

# Assessing disparity in the distribution of HIV and sexually transmitted infections in Australia: a retrospective cross-sectional study using Gini coefficients

Phyu Mon Latt <sup>1,2</sup> Nyi Nyi Soe,<sup>1,2</sup> Xianglong Xu,<sup>2,3</sup> Rashidur Rahman,<sup>4</sup> Eric P F Chow <sup>1,4,5</sup> Jason J Ong <sup>1,4</sup> Christopher Fairley,<sup>1,4</sup> Lei Zhang <sup>1,2,6</sup>

**To cite:** Latt PM, Soe NN, Xu X, *et al.* Assessing disparity in the distribution of HIV and sexually transmitted infections in Australia: a retrospective cross-sectional study using Gini coefficients. *BMJ Public Health* 2023;1:e000012. doi:10.1136/bmjph-2023-000012

Received 20 February 2023  
Accepted 11 August 2023

## ABSTRACT

**Introduction** The risk of HIV and sexually transmitted infections (STIs) varies substantially across population groups in Australia. We examined this disparity in HIV/STI distribution using Gini coefficients, where scores closer to one indicate greater disparity.

**Methods** We used demographic and sexual behaviour data from the Melbourne Sexual Health Centre, between 2015 and 2018. We examined 88 642 HIV consultations, 92 291 syphilis consultations, 97 473 gonorrhoea consultations and 115 845 chlamydia consultations. We applied a machine learning-based risk assessment tool, MySTIRisk, to determine the risk scores. Based on individuals' risk scores and HIV/STIs diagnoses, we calculated the Gini coefficients for these infections for different subgroups.

**Results** Overall, Gini coefficients were highest for syphilis (0.60, 95% CI 0.57 to 0.64) followed by HIV (0.57, 95% CI 0.52 to 0.62), gonorrhoea (0.38, 95% CI 0.36 to 0.42) and chlamydia (0.31, 95% CI 0.28 to 0.35). Gay, bisexual and other men who have sex with men (GBMSM) had lower Gini coefficients compared with heterosexual men or women; HIV (0.54 vs 0.94 vs 0.96), syphilis (0.50 vs 0.86 vs 0.93), gonorrhoea (0.24 vs 0.57 vs 0.57) and chlamydia (0.23 vs 0.42 vs 0.40), respectively. The Gini coefficient was lower among 25–34 years than in other age groups for HIV (0.66 vs 0.83–0.90) and gonorrhoea (0.38 vs 0.43–0.47). For syphilis, the oldest age group ( $\geq 45$  years) had a lower Gini coefficient than 18–24 years (0.61 vs 0.70).

**Conclusions** Our study demonstrated that HIV/STIs are more evenly distributed among GBMSM, suggesting widely disseminated interventions for GBMSM communities. In contrast, interventions for heterosexual men and women should be more targeted at individuals with higher risk scores.

## INTRODUCTION

The WHO declared the target of ending the pandemic of sexually transmitted infections (STIs) by 2030. This is an enormous task considering that the WHO projected

## WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Disparities exist in the distribution of HIV and sexually transmitted infections (STIs) among different population groups.

## WHAT THIS STUDY ADDS

⇒ Our analysis provides specific insights into the distribution of HIV and STIs among different groups. Notably, we found a higher Gini coefficient for heterosexual men and women, highlighting a concentrated risk, especially for HIV and syphilis. Conversely, lower Gini coefficients for gay, bisexual and other men who have sex with men (GBMSM) indicate a widespread risk, particularly for chlamydia and gonorrhoea, regardless of their risk scores among GBMSMs.

## HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Our findings indicate that the interventions will be most cost-effective if they are focused on small, high-risk groups of heterosexuals. In contrast, the interventions for GBMSM should involve the entire GBMSM community, particularly for chlamydia and gonorrhoea, where risk appears to play a minimal role in identifying those most at risk.

that in 2020 more than one million STIs were contracted each day, leading to a total of 374 million new infections with one of the four curable STIs: chlamydia, gonorrhoea, syphilis and trichomoniasis.<sup>1</sup> Australia has observed a rapid rise in the incidence of STIs over the last decade, particularly in gay and bisexual men.<sup>2–5</sup>

In response to the rise in STIs, governments have implemented interventions to control these infections.<sup>6 7</sup> A key part of implementing these interventions is deciding how best to target populations at particularly high



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. Published by BMJ.

For numbered affiliations see end of article.

## Correspondence to

Prof Lei Zhang;  
lei.zhang1@monash.edu

risk for these infections.<sup>8–10</sup> The effectiveness of interventions for HIV/STIs depends on the distribution of these infections within a population. If infections are widespread across all risk levels, yet interventions only target those at highest risk, these efforts may not yield desired results. Conversely, if infections primarily affect high-risk individuals, but interventions are distributed across all risk groups, resources might not be used effectively.

Gini coefficients have primarily been used to measure economic inequality in countries. More recently, however, several studies in Canada, the UK and the USA have used Gini coefficients to investigate the inequalities in the geographical distribution of STIs in different locations to facilitate geographically specific interventions.<sup>11–14</sup> Recently, limited studies have used the Gini coefficients to measure the disparity in the distribution of STIs by risk<sup>15 16</sup> and by geographical location.<sup>17</sup> Gsteiger *et al*<sup>15</sup> compared the disparity of chlamydia by risk over time using two population datasets in the UK and reported that the Gini coefficients for chlamydia among females and found they had not changed over time; 0.30 (1999–2001) and 0.33 (2010–2012).<sup>15</sup> van Wees *et al*<sup>16</sup> calculated the Gini coefficients for chlamydia, gonorrhoea and syphilis among men who have sex with men (MSM) before and after the implementation of pre-exposure prophylaxis (PrEP) for HIV in the Netherlands. van Wees *et al* demonstrated that the Gini coefficients had increased for chlamydia (0.37–0.43) and syphilis (0.50–0.66) after PrEP became available, while gonorrhoea remained stable.<sup>16</sup> To date, however, no study has assessed the Gini index within different risk populations in the same study to compare different infections within different populations or used a composite measure of risk.<sup>15 16</sup> These studies estimated the risk of HIV/STI infections based on only up to five risk measures.<sup>15 16</sup> We have recently developed machine-learning approaches to calculating a composite risk score based on many individual risks.<sup>18–20</sup>

We aimed to estimate the disparity in the distribution of four infections (HIV, chlamydia, gonorrhoea and syphilis) across different population groups (GBMSM, heterosexual men and women) and different age groups using a composite measure of risk.

