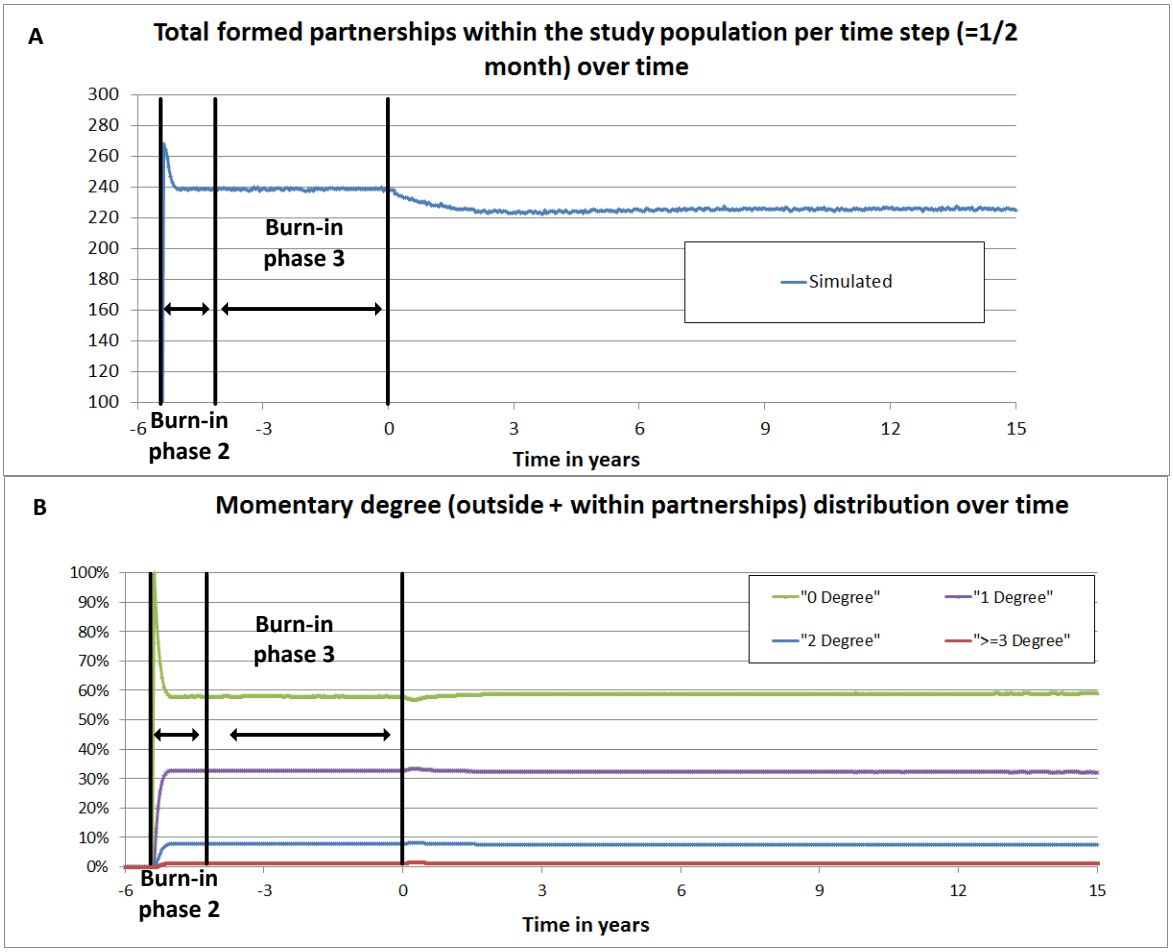


Figure 14: Total within partnerships formed per month (Figure A) and momentary degree distribution (outside + within partnerships) (Figure B) over time. Phase 2 of burn-in period to simulate partnership formation and dissolution process in steady-state. 1 time-step equals 0.5 months. T=0 actual start of the simulation. Burn-in phase 2 (partnership formation) lasting 30 time-steps from year t=-5.42 to year t=-4.17 before actual simulation starts. Phase 3 of burn-in period to simulate NG and CT

transmission model in steady-state. Burn-in phase 3 (NG and CT transmission) lasts 100 time-steps from year t=-4.17 before to the actual start of the simulation of HIV, NG and CT spread among n=4484 at t=0 YMSM in Chicago.

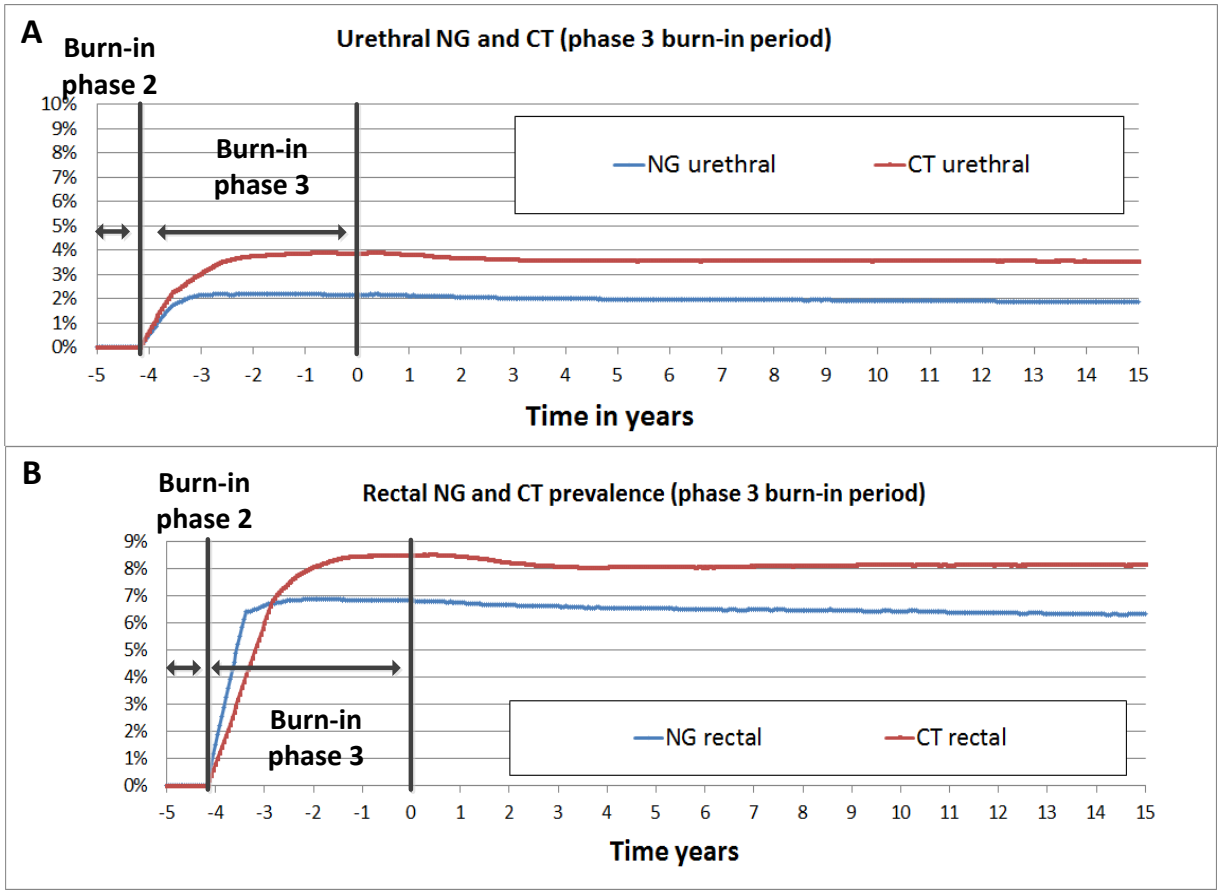


Figure 15: Urethral (Figure A) and rectal (Figure B) NG and CT prevalence over time. Phase 3 of burn-in period to simulate NG and CT transmission model in steady-state. 1 time-step equals 0.5 months. T=0 actual start of the simulation. Burn-in phase 2 (partnership formation) lasting 30 time-steps from year t=-5.42 to year t=-4.17 before actual simulation starts. Phase 3 of burn-in period to simulate NG and CT transmission model in steady-state. Burn-in phase 3 (NG and CT transmission) lasts 100 time-steps from year t=-4.17 before to the actual start of the simulation of HIV, NG and CT spread among n=4484 at t=0 YMSM in Chicago.

# 2. Validation

## 2.1 Partnership formation and dissolution

### 2.1.1 Partnership formation rates

The partnership formation and dissolution model was parameterized using data of the Crew 450 study at T1 and T2. Figure 16A, Figure 16B, and Figure 16C, show the simulated and empirical partnership formation rates for one-night-partnerships, outside-partnerships, and within-partnerships over the simulated time horizon of 15 years after the burn-in period (see discussion about burn-in period in section SDC 5.1). The empirical estimates are based on the original Crew 450 data and are adjusted due to the new simulated population size and race-mix, see also SDC 4.2.The simulated results for the formation rates of one-night-partnerships, outside-partnerships, and within-partnerships match the empirical results for T1 and T2 (i.e., are within the 95% confidence intervals; data collection of Crew 450 for waves T1 and T2 completed) which shows that the partnership formation and dissolution model was accurately parameterized. For a discussion of the accurate parameterization of the dissolution model see section SDC 2.3. The Crew 450 study follows up YMSM over time whereas we simulate a fixed age-range with constant average-age over time. Thus the adjusted empirical Crew 450 data of T1 and T2 itself provide the only empirical data of the Crew 450 study to validate the partnership formation model. The formation rates of one-night partnerships and outside-partnerships increase over 15 years by 2.8% and 3.5% respectively which corresponds to the overall population growth of the simulated YMSM population of 3.96% (see Figure 16A and Figure 16B). In case of within-partnerships (Figure 16C), we observe a small decline in the first 3 years which we reason that this is caused by the steep increase in HIV prevalence in these years (see also section SDC 2.2 on the regression model for within-partnership formation).

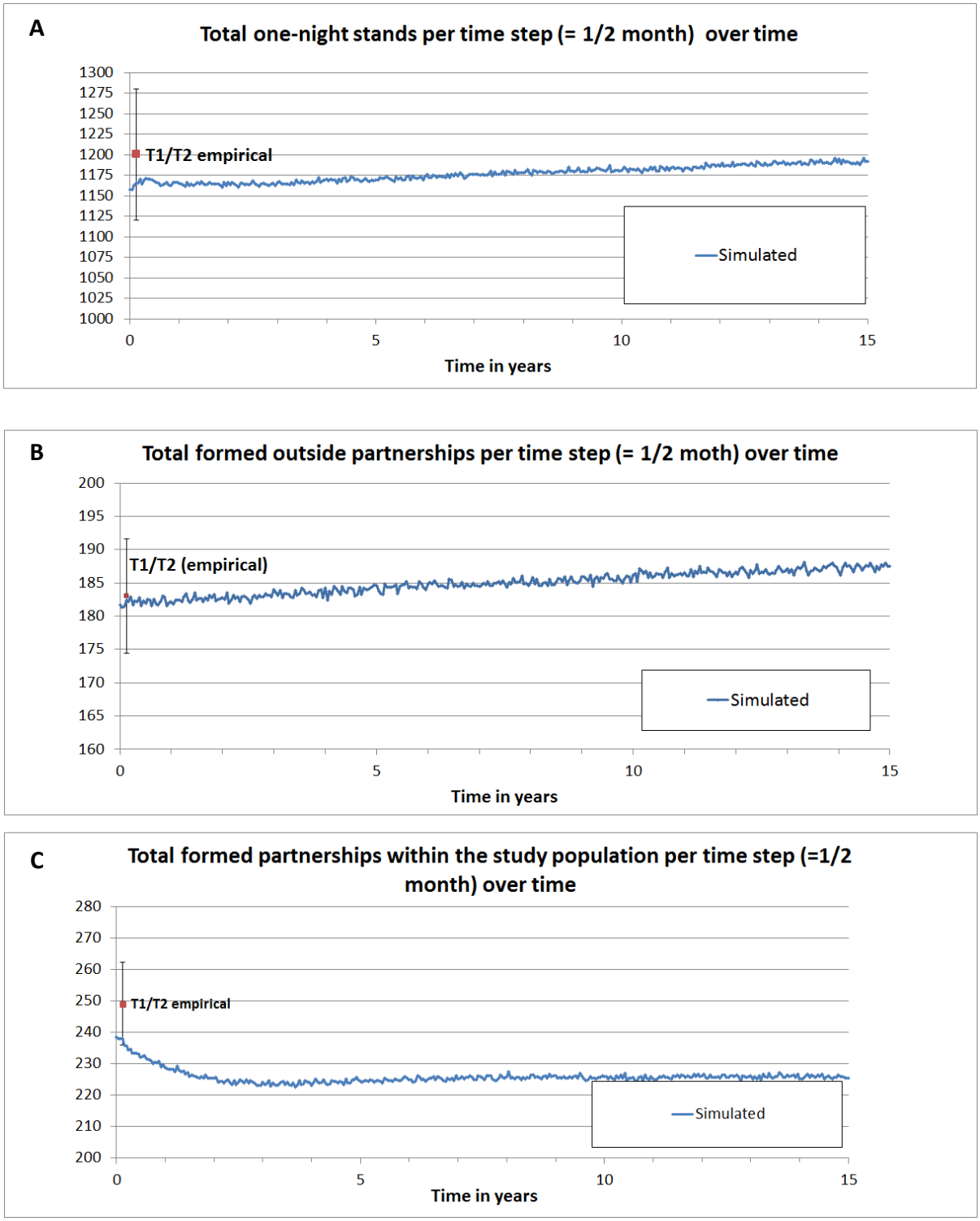


Figure 16: Simulated partnership formation rates for one-night-partnerships (Figure A), outside-partnerships (**Error! Reference source not found.**B) and within-partnerships (Figure C). Formed partnerships per time-step (i.e., 0.5 months) over simulated horizon of 15 years in comparison to empirical estimates of the Crew 450 study at completed data collection waves T1/T2 which were adjusted for the simulated population size of n=4484. Adjusted to the new population size, error bars show 95% confidence intervals of empirical estimates.

