



## OPEN ACCESS

## EDITED BY

Nilu Goonetilleke,  
University of North Carolina at Chapel  
Hill, United States

## REVIEWED BY

Haidong Lu,  
Yale University, United States  
Xiaojing Zheng,  
University of North Carolina at Chapel  
Hill, United States  
Bernice Hoenderboom,  
National Institute for Public Health and  
the Environment, Netherlands

## \*CORRESPONDENCE

Lei Zhang  
✉ lei.zhang1@monash.edu  
Jason J. Ong  
✉ jason.ong@monash.edu

†These authors have contributed  
equally to this work

## SPECIALTY SECTION

This article was submitted to  
Infectious Diseases: Epidemiology and  
Prevention,  
a section of the journal  
Frontiers in Public Health

RECEIVED 29 August 2022

ACCEPTED 23 December 2022

PUBLISHED 13 January 2023

## CITATION

Xu X, Chow EPF, Fairley CK, Chen M,  
Aguirre I, Goller J, Hocking J,  
Carvalho N, Zhang L and Ong JJ  
(2023) Determinants and prediction of  
*Chlamydia trachomatis* re-testing and  
re-infection within 1 year among  
heterosexuals with chlamydia  
attending a sexual health clinic.  
*Front. Public Health* 10:1031372.  
doi: 10.3389/fpubh.2022.1031372

## COPYRIGHT

© 2023 Xu, Chow, Fairley, Chen,  
Aguirre, Goller, Hocking, Carvalho,  
Zhang and Ong. This is an  
open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other  
forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the  
original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution  
or reproduction is permitted which  
does not comply with these terms.

# Determinants and prediction of *Chlamydia trachomatis* re-testing and re-infection within 1 year among heterosexuals with chlamydia attending a sexual health clinic

Xianglong Xu<sup>1,2,3</sup>, Eric P. F. Chow<sup>2,3,4</sup>, Christopher K. Fairley<sup>2,3</sup>,  
Marcus Chen<sup>2,3</sup>, Ivette Aguirre<sup>2</sup>, Jane Goller<sup>4</sup>, Jane Hocking<sup>4</sup>,  
Natalie Carvalho<sup>5</sup>, Lei Zhang<sup>2,3,6\*†</sup> and Jason J. Ong<sup>2,3,7\*†</sup>

<sup>1</sup>Department of Epidemiology and Health Statistics, School of Public Health, Shanghai University of Traditional Chinese Medicine, Shanghai, China, <sup>2</sup>Melbourne Sexual Health Centre, The Alfred, Melbourne, VIC, Australia, <sup>3</sup>Central Clinical School, Monash University, Melbourne, VIC, Australia, <sup>4</sup>Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, VIC, Australia, <sup>5</sup>Centre for Health Policy, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, VIC, Australia, <sup>6</sup>China Australia Joint Research Center for Infectious Diseases, School of Public Health, Xi'an Jiaotong University Health Science Centre, Xi'an, Shaanxi, China, <sup>7</sup>Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom

**Background:** *Chlamydia trachomatis* (chlamydia) is one of the most common sexually transmitted infections (STI) globally, and re-infections are common. Current Australian guidelines recommend re-testing for chlamydia 3 months after treatment to identify possible re-infection. Patient-delivered partner therapy (PDPT) has been proposed to control chlamydia re-infection among heterosexuals. We aimed to identify determinants and the prediction of chlamydia re-testing and re-infection within 1 year among heterosexuals with chlamydia to identify potential PDPT candidates.

**Methods:** Our baseline data included 5,806 heterosexuals with chlamydia aged  $\geq 18$  years and 2,070 re-tested for chlamydia within 1 year of their chlamydia diagnosis at the Melbourne Sexual Health Center from January 2, 2015, to May 15, 2020. We used routinely collected electronic health record (EHR) variables and machine-learning models to predict chlamydia re-testing and re-infection events. We also used logistic regression to investigate factors associated with chlamydia re-testing and re-infection.

**Results:** About 2,070 (36%) of 5,806 heterosexuals with chlamydia were re-tested for chlamydia within 1 year. Among those retested, 307 (15%) were re-infected. Multivariable logistic regression analysis showed that older age ( $\geq 35$  years old), female, living with HIV, being a current sex worker, patient-delivered partner therapy users, and higher numbers of sex partners were associated with an increased chlamydia re-testing within 1 year. Multivariable logistic regression analysis also showed that younger age (18–24 years), male

gender, and living with HIV were associated with an increased chlamydia re-infection within 1 year. The XGBoost model was the best model for predicting chlamydia re-testing and re-infection within 1 year among heterosexuals with chlamydia; however, machine learning approaches and these self-reported answers from clients did not provide a good predictive value (AUC < 60.0%).

**Conclusion:** The low rate of chlamydia re-testing and high rate of chlamydia re-infection among heterosexuals with chlamydia highlights the need for further interventions. Better targeting of individuals more likely to be re-infected is needed to optimize the provision of PDPT and encourage the test of re-infection at 3 months.

#### KEYWORDS

*Chlamydia trachomatis*, re-testing, re-infection, heterosexual, predictive model, machine learning, variable importance, risk factors

## Background

*Chlamydia trachomatis* (chlamydia) is the most common bacterial sexually transmitted infection (STI) globally (1, 2). An estimated 131 million incident cases of chlamydia infections are acquired globally among people aged 15–49 years annually (3). After treatment, chlamydia re-infections are common, occurring in about 10–20% of patients within 12 months (4–8). Identifying and treating chlamydia re-infection promptly among women is vital because re-infection is associated with a higher risk of pelvic inflammatory disease (9). A study in Australia reported that the chlamydia re-infection rate was 22.3 (95%CI: 13.2–37.6) per 100 person-years among women aged 16–25 years (8). Australian guidelines recommend chlamydia re-testing at 3 months after treatment to identify possible re-infection (10, 11). Chlamydia re-testing was relatively low (40%) within 1.5–12 months (12). Therefore, developing innovative measures that increase repeat testing for chlamydia following chlamydia infection is important.

Treatment of partners is necessary for heterosexuals with chlamydia to prevent re-infection and interrupt the chain of chlamydia transmission (13). Patient Delivery Partner Therapy (PDPT) has been proposed as a possible strategy for reducing re-infection and preventing the sequelae of STIs (14). PDPT involves providing antibiotic treatment to the sexual partners of patients diagnosed with a treatable sexually transmitted infection by giving the prescription or medication to the patient's sexual partner(s) *via* the diagnosed patient (15). The purpose of PDPT is to enable sexual partners who may be reluctant to attend a health provider to be treated earlier, thus reducing the risk of re-infecting the diagnosed patient. Two observational studies found that PDPT may lower women's re-infection risk (16, 17). In Australia, PDPT is permissible for heterosexual men and women diagnosed with chlamydia (18). Identifying

heterosexuals with chlamydia at high risk of re-infection who could benefit from PDPT is important for implementing PDPT.