## METHODS

The Melbourne Sexual Health Centre (MSHC) is Australia's largest public sexual health clinic and offers free HIV/STI services to the general public. At MSHC, the individuals' demographic and sexual behavioural data are recorded through computer-assisted self-reported interviews at the initial visit and any subsequent visit that is at least 3 months apart. Diagnoses are recorded in the clinic database using predetermined data fields.<sup>19</sup>

## Data

We conducted a retrospective cross-sectional study using data from the MSHC. We extracted demographic and behavioural data from the medical record for individuals attending

between 2 March 2015 and 31 December 2018. The datasets for each infection included consultations where individuals were tested for specific infections at that consultation. The HIV dataset included a total of 88642 consultations, the syphilis dataset included 92291 consultations, the gonorrhoea dataset included 97473 consultations and the chlamydia dataset had 115845 consultations.<sup>20</sup>

## Estimating the risk scores of HIV/STIs

We applied a risk prediction tool, MySTIRisk, developed from a previous study to the dataset.<sup>20</sup> MySTIRisk is a machine learning-based risk assessment tool that estimates a risk score, which ranges from 0 to 1 with 1 indicating the highest risk, for each infection based on the predictors.<sup>20</sup> During the development of MySTIRisk, rigorous training and testing procedures were conducted. For the development process (training and testing) of the tool, we used clinic consultations tested for HIV/STIs at the MSHC between 2015 and 2018. The results demonstrated that the tool performed at an acceptable to excellent level, as indicated by the area under the curve (AUC) values: HIV (0.78), syphilis (0.84), gonorrhoea (0.78) and chlamydia (0.70).<sup>20</sup>

To ensure the reliability and generalisability of the tool, external validation was performed using two separate datasets. The first external validation dataset, from 2019, showed consistent and stable performance with AUC values of HIV (0.79), syphilis (0.85), gonorrhoea (0.81) and chlamydia (0.69). The second external validation dataset, covering the years 2020–2021, also demonstrated reliable performance with AUC values of HIV (0.71), syphilis (0.84), gonorrhoea (0.79) and chlamydia (0.69).

The most important predictors in our models were gender, age, country of birth, men who reported having sex with men, presence of STI symptoms at the time of visit, the number of casual sexual partners, condom use, the last time of drug injection (if present), past STIs, contact with someone diagnosed with STIs, and having sex with someone outside Australia and New Zealand. Since infection risk scores generated by our models were not equivalent to the probability of infection, we calibrated and fitted the data using a logistic function to provide the prevalence for each infection for the clients. MSHC is working on deploying the MySTIRisk tool as a public-facing web app<sup>21</sup> that will allow people to determine their risks and encourage early STI testing and diagnosis. We use the infection risk scores in the study to calculate Gini coefficients.

## Estimating Gini coefficients

Based on the above risk scores, we used Lorenz curves to plot the cumulative proportion of HIV/STIs as a function of the cumulative proportion of the clinic consultations ranked from lowest to highest risk scores that we generated from MySTIRisk. By definition, the Gini coefficient is a single number that measures the disparity of the distribution of HIV/STIs. The coefficient is calculated as the area beneath the line of perfect equality minus the area beneath the Lorenz curve, divided by the area

beneath the line of perfect equality.<sup>22</sup> A Gini coefficient of 0 indicates perfect equality, with a homogeneous distribution of HIV/STIs over the infection risk scores of the population. In contrast, Gini coefficients closer to 1 indicate that STI diagnoses are concentrated in parts of the population with higher infection risk scores.

We computed Lorenz curves and estimated Gini coefficients along with their 95% bootstrap CIs for HIV, syphilis, gonorrhoea and chlamydia infections for each risk group (heterosexual men, women and GBMSM). The bootstrap method, a resampling technique where each sample is generated through random selection with replacement from the original dataset, was employed to account for variability and provide robustness in the estimation of Gini coefficients.<sup>11 12 15</sup> We calculated the 2.5th and 97.5th percentiles from 5000 bootstrapped Gini coefficients to form the 95% bootstrap CIs. This methodology allowed us to ascertain the precision of our estimates and understand the distribution of these infections across different risk groups. We also conducted similar Gini coefficient analyses for various age groups (18–24 years, 25–34 years, 35–44 years, and 45 years and above) and compared the corresponding Gini coefficients and the positivity percentages across different subsets of the population.

We conducted these analyses using MATLAB R2022a.

### Patient and public involvement

Given the retrospective nature of this study, patients and the public were not directly involved in research design, recruitment or conduct. However, the findings, particularly regarding the disparity of HIV/STIs as measured by Gini coefficients, contribute to the understanding of infection distribution, potentially informing public health policies and interventions. The results will be disseminated through academic publication, thereby contributing to the wider body of knowledge in the field.

## RESULTS

### Demographic characteristics of study participants

For each of the four infections, gay, bisexual and other men who have sex with men (GBMSM) accounted for the majority (43%–52%) of the clinic consultations, followed by women (29%–33%) and heterosexual men (16%–23%) (table 1). Younger age groups (18–34 years) accounted for over 70% of clinic consultations. More than 55% of consultations were from individuals born overseas, and the remaining were born in Australia and New Zealand. The sexual risk predictors for each of the four infections are displayed in table 1.

### Infection positivity in study participants

As shown in table 2, the positivity (the positive percentage of the HIV/STI diagnosis) for chlamydia was 8.82% (95% CI 8.66% to 8.98%), for gonorrhoea was 7.78% (95% CI 7.61% to 7.95%), for syphilis was 1.94% (95% CI 1.85% to 2.03%) and HIV was 0.24% (95% CI 0.21% to 0.28%). The positivity was highest among GBMSM for all four infections, while women had the lowest positivity across all HIV/STIs.

### The risk scores among STI-positive participants

Among STI-positive participants, the median risk scores and IQRs were 0.76 (0.48–0.88) for the syphilis dataset, 0.66 (0.54–0.77) for the gonorrhoea dataset and 0.56 (0.44–0.68) for the chlamydia dataset (see figure 1).