## 2.1.2 Momentary degree

The momentary degree distribution is an important characteristic of a network[84] and is widely used to characterize and construct models of networks[4]. Figure 17 shows the estimated empirical momentary degree distribution at T1 and T2 (see discussion in section SDC 2.2) and the simulated degree distribution at T1, i.e. t=0 (see also Figure 14B). We observe that the simulated fraction of individuals having no partnership and 2 partnerships is within the 95% CI of the corresponding fractions of the estimated empirical degree distribution at both t=0 and over the total simulated time horizon. Further, the simulated fractions of individuals having 1 or at least 3 partnerships is close to the upper and lower 95% CI bounds of the corresponding empirical estimates. Both the simulated and the estimated empirical degree distribution of the Crew 450 study at T1 and T2 are comparable to the momentary degree distributions of the Explore study used in the model of Goodreau et al. [4], where 60.0% percent of MSM have no main partnership, 38.3% have 1 main partnership and 1.7% have two main partnerships (for detailed discussion see section 2.2.1.1 in SDC 2).

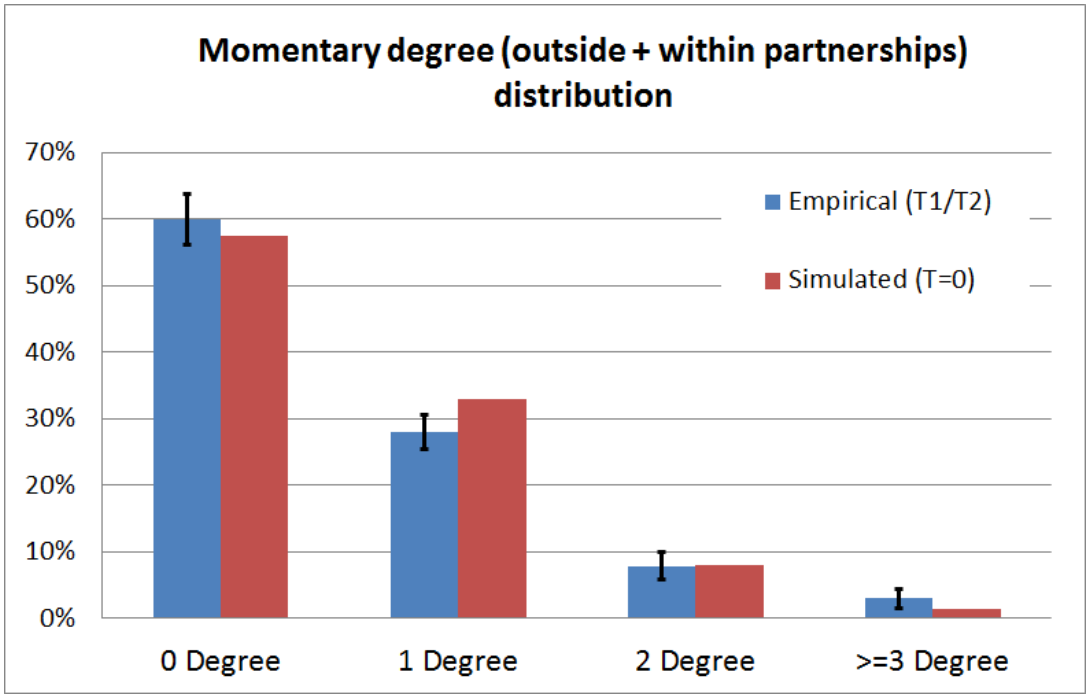


Figure 17: Empirical and simulated momentary degree distributions. Estimated empirical momentary degree distribution at T1 and T2 where error bars show 95% confidence intervals. Simulated momentary degree distribution for T1 and T2 (T=0 months).

# 2.2 Disease transmission

## 2.2.1 Simulated results and empirical findings

### 2.2.1.1 HIV infections per 100 person-years

#### *Years 1-3.5*

Figure 18 shows the simulated and empirical HIV infections per 100 person-years stratified by race within the first 3.5 years. We compared this to the HIV infections per 100 person-years of the Crew 450 study. The simulation time horizon of 3.5 years was chosen because the empirical HIV incidence data of the Crew 450 study were only available for a follow-up period of 3.5 years (i.e., 7 consecutive data collection waves). The Crew 450 study followed-up individuals over time, thus the empirical incidence presented in Figure 18 does methodologically-wise not exactly equal the simulated HIV incidence because we simulate a population of fixed age-range over time, i.e. in the Crew 450 YMSM get older whereas in the simulation the average-age is maintained. However, in the absence of specific empirical data of the HIV incidence per 100 person-years and the relatively short duration of the follow-up (i.e., 3.5 years) of the Crew 450 study, the empirical data presented in Figure 18 are highly relevant for the validation of our simulation model.

We observe that the simulated HIV incidence is close to the preliminary estimates of the HIV incidence of the Crew 450 study as of 6/6/2014, i.e. the simulated overall HIV incidence (i.e., not stratified by race) and simulated HIV incidence stratified by race are within the bounds of the 95% confidence intervals of the preliminary empirical results. Also, the 'overall' simulated results show a value close to the preliminary empirical results of the Crew 450 study (i.e., 4.9 and 4.1 HIV infections per 100-person years respectively). Further, empirical and simulated results show significant racial differences in HIV incidence among YMSM. In particular, Blacks and Others have a significantly increased HIV incidence compared to Whites and Latinos where Latinos have a higher HIV incidence compared to Whites. These simulated and empirical racial differences in HIV incidence among YMSM in Chicago align with the findings of other studies among MSM in the US shown in Table 36.

The simulated HIV incidence for Blacks and Others tends to be lower than the empirical HIV incidence observed in the Crew 450 study whereas the simulated HIV incidence for Latinos and Whites tend to be higher than the empirical estimates. As a result, the HIV incidence ratio of Blacks vs. Whites in the simulation is 1.75 which is lower than the HIV incidence ratios shown in Table 36. We reason that racial differences in the HIV prevalence on the community area level among MSM in Chicago might be bigger than the racial differences in HIV prevalence incorporated in the model[85] (i.e., we parameterized our model with the Chicago wide HIV prevalence data among MSM stratified by race[11] which were the only available data about HIV prevalence of Chicago MSM; an analysis of the male HIV prevalence on the community area level of communities, where YMSM who participate in the Crew 450 study which got newly infect with HIV are living, show HIV prevalence ratios of 5-10 of Blacks vs. Whites whereas the HIV prevalence ratio of Blacks vs. Whites MSM in our model is 2.76) (see Table 27 in SDC 3). Further, racial differences in HIV infectivity due to different levels in HIV testing, full and partial suppression and ART coverage among races (see Table 27) might be bigger than the racial differences in HIV infectivity incorporated in the model. Because the actual differences in HIV prevalence and HIV infectivity might be bigger, differences in the HIV incidence might be bigger and thus may lead to ratios of Black and White HIV incidence observed in other studies (see also sensitivity analysis in section 2.2.2 in SDC 5).

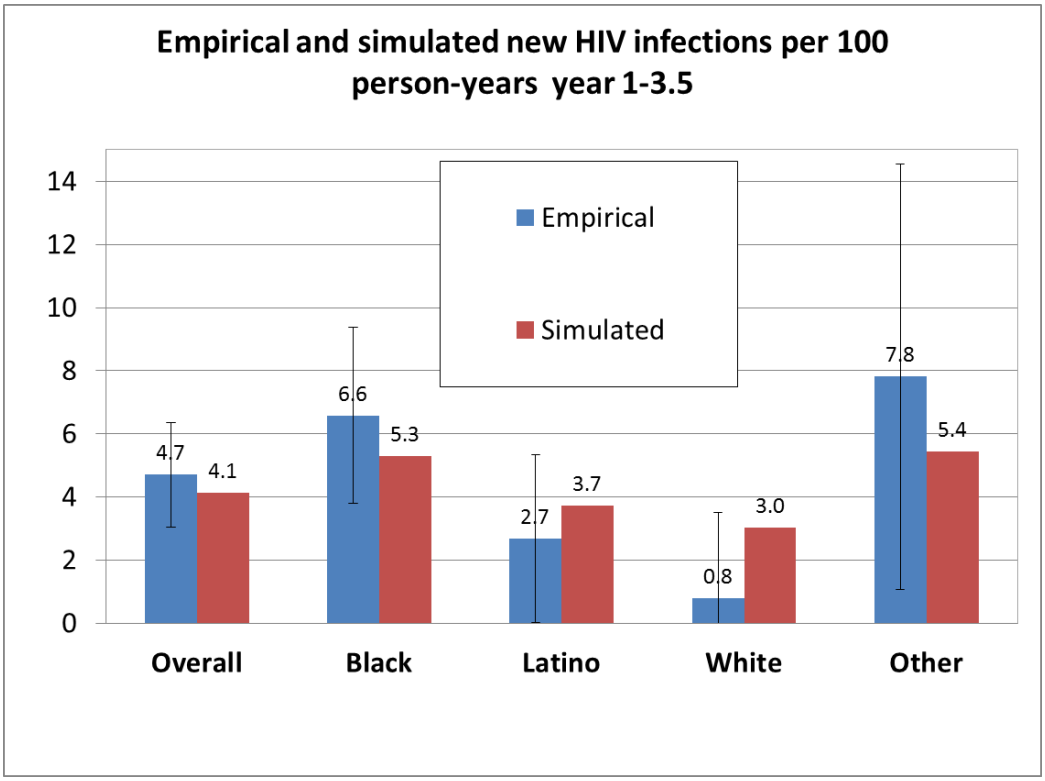


Figure 18: Empirical (Crew 450 study, n=450 at baseline, as of 6/6/2014) and simulated (n=4484 at t=0) HIV infections per 100 person-years. Duration of study and simulated time was 3.5 years. Error bars show 95% confidence intervals for preliminary empirical estimates of Crew 450 study as of 6/6/2014.

Table 36: HIV incidence ratios of YMSM simulation, Crew 450 study and other studies of MSM in the US.

| Study/ Reference | Sample size (n) | Age range (years) | HIV incidence ratio[c] with baseline Whites=1.0 | | |
|---|---|---|---|---|---|
| | | | Blacks | Latinos | Others |
| YMSM simulation (year 1-3.5) | 4484 (Chicago) | 16-21.8[b] | 1.75 | 1.23 | 1.80 |
| Crew 450[a] | 450[b] (Chicago) | 16-21[b] | 8.25 | 3.34 | 9.75 |
| CDC [86] | - (US wide) | 13-29 | 2.7 | - | - |
| Rosenberg et al. [87] | 803 (Atlanta) | 18-39 | 3.88 | - | - |
| Lieb et al. [88] | - (Florida, State) | ≥18 | 5.5 | 2.0 | - |
| Neaigus et al.[89] | 550 (NYC) | 18-68 | 3.16 | 0.836 | 1.22 |
| Sifakis et al. [90] | 948 (Baltimore) | 15-29 | 18.33 | 0 | 11.83 |
| Balaji et al.[91] | 28,468 (NHBS 2008) | 18-24 | 3.19 | 1.19 | 1.81 |

[a] Preliminary estimates of Crew 450 study as of 6/6/2014.

[b] At baseline of study (i.e., time of enrollment)

[c] HIV incidence ratio calculated based upon study individual HIV incidence measurements (e.g., HIV incidence per 100 person-years or HIV incidence per person year etc.)

### *Year 1-15*

Over the total simulated time horizon of 15 years Black YMSM experienced the highest HIV incidence per 100 person-years compared to Latino YMSM and White YMSM (Figure 1a and 1b in the manuscript). HIV incidence per 100 person-years and HIV prevalence among Black YMSM increased moderately over time (i.e., HIV incidence per 100 person-years and HIV prevalence increased both 1.59 fold) whereas Latino YMSM and White YMSM experienced a steeper increase in HIV incidence and HIV prevalence (i.e., HIV incidence per 100 person-years increased 1.97 fold for Latino YMSM and 2.03 fold for White YMSM, HIV prevalence 2.73 for Latino YMSM and 16.02 fold for White YMSM). Thus,

racial disparities in HIV incidence per 100 person-years and HIV prevalence decreased over time but remained significant (i.e., HIV prevalence ratio for Black-Latino and Black-White YMSM and HIV incidence ratio for Black-White YMSM where the HIV incidence ratio for Black-Latino and Black-White YMSM was within year 1 1.61 and 2.01 respectively whereas in year 15 1.05 and 1.25 respectively; for HIV prevalence ratios the Black-Latino and Black-White HIV prevalence ratio at t=0 was 1.94 and 13.92 respectively and at t=15 years 1.13 and 1.38 respectively). The steep increase in HIV prevalence and incidence among White YMSM in our simulation aligns with findings of the 2011 MSM Chicago surveillance report[11] which showed an increase in HIV prevalence from 1.0% among White YMSM age 18 to 29 in 2008 to 10.0% in 2011 whereas HIV prevalence among Black YMSM of same age showed a small decrease from 28.4% to 27.9%. Further, HIV diagnoses data of 15 to 24 year old MSM in Chicago (see Table 38, personal communication Chicago Department of Public Health (CDPH), March 31st, 2015), showed an increase in HIV diagnoses from 2009 to 2013 among all YMSM (i.e., 1.19 fold) and stratified by race (i.e., Black YMSM 1.16 fold, Latino YMSM 1.54 fold, and White YMSM 1.82 fold). This aligns with our findings that HIV incidence and prevalence among YMSM continue to rise over 15 years with Latino and White YMSM experiencing a steeper increase in the HIV burden compared to Black YMSM as well as White YMSM experiencing a marginally steeper increase compared to Latino YMSM.