Predictive models can assist clinical decision-making (19). Developing a predictive model for predicting chlamydia re-testing and re-infection within 1 year among heterosexuals could be used to identify potential PDPT candidates and prioritize chlamydia re-infection screening and sexual health service planning. Machine learning approaches have some advantages in prediction, including not requiring statistical inferences and assumptions (20), improving accuracy by exploiting complex interactions between risk factors (21), handling a mass of predictors and combining them in a non-linear and highly interactive way (22). Despite these advantages, to our knowledge, only one study used machine learning to predict STI re-infection. Our study used machine learning algorithms to predict the risk of acquiring  $\geq 1$  or  $\geq 2$  additional STIs (combinations of gonorrhea, chlamydia, and syphilis) diagnoses within 1 or 2 years after the initial diagnosis among patients (not limited to heterosexuals) in Massachusetts, USA (23). No studies have used machine learning to predict chlamydia re-testing or re-infection among heterosexuals. Our research team has also used machine learning approaches to predict the uptake of HIV/STI testing among men who have sex with men receiving reminders for testing (24).

We aimed to predict chlamydia re-testing and re-infection within 1 year among heterosexuals with chlamydia using machine learning algorithms. We also used univariable and multivariable logistic regression to identify variables associated with chlamydia re-testing and re-infection.

## Methods

Our analysis used electronic health records (EHRs) data between January 2, 2015, and May 15, 2020, at the Melbourne

Sexual Health Center (MSHC). We included individuals who were (1) heterosexuals with chlamydia (including heterosexual males and females) and (2) aged 18 years and above.

## Predictors

Based on the literature review and expert discussion, we selected potential predictors. Potential predictors were selected from data routinely collected through a computer-assisted self-interviewing system used when patients enter the clinic (25) and from other routinely collected data items in the clinical setting. Self-reported predictors at baseline (i.e., at the time of their diagnosis of chlamydia) were included, such as demographics (age at the consultation, gender, country of birth, access to Medicare, ever sex worker, current sex worker), sexual practices (numbers of sex partners in the last 3 months, sex overseas in the last 12 months, condoms used with sex partners in the last 3 months), self-reported past STI (i.e., chlamydia, genital herpes, genital warts, gonorrhea, syphilis), HIV infection, use of PDPT (i.e., the diagnosed client receiving an extra prescription of antibiotics to give to their sexual partner), injecting drug use in the last 3 months, triage reason as asymptomatic screening, triage reason as STI symptoms, and triage reason as contact of any STI (summarized in Table 1). In the machine learning analysis, we did not include PDPT users.

## Outcome

The primary outcome was chlamydia re-infection, defined as the first new chlamydia diagnosis using nucleic acid amplification testing from any anatomical site, including the oropharynx, urethra/urine, or anorectum, at least 30 days after and within 365 days after a positive chlamydia diagnosis. The selection of chlamydia re-infection from 30 to 365 days after a positive chlamydia diagnosis is consistent with a *Chlamydia trachomatis* re-infection study among female adolescents (26). In this study, chlamydia re-testing was defined as those heterosexuals who had new chlamydia diagnosis results (either positive or negative) at least 30 days after and within 365 days after a positive chlamydia diagnosis.

## Analysis

Statistical analysis was conducted with R 4.0.3. Descriptive statistics were used to summarize the patient characteristics. We also performed risk factor analysis using univariable and multivariable logistic regression analyses to calculate the unadjusted odds ratio (OR) and adjusted

OR. Variables with  $p < 0.20$  in the univariable logistic regression analysis were included in the multivariable logistic regression analysis to identify independent risk factors. All statistics were performed using a two-sided test, and statistical significance was considered at  $p < 0.05$ .

Machine learning algorithms were conducted with Python 3.9.7. We developed 10 commonly utilized machine learning models, including Logistic Regression (LR), K-Nearest Neighbors (KNN), Adaptive Boosting Classifier (AdaBoost), SVM with a Radial Basis Function Kernel (SVM), Gaussian Naive Bayes (GaussianNB), Gradient Boosting Machine (GBM), Light Gradient Boosting Decision Machine (LightGBM), Extreme Gradient Boosting (XGBoost), Random Forest (RF), and Multi-Layer Perceptron (MLP). LR, GBM, AdaBoost, GaussianNB, KNN, SVM, RF, and MLP was built using the *scikit-learn* library. LightGBM was built using the *LightGBM* library. XGBoost was built using the *xgboost* library. We performed a  $3 \times 10$  (3 outer folds, 10 inner folds) nested cross-validation to avoid overfitting and improve generalizability (25, 27). We used the Bayesian optimisation method for tuning the hyperparameters. The performance of machine learning models was measured by the area under the receiver operating characteristics curve (AUC) and accuracy on the independent testing dataset. We also calculated the variable importance of chlamydia re-testing and re-infection within 1 year to investigate the effect of different predictors on prediction (24, 25, 28). First, we obtained the results of variable importance analysis from each fold in the outer loop of nested cross-validation. Then based on the value of AUC, we chose the one closest to the mean AUC and reported the results of this variable importance analysis.

## Ethics approval

Ethical approval was granted by the Alfred Hospital Ethics Committee, Australia (project number: 277/20).

## Results

### Characteristics of the study population

Our study data included 5,806 heterosexual patients diagnosed with chlamydia. Among 5,806 patients, 35.7% (2,070/5,806) were re-tested for chlamydia within 1 year. A total of 14.8% (307/2,070) heterosexuals with chlamydia were re-infected with chlamydia within 1 year. Further details of the sociodemographic characteristics and sexual practices are provided in Tables 1, 2.

TABLE 1 Factors associated with chlamydia re-testing within 1 year among 5,806 heterosexual individuals with chlamydia.

Characteristic	Event rate	Crude OR	95% CI	p-value	Adjusted OR*	95% CI	p-value
<b>Gender</b>							
Female	1,089/2,752 (39.6%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
Male	981/3,054 (32.1%)	0.72	0.65, 0.80	<b>&lt;0.001</b>	0.74	0.66, 0.84	<b>&lt;0.001</b>
<b>HIV infection</b>							
HIV-negative	2,042/5,762 (35.4%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
Living with HIV	28/44 (63.6%)	3.19	1.74, 6.04	<b>&lt;0.001</b>	2.27	1.18, 4.50	<b>0.016</b>
<b>Triage reason as asymptomatic screening</b>							
No	1,910/5,404 (35.3%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
Yes	160/402 (39.8%)	1.21	0.98, 1.49	<b>0.072</b>	1.04	0.83, 1.29	0.749
<b>Triage reason as STI symptoms</b>							
No	1,139/3,235 (35.2%)	1 (ref)	1 (ref)				
Yes	931/2,571 (36.2%)	1.04	0.94, 1.16	0.428			
<b>Triage reason as contact of STI</b>							
No	1,703/4,547 (37.5%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
Yes	367/1,259 (29.2%)	0.69	0.60, 0.79	<b>&lt;0.001</b>	0.79	0.69, 0.91	<b>0.001</b>
<b>Past chlamydia</b>							
No	1,663/4,678 (35.5%)	1 (ref)	1 (ref)				
Yes	407/1,128 (36.1%)	1.02	0.89, 1.17	0.738			
<b>Past genital herpes</b>							
No	2,011/5,655 (35.6%)	1 (ref)	1 (ref)				
Yes	59/151 (39.1%)	1.16	0.83, 1.61	0.374			
<b>Past genital warts</b>							
No	1,991/5,610 (35.5%)	1 (ref)	1 (ref)				
Yes	79/196 (40.3%)	1.23	0.92, 1.64	<b>0.167</b>			
<b>Past gonorrhea</b>							
No	2,015/5,661 (35.6%)	1 (ref)	1 (ref)				
Yes	55/145 (37.9%)	1.11	0.78, 1.55	0.562			