### Gini coefficients for HIV/STIs, in all participants

The Lorenz curves for HIV, syphilis, gonorrhoea and chlamydia infections are shown in figure 2. The chlamydia curve was closest to the diagonal line, indicating that chlamydia diagnoses were more homogeneously distributed over the population regardless of their risk scores. In contrast, the syphilis curve was furthest away from the diagonal line, indicating that syphilis diagnoses were strongly associated with the individuals with higher risk scores. The Gini coefficients were lower for chlamydia (0.31, 95% CI 0.28 to 0.35) and gonorrhoea (0.38, 95% CI 0.36 to 0.42) and higher for HIV (0.57, 95% CI 0.52 to 0.62) and syphilis (0.60, 95% CI 0.57 to 0.64).

### Gini coefficients for HIV/STIs, by sexual orientation and age groups

Table 3 and figure 3 show the Gini coefficients in GBMSM, heterosexual men and women. The Gini coefficients for each infection were lowest for GBMSM, indicating that each infection was more homogeneously distributed in GBMSM populations than heterosexual men or women. This difference in Gini coefficients between GBMSM and heterosexual men or women was more pronounced for HIV and syphilis than for gonorrhoea or chlamydia.

The differences in Gini coefficients across all age groups were less marked than the differences by risk group. Those aged 25–34 had the lowest Gini coefficients of 0.66 (95% CI 0.61 to 0.72) for HIV, 0.38 (95% CI 0.35 to 0.41) for gonorrhoea and 0.31 (95% CI 0.28 to 0.35) for chlamydia, respectively. For syphilis, those aged 45 years and above had the lowest Gini coefficient of 0.61 (95% CI 0.56 to 0.66) (see figure 3).

## DISCUSSION

We found that there were substantial differences in the Gini coefficients across the four HIV/STIs and across the risk groups. The highest Gini coefficients were among heterosexual males and females for HIV or syphilis, suggesting these infections are concentrated among heterosexuals with higher infection risk scores. These risk scores were calculated using our machine learning based risk assessment tool, MySTIRisk, which considers a variety of factors such as age, gender, country of birth, sexual behaviour and history of previous STIs. Therefore, a higher risk score suggests a higher probability of acquiring HIV and STIs. In contrast, the Gini coefficients were lower among GBMSM, particularly for chlamydia or gonorrhoea, suggesting these infections are widely distributed throughout the GBMSM population. The significance of these findings is that for HIV and syphilis, the interventions will be most cost-effective if they are focused on small, high-risk groups of heterosexuals.

**Table 1** Characteristics of clinic consultations in individual datasets

Predictors	HIV (n=88 642 consultations)	Syphilis (n=92 291 consultations)	Gonorrhoea (n=97 473 consultations)	Chlamydia (n=115 845 consultations)
Age groups at consultation (years)				
18–24	22 233 (25.1%)	22 747 (24.6%)	25 238 (25.9%)	31 383 (27.1%)
25–34	43 400 (49.0%)	44 924 (48.7%)	47 052 (48.3%)	56 306 (48.6%)
35–44	14 486 (16.3%)	15 346 (16.6%)	15 631 (16.0%)	17 607 (15.2%)
>44	8523 (9.6%)	9274 (10.0%)	9552 (9.8%)	10 549 (9.1%)
Country of birth				
Australia and New Zealand	39 148 (44.2%)	40 990 (44.4%)	43 881 (45.0%)	51 162 (44.2%)
Overseas	46 003 (51.9%)	47 670 (51.7%)	49 835 (51.1%)	60 272 (52.0%)
Missing	3491 (3.9%)	3631 (3.9%)	3757 (3.9%)	4411 (3.8%)
Population type				
GBMSM	45 483 (51.3%)	48 068 (52.1%)	50 946 (52.3%)	50 322 (43.4%)
Heterosexual male	16 508 (18.6%)	17 089 (18.5%)	15 245 (15.6%)	26 975 (23.3%)
Female	26 651 (30.1%)	27 134 (29.4%)	31 282 (32.1%)	38 548 (33.3%)
Condom use with casual male partner				
Always	18 330 (20.7%)	18 825 (20.4%)	19 895 (20.4%)	20 490 (17.7%)
Sometimes	37 819 (42.7%)	39 485 (42.8%)	43 943 (45.1%)	48 764 (42.1%)
Never	5342 (6.0%)	5638 (6.1%)	6518 (6.7%)	7128 (6.2%)
Not applicable	1318 (1.5%)	1378 (1.5%)	1479 (1.5%)	1488 (1.3%)
Unsure/decline	3164 (3.6%)	3540 (3.8%)	3826 (3.9%)	4009 (3.5%)
Missing	22 669 (25.6%)	23 425 (25.4%)	21 812 (22.4%)	33 966 (29.3%)
Last time injected drugs not prescribed by doctor				
Never	83 634 (94.4%)	86 658 (93.9%)	91 452 (93.8%)	109 480 (94.5%)
Less than 3 months	1008 (1.1%)	1154 (1.3%)	1243 (1.3%)	1284 (1.1%)
within 3–12 months	372 (0.4%)	419 (0.5%)	455 (0.5%)	486 (0.4%)
More than 12 months	976 (1.1%)	1061 (1.1%)	1089 (1.1%)	1217 (1.1%)
Decline/unsure	1783 (2.0%)	2060 (2.2%)	2235 (2.3%)	2342 (2.0%)
Missing	869 (1.0%)	939 (1.0%)	999 (1.0%)	1036 (0.9%)
Having sex at overseas				
Yes	30 968 (34.9%)	31 962 (34.6%)	32 689 (33.5%)	40 917 (35.3%)
No	49 436 (55.8%)	51 496 (55.8%)	55 300 (56.7%)	64 794 (55.9%)
Unsure	4617 (5.2%)	4813 (5.2%)	5199 (5.3%)	5552 (4.8%)
Missing	3621 (4.1%)	4020 (4.4%)	4285 (4.4%)	4582 (4.0%)
Past history of gonorrhoea				
No	19 111 (21.6%)	20 250 (21.9%)	22 495 (23.1%)	27 349 (23.6%)
Yes	12 559 (14.2%)	13 715 (14.9%)	15 480 (15.9%)	15 198 (13.1%)
Unsure	1644 (1.9%)	1818 (2.0%)	1985 (2.0%)	2128 (1.8%)
Missing	55 328 (62.4%)	56 508 (61.2%)	57 513 (59.0%)	71 170 (61.4%)
History of non-specific urethritis				
No	24 050 (27.1%)	26 048 (28.2%)	28 180 (28.9%)	30 627 (26.4%)
Yes	857 (1.0%)	952 (1.0%)	1100 (1.1%)	1151 (1.0%)
Unsure	1254 (1.4%)	1420 (1.5%)	1486 (1.5%)	1594 (1.4%)
Missing	62 481 (70.5%)	63 871 (69.2%)	66 707 (68.4%)	82 473 (71.2%)
Past history of syphilis				
No	27 119 (30.6%)	28 478 (30.9%)	32 212 (33%)	36 846 (31.8%)