Given the initial HIV prevalence data at t=0, we observe that racial disparities decrease over time but remain significant. CDPH data show similar trends, but given the fact that our simulation model tends to shows smaller racial disparities in HIV infections per 100 person-years within year 1 to 3.5 in comparison to estimates of the Crew 450 study, the CDC, and other publications listed in Table 36, the magnitude of racial disparities over time shown in our simulation model might be influenced by the same reasons as discussed in the above section (i.e., section SDC 5.2.2.1.1 year 1-3.5) and thus potentially lead to an underestimation of racial disparities on the long run in our simulation model.

### 2.2.1.2 Total number of HIV infections

Simulating HIV, NG, and CT spread over 15 years, we estimated 3076 new HIV infections (95% CI 3066-3086) to occur among YMSM age 16 to 21.8 years in Chicago with 1220 new HIV infections (95% CI 1215-1225) occurring among Black YMSM, 836 new HIV infections (95% CI: -830-842 ) among Latino YMSM, and 770 new HIV infections (95%CI 766-774) occurring among White YMSM. Table 37 shows the simulated number of new HIV infections per year across all YMSM and stratified by race.

| HIV infection | Year1 | Year2 | Year3 | Year4 | Year5 | Year6 | Year7 | Year8 | Year9 | Year10 | Year11 | Year12 | Year13 | Year14 | Year15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Total | 161 | 168 | 169 | 174 | 181 | 189 | 197 | 203 | 211 | 217 | 225 | 231 | 237 | 244 | 252 |
| Black | 71 | 71 | 70 | 71 | 73 | 76 | 78 | 80 | 82 | 84 | 87 | 88 | 89 | 92 | 94 |
| Latino | 40 | 43 | 44 | 46 | 49 | 51 | 54 | 56 | 58 | 61 | 63 | 65 | 67 | 69 | 72 |
| White | 36 | 40 | 41 | 42 | 44 | 47 | 49 | 51 | 53 | 55 | 58 | 60 | 62 | 64 | 67 |

Table 37: Simulated absolute number of new HIV infections per year over time horizon of 15 years. 'Total' denotes total number of new HIV infections among YMSM, 'Black' denotes number of new HIV infections among Black YMSM, 'Latino' denotes number of new HIV infections among Latino YMSM, and 'White' denotes number of new HIV infections among Latino YMSM. Numbers represent mean of simulation runs and the half-width of the 95% CI are within ≤ 1.5% of the mean.

We observe that overall the total number of HIV infections increases 1.12 fold in the first 5 years and 1.57 fold over 15 years. Stratified by race, the absolute number of HIV infections increases 1.07 fold among Black YMSM in the first 5 years and 1.32 fold over 15 years. For Latino YMSM, we observe a 1.23 fold increase in the first 5 years and a 2.35 fold increase over 15 years. Finally, for White YMSM we observe a 1.22 fold increase in the first 5 years and a 1.86 fold increase over 15 years.

Table 38: HIV diagnoses data among 15 to 24 year old MSM in Chicago from 2009 to 2013 (Chicago Department of Public Health, personal communication, March 31st, 2015).

| **Age group and race** | **HIV infections per year[b]** | | | | |
|---|---|---|---|---|---|
| | **2009** | **2010** | **2011** | **2012** | **2013** |
| **Overall** | | | | | |
| 15-19 years | 46 | 40 | 42 | 52 | 41 |
| 20-24 years | 162 | 160 | 155 | 152 | 208 |
| Total | 208 | 200 | 197 | 204 | 249 |
| **Race-stratified** | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| *15-19 years* | | | | | |
| Non Hispanic, Black | 42 | 33 | 34 | 45 | 34 |
| Non Hispanic, White | 0 | 0 | 0 | <5 | 0 |
| Hispanic | <5[a] | 5 | <5 | <5 | 6 |
| Non Hispanic, Other/Unk[c] | <5 | <5 | <5 | <5 | <5 |
| *20-24 years* | | | | | |
| Non Hispanic, Black | 106 | 104 | 105 | 93 | 138 |
| Non Hispanic, White | 17 | 25 | 12 | 13 | 31 |
| Hispanic | 26 | 25 | 29 | 30 | 34 |
| Non Hispanic, Other/Unk[c] | 13 | 6 | 9 | 16 | 5 |

[a] Cell sizes <5 were suppressed
[b] MSM includes MSM+IDUs
[c] Non Hispanic, Other/Unk includes: Asian, Pacific Islanders, Multiple Races, Asian American/Alaskan Native, Other, Unknown
Prepared by the Chicago Department of Public Health, 3/31/2015

Table 38 shows the HIV diagnoses data among YMSM age 15 to 24 years in Chicago from 2009 throughout 2013 (personal communication, Chicago Department of Public Health (CDPH), March 31st, 2015). To be able to compare the simulated HIV incidence in our model with the HIV diagnoses data among YMSM age 15 to 24 years in Chicago, we first adjust the CDPH data for the age-range. Assuming HIV diagnoses to be uniformly distributed within each age-range, i.e. within 15 to 19 year olds and 20 to 24 year olds, we calculate the age-adjusted HIV incidence for YMSM age 16 to 21.8 years in Chicago stratified by race. HIV diagnoses data represent only positive cases tested at a specific age. Thus, the actual number of HIV infections among 15 to 24 year old YMSM in Chicago might differ from the values shown in Table 38 because the actual HIV infection might have happened at a different age than the YMSM was tested positive (e.g., a YMSM got infected at age 17 and was tested positive at age 19). Further, HIV diagnoses data representing positive test case do not take into account those unaware of their HIV infection, i.e. the fraction of HIV infected among YMSM unaware of their infection. In our adjustment of the presented CDPH data in Table 38 we account for the fraction of HIV infected unaware of their infection. In 2011, among YMSM age 18 to 29, 33.33% of Black YMSM were unaware of their

HIV infection, 20% of White YMSM, and 40% of Latino YMSM[11]. Using binomial proportion confidence intervals, we calculated the 95% corresponding confidence intervals of the fraction of HIV infected unaware of their HIV infection (lower bounds assumed to be not negative). Thus, we determined the 95% CI of the HIV infected unaware of their HIV infection for Black YMSM to be 0.144 to 0.52, for White YMSM 0 to 0.55, and for Latino YMSM 0 to 0.83. Using these 95% confidence intervals we calculated across races as well as specific for each race the total number of HIV infected among YMSM in Chicago as the number of HIV incident cases of the CDPH adjusted for age (i.e., positive tested cases) divided by 1 minus the fraction of HIV infected unaware of their HIV infection. Table 39 shows the final estimates of the mean and the 95% CI of the number of new HIV infections among 16 to 21.8 year olds MSM in Chicago which represent the CDPH data of the number of new HIV infections among 15 to 24 year old YMSM in Chicago adjusted for age and fraction unaware of their HIV infection.

Table 39: Estimated mean and 95% CI of the number of new HIV infections among 16 to 21.8 year olds MSM in Chicago which represent the CDPH data of the number of new HIV infections among 15 to 24 year old YMSM in Chicago adjusted for age and fraction unaware of their HIV infection.

| **Age 16 to 21.8 years** | **Estimated HIV infections per year[a] and 95% CI (in parenthesis)** | | | | |
|---|---|---|---|---|---|
| | **2009** | **2010** | **2011** | **2012** | **2013** |
| **Overall[b,c]** | | | | | |
| Total | 140.6 (114.1-183.1) | 132.4 (107.4-172.5) | 132.1 (107.2-172.1) | 142.3 (115.5-185.4) | 159.1 (129.1-207.3) |
| **Race-stratified** | | | | | |
| Non Hispanic, Black | 107.6 (83.9-150.0) | 95.7 (74.6-133.4) | 97.5 (75.9-135.9) | 104.17 (81.2-145.2) | 115.3 (89.9-160.7) |
| Non Hispanic, White | 7.7 (6.1-13.6) | 11.3 (9.0-20.0) | 5.4 (4.3-9.6) | 8.9 (7.1-15.8) | 14.0 (11.2-24.8) |
| Hispanic | 16.9 (10.16-59.6) | 21.7 (13.0-76.2) | 22.7 (13.6-80.0) | 22.0 (13.2-77.4) | 28.4 (17.0-99.9) |

[a] Estimated mean and 95% CI of the number of HIV infections among 16 to 21.8 year olds MSM in Chicago which represent the CDPH data of the number of HIV infections among 15 to 24 year old YMSM in Chicago (Table 38) adjusted for age and fraction unaware of their HIV infection.
[b] Mean overall HIV incidence and corresponding 95% CI were calculated using weighted average of fraction unaware of HIV infection across races Black, White, and Latino being 0.3233 with corresponding 95% confidence interval (0.4805 – 0.167)
[c] We assume that the CDPH estimates[11] of the fraction of HIV infected YMSM age 18 to 29 which are unaware of their HIV infection applies both to YMSM age 15 to 19 and 20 to 24 years.

Comparing the simulated mean number of new HIV infections shown in Table 37 and the estimated number of new HIV infections among 16 to 21.8 year olds MSM in Chicago shown in Table 39, we observe that the simulated total number of new HIV infections of year 1 and year 5 are within the 95% CI of the estimated number of new HIV infections of the empirical estimates. Further, the absolute increase in the total number of new HIV infections from 2009 to 2013 is almost the same as the absolute increase in the simulated total number of new HIV infections from year 1 to year 5. For Latino YMSM, the simulation estimates of the mean number of new HIV infections are also within the 95% CI of the empirical estimates for year 1 (empirical 2009) and year 5 (empirical 2013) and for Black YMSM our estimates of the mean number of new HIV infections is close to the lower bound of the corresponding 95% CI of the empirical estimates. We refer to section SDC 5.2.2.1.1 for an explanation about why our simulated results do not match the empirical estimates for HIV incidence among White YMSM and Black YMSM. Additionally, we want to mention that the empirical estimates shown in Table 39 heavily depend on the fraction of HIV infected unaware of their HIV infection and the only estimates available for Chicago YMSM age 18 to 29 rely on small sample sizes (i.e., Black YMSM n=24, White YMSM n=5, Latino YMSM n=5). Further, we did not take into account backtracking and thus did not consider that the actual time of HIV infection might have differed from the testing date of those YMSM within the age of 15 to 24 who tested positive. Given, that the majority of new HIV infections occurred among 20 to 24 year olds in comparison to 15 to 19 year olds (see Table 38), backtracking of the actual date of the infection might lead to an increase in the estimated number of new HIV infected among YMSM age 16 to 21.8 years shown in Table 39.

### 2.2.1.3 NG and CT prevalence

Figure 19 shows the simulated and empirical urethral and rectal NG and CT prevalence over the simulated time horizon of 15 years. For urethral NG and CT the simulated results are within the 95% confidence interval bounds of the empirical estimates (i.e., empirical NG and CT is urethral test

positivity). In the urethral case (Figure 19A), these findings align with the findings of Sullivan et al.[58] who estimated the prevalence (i.e., test positivity) of urethral NG and CT among MSM age 18-39 in Atlanta, GA to be 1.5% and 2.7% respectively. Further, NG and CT testing data of the Howard Brown Health Center in Chicago[92] show approximately 6% and 4% prevalence (i.e., test positivity) for urethral NG and CT. Because YMSM enrolled in the Crew 450 study were not tested for rectal NG and CT and rectal NG and CT testing is not common in Chicago[54] we rely on the few published estimates of rectal NG and CT prevalence among MSM[58,92] to validate the rectal NG and CT transmission model. Hotton and Gratzer [58,92] estimate the rectal NG and CT prevalence among MSM to be 2-3 times higher than urethral NG and CT prevalence. Figure 19B shows the simulated rectal NG and CT prevalence among YMSM and empirical estimates which are 2 times the urethral NG and CT prevalence respectively with adjusted 95% confidence intervals. We observe the simulated NG and CT prevalence to be within the range of 2-3 times the urethral NG and CT prevalence over the total simulated time horizon of 15 years. Further, rectal NG prevalence is lower compared to rectal CT prevalence over the total time horizon, matching the empirical findings of Hotton and Grazer[92] and Patton et al.[54].

The simulated urethral and rectal NG and CT prevalence declines marginally in the first three years (see Figure 19) and then remains stable. We reason that a decline in within-partnership formation rates and the steep increase in HIV prevalence (i.e., higher HIV prevalence implies increased testing rates for STIs) are likely to contribute to the decline in NG and CT prevalence in the YMSM cohort within the first three years.

Finally, our assumptions regarding NG and CT and in particular rectal NG and CT might contribute to an underestimation of the spread of NG and CT (i.e., might cause lower prevalence), and thus to a lower estimated impact of NG and CT on HIV.