(Continued)

TABLE 1 (Continued)

Characteristic	Event rate	Crude OR	95% CI	<i>p</i> -value	Adjusted OR*	95% CI	<i>p</i> -value
<b>Past syphilis</b>							
No	2,061/5,779 (35.7%)	1 (ref)	1 (ref)				
Yes	9/27 (33.3%)	0.90	0.39, 1.96	0.801			
<b>Current sex worker</b>							
No	1,773/5,255 (33.7%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
Unknown/missing	123/280 (43.9%)	1.54	1.21, 1.96	<0.001	1.35	0.81, 2.28	0.250
Yes	174/271 (64.2%)	3.52	2.74, 4.56	<0.001	2.50	1.88, 3.33	<0.001
<b>Sex overseas in the past 12 months</b>							
No	942/2,517 (37.4%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
Unknown/missing	188/450 (41.8%)	1.20	0.98, 1.47	0.080	1.06	0.78, 1.43	0.703
Yes	940/2,839 (33.1%)	0.83	0.74, 0.93	<0.001	0.90	0.79, 1.01	0.077
<b>Use of PDPT</b>							
No	1,947/5,540 (35.1%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
Yes	123/266 (46.2%)	1.59	1.24, 2.03	<0.001	1.32	1.02, 1.70	0.033
<b>Country of birth</b>							
Australia	587/1,610 (36.5%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
Not Australia	1,414/3,987 (35.5%)	0.96	0.85, 1.08	0.482	0.92	0.80, 1.06	0.263
Unknown/missing	69/209 (33.0%)	0.86	0.63, 1.16	0.329	0.79	0.57, 1.09	0.152
<b>Access to Medicare</b>							
No	786/1,995 (39.4%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
Unknown/missing	155/780 (19.9%)	0.38	0.31, 0.46	<0.001	0.41	0.34, 0.50	<0.001
Yes	1,129/3,031 (37.2%)	0.91	0.81, 1.03	0.125	0.91	0.79, 1.04	0.182
<b>Age</b>							
18–24 years	783/2,331 (33.6%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
25–34 years	1,044/2,895 (36.1%)	1.12	0.99, 1.25	0.063	1.13	1.00, 1.27	0.052
35 years old and above	243/580 (41.9%)	1.43	1.18, 1.72	<0.001	1.27	1.04, 1.55	0.019

(Continued)

TABLE 1 (Continued)

Characteristic	Event rate	Crude OR	95% CI	<i>p</i> -value	Adjusted OR*	95% CI	<i>p</i> -value
<b>Injecting drug use in the past 3 months</b>							
No	1,926/5,440 (35.4%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
Unknown/missing	132/320 (41.2%)	1.28	1.02, 1.61	<b>0.034</b>	0.83	0.52, 1.29	0.404
Yes	12/46 (26.1%)	0.64	0.32, 1.21	<b>0.191</b>	0.54	0.26, 1.06	0.085
<b>Numbers of sex partners in the past 3 months</b>							
1 sex partner	497/1,446 (34.4%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
2–5 sex partners	983/2,816 (34.9%)	1.02	0.90, 1.17	0.727	1.13	0.98, 1.30	0.083
5–10 sex partners	265/768 (34.5%)	1.01	0.84, 1.21	0.949	1.20	0.99, 1.46	0.064
More than 10 sex partners	76/196 (38.8%)	1.21	0.89, 1.64	0.225	1.46	1.06, 2.01	<b>0.020</b>
Unknown/missing	41/137 (29.9%)	0.82	0.55, 1.19	0.295	0.89	0.57, 1.39	0.626
N/A	443						
<b>Condoms used with sex partners in the last 3 months</b>							
Always use condom	168/395 (42.5%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
Not Always	1643/4,793 (34.3%)	0.70	0.57, 0.87	<b>&lt;0.001</b>	0.80	0.65, 1.00	0.052
Unknown/missing	1643/4,793 (34.3%)	0.56	0.38, 0.81	<b>&lt;0.003</b>	0.65	0.42, 1.00	0.052
N/A	208/443 (47.0%)						

OR, Odds Ratio; CI, Confidence Interval; N/A, no sex in the past 3 months; PDPT, patient delivery partner therapy. \* Variables with  $p < 0.20$  in the univariable logistic regression analysis were included in the multivariable logistic regression analysis. The bold values indicate variables with  $p < 0.20$  in the univariable logistic regression analysis.

TABLE 2 Factors associated with chlamydia re-infection within 1 year among 2,070 re-tested heterosexual individuals.

Characteristic	Event rate	Crude OR	95% CI	p-value	Adjusted OR*	95% CI	p-value
<b>Gender</b>							
Female	132/1,089 (12.1%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
Male	175/981 (17.8%)	1.57	1.23, 2.01	<b>&lt;0.001</b>	1.55	1.16, 2.08	<b>0.003</b>
<b>HIV infection</b>							
HIV-negative	294/2,042 (14.4%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
Living with HIV	13/28 (46.4%)	5.15	2.39, 11.0	<b>&lt;0.001</b>	4.02	1.64, 9.87	<b>0.002</b>
<b>Triage reason as asymptomatic screening</b>							
No	289/1,910 (15.1%)	1 (ref)	1 (ref)				
Yes	18/160 (11.2%)	0.71	0.42, 1.15	<b>0.187</b>			
<b>Triage reason as STI symptoms</b>							
No	162/1,139 (14.2%)	1 (ref)	1 (ref)				
Yes	145/931 (15.6%)	1.11	0.87, 1.42	0.390			
<b>Triage reason as contact of STI</b>							
No	251/1,703 (14.7%)	1 (ref)	1 (ref)				
Yes	56/367 (15.3%)	1.04	0.75, 1.42	0.799			
<b>Past chlamydia</b>							
No	247/1,663 (14.9%)	1 (ref)	1 (ref)				
Yes	60/407 (14.7%)	0.99	0.73, 1.34	0.955			
<b>Past genital herpes</b>							
No	303/2,011 (15.1%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
Yes	4/59 (6.78%)	0.41	0.12, 1.01	<b>0.087</b>	0.40	0.12, 1.00	0.085
<b>Past genital Warts</b>							
No	293/1,991 (14.7%)	1 (ref)	1 (ref)				
Yes	14/79 (17.7%)	1.25	0.66, 2.19	0.462			
<b>Past gonorrhea</b>							
No	295/2,015 (14.6%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
Yes	12/55 (21.8%)	1.63	0.81, 3.03	<b>0.143</b>	1.65	0.80, 3.18	0.150

(Continued)

TABLE 2 (Continued)