Continued

**Table 1** Continued

Predictors	HIV (n=88 642 consultations)	Syphilis (n=92 291 consultations)	Gonorrhoea (n=97 473 consultations)	Chlamydia (n=115 845 consultations)
Yes	4551 (5.1%)	5487 (5.9%)	5763 (5.9%)	5701 (4.9%)
Unsure	1644 (1.9%)	1818 (2.0%)	1985 (2.0%)	2128 (1.8%)
Missing	55 328 (62.4%)	56 508 (61.2%)	57 513 (59%)	71 170 (61.4%)
Sexual contact with someone diagnosed with gonorrhoea				
No	86 552 (97.6%)	89 992 (97.5%)	94 662 (97.1%)	113 074 (97.6%)
Yes	2090 (2.4%)	2299 (2.5%)	2811 (2.9%)	2771 (2.4%)
Sexual contact with someone diagnosed with chlamydia				
No	86 138 (97.2%)	89 628 (97.1%)	96 502 (99.0%)	111 353 (96.1%)
Yes	2504 (2.8%)	2663 (2.9%)	971 (1.0%)	4492 (3.9%)
Sexual contact with someone diagnosed with syphilis				
No	87 813 (99.1%)	91 287 (98.9%)	96 502 (99.0%)	114 867 (99.2%)
Yes	829 (0.9%)	1004 (1.1%)	971 (1.0%)	978 (0.8%)

GBMSM, gay, bisexual and other men who have sex with men.

In comparison, the interventions for GBMSM should involve the entire GBMSM community, particularly for chlamydia and gonorrhoea, where risk appears to play a minimal role in identifying those most at risk.

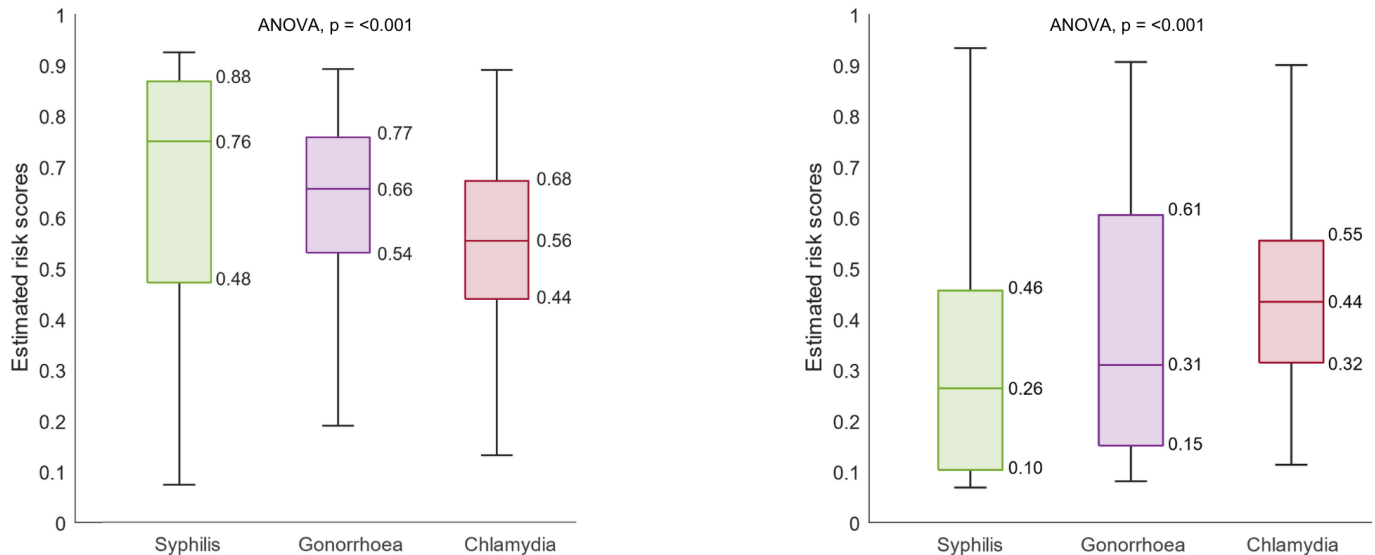
Our study demonstrated that Gini coefficients for syphilis and HIV were higher than those for chlamydia and gonorrhoea. In particular, the finding that syphilis demonstrated the highest disparity followed by gonorrhoea and chlamydia is consistent with another study conducted in the Netherlands.<sup>16</sup> Specifically, higher Gini coefficients indicate that infections are more concentrated in

a smaller group of individuals with higher risk scores. This finding implies that more targeted interventions are necessary for HIV and syphilis infection, while more widespread programmes would be suitable for chlamydia and gonorrhoea. Higher Gini coefficients among HIV and syphilis may be explained by the known fact that those infections are more concentrated among certain subpopulations such as men who have sex with men, and sex workers leading to a higher degree of inequality in terms of disease prevalence.<sup>23–25</sup> It means that a relatively small proportion of people disproportionately affected

**Table 2** Positivity (the positive percentage of the HIV/STI diagnoses) with corresponding 95% CI per total consultations