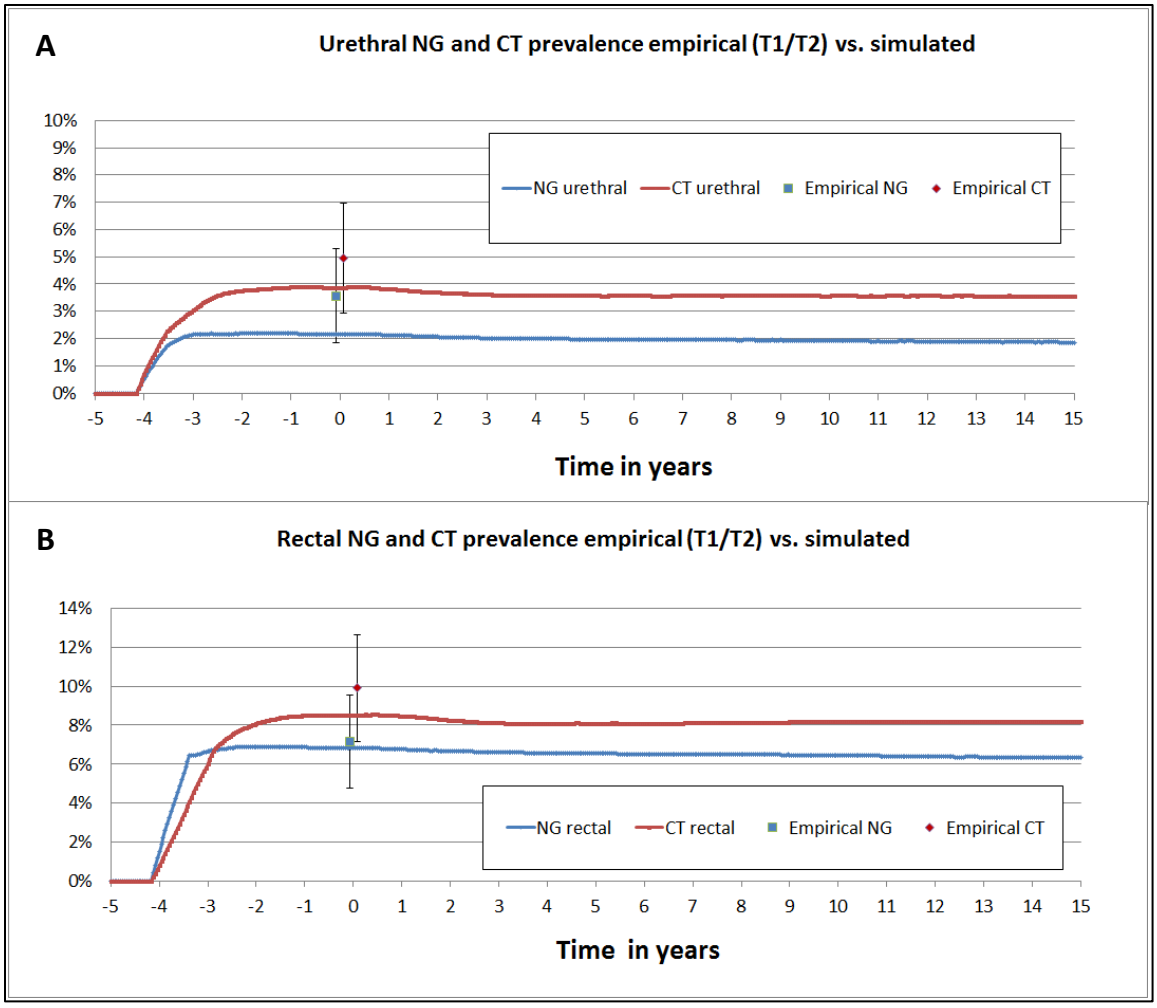


Figure 19: Simulated and empirical gonorrhea (NG) and chlamydia (CT) urethral (Figure A) and rectal (Figure B) prevalence among YMSM in simulated cohort. In case of rectal NG and CT the empirical prevalence are estimates which equal 2 times the urethral empirical prevalence. Error bars show 95% confidence intervals of the empirical Crew 450 study data of data collection waves T1 and T2. Rectal prevalence 95% confidence intervals were adjusted to increased prevalence levels compared to urethral prevalence. Prevalence is defined as the percentage of YMSM who tested positive.

### 2.2.1.3 HIV incidence stratified by partnership-type

Figure 20 shows the simulated HIV infections per 100 person-years stratified by partnership-type over 15 years. Of all HIV infections, 21.1% were attributable to one-night-partnerships, 34.5% to outside-partnerships and 44.4% to within-partnerships. This pattern shows only marginal variations over time and across races (changes within 5 percentage-points). In comparison, Xiridou et al.[93] estimated the percentage of HIV infections attributable to one-night-partnerships (i.e., a casual partnership with one

sexual encounter) to be 14% among a cohort of adult MSM in Amsterdam. Taking into account that younger MSM have more one-night stands than older MSM (age was a significant covariate in the regression model for one-night stand formation in SDC 2.2), our findings align with those of Xiridou et al.[93]. The majority of studies which investigate the impact of partnership-type on HIV infections among MSM[1,2,4,94] classify partnerships into either casual partnerships or main partnerships. Assuming a 'main partnership' in these studies to equal a serious outside- or within-partnership in our model, we estimate the percentage of HIV infections among YMSM attributable to 'main partnerships' to be 33.5% over 15 years. This aligns with the findings of Goodreau et al.[4] and Jansen et al.[94] who estimate the percentage of HIV infections attributable to main partnerships to be 36%-39% among MSM in the US and 26% among MSM in Amsterdam respectively.

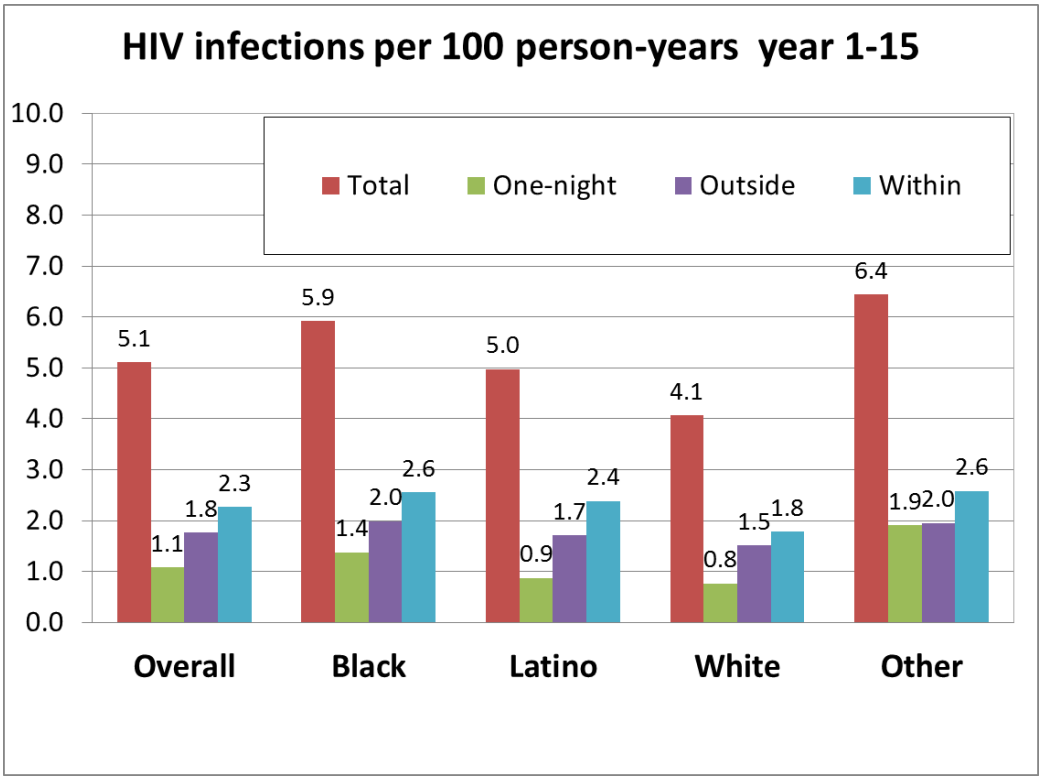


Figure 20: Simulated new HIV infections per 100 person-years stratified by partnership-type and race (same as Figure 2 in manuscript).

## 2.2.2 Sensitivity analysis

### 2.2.2.1 HIV incidence among (Y)MSM

To fully validate the simulation model of HIV, NG, and CT transmission and partnership formation and dissolution among YMSM in Chicago we perform sensitivity analysis on specific parameters of the HIV, NG and CT disease transmission model in addition to the comparison of simulated and empirical results (see sections SDC 5.2.1 and 5.2.2.1).

As hypothesized in section SDC 5.2.2.1 the magnitude of racial differences in HIV incidence might depend on the magnitude of racial differences in HIV prevalence among MSM. To evaluate the sensitivity of the outcome HIV incidence on the input parameters, HIV prevalence of older MSM partners stratified by race (i.e., male MSM partner in outside-partnerships) as well as the HIV prevalence among YMSM at baseline (T1) we simulate two counterfactual scenarios:

- *Outside HIV prevalence 2008:* We assume the HIV prevalence for older MSM partners stratified by race (i.e., male partners of YMSM in outside-partnerships) to equal the CDPH MSM surveillance data of 2008[11] at the beginning of the simulation at t=0.
- *No difference in HIV prevalence:* We assume no differences in HIV prevalence among older MSM partners (i.e., male partners of YMSM in outside-partnerships) at the beginning of the simulation at t=0, i.e. older MSM across races have the same HIV prevalence which is the average HIV prevalence among Chicago MSM in 2011[11]. Further, we assume no differences in HIV prevalence among YMSM at the beginning of the simulation at t=0, i.e. YMSM across races have the same HIV prevalence. This counterfactual scenario corresponds to Figure 3C in the manuscript.

Table 40 shows the HIV prevalence of the base case, the counterfactual scenario *Outside HIV prevalence 2008*, and the counterfactual scenario *No difference in HIV prevalence*. Figure 21 shows the HIV incidence per 100 person-years for the three scenarios.

Table 40: HIV prevalence for scenarios base case, 'Outside HIV prevalence 2008' and 'No difference in HIV prevalence' at t=0.

| HIV prevalence at t=0 | Parameter value in % | | | | | |
|---|---|---|---|---|---|---|
| | Base case | | Outside HIV prevalence 2008 | | No difference in HIV prevalence | |
| | MSM age 21.8 to 29 | MSM age 30 to 39 | MSM age 21.8 to 29 | MSM age 30 to 39 | MSM age 21.8 to 29 | MSM age 30 to 39 |
| | **YMSM (at baseline age 16 to 21.8, above age stratifications do not apply)** | | | | | |
| Blacks | 10.67 | 10.67 | 10.67 | 10.67 | 5.61 | 5.61 |
| Latinos | 3.61 | 3.61 | 3.61 | 3.61 | 5.61 | 5.61 |
| Whites | 1.32 | 1.32 | 1.32 | 1.32 | 5.61 | 5.61 |
| Others | 5.41 | 5.41 | 5.41 | 5.41 | 5.61 | 5.61 |
| | **Older MSM (i.e., outside male partners, above age stratifications do apply)** | | | | | |
| Blacks | 27.9 | 47.4 | 28.4 | 40.6 | 17.18 | 24.13 |
| Latinos | 8.1 | 16.0 | 9.4 | 15.4 | 17.18 | 24.13 |
| Whites | 9.4 | 18.9 | 1.1 | 17.7 | 17.18 | 24.13 |
| Others | 17.18 | 24.13 | 13.70 | 22.54 | 17.18 | 24.13 |

In Figure 21 we observe a greater magnitude of the racial differences in HIV incidence among YMSM with the increase in racial differences in the outside HIV prevalence among MSM at the beginning of the simulation at t=0. In the base-case scenario the HIV prevalence ratio (PR) of Black older MSM to White older MSM is 2.76 for MSM age 21.8 to 29 years and 2.51 for older MSM age 30 to 39 years at t=0 and the HIV incidence ratio (IR) of Black YMSM to White YMSM is 1.46. The HIV PRs of Black older MSM to White older MSM at t=0 for the scenarios *Outside HIV prevalence 2008* and *No difference in HIV prevalence* are 25.8 for older MSM age 21.8 to 29 years and 2.29 for older MSM age 30 to 39 years and 1 for both older MSM age 21.8 to 29 years and MSM age 30 to 39 years respectively. Corresponding HIV IRs over 15 years of Black YMSM to White YMSM are 1.71 and 1.05 for the counterfactual scenarios *Outside HIV prevalence 2008* and *No difference in HIV prevalence* respectively. Thus, racial

differences in HIV prevalence among (older) MSM account for racial differences in HIV incidence among YMSM. However, they do not account solely for these differences because other partnership factors such as partnership specific sex risk behavior and individual factors such as ART coverage do also contribute to racial differences in HIV incidence among YMSM. These findings align with current hypotheses in the corresponding literature[5].