Characteristic	Event rate	Crude OR	95% CI	<i>p</i> -value	Adjusted OR*	95% CI	<i>p</i> -value
<b>Past syphilis</b>							
No	306/2,061 (14.8%)	1 (ref)	1 (ref)				
Yes	1/9 (11.1%)	0.72	0.04, 3.93	0.754			
<b>Current sex worker</b>							
No	249/1,773 (14.0%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
Unknown/missing	30/123 (24.4%)	1.97	1.26, 3.01	<b>0.002</b>	0.55	0.17, 1.81	0.334
Yes	28/174 (16.1%)	1.17	0.75, 1.77	0.461	1.22	0.71, 2.04	0.456
<b>Sex overseas in the past 12 months</b>							
No	137/942 (14.5%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
Unknown/missing	37/188 (19.7%)	1.44	0.95, 2.13	<b>0.076</b>	0.61	0.28, 1.24	0.199
Yes	133/940 (14.1%)	0.97	0.75, 1.25	0.807	0.91	0.69, 1.20	0.501
<b>Use of PDPT</b>							
No	285/1,947 (14.6%)	1 (ref)	1 (ref)				
Yes	22/123 (17.9%)	1.27	0.77, 2.01	0.327			
<b>Country of birth</b>							
Australia	97/587 (16.5%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
Not Australia	199/1,414 (14.1%)	0.83	0.64, 1.08	<b>0.160</b>	0.87	0.63, 1.20	0.387
Unknown/missing	11/69 (15.9%)	0.96	0.46, 1.82	0.902	1.11	0.52, 2.18	0.772
<b>Access to Medicare</b>							
No	116/786 (14.8%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
Unknown/missing	16/155 (10.3%)	0.66	0.37, 1.13	<b>0.148</b>	0.61	0.34, 1.05	0.088
Yes	175/1,129 (15.5%)	1.06	0.82, 1.37	0.656	0.91	0.66, 1.23	0.533
<b>Age</b>							
18–24 years	125/783 (16.0%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
25–34 years	148/1,044 (14.2%)	0.87	0.67, 1.13	0.289	0.78	0.60, 1.02	0.074
35 years old and above	34/243 (14.0%)	0.86	0.56, 1.28	0.458	0.58	0.37, 0.89	<b>0.016</b>

(Continued)



TABLE 2 (Continued)

Characteristic	Event rate	Crude OR	95% CI	<i>p</i> -value	Adjusted OR*	95% CI	<i>p</i> -value
<b>Injecting drug use in the past 3 months</b>							
No	272/1,926 (14.1%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
Unknown/missing	34/132 (25.8%)	2.11	1.38, 3.15	<b>&lt;0.001</b>	2.11	0.73, 5.55	0.144
Yes	1/12 (8.33%)	0.55	0.03, 2.86	0.571	0.30	0.01, 1.95	0.306
<b>Numbers of sex partners in the past 3 months</b>							
1 sex partner	64/497 (12.9%)	1 (ref)	1 (ref)		1 (ref)	1 (ref)	
2–5 sex partners	129/983 (13.1%)	1.02	0.74, 1.42	0.894	1.02	0.73, 1.42	0.927
5–10 sex partners	47/265 (17.7%)	1.46	0.96, 2.19	<b>0.071</b>	1.38	0.90, 2.11	0.140
More than 10 sex partners	15/76 (19.7%)	1.66	0.87, 3.04	<b>0.109</b>	1.45	0.74, 2.71	0.261
Unknown/missing	3/41 (7.32%)	0.53	0.13, 1.53	0.307	0.51	0.12, 1.51	0.287
N/A	208						
<b>Condoms used with sex partners in the last 3 months</b>							
Always use condom	19/168 (11.3%)	1 (ref)	1 (ref)				
Not Always	230/1,643 (14.0%)	1.28	0.80, 2.16	0.336			
Unknown/missing	9/51 (17.6%)	1.68	0.68, 3.90	0.239			
N/A	208						

OR, Odds Ratio; CI, Confidence Interval; N/A, no sex in the past 3 months; PDPT, patient delivery partner therapy. \*Variables with  $p < 0.20$  in the univariable logistic regression analysis were included in the multivariable logistic regression analysis. The bold values indicate variables with  $p < 0.20$  in the univariable logistic regression analysis.

## Factors associated with chlamydia re-testing and chlamydia re-infection within 1 year using univariable and multivariable logistic regression analyses

The potential factors that were included in the multivariable logistic regression analysis of chlamydia re-testing within 1 year among heterosexuals with chlamydia were age, gender, HIV infection, triage reason as asymptomatic screening, triage reason as contact of STI infection, current sex work, use of PDPT, country of birth, access to Medicare, injecting drug use in the past 3 months, had sex with a partner in the past 3 months, numbers of sex partners in the past 3 months, condoms used with sex partners in the last 3 months. Chlamydia re-testing within 1 year among heterosexuals was associated with male gender [adjusted odds ratios (aOR) = 0.74, 95%CI 0.66, 0.84], living with HIV (aOR = 2.27, 95%CI 1.18, 4.50), presentation as contact of STI infection (aOR = 0.79, 95%CI 0.69, 0.91), being a current sex worker (aOR = 2.50, 95%CI 1.88, 3.33), users of patient delivered partner therapy (aOR = 1.32, 95%CI 1.02, 1.70), older age ( $\geq 35$  years old vs. 18–24 years, aOR = 1.27, 95%CI 1.04, 1.55), unknown/missing on accessing to Medicare (aOR = 0.41, 95%CI 0.34, 0.50), and higher numbers of partners in the last 3 months ( $>10$  vs. 1, aOR = 1.46, 95%CI 1.06, 2.01) (Table 1).

The potential factors that were included in the multivariable logistic regression analysis of chlamydia re-infection within 1 year among heterosexuals with chlamydia were age, gender, HIV infection, past genital herpes, past gonorrhea, current sex worker, sex overseas in the past 12 months, country of birth, access to Medicare, injecting drug use in the past 3 months, had sex with a partner in the past 3 months, numbers of sex partners in the past 3 months, and condoms used with sex partners in the past 3 months. Chlamydia re-infection within 1 year among heterosexuals was associated with male gender (aOR = 1.55, 95%CI 1.16, 2.08), living with HIV (aOR = 4.02, 95%CI 1.64, 9.87), and older age ( $\geq 35$  years old vs. 18–24 years, aOR = 0.58, 95%CI 0.37, 0.89) (Table 2).

## Prediction of chlamydia re-testing and chlamydia re-infection within 1 year using machine learning approaches

The best model for predicting chlamydia re-testing within 1 year among heterosexuals with chlamydia was the XGBoost [mean 58.5% (SD 0.1%)]. Adaboost, SVM, LightGBM, and XGBoost performed better than LR in predicting chlamydia re-testing within 1 year among heterosexuals with chlamydia. However, machine learning approaches did not provide a good predictive value for chlamydia re-testing on the testing data (AUC < 60.0%). Further details of model performance

are provided in Table 3 and Supplementary Table 1. Variable importance analysis using XGBoost showed that the top 10 identified predictors for chlamydia re-testing within 1 year among heterosexuals with chlamydia included current sex worker, access to Medicare, condoms used with sex partners, age, gender, triage reason as contact of STI infection, injecting drug use, sex overseas, numbers of sex partners, and triage reason as STI symptoms (Figure 1).