	HIV	Syphilis	Gonorrhoea	Chlamydia
<b>All</b>	213/88 642 0.2% (0.2% to 0.3%)	1790/92 291 1.9% (1.9% to 2.0%)	7583/97 473 7.8% (7.6% to 8.0%)	10 218/115 845 8.8% (8.7% to 9.0%)
Risk groups				
GBMSM	196/45 483 0.4% (0.4% to 0.5%)	1553/48 068 3.2% (3.1% to 3.4%)	6598/50 946 13% (12.7% to 13.2%)	5470/50 322 10.9% (10.6% to 11.1%)
Heterosexual male	11/16 508 0.1% (0.0% to 0.1%)	144/17 089 0.8% (0.7% to 1.0%)	355/15 245 2.3% (2.1% to 2.6%)	2236/26 975 8.3% (8.0% to 8.6%)
Female	9/26 651 0.0% (0.0% to 0.1%)	90/27 134 0.3% (0.3% to 0.4%)	629/31 282 2.0% (1.9% to 2.2%)	2513/38 548 6.5% (6.2% to 6.8%)
Age groups				
18–24 years	40/22 233 0.2% (0.1% to 0.2%)	275/22 747 1.2% (1.1% to 1.4%)	1880/25 238 7.5% (7.2% to 7.8%)	3251/31 383 10.4% (10.0% to 10.7%)
25–34 years	117/43 400 0.3% (0.2% to 0.3%)	840/44 924 1.9% (1.8% to 2.0%)	3891/47 052 8.3% (8.0% to 8.5%)	5113/56 306 9.1% (8.9% to 9.3%)
35–44 years	39/14 486 0.3% (0.2% to 0.4%)	388/15 346 2.5% (2.3% to 2.8%)	1171/15 631 7.5% (7.1% to 7.9%)	1148/17 607 6.5% (6.2% to 6.9%)
45 and above	20/8523 0.2% (0.1% to 0.4%)	282/9274 3.0% (2.7% to 3.4%)	639/9552 6.7% (6.2% to 7.2%)	704/10 549 6.7% (6.2% to 7.2%)

GBMSM, gay, bisexual and other men who have sex with men; STI, sexually transmitted infection.

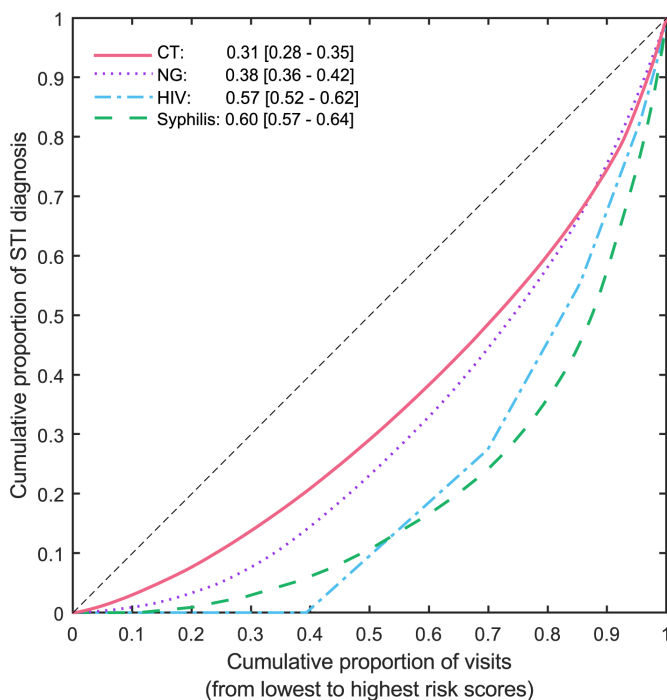


**Figure 1** Risk scores’ distribution for STI-positive versus STI-negative consultations. ANOVA, analysis of variance; STI, sexually transmitted infection.

by syphilis and HIV, can lead to higher Gini coefficients than more evenly distributed STIs such as gonorrhoea and chlamydia. Additionally, HIV and syphilis are more severe and chronic conditions with a social and cultural stigma attached to them, which can prevent the infected individuals from seeking the testing and treatment and may potentially lead to higher Gini coefficients.<sup>26–28</sup>

In another study, Gsteiger *et al*<sup>15</sup> used population data in the UK to calculate the Gini coefficients to measure the distribution of chlamydia. The authors compared

the Gini coefficients for STIs for two survey periods of Natsal-2 (1999–2001) and Natsal-3 (2010–2012) and found similar results for chlamydia with 0.30 (95% CI 0.12 to 0.50) in Natsal-2 and 0.33 (95% CI 0.18 to 0.49) in Natsal-3 among females. However, our study has a slightly higher Gini coefficient of 0.40 (95% CI 0.35 to 0.45) in women with chlamydia. First, the difference in study population could have contributed to the discrepancy. The Natsal studies employed general population data, whereas our study focused on data from a sexual health clinic. It is well known that individuals attending such clinics may have a higher risk profile for STIs than individuals from the community which can consequently influence the Gini coefficient. Second, our definition of the exposure variable differed from the Natsal studies. The latter used a single risk factor—the number of new opposite-sex partners in the previous year—while we adopted a broader approach. By employing machine-learning algorithms, we generated a composite risk score encompassing multiple exposure variables. This more comprehensive risk assessment likely impacted our Gini coefficient. Lastly, unlike the Natsal studies, which focused solely on heterosexual females, we included heterosexual, bisexual and women who have sex with women (WSW). Similarly, it is important to note that the Natsal studies provided limited analysis on subgroups for different age groups and population groups as it only included the opposite-sex contacts for sexual behaviour variables population because of a limited proportion of individuals having same-sex partner in the dataset.



**Figure 2** Lorenz curve and Gini coefficients for HIV/STIs. CT, chlamydia; NG, gonorrhoea; STIs, sexually transmitted infections.