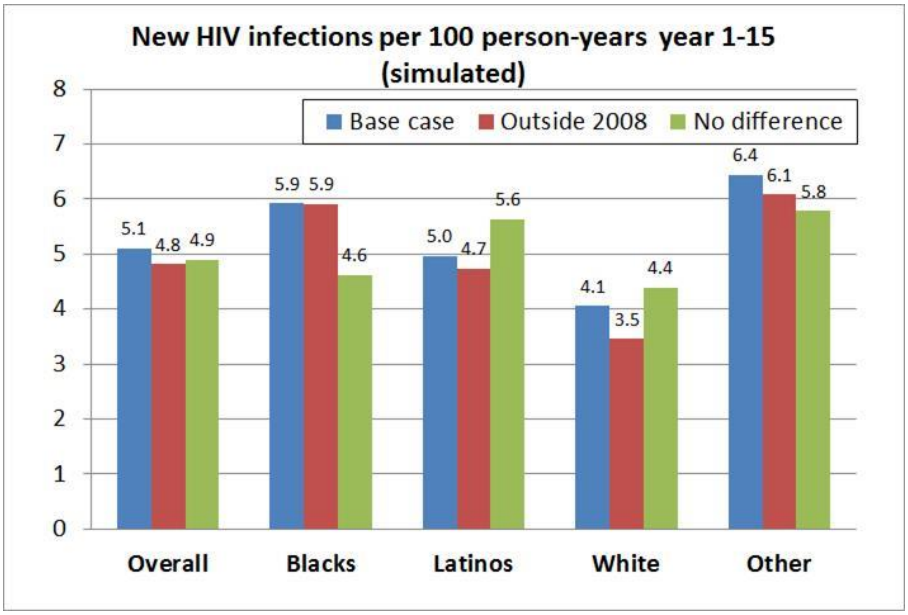


Figure 21: Simulated HIV infections per 100 person years stratified by race over 15 years for base-case scenario (see Figure 20), counterfactual scenario 'Outside HIV prevalence 2008', and counterfactual scenario 'No difference in HIV prevalence'.

To study the impact of HIV prevalence among those aging-in on HIV spread and racial disparities among YMSM over the total simulated time horizon of 15 years we studied a counterfactual scenario where all YMSM aging into the study population are not infected, i.e. the HIV prevalence among YMSM of age 16,17, and 18 years aging into the simulated population was set to 0.0%. HIV prevalence for YMSM of age 16, 17, and 18 who age-in in the base-case scenario is shown in Table 35 in section SDC 4.2. The total number of new HIV infections which occurred over 15 years in the counterfactual scenario with no aging-in HIV prevalence (i.e., 0.0% HIV prevalence for those aging-in) was 2846 with 1163 cases occurring among Black YMSM, 742 among Latino YMSM, and 705 among White YMSM. In comparison, 3076 new HIV infections occurred overall in the base-case scenario, 1220 new HIV

infections occurred among Black YMSM, 836 new HIV infections among Latino YMSM, and 770 among White YMSM. Figure 22 shows the HIV prevalence stratified by race over time for both scenarios. We observe a small decrease in HIV incidence when there are no HIV infected YMSM entering the simulated population: Overall it decreased by 7.5%, among Black YMSM by 4.7%, among Latinos by 11.2%, and among Whites by 8.4%. Differences in HIV prevalence in year 15 where similar in both scenarios (i.e., Black-White and Black-Latino HIV prevalence ratios in year 15 in the base-case scenario are 1.33 and 1.12 vs. Black-White and Black-Latino HIV prevalence ratios in year 15 in the counterfactual scenario are 1.35 and 1.12 respectively). Thus, we expect differences in the aging-in HIV prevalence for 16, 17, and 18 year old YMSM entering the simulated YMSM population to only have a marginal impact on HIV incidence and prevalence and racial disparities in our simulation model.

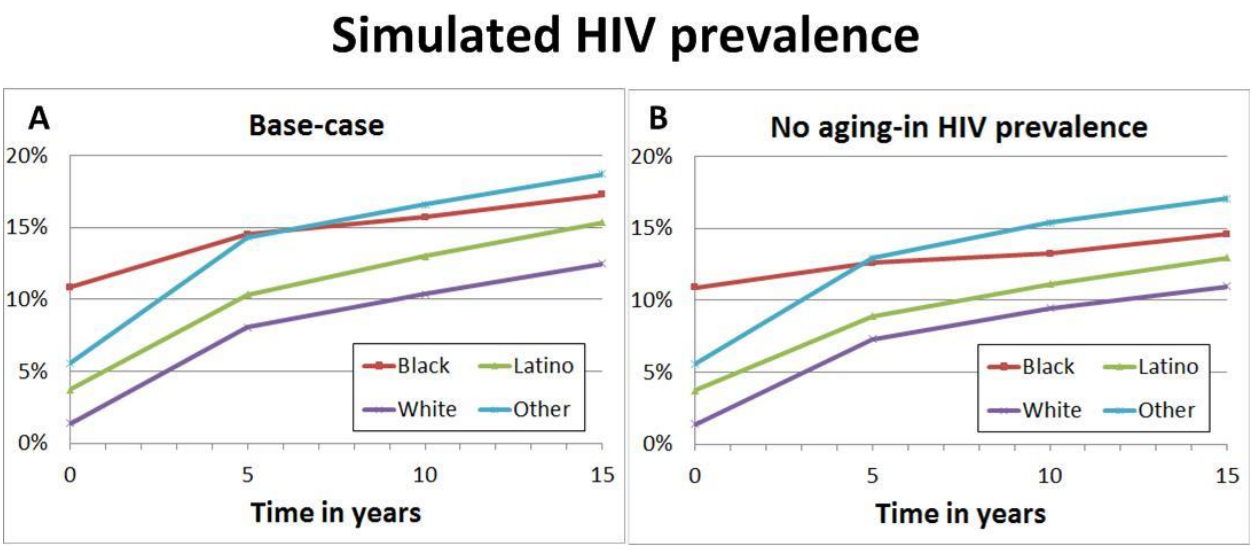


Figure 22: Simulated HIV prevalence among YMSM age 16 to 21.8 years over 15 years. Figure A shows HIV prevalence in base-case scenario and Figure B shows HIV prevalence in counterfactual where the HIV prevalence among YMSM aging into the YMSM population is 0.0%.

### 2.2.2.2 NG and CT prevalence

We analyze the sensitivity of the NG and CT transmission model comparing the base case scenario to three counterfactual scenarios where NG and CT prevalence among women is low (i.e., counterfactual

scenario ‘women low’) and the average duration of the rectal asymptomatic NG and CT infections is both decreased and increased by 50% (i.e., counterfactual scenarios ‘decreased rectal asymptomatic’ and ‘increased rectal asymptomatic’).

In case of the counterfactual scenario *women low*, we reduce the NG and CT prevalence among women to 0.2842% and 1% in comparison to 2.7% and 9.5% in the base case scenario (see section 3.2.3 in SDC 3). Figure 23C and Figure 23D show the simulated NG and CT prevalence of the counterfactual scenario *women low* over time. In comparison, Figure 23A and Figure 23B show the NG and CT prevalence for the base case scenario over time. We observe a marginally lower urethral and rectal NG and CT prevalence in the counterfactual scenario *women low* compared to the base case scenario. We reason that the number of partnerships and sex-acts with females do account for approximately 10% of all partnerships and sex-acts among YMSM, the male-to-female transmission risk for penile-vaginal sex being low compared to the transmission risk for male-male anal insertive and receptive sex (see section 3.2 in SDC 3), and only the urethra of YMSM can get infected during sex with a female. Therefore, changes in the NG and CT prevalence among women are expected to only marginally impact NG, CT and HIV transmission among YMSM, a result we see in Figure 23A-D.

Only few estimates of the duration of asymptomatic NG and CT infections are published and the estimated durations vary widely[52]. In particular rectal asymptomatic NG and CT infections are hypothesized to act as a hidden reservoir and thus might significantly contribute to the ongoing STI epidemic among MSM[95] (add two references). To study the impact of rectal asymptomatic infections on NG and CT transmission among YMSM we simulate two counterfactual scenarios where the duration of the rectal asymptomatic NG and CT infections is decreased and increased (i.e., counterfactual scenarios *decreased and increased rectal asymptomatic*). In these scenarios, we decreased and increased the original average duration of rectal asymptomatic NG and CT infections arbitrarily by 50% (for original values see Table 29 in section SDC 3.2). The decrease in the average duration of the rectal asymptomatic NG and CT infections could be seen as the result of increased testing rates in particular for rectal NG and

CT which cause an increased detection and treatment of rectal asymptomatic NG and CT infections and thus a decrease of the average duration of the rectal asymptomatic NG and CT infections. For the counterfactual scenario with decreased duration of rectal asymptomatic NG and CT infections, Figure 23E and Figure 23F show the urethral and rectal NG and CT prevalence among YMSM over time. We observe a reduction in the urethral and rectal NG and CT prevalence of approximately a 1/4 and a 1/3 (for urethral NG and CT prevalence it is slightly smaller than a 1/4 and a 1/3 reduction, for rectal NG and CT prevalence it is slightly greater than a 1/4 and a 1/3 reduction) in comparison to the urethral and rectal NG and CT prevalence in the base case scenario (i.e., Figure 23A and Figure 23B). For the counterfactual scenario with increased duration of rectal asymptomatic NG and CT infections, Figure 23G and Figure 23H shows the urethral and rectal NG and CT prevalence among YMSM over time. We observe that the rectal NG and CT prevalence increased by approximately 50% in comparison to the base-case scenario whereas urethral NG and CT prevalence increased only marginally. We conclude that NG and CT transmission among YMSM in our simulation model significantly depends on the duration of rectal asymptomatic infections and thus our simulated results align with the hypothesis that rectal asymptomatic NG and CT infections act as a hidden reservoir and thus significantly contribute to the spread of NG and CT among YMSM[95,96]. Further, we observe a bigger absolute increase in the rectal NG and CT prevalence than the absolute decrease in NG and CT prevalence when increasing and decreasing the duration of the average rectal asymptomatic NG and CT infection both by 50%. Taking into account that our conservative estimates of the average duration of rectal asymptomatic NG and CT infections used in the simulation model (see also Table 29) are in the lower range of the available estimates of the duration of (rectal) asymptomatic NG and CT infections, our simulation model is likely to provide a very conservative estimate und thus is likely to underestimate the spread an impact of NG and CT infections among YMSM.

Further, we observe for the *decreased rectal asymptomatic* scenario a steeper decline in both urethral and rectal NG and CT prevalence over time in comparison to the base-case scenario. For the *increased rectal*

*asymptomatic* scenario we observe an increase in both urethral and rectal NG and CT prevalence. Again, this finding confirms the importance of rectal infections in NG and CT and thus HIV spread among YMSM. It also suggests that if the average duration of rectal asymptomatic NG and CT infections is above a certain threshold, rectal asymptomatic NG and CT infections might be the reason why NG and CT prevalence might actually increase despite an increase in the HIV prevalence and thus testing for STI (note that in our simulation model HIV positive YMSM have a higher likelihood for being tested for STIs annually compared to HIV negative YMSM).

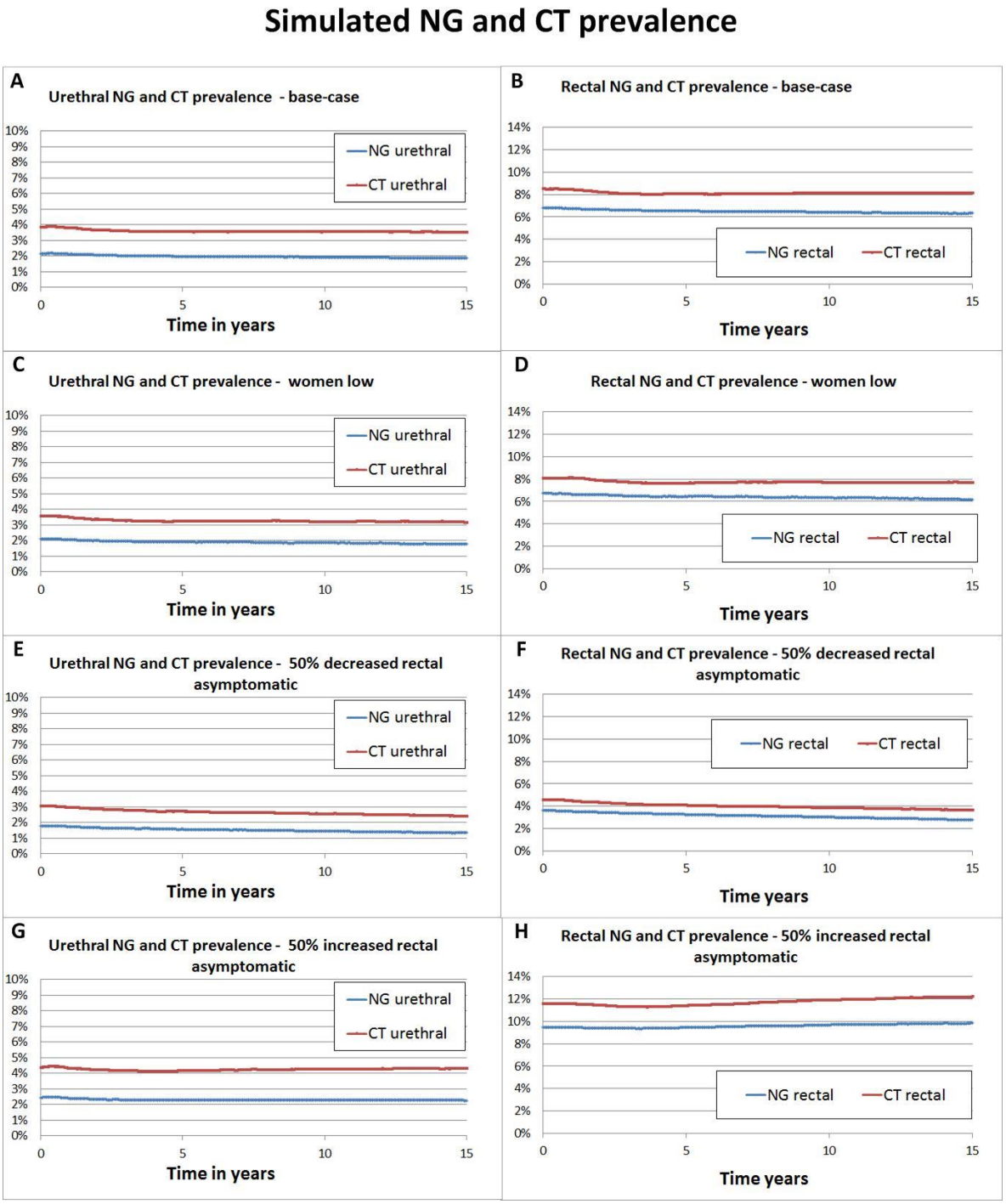


Figure 23: Urethral and rectal NG and CT prevalence among YMSM over simulated time horizon of 15 years. Figure A and Figure B show the urethral and rectal NG and CT prevalence for the base case scenario, Figure C and Figure2 D show the urethral and rectal NG and CT prevalence for the counterfactual scenario 'women low' (i.e., reduced NG and CT prevalence among women). Figure E and Figure F show the urethral and rectal NG and CT prevalence for the counterfactual scenario 'reduced rectal asymptomatic' (i.e., reduced average duration of the rectal asymptomatic NG and CT infections by 50%). Figure G and Figure H show the urethral and rectal NG and CT prevalence for the counterfactual scenario 'increased rectal asymptomatic' (i.e., increased average duration of the rectal asymptomatic NG and CT infections by 50%).