The best model for predicting chlamydia re-infection within 1 year among heterosexuals with chlamydia was the XGBoost [mean 58.5% (SD 1.1%)]. Adaboost and XGBoost performed better than LR in predicting chlamydia re-infection within 1 year among heterosexuals with chlamydia. However, machine learning approaches did not provide a good predictive value for chlamydia re-infection on the testing data (all AUC < 60.0%). Further details of model performance are provided in Table 3 and Supplementary Table 1. Variable importance analysis using XGBoost showed that the top 10 identified predictors for chlamydia re-infection within 1 year among heterosexuals with chlamydia included gender, number of sex partners, age, past gonorrhea, past chlamydia, injecting drug use, past genital warts, condoms used with sex partners, sex overseas, and country of birth (Figure 2).

## Discussion

To our knowledge, this is the first study using machine learning approaches to predict chlamydia re-testing and re-infection within 1 year among heterosexuals with chlamydia. Our study found a relatively low chlamydia re-testing rate (36%) within 1 year among heterosexuals with chlamydia. Our finding was similar to a previous study reporting that about 40% were re-tested within 12 months in 25 general practice clinics in the Australian Collaboration for Chlamydia Enhanced Sentinel Surveillance system (12). Our study also found a relatively high rate of chlamydia re-infection (15%) within 1 year among heterosexuals with chlamydia. Our result is consistent with previous studies reporting that the chlamydia re-infection rate was about 10–20% of patients within 12 months (4–8). The low rate of chlamydia re-testing and high rate of chlamydia re-infection suggest that further targeted and personalized interventions are warranted among heterosexuals with chlamydia. Our logistic regression analysis showed the determinants of chlamydia re-testing and re-infection within 1 year among heterosexuals with chlamydia. Our machine learning predictive models proposed a new way of using risk prediction models to increase the low chlamydia re-testing rate and reduce the high rate of chlamydia re-infection. Despite the relatively poor prediction of chlamydia re-testing and re-infection among heterosexuals, our findings suggest that researchers could develop a wide range of machine learning algorithms to find the best algorithm for prediction. Our study

**TABLE 3** Machine learning model evaluation of chlamydia re-testing and re-infection within 1 year among heterosexual individuals with chlamydia on the testing data set (mean/SD).

	Chlamydia re-testing ( <i>n</i> = 5,806)		Chlamydia re-infection ( <i>n</i> = 2,070)	
	AUC, %	Accuracy, %	AUC, %	Accuracy, %
LR	57.6 (1.3)	65.3 (4.1)	57.3 (1.5)	85.2 (1.6)
Adaboost	57.8 (0.4)	65.4 (3.9)	58.3 (1.7)	85.0 (1.7)
SVM	58.1 (2.4)	65.3 (4.4)	47.8 (0.9)	85.2 (1.6)
XGBoost	58.5 (0.1)	64.9 (3.1)	58.5 (1.1)	85.2 (1.6)
RF	57.0 (1.1)	64.1 (2.6)	56.5 (2.0)	85.2 (1.6)
KNN	53.7 (0.1)	63.0 (3.2)	49.5 (2.5)	84.6 (2.3)
Gaussian NB	56.2 (0.6)	61.1 (2.1)	56.7 (1.6)	59.0 (29.7)
GBM	56.9 (1.5)	64.9 (4.4)	56.6 (3.3)	79.2 (9.9)
LightGBM	58.5 (0.2)	63.3 (2.5)	55.3 (1.7)	85.2 (1.6)
MLP	54.7 (3.8)	64.2 (4.7)	55.8 (1.4)	85.3 (1.3)

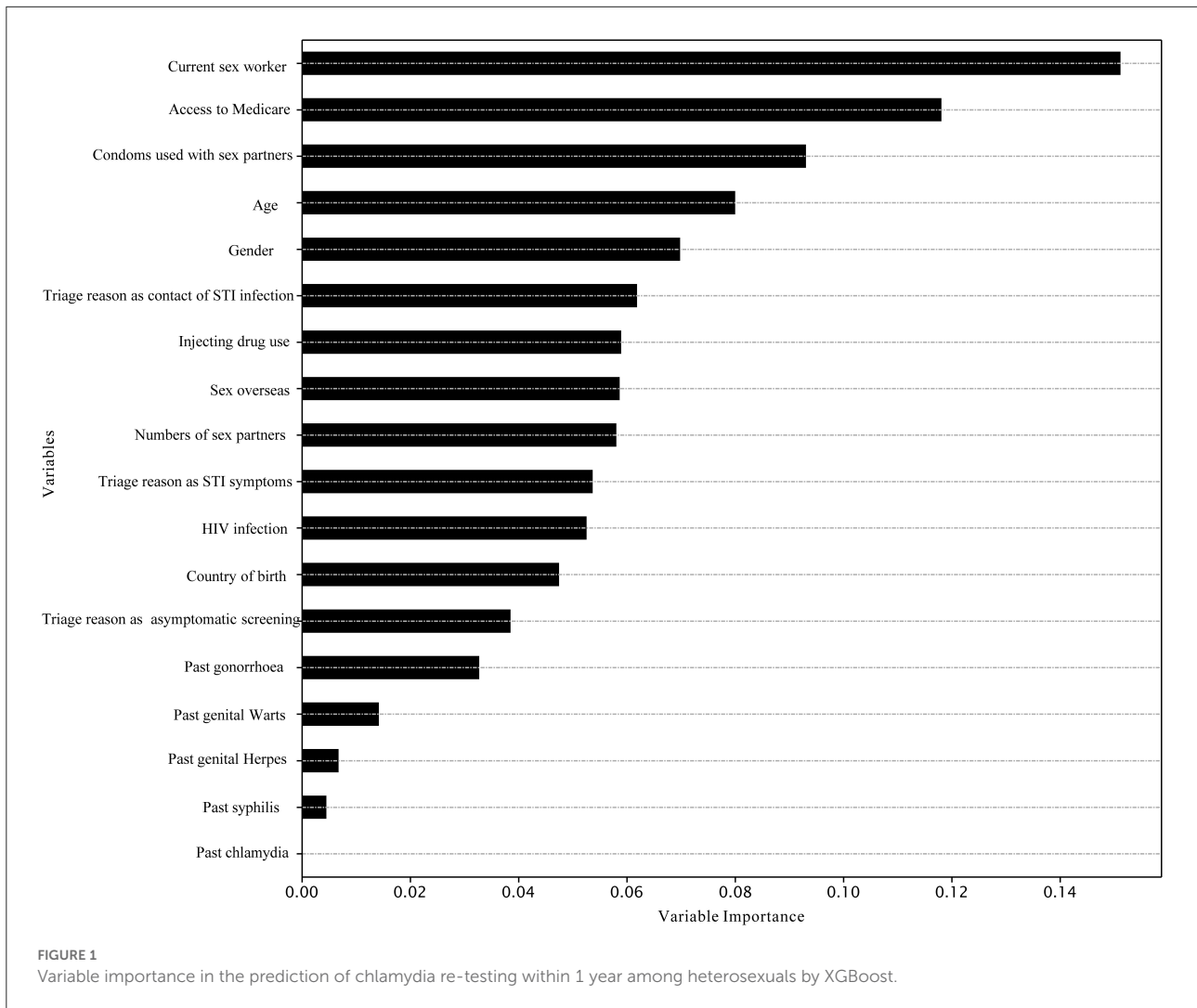
SD, standard deviation; AUC, area under the receiver operating characteristics curve; LR, Logistic Regression; KNN, K-Nearest Neighbor; AdaBoost, AdaBoost classifier; SVM, SVM with a Radial Basis Function Kernel; GaussianNB, Gaussian Naive Bayes; GBM, Gradient Boosting Machine; LightGBM, Light Gradient Boosting Machine, XGBoost, Extreme Gradient Boosting; RF, Random Forest; MLP, and multi-layer perceptron.

also provides implications for further data collection in machine learning analysis. In this study, we used routinely collected data from the clinical setting. To improve the performance of learning models in predicting chlamydia re-testing and re-infection, researchers could pay more attention to the quality, comprehensiveness, and accuracy of data collection to avoid “garbage in, garbage out” (29). Our study provides implications for selecting machine learning algorithms and predictors for future machine learning studies in predicting chlamydia re-testing and re-infection among heterosexuals with chlamydia. Better targeting interventions for those at increased risk of re-infection may help reduce their risk or encourage re-testing to diagnose and treat their infection earlier.