Our finding that the Gini coefficients were significantly lower in GBMSM is consistent with a previous study.<sup>16</sup> van Wees *et al*<sup>16</sup> developed a risk score calculator using multivariable logistic regression for HIV/STIs and used this to calculate the Gini coefficients for chlamydia, gonorrhoea and syphilis during the period before PrEP (2009 to mid-2015)

**Table 3** Gini coefficients with 95% bootstrap CIs for HIV, syphilis, gonorrhoea and chlamydia across different risk and age groups

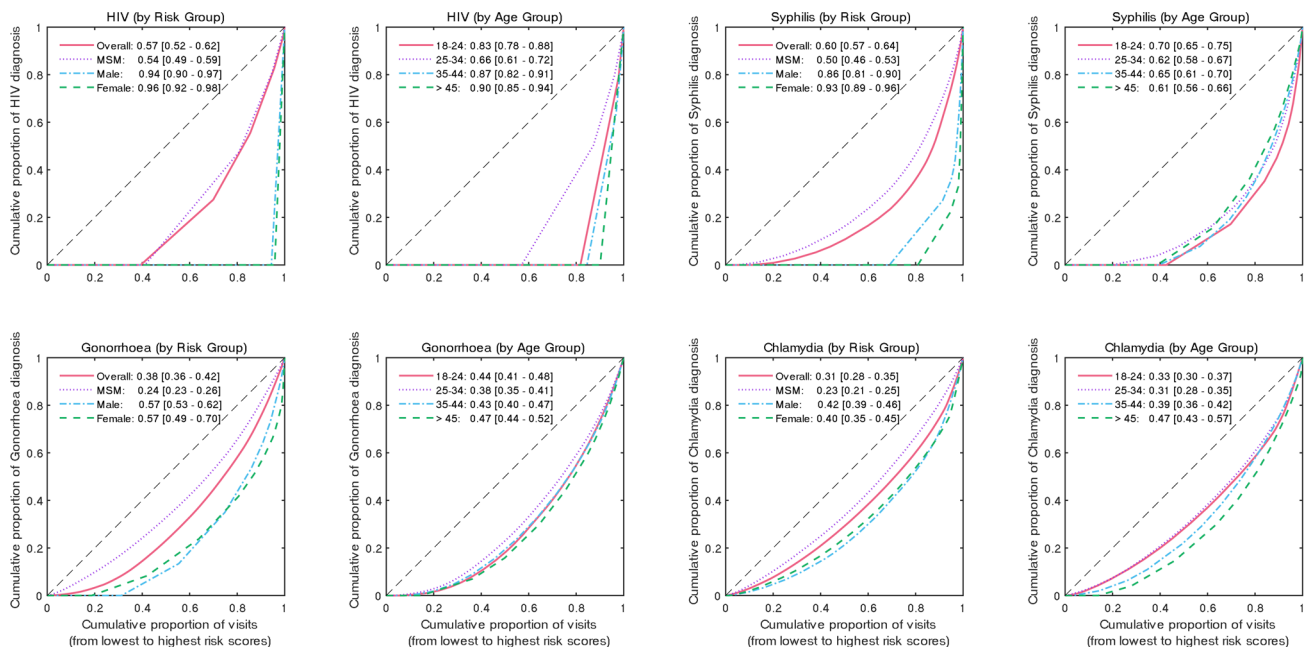
	HIV	Syphilis	Gonorrhoea	Chlamydia
Overall	0.57 (0.52 to 0.62)	0.60 (0.57 to 0.64)	0.38 (0.36 to 0.42)	0.31 (0.28 to 0.35)
Risk groups				
GBMSM	0.54 (0.49 to 0.59)	0.50 (0.46 to 0.53)	0.24 (0.23 to 0.26)	0.23 (0.21 to 0.25)
Heterosexual male	0.94 (0.90 to 0.97)	0.86 (0.81 to 0.90)	0.57 (0.53 to 0.62)	0.42 (0.39 to 0.46)
Female	0.96 (0.92 to 0.98)	0.93 (0.89 to 0.96)	0.57 (0.49 to 0.70)	0.40 (0.35 to 0.45)
Age groups				
18–24 years	0.83 (0.78 to 0.88)	0.70 (0.65 to 0.75)	0.44 (0.41 to 0.48)	0.33 (0.30 to 0.37)
25–34 years	0.66 (0.61 to 0.72)	0.62 (0.58 to 0.67)	0.38 (0.35 to 0.41)	0.31 (0.28 to 0.35)
35–44 years	0.87 (0.82 to 0.91)	0.65 (0.61 to 0.70)	0.43 (0.40 to 0.47)	0.39 (0.36 to 0.42)
45 and above	0.90 (0.85 to 0.94)	0.61 (0.56 to 0.66)	0.47 (0.44 to 0.52)	0.47 (0.43 to 0.57)

GBMSM, gay, bisexual and other men who have sex with men.

and after PrEP (mid-2015 to 2019) was introduced. This study examined the distribution of STIs among HIV-negative MSM in Amsterdam Cohort Studies and found a similar pattern to our study, but with a slightly higher Gini coefficients for gonorrhoea (0.46) and chlamydia (0.43) and lower Gini coefficient for syphilis (0.50). These higher Gini coefficients for chlamydia and gonorrhoea in the Netherlands study may be explained by the lower positivity of chlamydia and gonorrhoea in their dataset compared with our study (4.6% vs 8.8% for chlamydia and 5.1% vs 7.8% for gonorrhoea). However, the Gini coefficient for syphilis was lower in the Netherlands study despite the lower positivity (0.7% vs 1.9%). This variation may partly be explained by the

difference in sampling frame as the study only included HIV-negative MSM while our study included all GBMSM, regardless of the HIV status and the difference in the nature of risk prediction tools in calculating infection risk scores. While direct comparisons are not currently available, variations in healthcare systems, sexual health education and societal attitudes towards GBMSM populations between different countries such as the Netherlands and Australia could potentially lead to differences in the distribution of STIs such as syphilis. Further research is required to confirm this.

Based on our findings, we propose the need for both targeted and widespread intervention strategies for the control of HIV/STIs. For HIV and syphilis, which



**Figure 3** Lorenz curves showing the cumulative proportion of STI diagnoses among patients who visited the MSHC as a function of the cumulative proportion of all visits from lowest to highest risk score. In the figure, the diagonal line (black dash line) denotes perfect equality, which means an equal dispersion of the infection across the population. Note: MSM, male and female refer to GBMSM, heterosexual men and women, respectively. GBMSM, gay, bisexual and other men who have sex with men; MSM, men who have sex with men; MSHC, Melbourne Sexual Health Centre.

show a high degree of concentration in individuals with higher risk scores, targeted interventions are essential. Such interventions can be integrated into existing services and may include initiatives focused on testing, treatment adherence and education about safe sex practices. For chlamydia and gonorrhoea, which have a more widespread distribution across the population, broader public health strategies are needed. These could include regular STI screening programmes, public awareness campaigns and improving access to treatment, which can be integrated into general healthcare services. The strategic integration of these interventions into existing public health programmes and policies could contribute significantly to the control of HIV/STIs.