## 2.3 Validation summary

Simulated results match empirical results of the Crew 450 study for the partnership formation and dissolution model, HIV incidence, and urethral NG and CT prevalence. Additionally, our simulation based estimates of the overall number of HIV infections per year match relevant CDPH estimates. Simulated rectal NG and CT prevalence match published estimates and sensitivity analysis conducted show high face validity of the established simulation model. Thus, we conclude that the discrete-time agent-based network simulation model of partnership formation and dissolution and HIV transmission among YMSM is properly validated.

# VI. Supplemental Digital Content 6: Counterfactual Scenarios

# 1. No race-assortative mixing

In this counterfactual scenario (Figure 3B in the manuscript) we assume YMSM to choose partners not dependent on the race of the partner, i.e. YMSM of one race are equally likely to mix with YMSM of other races.

For one-night-partnerships and outside-partnerships we set the entries of the race mixing matrices (see Table 6, Table 7, Table 11 in SDC 2) to 0.25. For within-partnerships we removed the independent variables racial mixes (i.e., race-mix combinations BB, BL, BW, etc) from the Logistic regression model for partnership formation shown in Table 18 in SDC 2 such that the outcome newly formed within-partnership does not depend on the outcome race-mix of the partnership. Table 41 shows the coefficient estimates of the adjusted multivariate Logistic regression model introduced in Table 18 without the race-mix related independent variables.

| Covariates | Average Estimate* | p-value |
|---|---|---|
| **Intercept** | -4.66867 | <0.001 |
| **page(abs)** | -0.15378 | <0.001 |
| **avgage** | -0.1444 | <0.01 |
| **hivstatusp_p+hivstatusp_d** | -1.00179 | <0.01 |
| **partcon_p+partcon_d** | -0.86536 | <0.05 |

Table 41: Multivariate adjusted Logistic regression model shown in Table 18 in SDC 2 without race-mix as independent variable on the outcome newly formed partnerships. For description of dependent variables/covariates and intercept see Table 18 in SDC 2. *The combined variable (partcon_p+partcon_d), i.e. at least one of the partners has one ongoing/concurrent partnership, was included because it was statistically significant (i.e., $p<0.05$) in 6 out of 10 realizations.

# 2. No difference in HIV prevalence

In the counterfactual scenario *No difference in HIV prevalence* (Figure 3C in the manuscript) we assume no racial differences in HIV prevalence among older MSM partners (i.e., male partners of YMSM in outside-partnerships) and among YMSM at baseline. Thus, all races of older MSM have the same HIV prevalence as well as all races of YMSM have the same HIV prevalence in this counterfactual scenario at baseline. We set the HIV prevalence for all race groups among older MSM partners to the average HIV prevalence among Chicagoan MSM in 2011[11] being 17.18% for MSM age 21.8 to 29 years and 24.13% for MSM age 30 to 39 years. The HIV prevalence for all race groups among YMSM at baseline of the simulation (T1) was set to 5.61%. (Y)MSM were randomly assigned (at baseline) to be HIV infected in this counterfactual scenario.

# 3. No racial assortativity and no difference in HIV prevalence

The counterfactual scenario *No racial assortativity and no difference in HIV prevalence* (corresponds to Figure 3D in the manuscript) combines both the counterfactual scenarios *No race-assortative mixing* and *No difference in HIV prevalence*.

# 4. No increased HIV transmissibility and susceptibility due to NG and CT

To quantify the impact of NG and CT on HIV transmission among YMSM we simulate a counterfactual scenario where NG and CT infections do not cause increased HIV transmissibility and susceptibility, i.e.

we set the increased HIV transmission factors (IT) and increased HIV susceptibility factors (IS) shown in Table 31 in SDC 3 to 1. The difference between the HIV incidence in the base-case scenario with increased IT and IS due to NG and CT infections and the HIV incidence in this counterfactual denotes the number of HIV infections among YMSM attributable to NG and CT.

# 5. No HIV transmission in outside partnerships and 50% reduction in HIV transmission risk in outside partnerships

To study the impact of HIV transmission in outside partnerships and thus the elevated HIV risk of YMSM being in a partnership with an older MSM on HIV transmission among YMSM, we simulate two counterfactual scenarios: first, there happens no HIV transmission in outside partnerships of YMSM with both older MSM and females, , i.e. YMSM can't get infected in these partnerships. Second, we assume a 50% HIV transmission risk reduction in outside partnership compared to the base-case scenario, i.e. we divide HIV risk for YMSM in each sex-act in an outside partnership by 2. These counterfactual scenarios allow to quantify the impact of the reduction in HIV transmission risk in outside partnerships and thus allow to quantify the impact of HIV prevention efforts targeting partnerships of YMSM with older MSM partners[85].

# VII. Supplemental Digital Content 7: Addendum to results

# 1. HIV, NG, and CT prevalence among outside (older) MSM over time

Figure 24 shows the HIV prevalence among outside (i.e., older) MSM stratified by race an age. Overall, we observe an increase in HIV prevalence among older MSM over time due to the increase in the HIV prevalence among YMSM and aging-out of YMSM over the simulated time horizon. The HIV prevalence among MSM age 21.8 to 29 years starts to increase later (i.e., time-lag of about 1 year) compared to the HIV prevalence among YMSM (see also Figure 1A in the manuscript) and the HIV prevalence among MSM age 30 to 39 years starts to increase about 4 years after the simulation starts. This time-lag is caused by the time-lag of the increase in HIV prevalence in the MSM populations preceding the specific MSM age-group as well as the aging-out process of each MSM age-group and the different population sizes (for details about the update mechanisms see section SDC 3.1). Further, we observe that the HIV prevalence among Latino and White MSM age 21.8 to 29 years shows a steeper increase compared to Black MSM and thus the outside Latino and White MSM age 21.8 to 29 year HIV prevalence reflects the steeper increase in HIV prevalence and incidence among Latino and White YMSM in the simulated YMSM population over time (i.e., 3.1 and 3.4 fold increase in HIV prevalence among Latino and White MSM age 21.8 to 29 years compared to 1.7 fold increase in HIV prevalence among Black MSM age 21.8 to 29 years over 15 years). Similar, we observe a steeper increase in HIV prevalence among Latino and White MSM age 30 to 39 years compared to Black MSM age 30 to 39 years. However, the relative increase within this age-group is smaller compared to the increase in HIV prevalence among MSM age 21.8 to 29 years due to the aging-out processes and corresponding time-lag of the increase in HIV prevalence.

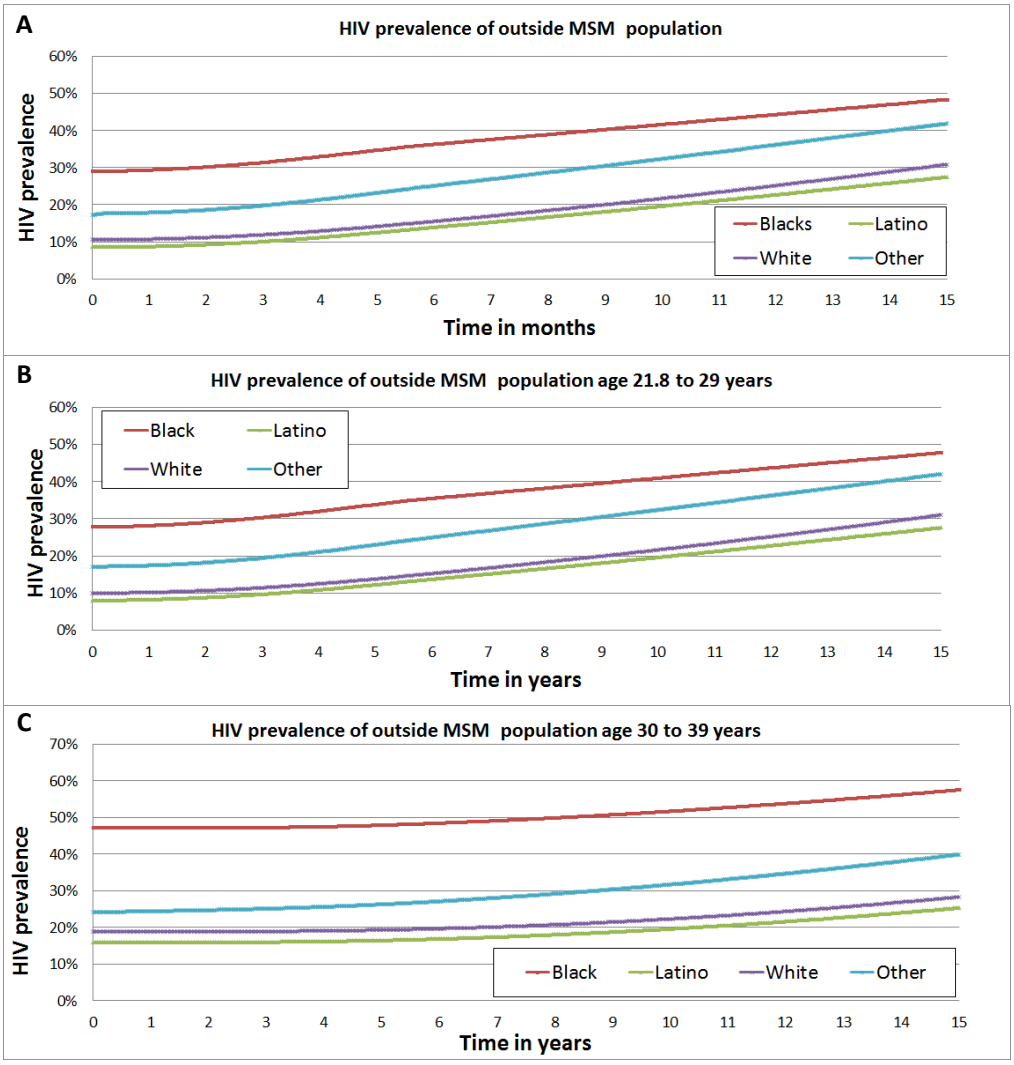


Figure 24: HIV prevalence among outside (older) MSM stratified by race over simulated time horizon of 15 years. Figure A shows overall HIV prevalence among outside MSM age 21.8 to 39 years stratified by race over 15 years. Figure B shows HIV prevalence among outside MSM age 21.8 to 29 years stratified by race over 15 years and Figure C shows HIV prevalence among outside MSM age 30 to 39 years stratified by race over 15 years.

Figure 25 shows the NG and CT prevalence among outside (i.e., older) MSM age 21.8 to 39 years over the total simulated time horizon. For details about the update mechanism of the NG and CT prevalence among outside MSM see section SDC 3.2. Overall, we observe small decreases and increases in NG and CT prevalence respectively because NG and CT prevalence among YMSM changes only marginally over time, see also Figure 19. CT prevalence among outside MSM increases marginally over time because of

the CDPH CT prevalence estimates[11] were relatively low compared to the CT prevalence among YMSM in the beginning of the simulation, see also section SDC 3.2.