We aimed to develop machine learning models to predict chlamydia re-testing behavior within 1 year among heterosexuals with chlamydia using routine questions in the clinical setting. There are already some studies that have utilized machine learning algorithms to predict sexual-health behaviors, including sexual recidivism (30), sexual desire (31), reasons for not using a condom (32), and HIV risk behaviors (33). However, limited studies use machine learning algorithms to predict sexual health service utilization behaviors. We noted a previous study that reported that machine learning approaches could predict HIV/STI re-testing after receiving clinic reminder messages among MSM (24). Another study used machine learning algorithms to predict HIV testing behavior among participants from substance use disorder treatment programs (34). Our study is the first to use machine learning approaches to predict chlamydia re-testing. Therefore, it is difficult to compare our results to others. Despite the poor prediction of chlamydia re-testing behavior, developing good predictive models for this is still necessary and important. Predicting

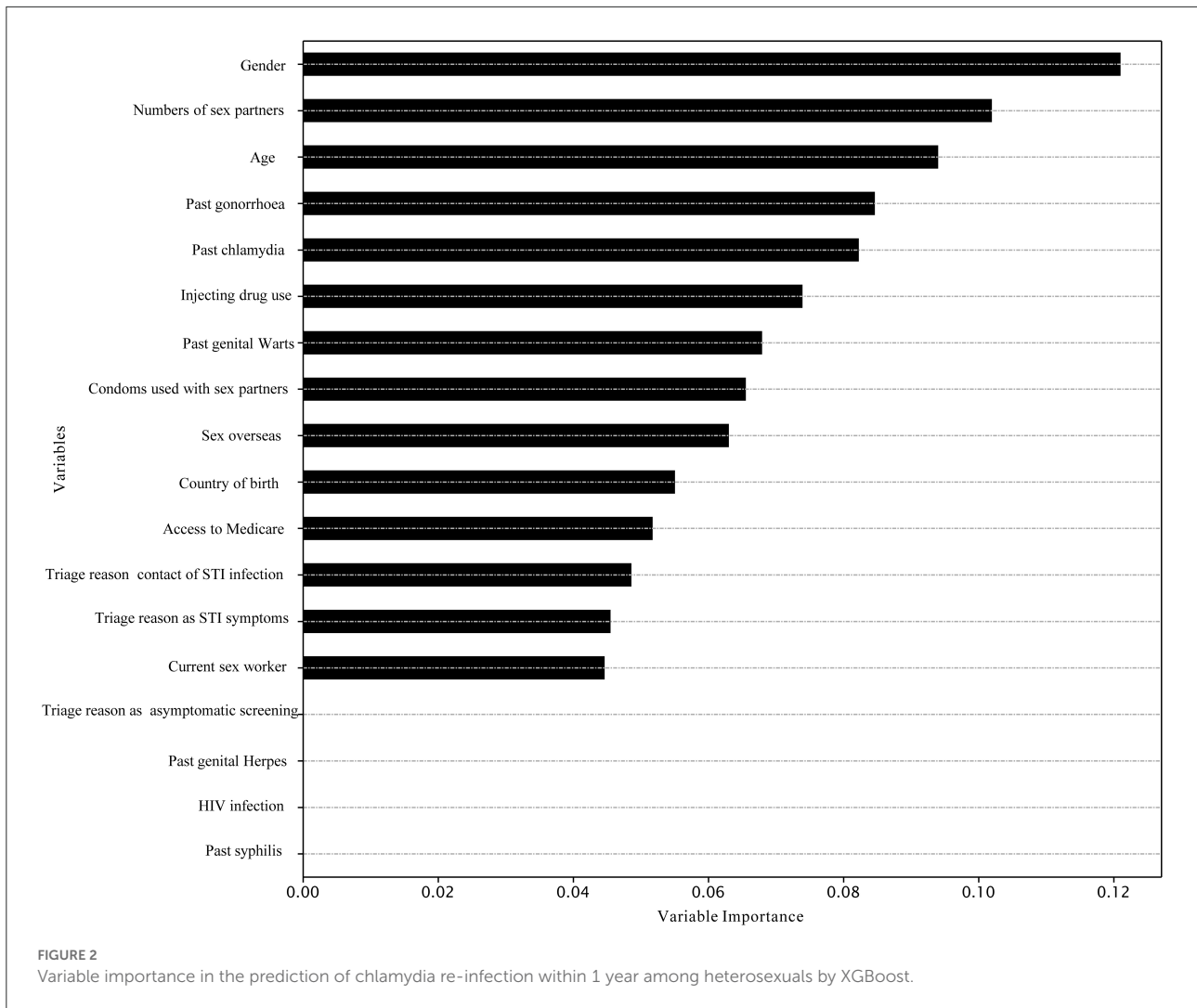
sexual healthcare utilization behaviors (e.g., future chlamydia re-testing) may benefit the management and planning of sexual health services for high-risk populations or focus prevention interventions in advance. We hope our work could encourage more machine learning research to predict chlamydia re-testing behavior among heterosexuals with chlamydia.

We intended to develop predictive models for chlamydia re-infection within 1 year among heterosexuals using machine learning approaches. Predicting future chlamydia re-infection could inform prioritizing interventions. A previous study used machine learning algorithms that could predict chlamydia (AUC = 0.67) acquisition within 1 year among males and females (28). However, our machine learning predictive models indicated that existing self-reported and routinely collected EHRs data could not provide a high predictive value for chlamydia re-infection events within 1 year among heterosexuals with chlamydia (AUC < 0.6). Despite this, our study still provides implications for future machine learning studies focusing on predicting chlamydia re-infection. For example, our results showed that Adaboost, and XGBoost performed better than conventional logistic regression in predicting chlamydia re-infection within 1 year among heterosexuals. Besides, the low predictive performances suggest that existing self-reported and routinely collected EHR data may not include some important predictors for chlamydia re-infection, such as sexual networks and background chlamydia prevalence. Another machine learning study indicated that their models lacked data for specific sexual practices (23). Our models did not include other potential factors because these data were not routinely collected EHRs data in the clinic, such as employment status (8), cervical infection, contact with uncircumcised partners and resuming sex with an untreated partner (35, 36).