To our knowledge, this is the first research from Australia to examine the distribution of four STIs across different risk populations (GBMSM, heterosexual men and women) and different age groups. The main strength of the study is the use of a composite risk score that was only available because of the extensive data on sexual risk from attendees at MSHC. This composite score was generated using a machine learning approach<sup>20</sup> and is likely to be a better representation of overall risk than a single epidemiological measure.

This study has several limitations. First, the predictive criteria are based on the clients' self-reported information, which is subject to recall, non-response and social-desirability biases. However, there is no other way to collect risk information and we have previously shown that the self-interview method which we used is the least influenced by social-desirability bias.<sup>29</sup> Second, the datasets only included MSHC clients, who are at higher risk than the general population that includes lower-risk individuals. This may lead to an underestimate of the Gini index because our dataset likely under-represents the lower-risk individuals. While our focus was on the disparities within the clinic population, we acknowledge that this may limit the generalisability of our findings to other populations or settings. Therefore, it is important for policymakers and public health officials to consider these limitations when applying our findings to their respective contexts. Caution must be applied when extrapolating these results to broader contexts, as different populations may present unique risk profiles. Further studies in diverse settings are necessary to validate and extend our findings. Third, our unit of analysis was client consultations, not individual clients. This means our study reflects the number of consultations rather than the number of unique individuals. This approach could potentially over-represent individuals who had multiple consultations, thereby skewing the positivity rates of STIs and creating potential bias in estimating Gini coefficients. Nevertheless, given our large sample size and our focus on internal population disparities, we believe this approach's impact on our findings is minimal. Fourth, we only used data from 2015 to 2018 because in 2015 we moved from culture to nucleic acid amplification testing for gonorrhoea<sup>30</sup> and this period was too brief to identify a changing trend

in HIV/STIs in Australia. Fifth, although our dataset was large enough for machine learning training and testing, the number of HIV and syphilis-positive cases was notably low, which may affect MySTIRisk prediction risk scores and Gini coefficients. Sixth, as the PrEP uptake status was not included as a predictor, we were unable to examine the changes in Gini coefficients before and after PrEP utilisation. Seventh, for gonorrhoea and syphilis, we did not include the anatomical site of infection as a predictor, for which we could not identify the distributions of STIs at different anatomical sites. Looking forward, we propose several directions for future research that could address our study's limitations and enrich our understanding of HIV/STI disparities. Longitudinal studies are crucial for capturing shifts in HIV/STI distribution over time and evaluating the impact of interventions like PrEP. Additionally, exploring STI distribution across various anatomical sites could offer insights for targeted prevention strategies.

While our current study explores disparities in the distribution of HIV/STIs, we have not distinguished between high-risk and low-risk individuals based on explicit thresholds; a more nuanced approach is planned for future research. In anticipation of these future investigations, we intend to define these thresholds and risk subgroups, balancing factors such as disease burden, healthcare capacity and government funding. Therefore, the objective of our ongoing research is to refine our existing findings, thereby guiding the development of HIV/STIs intervention strategies and policy.

In conclusion, our study demonstrates that disparities exist in the distribution of HIV/STIs among different population groups, with implications for policy and interventions. The higher concentration of HIV and syphilis among heterosexual men and women indicates the need to identify and target high-risk subsets with focused testing and treatment. In contrast, the widespread distribution of chlamydia and gonorrhoea among GBMSM reinforces implementing broader screening and prevention. Estimating Gini coefficients enables tailored, data-driven approaches to early HIV/STI testing and treatment for those most affected. Our findings highlight the potential of Gini coefficients to inform resource allocation and policies aimed at HIV/STI control through precision public health strategies.

#### Author affiliations

<sup>1</sup>Central Clinical School, Faculty of Medicine, Nursing and Health Sciences, Monash University, Clayton, Victoria, Australia

<sup>2</sup>Artificial Intelligence and Modelling in Epidemiology Program, Melbourne Sexual Health Centre, Alfred Health, Carlton, Victoria, Australia

<sup>3</sup>School of Public Health, Shanghai University of Traditional Chinese Medicine, Shanghai, China

<sup>4</sup>Melbourne Sexual Health Centre, Alfred Health, Carlton, Victoria, Australia

<sup>5</sup>Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Victoria, Australia

<sup>6</sup>China-Australia Joint Research Center for Infectious Diseases, School of Public Health, Xi'an Jiaotong University Health Science Center, Xi'an, Shaanxi Province, People's Republic of China



**Acknowledgements** The authors would like to thank Monash University for supporting PhD scholarship for PL. The authors also express their gratitude to all those who have contributed to this study.

**Contributors** LZ conceived of the study. PML and NNS conducted the data analysis. XX, RR, JJO and EPFC and CF contributed to the design of the study. PML wrote the manuscript with input from all authors. LZ acted as a guarantor.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Patient and public involvement** Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

**Patient consent for publication** Not applicable.

**Ethics approval** The Alfred Hospital Ethics Committee in Melbourne, Australia granted ethical approval (project number: 78/22). The Alfred Hospital Ethics Committee's pertinent standards and regulations were followed for all procedures. As this is a retrospective study with low risk to the individual's privacy, the Alfred Hospital Ethics Committee waived the informed consent. Before computational analysis, all identifying information about the study participants was removed. The deidentified data file was securely stored on the computer servers of Melbourne Sexual Health Centre, within Alfred Health and Monash University's HELIX computational platform. Access to the database was restricted to the researchers named in the ethics application through password protection and stringent data handling protocols were followed to ensure data privacy and confidentiality.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** No data are available. No data are available due to the sensitive nature of the study. The data analysed during the current study are not publicly available due to their sensitive nature. These data encompass specific demographic and sexual behavior details that are subject to strict privacy regulations and ethical considerations.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iDs