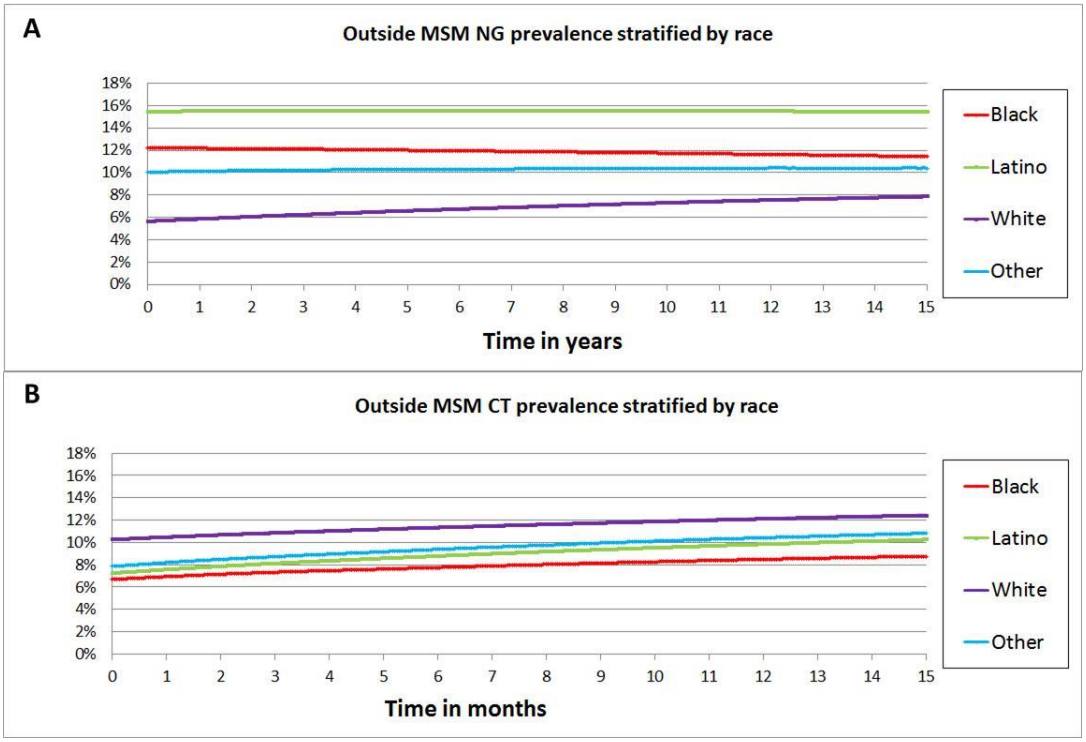


Figure 25: NG and CT prevalence among outside (older) MSM age 21.8 to 39 years stratified by race over simulated time horizon of 15 years. Figure A shows NG prevalence among outside MSM age 21.8 to 39 years stratified by race over 15 years. Figure B shows CT prevalence among outside MSM age 21.8 to 39 years stratified by race over 15 years.

# 2. HIV infections per 100 partnership years, male-male partnerships only

Figure 26 shows the HIV infections per 100 male-male partnership-years stratified by race-mix for the base-case scenario, which is described in the corresponding manuscript. For one-night-partnerships, HIV infections of YMSM per male-male partnership-year denote the number of HIV infections which occurred per average number of one-night-partnerships per year. For outside-partnerships, HIV infections of YMSM per male-male partnership-year denote the number of HIV infections which occurred in outside-partnerships where the YMSM was HIV-negative until the time of HIV infection. In within-partnerships (i.e. partnerships among YMSM), we define the number of male-male partnership-years (i.e., the denominator of the ratio HIV infections per 100 male-male partnership years) as the sum of the number of susceptible-infected partnership-years plus two times the number of susceptible-susceptible partnership-years. We doubled the number of susceptible-susceptible partnership-years in this calculation because the outcome HIV infections per 100 partnership-years denotes the HIV infection probability of a HIV-susceptible YMSM.

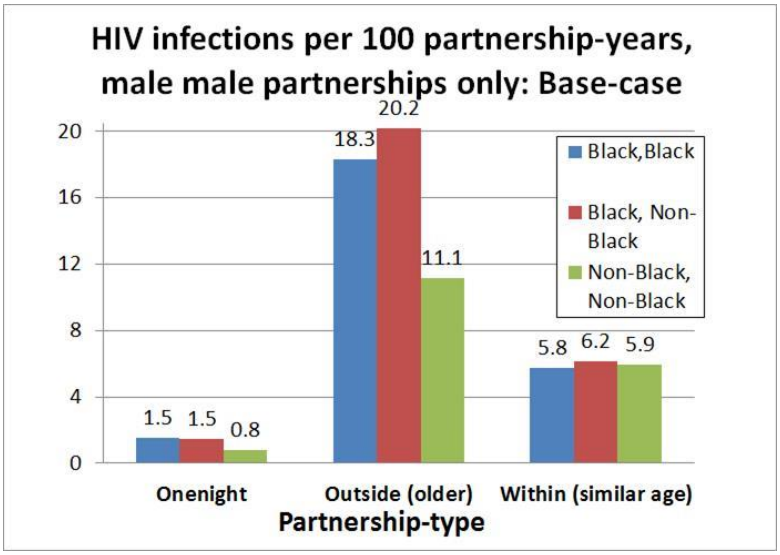


Figure 26: HIV infections per 100 partnership years, male-male partnerships only. Base-case scenario corresponding to Figures 1 and 2 in the manuscript.

In case of male-male outside partnerships, the race-mix Black,Black denotes a partnership of a Black YMSM with an older Black MSM. Male-male outside Black,Non-Black partnerships denote both outside partnerships of a Black YMSM with an older Non-Black MSM (green 'Black, (older)Non-Black', Figure 27) and outside partnerships of a Non-Black YMSM with an older Black MSM (purple 'Non-Black, (older)Black', Figure 27). As shown in Figure 27, HIV incidence per 100 partnership-years in outside partnerships where Non-Black YMSM have a partnership with older Black MSM is more than double the HIV incidence per 100 partnership-years in outside partnerships where Black YMSM have a partnership with older Non-Black MSM. The high HIV incidence for Non-Black YMSM with older Black MSM partnerships is mostly driven by the high HIV prevalence among older Black MSM and Other YMSM who have an older Black MSM partner because these partnerships of Other YMSM with an older Black MSM partner are more likely to be a serious partnership and thus Other YMSM have a higher likelihood of having unprotected anal intercourse (see also Table 15 in SDC 2.2).

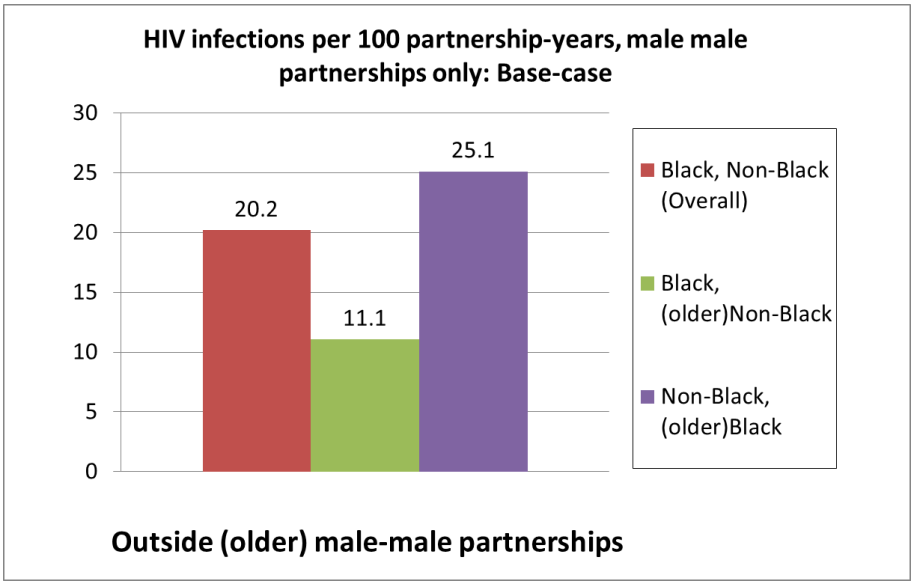


Figure 27: HIV infections per 100 partnership-years, outside male-male partnerships only for the base-case scenario over 15 years. Red denotes overall male-male outside partnerships where either the YMSM or the older MSM partner is Black. Green denotes outside partnerships of Black YMSM with older Non-Black MSM. Purple denotes outside male-male partnerships of Non-Black YMSM with older Black MSM.

# References


1. Sullivan PS, Salazar L, Buchbinder S, et al. Estimating the proportion of HIV transmissions from main sex partners among men who have sex with men in five US cities. *AIDS.* 2009;23(9):1153-1162.
2. Beyrer C, Baral SD, van Griensven F, et al. Global epidemiology of HIV infection in men who have sex with men. *Lancet.* 2012;380(9839):367-377.
3. Morris M, Goodreau S, Moody J. Sexual networks, concurrency, and STD/HIV. In: *Sexually transmitted diseases.* Holmes K(editor). New York: McGraw-Hill; 2007:109-126.
4. Goodreau SM, Carnegie NB, Vittinghoff E, et al. What drives the US and Peruvian HIV epidemics in men who have sex with men (MSM)? *PloS ONE.* 2012;7(11):e50522.
5. Maulsby C, Millett G, Lindsey K, et al. HIV among Black men who have sex with men (MSM) in the United States: a review of the literature. *AIDS Behav.* 2014;18(1):10-25.
6. Koblin BA, Chesney MA, Husnik MJ, et al. High-risk behaviors among men who have sex with men in 6 US cities: baseline data from the EXPLORE Study. *Am J Public Health.* 2003;93(6):926-932.
7. *R: A language and environment for statistical computing* [computer program]. Vienna, Austria: R Foundation for Statistical Computing; 2010.
8. Vittinghoff E, Douglas J, Judson F, et al. Per-Contact Risk of Human Immunodeficiency Virus Transmission between Male Sexual Partners. *Am J Epidemiol.* 1999; 150:306-311.
9. Friedman MR, Wei C, Klem ML, et al. HIV infection and sexual risk among men who have sex with men and women (MSMW): a systematic review and meta-analysis. *PloS ONE.* 2014;9(1):e87139.
10. Herbenick D, Reece M, Schick V, et al. Sexual behavior in the United States: results from a national probability sample of men and women ages 14-94. *J Sex Med.* 2010;7 Suppl 5:255-265.
11. Chicago Department of Public Health. HIV Risk and Prevention Behaviors Among Men Who Have Sex With Men, Chicago, 2008 and 2011. Chicago, IL: *City of Chicago*; 2012.
12. Serovich JM, Mosack KE. Reasons for HIV disclosure or nondisclosure to casual sexual partners. *AIDS Educ Prev.* 2003; 15:70-80.
13. Perry SW, Card CA, Moffatt M, Jr., et al. Self-disclosure of HIV infection to sexual partners after repeated counseling. *AIDS Educ Prev.* 1994;6(5):403-411.
14. Krivitsky P, Handcock M. A separable model for dynamic networks. *J R Stat Soc: Series B.* 2014;76(1):39-46.
15. Pathela P, Hajat A, Schillinger J, et al. Discordance between sexual behavior and self-reported sexual identity: a population-based survey of New York City men. *Ann Intern Med.* 2006;145(6):416-425.
16. Tieu HV, Li X, Donnell D, et al. Anal Sex Role Segregation and Versatility Among Men Who Have Sex With Men: EXPLORE Study. *J Acquir Immune Defic Syndr.* 2013;64(1):121-125.
17. Chicago Department of Public Health. HIV/HCV Risk Behaviors, Testing, Prevention and Care. IDU in Chicago, 2009. Chicago, IL: *City of Chicago*; 2009.
18. Chicago Department of Public Health. HIV/STI Surveillance Report, 2013. Chicago, IL: *City of Chicago*; 2013.
19. Chicago Department of Public Health. The Heterosexual HIV Epidemic in Chicago: Insights into the Social Determinants of HIV. Chicago, IL: *City of Chicago*; 2011.

20. Centers for Disease Control and Prevention. Regional Infertility Prevention Projects. All 2011 Chlamydia National, Regional and State Profiles. Accessed at: http://www.cdc.gov/std/chlamydia2011/CT_Supplement_2011.pdf. Access Date: June 26th, 2014.
21. Centers for Disease Control and Prevention. Gonorrhea-Prevalence Among Women Aged 16-24 Years Entering the National Training Program, by State of Residence, United States and Outlying Areas, 2011. Accessed at: http://www.cdc.gov/std/stats11/figures/m.htm. Access Date: June 26th, 2014.
22. Hollingsworth TD, Anderson RM, Fraser C. HIV-1 transmission, by stage of infection. *J Inf Dis.* 2008;198(5):687-693.
23. Armbruster B, Beck EC, Waheed M. The importance of extended high viremics in models of HIV spread in South Africa. *Health Care Manag Sci.* 2014;17:182-193.
24. Todd J, Glynn JR, Marston M, et al. Time from HIV seroconversion to death: a collaborative analysis of eight studies in six low and middle-income countries before highly active antiretroviral therapy. *AIDS.* 2007;21 Suppl 6:S55-63.
25. Mills EJ, Bakanda C, Birungi J, et al. Life expectancy of persons receiving combination antiretroviral therapy in low-income countries: a cohort analysis from Uganda. *Ann Int Med.* 2011;155(4):209-216.
26. Cohen MS, Chen YQ, McCauley M, et al. Prevention of HIV-1 infection with early antiretroviral therapy. *New Engl J Med.* 2011;365(6):493-505.
27. May M, Gompels M, Sabin C. Life expectancy of HIV-1-positive individuals approaches normal conditional on response to antiretroviral therapy: UK Collaborative HIV Cohort Study. *Journal of the International AIDS Society.* 2012;15(6)Suppl 4.
28. Swindells S, Cobos DG, Lee N, et al. Racial/ethnic differences in CD4 T cell count and viral load at presentation for medical care and in follow-up after HIV-1 infection. *AIDS.* 2002;16(13):1832-1834.
29. Weintrob AC, Grandits GA, Agan BK, et al. Virologic response differences between African Americans and European Americans initiating highly active antiretroviral therapy with equal access to care. *J Acquir Immune Defic Syndr.* 2009;52(5):574-580.
30. Cohen MS, Shugars DC, Fiscus SA. Limits on oral transmission of HIV-1. *Lancet.* 2000;356(9226):272.
31. Buchbinder SP, Vittinghoff E, Heagerty PJ, et al. Sexual risk, nitrite inhalant use, and lack of circumcision associated with HIV seroconversion in men who have sex with men in the United States. *J Acquir Immune Defic Syndr.* 2005;39(1):82-89.
32. Vermund SH, Qian HZ. Circumcision and HIV prevention among men who have sex with men: no final word. *JAMA.* 2008;300(14):1698-1700.
33. Leynaert B, Downs AM, de Vincenzi I. Heterosexual transmission of human immunodeficiency virus: variability of infectivity throughout the course of infection. European Study Group on Heterosexual Transmission of HIV. *Am J Epidemiol.* 1998;148(1):88-96.
34. Dave SS, Fenton KA, Mercer CH, et al. Male circumcision in Britain: findings from a national probability sample survey. *Sex Transm Infect.* 2003;79(6):499-500.
35. Bailey RC, Moses S, Parker CB, et al. Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomised controlled trial. *Lancet.* 2007;369(9562):643-656.
36. Wilson DP, Law MG, Grulich AE, et al. Relation between HIV viral load and infectiousness: a model-based analysis. *Lancet.* 2008;372(9635):314-320.
37. Chu H, Gange SJ, Li X, et al. The effect of HAART on HIV RNA trajectory among treatment-naïve men and women: a segmental Bernoulli/lognormal random effects model with left censoring. *Epidemiology.* 2010;21:S25-34.