Better targeting of interventions to improve re-testing is needed, especially for those with an increased likelihood of chlamydia re-infection. Our study findings are consistent with previous studies, which also found chlamydia re-testing was associated with higher numbers of sex partners (37) and among heterosexual females (38). We also observed that sex workers showed a higher re-testing rate, probably because they had a legal requirement for 3 monthly testing. Our variable importance analyses suggest other possible predictors for future predictive models of chlamydia re-testing using machine learning approaches. The top identified predictors for chlamydia re-testing were being a current sex worker, access to Medicare, condoms used with sex partners, age, gender, triage reason as contact of STI infection, injecting drug use, sex overseas, number of sex partners, and presentation as symptomatic. We hope our preliminary work encourages more machine learning research to explore the effect of introducing additional predictors on predicting chlamydia re-testing among heterosexuals.

Our study found that the chlamydia re-infection rate was greater among individuals of younger age (18–24 years), male gender, and living with HIV. These factors were consistent with previous studies on risk factors for chlamydia infection. A previous study showed that people living with HIV had more repeat chlamydia infections in the Netherlands (39). Young age at first infection was associated with an increased risk of subsequent chlamydia infection within 1 year among female adolescents in the US (26). A previous study found that being a sex worker was associated with chlamydia (40). These findings may provide clinical and public health researchers and policymakers with a deeper understanding of the drivers of chlamydia re-infection among heterosexuals at the population level. Besides, our variable importance analyses suggest possible important predictors for future predictive models of chlamydia re-infection using machine learning approaches. Consistent with previous studies on risk factors of chlamydia infection, we found that factors such as condom use (41), having sex



with a female or male partner (42), and sex workers (40) were among the important predictors included in the machine learning models.

Our study has some limitations. First, only about 36% of individuals diagnosed with chlamydia returned for another chlamydia test within 1 year. We did not have information on whether the remaining individuals were tested elsewhere, so this may be an underestimate and may lead to a selection bias in our study. Among those retested, about 15% were re-infected. It is possible that those at higher risk of re-infection may be more likely to be retested. Therefore, the chlamydia re-infection (15%) may be overestimated. Second, our findings were from a single sexual clinic; therefore, studies from different settings (e.g., those attending general practice) would be needed to verify our results. Third, the validity of the predictive factors depends on the accuracy of the self-reported information, subject to participants’ recall and non-response bias. There has been substantial work on the CASI system to ensure

its validity and accuracy (43). Fourth, we did not include gonorrhea and chlamydia coinfection status in this study. A previous study suggested that gonorrhea and chlamydia coinfection may increase the risk of chlamydia re-infection (36). Last, we included data up to May 2020, which might introduce selection bias due to the COVID-19 lockdown in Melbourne. The first COVID-19 lockdown started on March 30 and ended on May 12, 2020. Thus, our study included a small number of clinic consultations during the COVID-19 lockdown period. A previous study showed that the re-testing patterns and sexual practices might have changed due to COVID-19 (44).

## Conclusions

The chlamydia re-testing rate within 1 year was relatively low among heterosexuals with chlamydia; however, the

re-infection rate of chlamydia was relatively high. Our study highlights the need for innovative interventions to increase chlamydia re-testing and reduce re-infection among heterosexuals with chlamydia. Our findings elaborate on the determinants of chlamydia re-testing and re-infection within 1 year among heterosexuals with chlamydia. To our knowledge, it is the first demonstration of machine learning algorithms to identify heterosexuals with chlamydia at high risk of re-infection who could benefit from patient-delivered partner therapy. XGBoost model can improve the prediction of chlamydia re-testing and re-infection compared with traditional logistic regression. Our machine learning predictive models may provide a promising way to develop an innovative method to increase the low chlamydia re-testing rate and reduce the high rate of chlamydia re-infection.

## Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: The data is not publicly available due to privacy or ethical restrictions but will be made available on reasonable request to the corresponding author, with the permission of the Alfred Hospital Ethics Committee. Restrictions apply to the availability of these data, which were used under ethical approval for this study. Requests to access these datasets should be directed to JO, [jason.ong@monash.edu](mailto:jason.ong@monash.edu).

## Ethics statement

The studies involving human participants were reviewed and approved by the Alfred Hospital Ethics Committee, Australia (project number: 277/20). As this was a retrospective study involving minimal risk to the privacy of the study subjects, informed consent was waived by the Alfred Hospital Ethics Committee. All identifying details of the study subjects were removed before any computational analysis.

## Author contributions

JO, XX, and LZ conceived and designed the study. XX did data cleaning, established the models and coding, and wrote the first draft and editing. EC, JO, and LZ contributed to data cleaning. EC, CF, MC, IA, JG, JH, NC, LZ, and JO contributed to the interpretation of data and manuscript revision. All authors contributed to the preparation of the manuscript and approved the final manuscript.

## Funding

This work was supported by Australian National Health and Medical Research Council Emerging Leadership Investigator Grant (Grant GNT1172873 to EC; GNT1193955 to JO), the Australian National Health and Medical Research Council Leadership Investigator Grant (Grant GNT1172900 to CF), National Natural Science Foundation of China (Grant 81950410639 to LZ), Outstanding Young Scholars Funding (Grant 3111500001 to LZ), Xi'an Jiaotong University Basic Research and Profession Grant (Grant xtr022019003 to LZ, Grant xzy032020032 to LZ), Epidemiology modeling and risk assessment (Grant 20200344 to LZ), and Xi'an Jiaotong University Young Talent Support Grant (Grant YX6J004 to LZ).

## Acknowledgments

The authors would like to acknowledge Afrizal Afrizal from Melbourne Sexual Health Center for data extraction. The authors would also like to thank Zhen Yu and Zongyuan Ge at Monash University for their insightful discussion of machine learning techniques.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpubh.2022.1031372/full#supplementary-material>