Phyu Mon Latt <http://orcid.org/0000-0001-5880-5731>

Eric P F Chow <http://orcid.org/0000-0003-1766-0657>

Jason J Ong <http://orcid.org/0000-0001-5784-7403>

Lei Zhang <http://orcid.org/0000-0003-2343-084X>

#### REFERENCES

- World Health Organization. Sexually transmitted infections (STIs) fact sheets Geneva 22 November 2021. Available: [https://www.who.int/news-room/fact-sheets/detail/sexually-transmitted-infections-\(stis/](https://www.who.int/news-room/fact-sheets/detail/sexually-transmitted-infections-(stis/) [Accessed 06 Jun 2022].
- Traeger MW, Cornelisse VJ, Asselin J, et al. Association of HIV preexposure prophylaxis with incidence of sexually transmitted infections among individuals at high risk of HIV infection. *JAMA* 2019;321:1380–90.
- Traeger MW, Guy R, Asselin J, et al. Real-world trends in incidence of bacterial sexually transmissible infections among gay and bisexual men using HIV pre-exposure prophylaxis (PREP) in Australia following nationwide prep implementation: an analysis of sentinel surveillance data. *Lancet Infect Dis* 2022;22:1231–41.
- McManus H, Grulich AE, Amin J, et al. Comparison of trends in rates of sexually transmitted infections before vs after initiation of HIV preexposure prophylaxis among men who have sex with men. *JAMA Netw Open* 2020;3:e2030806.
- The Department of Health and Aged Care. In: *National Notifiable Diseases Surveillance System (NNDSS)*. Available: <https://www.health.gov.au/our-work/nndss#diseases-on-the-national-notifiable-disease-list/> [accessed 06 Jun 2022].
- World Health Organization. Global health sector strategies on, respectively, HIV, viral hepatitis and sexually transmitted infections for the period 2022–2030. Geneva World Health Organization; 2022.
- Australian Government DoH. *Fourth national sexually transmissible infections strategy*. Canberra: Department of Health, 2018.
- Nguyen LH, Tran BX, Rocha LEC, et al. A systematic review of eHealth interventions addressing HIV/STI prevention among men who have sex with men. *AIDS Behav* 2019;23:2253–72.
- Kumar R, Mehendale SM, Panda S, et al. Impact of targeted interventions on heterosexual transmission of HIV in India. *BMC Public Health* 2011;11:549.
- Clift S, Anemona A, Watson-Jones D, et al. Variations of HIV and STI prevalences within communities neighbouring new Goldmines in Tanzania: importance for intervention design. *Sex Transm Infect* 2003;79:307–12.
- Chesson HW, Sternberg M, Leichter JS, et al. Changes in the state-level distribution of primary and secondary syphilis in the USA, 1985–2007. *Sex Transm Infect* 2010;86 Suppl 3:iii58–62.
- Kerani RP, Handcock MS, Handsfield HH, et al. Comparative geographic concentrations of 4 sexually transmitted infections. *Am J Public Health* 2005;95:324–30.
- Monteiro EF, Lacey CJN, Merrick D. The interrelation of demographic and geospatial risk factors between four common sexually transmitted diseases. *Sex Transm Infect* 2005;81:41–6.
- Elliott LJ, Blanchard JF, Beaudoin CM, et al. Geographical variations in the epidemiology of bacterial sexually transmitted infections in Manitoba, Canada. *Sex Transm Infect* 2002;78 Suppl 1:i139–44.
- Gsteiger S, Low N, Sonnenberg P, et al. Gini coefficients for measuring the distribution of sexually transmitted infections among individuals with different levels of sexual activity. *PeerJ* 2020;8:e8434.
- van Wees DA, Diexer S, Rozhnova G, et al. Quantifying heterogeneity in sexual behaviour and distribution of STIs before and after pre-exposure prophylaxis among men who have sex with men. *Sex Transm Infect* 2022;98:395–400.
- Gibney KB, Cheng AC, Hall R, et al. Sociodemographic and geographical inequalities in notifiable infectious diseases in Australia: a retrospective analysis of 21 years of national disease surveillance data. *Lancet Infect Dis* 2017;17:86–97.
- Bao Y, Medland NA, Fairley CK, et al. Predicting the diagnosis of HIV and sexually transmitted infections among men who have sex with men using machine learning approaches. *J Infect* 2021;82:48–59.
- Xu X, Ge Z, Chow EPF, et al. A machine-learning-based risk-prediction tool for HIV and sexually transmitted infections acquisition over the next 12 months. *J Clin Med* 2022;11:1818.
- Xu X, Yu Z, Ge Z, et al. Web-based risk prediction tool for an individual's risk of HIV and sexually transmitted infections using machine learning algorithms: development and external validation study. *J Med Internet Res* 2022;24:e37850.
- Melbourne sexual health centre [MySTIRisk]. Available: <https://mystirisk.mshc.org.au/> [Accessed 06 Jun 2022].
- The World Bank. Metadata glossary of Gini index. Available: <https://databank.worldbank.org/metadataglossary/gender-statistics/series/SI.POV.GINI> [Accessed 06 Jun 2022].
- Tsuboi M, Evans J, Davies EP, et al. Prevalence of syphilis among men who have sex with men: a global systematic review and meta-analysis from 2000–20. *Lancet Glob Health* 2021;9:e1110–8.
- Kirby Institute. Annual surveillance report 2021 - sexually transmissible infections. Sydney UNSW; 2021.
- Kirby Institute. *Annual surveillance report 2021 - HIV*. Sydney: UNSW, 2021.
- Mahajan AP, Sayles JN, Patel VA, et al. Stigma in the HIV/AIDS epidemic: a review of the literature and recommendations for the way forward. *AIDS* 2008;22 Suppl 2:S67–79.
- Turan B, Budhwani H, Fazeli PL, et al. How does stigma affect people living with HIV? The mediating roles of internalized and anticipated HIV stigma in the effects of perceived community stigma on health and Psychosocial outcomes. *AIDS Behav* 2017;21:283–91.
- Ziersch A, Walsh M, Baak M, et al. "It is not an acceptable disease": a qualitative study of HIV-related stigma and discrimination and impacts on health and wellbeing for people from ethnically diverse backgrounds in Australia. *BMC Public Health* 2021;21:779.
- Fairley CK, Sze JK, Vodstrcil LA, et al. Computer-assisted self interviewing in sexual health clinics. *Sex Transm Dis* 2010;37:665–8.
- Cornelisse VJ, Chow EPF, Huffam S, et al. Increased detection of pharyngeal and rectal gonorrhoea in men who have sex with men after transition from culture to nucleic acid amplification testing. *Sex Transm Dis* 2017;44:114–7.