38. Granich RM, Gilks CF, Dye C, et al. Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. *Lancet.* 2009;373(9657):48-57.
39. Weller S, Davis K. Condom effectiveness in reducing heterosexual HIV transmission. *The Cochrane Database Sys Rev.* 2002(1):CD003255.
40. U.S. Census Bureau. Chicago Region. 2014; Accessed at: http://www.census.gov/regions/chicago/. Accessed June 14, 2014.
41. McCormick AW, Walensky RP, Lipsitch M, et al. The effect of antiretroviral therapy on secondary transmission of HIV among men who have sex with men. *Clin Inf Dis.* 2007;44(8):1115-1122.
42. Fleming DT, Wasserheit JN. From epidemiological synergy to public health policy and practice: the contribution of other sexually transmitted diseases to sexual transmission of HIV infection. *Sex Transm Inf.* 1999;75(1):3-17.
43. Bernstein KT, Marcus JL, Nieri G, et al. Rectal gonorrhea and chlamydia reinfection is associated with increased risk of HIV seroconversion. *J Acquir Immune Defic Syndr.* 2010;53(4):537-543.
44. Jin F, Prestage GP, Imrie J, et al. Anal sexually transmitted infections and risk of HIV infection in homosexual men. *J Acquir Immune Defic Syndr.* 2010;53(1):144-149.
45. Koblin BA, Husnik MJ, Colfax G, et al. Risk factors for HIV infection among men who have sex with men. *AIDS.* 2006;20(5):731-739.
46. Chesson HW, Pinkerton SD. Sexually transmitted diseases and the increased risk for HIV transmission: implications for cost-effectiveness analyses of sexually transmitted disease prevention interventions. *J Acquir Immune Defic Syndr.* 2000;24(1):48-56.
47. Vriend HJ, Lugner AK, Xiridou M, et al. Sexually transmitted infections screening at HIV treatment centers for MSM can be cost-effective. *AIDS* 2013;27(14):2281-2290.
48. Mushayabasa S, Tchuenche JM, Bhunu CP, et al. Modeling gonorrhea and HIV co-interaction. *Bio Systems.* 2011;103(1):27-37.
49. Chen MI, Ghani AC, Edmunds WJ. A metapopulation modelling framework for gonorrhoea and other sexually transmitted infections in heterosexual populations. *J R Soc Inf.* 2009;6(38):775-791.
50. Kretzschmar M, Welte R, van den Hoek A, Postma MJ. Comparative model-based analysis of screening programs for Chlamydia trachomatis infections. *Am J Epidemiol.* 2001;153(1):90-101.
51. Low N, McCarthy A, Macleod J, et al. Epidemiological, social, diagnostic and economic evaluation of population screening for genital chlamydial infection. *Health Technol Assess.* 2007;11(8):iii-iv, ix-xii, 1-165.
52. Davies B, Anderson SJ, Turner KM, et al. How robust are the natural history parameters used in chlamydia transmission dynamic models? A systematic review. *Theor Biol and Med Model.* 2014;11:8.
53. Kent CK, Chaw JK, Wong W, et al. Prevalence of rectal, urethral, and pharyngeal chlamydia and gonorrhea detected in 2 clinical settings among men who have sex with men: San Francisco, California, 2003. *Clin Infec Dis.* 2005;41(1):67-74.
54. Patton ME, Kidd S, Llata E, et al. Extragenital gonorrhea and chlamydia testing and infection among men who have sex with men--STD Surveillance Network, United States, 2010-2012. *Clin Infect Dis.* 2014;58(11):1564-1570.
55. Bissessor M, Tabrizi SN, Fairley CK, et al. Differing Neisseria gonorrhoeae bacterial loads in the pharynx and rectum in men who have sex with men: implications for gonococcal detection, transmission, and control. *J Clin Microbiol.* 2011;49(12):4304-4306.
56. Kramer MA, Aral SO, Curran JW. Self-reported behavior patterns of patients attending a sexually transmitted disease clinic. *Am J Public Health.* 1980;70(9):997-1000.

57. Geisler WM. Duration of untreated, uncomplicated Chlamydia trachomatis genital infection and factors associated with chlamydia resolution: a review of human studies. *J Infec Dis.* 2010;201 Suppl 2:S104-113.
58. Sullivan PS, Peterson J, Rosenberg ES, et al. Understanding racial HIV/STI disparities in black and white men who have sex with men: a multilevel approach. *PloS ONE.* 2014;9(3):e90514.
59. Farley TA, Cohen DA, Elkins W. Asymptomatic sexually transmitted diseases: the case for screening. *Prev Med.* 2003;36(4):502-509.
60. Turner KM, Adams EJ, Gay N, et al. Developing a realistic sexual network model of chlamydia transmission in Britain. *Theor Biol Med Model.* 2006;3:3.
61. Hooper RR, Reynolds GH, Jones OG, et al. Cohort study of venereal disease. I: the risk of gonorrhea transmission from infected women to men. *Am J Epidemiol.* 1978;108(2):136-144.
62. Althaus CL, Heijne JC, Low N. Towards more robust estimates of the transmissibility of Chlamydia trachomatis. *Sex Transm Dis.* 2012;39(5):402-404.
63. Morin BR, Medina-Rios L, Camacho ET, et al. Static behavioral effects on gonorrhea transmission dynamics in a MSM population. *J Theor Biol.* 2010;267(1):35-40.
64. Philipps AJ. Chlamydial Infections. In: Nelson A.L. WJA, ed. *Sexually Transmitted Diseases: A Practical Guide for Primary Care*. Totowa, NJ: Humana Press; 2007:127-151.
65. Sweet RL, Gibbs RS. *Infectious Diseases of the Female Genital Tract.* 5 ed. Philadelphia: Wolters Kluwer, Lippincott Williams & Wilkins; 2009.
66. National Institute of Allergy and Infectious Diseases, National Institute of Health, Department of Health and Human Services. Workshop Summary: Scientific Evidence on Condom Effectiveness for Sexually Transmitted Disease (STD) Prevention. Hendorn, VA; 2000. Accessed at: http://www.niaid.nih.gov/about/organization/dmid/documents/condomreport.pdf. Access Date: June 27th, 2014.
67. Mehta SD, Moses S, Agot K, et al. Adult male circumcision does not reduce the risk of incident Neisseria gonorrhoeae, Chlamydia trachomatis, or Trichomonas vaginalis infection: results from a randomized, controlled trial in Kenya. *J Infec Dis.* 2009;200(3):370-378.
68. Van Howe RS. Sexually transmitted infections and male circumcision: a systematic review and meta-analysis. *ISRN Urol.* 2013;2013:109846.
69. Lau CY, Qureshi AK. Azithromycin versus doxycycline for genital chlamydial infections: a meta-analysis of randomized clinical trials. *Sex Transm Dis.* 2002;29(9):497-502.
70. Moran JS, Levine WC. Drugs of choice for the treatment of uncomplicated gonococcal infections. *Clin Infec Dis.* 1995;20 Suppl 1:S47-65.
71. Newman LM, Moran JS, Workowski KA. Update on the management of gonorrhea in adults in the United States. *Clin Infec Dis.* 2007;44 Suppl 3:S84-101.
72. Darrow WW. Condom use and use-effectiveness in high-risk populations. 1989;16(3):157-160.
73. Joesoef MR, Linnan M, Barakbah Y, et al. Patterns of sexually transmitted diseases in female sex workers in Surabaya, Indonesia. *Int J STD AIDS.* 1997;8(9):576-580.
74. Rotchford K, Strum AW, Wilkinson D. Effect of coinfection with STDs and of STD treatment on HIV shedding in genital-tract secretions: systematic review and data synthesis. *Sex Transm Dis.* 2000;27(5):243-248.
75. Modjarrad K, Vermund SH. Effect of treating co-infections on HIV-1 viral load: a systematic review. *Lancet Infec Dis.* 2010;10(7):455-463.
76. U.S. Census Bureau. State and County QuickFacts. Chicago (city), Illinois. 2015; Accessed at: http://quickfacts.census.gov/qfd/states/17/1714000.html. Accessed 04/01, 2015.
77. U.S. Census Bureau. State and County QuickFacts. Chicago city, Illinois QuickLinks. 2015; Accessed at: http://quickfacts.census.gov/qfd/states/17/1714000lk.html. Accessed 04/01, 2015.

78. Purcell DW, Johnson CH, Lansky A, et al. Estimating the population size of men who have sex with men in the United States to obtain HIV and syphilis rates. *Open AIDS J.* 2012;6:98-107.
79. Nakagawa F, Lodwick RK, Smith CJ, et al. Projected life expectancy of people with HIV according to timing of diagnosis. *AIDS.* 2012;26(3):335-343.
80. Centers for Disease Control and Prevention (CDC), National Center for Injury Prevention and Control. WISQARS (Web-based Injury Statistics Query and Reporting System). 2015. Accessed at: http://www.cdc.gov/injury/wisqars/. Accessed 04/01/2015.
81. Jones RC, Harper-Jemison DM, Clark J.,et al. Leading Causes of Death in Chicago, 2007-2009. City of Chicago. 2013.
82. NetLogo itself: Wilensky U. NetLogo. http://ccl.northwestern.edu/netlogo/. *Center for Connected Learning and Computer Based Modeling, Northwestern University, Evanston, IL.* 1999.
83. Krivitsky P. Modeling of Dynamic Networks based on Egocentric Data with Durational Information. *Upenn Department of Statistics: Technical Reports and Preprints.* 2012;12(01):1-32.
84. Pastor-Satorras R, Rubí JM, Diaz-Guilera A. *Statistical mechanics of complex networks.* New York: Springer; 2003.
85. Mustanski B, Birkett M, Kuhns LM, et al. The Role of Geographic and Network Factors in Racial Disparities in HIV Among Young Men Who have Sex with Men: An Egocentric Network Study. *AIDS Behav.* 2014.
86. Centers for Disease Control and Prevention. HIV Surveillance Report, 2010; vol.22. Published March 2012. Accessed at: http://www.cdc.gov/hiv/topics/surveillance/resources/reports/. Accessed Date: June 27th, 2014.
87. Rosenberg E SP, Kelley C, Sanchez T, et al. Race and Age Disparities in HIV Incidence and Prevalence Among MSM in Atlanta, GA [CROI abstract 69]. In Special Issue: Abstracts From the 2014 Conference on Retroviruses and Opportunistic Infections. *Top Antivir Med.* 2014;22(e-1):34.
88. Lieb S, White S, Grigg BL, et al. Estimated HIV incidence, prevalence, and mortality rates among racial/ethnic populations of men who have sex with men, Florida. *J Acquir Immune Defic Syndr.* 2010;54(4):398-405.
89. Neaigus A, Jenness SM, Hagan H, et al. Estimating HIV incidence and the correlates of recent infection in venue-sampled men who have sex with men in New York City. *AIDS Behav.* 2012;16(3):516-524.
90. Sifakis F, Hylton JB, Flynn C, et al. Racial disparities in HIV incidence among young men who have sex with men: the Baltimore Young Men's Survey. *J Acquir Immune Defic Syndr.* 2007;46(3):343-348.
91. Balaji AB, Bowles KE, Le BC, Paz-Bailey G, Oster AM, Grp NS. High HIV incidence and prevalence and associated factors among young MSM, 2008. *AIDS.* 2013;27(2):269-278.
92. Hotton A, Gratzer B. Howard Brown Health Center: STI Annual Report, 2011. Chicago, IL: Howard Brown Center; 2012.
93. Xiridou M, Geskus R, De Wit J, et al. The contribution of steady and casual partnerships to the incidence of HIV infection among homosexual men in Amsterdam. *AIDS.* 2003;17(7):1029-1038.
94. Jansen IA, Geskus RB, Davidovich U, et al. Ongoing HIV-1 transmission among men who have sex with men in Amsterdam: a 25-year prospective cohort study. *Aids.* 2011;25(4):493-501.
95. Rieg G, Lewis RJ, Miller LG, et al. Asymptomatic sexually transmitted infections in HIV-infected men who have sex with men: prevalence, incidence, predictors, and screening strategies. *AIDS patient care and STDs.* 2008;22(12):947-954.
96. Lutz AR. Screening for asymptomatic extragential gonorrhea and chlamydia in men who have sex with men: significance, recommendations, and options for overcoming barriers to testing. *J LGBT Health.* 2015;2(1):1-8.