## References

- Institute. TK. *HIV, viral hepatitis and sexually transmissible infections in Australia: Annual surveillance report 2018*. Available online at: <https://kirby.unsw.edu.au/report/hiv-viral-hepatitis-and-sexually-transmissible-infections-australia-annual-surveillance> (accessed May 07, 2017).
- Rowley J, Vander Hoorn S, Korenromp E, Low N, Unemo M, Abu-Raddad LJ, et al. Chlamydia, gonorrhoea, trichomoniasis and syphilis: global prevalence and incidence estimates, 2016. *Bul World Health Organ.* (2019) 97:548–62. doi: 10.2471/BLT.18.228486
- World Health Organization. *Global Health Sector Strategy on Sexually Transmitted Infections 2016–2021: Toward Ending STIs*. Geneva: World Health Organization (2016).
- Geisler WM, Lensing SY, Press CG, Hook EW III. Spontaneous resolution of genital *Chlamydia trachomatis* infection in women and protection from reinfection. *J Infect Dis.* (2013) 207:1850–6. doi: 10.1093/infdis/jit094
- Hosenfeld CB, Workowski KA, Berman S, Zaidi A, Dyson J, Mosure D, et al. Repeat infection with Chlamydia and gonorrhoea among females: a systematic review of the literature. *Sex Transm Dis.* (2009) 36:478–89. doi: 10.1097/OLQ.0b013e3181a2a933
- Xu F, Stoner BP, Taylor SN, Mena L, Tian LH, Papp J, et al. Use of home-obtained vaginal swabs to facilitate rescreening for *Chlamydia trachomatis* infections: two randomized controlled trials. *Obstet Gynecol.* (2011) 118(2 Pt 1):231–9. doi: 10.1097/AOG.0b013e3182246a83
- Dunne EF, Chapin JB, Rietmeijer CA, Kent CK, Ellen JM, Gaydos CA, et al. Rate and predictors of repeat *Chlamydia trachomatis* infection among men. *Sex Transm Dis.* (2008) 35(11 Suppl.):S40–4. doi: 10.1097/OLQ.0b013e31817247b2
- Walker J, Tabrizi SN, Fairley CK, Chen MY, Bradshaw CS, Twin J, et al. Chlamydia trachomatis incidence and re-infection among young women—behavioural and microbiological characteristics. *PLoS ONE.* (2012) 7:e37778. doi: 10.1371/journal.pone.0037778
- Price MJ, Ades AE, De Angelis D, Welton NJ, Macleod J, Soldan K, et al. Risk of pelvic inflammatory disease following Chlamydia trachomatis infection: analysis of prospective studies with a multistate model. *Am J Epidemiol.* (2013). 178:484–92. doi: 10.1093/aje/kws583
- Munari SC, Goller JL, Hellard ME, Hocking JS. Chlamydia prevention and management in Australia: reducing the burden of disease. *Med J Aust.* (2022) 217:499–501. doi: 10.5694/mja2.51749
- Ong JJ, Bourne C, Dean JA, Ryder N, Cornelisse VJ, Murray S, et al. Australian sexually transmitted infection (STI) management guidelines for use in primary care 2022 update. *Sex Health.* (2022). doi: 10.1071/SH22134. [Epub ahead of print].
- Bowring AL, Gouillou M, Guy R, Kong FYS, Hocking J, Pirodda M, et al. Missed opportunities—low levels of chlamydia retesting at Australian general practices, 2008–2009. *Sex Transm Infect.* (2012) 88:330–4. doi: 10.1136/sixtrans-2011-050422
- Lang AS, Jansen K, Sailer A, Bremer V, Dudareva S. Not again! Effect of previous test results, age group and reason for testing on (re-) infection with Chlamydia trachomatis in Germany. *BMC Infect Dis.* (2018) 18:424. doi: 10.1186/s12879-018-3323-2
- Khan A, Fortenberry JD, Juliar BE, Tu W, Orr DP, Batteiger BE. The prevalence of chlamydia, gonorrhoea, and trichomonas in sexual partnerships: implications for partner notification and treatment. *Sex Transm Dis.* (2005) 32:260–4. doi: 10.1097/01.olq.0000161089.53411.cb
- Layton E, Goller JL, Coombe J, Temple-Smith M, Tomnay J, Vaisey A, et al. 'It's literally giving them a solution in their hands': the views of young Australians towards patient-delivered partner therapy for treating chlamydia. *Sex Transm Infect.* (2021) 97:256–60. doi: 10.1136/sixtrans-2020-054820
- Kissinger P, Brown R, Reed K, Salifou J, Drake A, Farley TA, et al. Effectiveness of patient delivered partner medication for preventing recurrent *Chlamydia trachomatis*. *Sex Transm Infect.* (1998) 74:331–3. doi: 10.1136/sti.74.5.331
- Ramstedt K, Forsman L, Johannisson G. Contact tracing in the control of genital *Chlamydia trachomatis* infection. *Int J STD AIDS.* (1991) 2:116–8. doi: 10.1177/095646249100200208
- Goller JL, Coombe J, Bourne C, Bateson D, Temple-Smith M, Tomnay J, et al. Patient-delivered partner therapy for chlamydia in Australia: can it become part of routine care? *Sex Health.* (2020) 17:321–9. doi: 10.1071/SH20024
- Steyerberg EW. *Clinical Prediction Models*. New York, NY: Springer (2019).
- Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. *Nat Methods.* (2018) 15:233–4. doi: 10.1038/nmeth.4642
- Weng SF, Reys J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS ONE.* (2017) 12:e0174944. doi: 10.1371/journal.pone.0174944
- Mullainathan S, Spiess J. Machine learning: an applied econometric approach. *J Econ Perspect.* (2017) 31:87–106. doi: 10.1257/jep.31.2.87
- Elder HR, Gruber S, Willis SJ, Cocoros N, Callahan M, Flagg EW, et al. Can machine learning help identify patients at risk for recurrent sexually transmitted infections? *Sex Transm Dis.* (2021) 48:56–62. doi: 10.1097/OLQ.0000000000001264
- Xu X, Fairley CK, Chow EPF, Lee D, Aung ET, Zhang L, et al. Using machine learning approaches to predict timely clinic attendance and the uptake of HIV/STI testing post clinic reminder messages. *Sci Rep.* (2022) 12:8757. doi: 10.1038/s41598-022-12033-7
- Bao Y, Medland NA, Fairley CK, Wu J, Shang X, Chow EPF, et al. Predicting the diagnosis of HIV and sexually transmitted infections among men who have sex with men using machine learning approaches. *J Infect.* (2020) 82:48–59. doi: 10.12139/ssrn.3550064
- Gaydos CA, Wright C, Wood BJ, Waterfield G, Hobson S, Quinn TC. *Chlamydia trachomatis* reinfection rates among female adolescents seeking rescreening in school-based health centers. *Sex Transm Dis.* (2008) 35:233–7. doi: 10.1097/OLQ.0b013e31815c11fe
- Parvande S, Yeh HW, Paulus MP, McKinney BA. Consensus features nested cross-validation. *Bioinformatics.* (2020) 36:3093–8. doi: 10.1093/bioinformatics/btaa046
- Xu X, Ge Z, Chow EP, Yu Z, Lee D, Wu J, 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. doi: 10.3390/jcm11071818
- Rose LT, Fischer KW. Garbage in, garbage out: having useful data is everything. *Measur Interdiscipl Res Perspect.* (2011) 9:222–6. doi: 10.1080/15366367.2011.632338
- Ozkan T, Clipper SJ, Piquero AR, Baglivio M, Wolff K. Predicting sexual recidivism. *Sex Abuse.* (2020) 32:375–99. doi: 10.1177/1079063219852944
- Vowels LM, Vowels MJ, Mark KP. Uncovering the most important factors for predicting sexual desire using explainable machine learning. *J Sex Med.* (2021) 18:1198–216. doi: 10.1016/j.jsxm.2021.04.010
- Govindan B, Maduravasagam K. Deploying machine learning to find out the reasons for not using condom in a questionnaire-based study of 120 patients. *Indian J Sex Transm Dis AIDS.* (2018) 39:50–4. doi: 10.4103/ijstd.ijstd\_64\_17
- Wray TB, Luo X, Ke J, Pérez AE, Carr DJ, Monti PM. Using smartphone survey data and machine learning to identify situational and contextual risk factors for HIV risk behavior among men who have sex with men who are not on PrEP. *Prev Sci.* (2019) 20:904–13. doi: 10.1007/s11211-019-01019-z
- Pan Y, Liu H, Metsch LR, Feaster DJ. Factors associated with HIV testing among participants from substance use disorder treatment programs in the US: a machine learning approach. *AIDS Behav.* (2017) 21:534–46. doi: 10.1007/s10461-016-1628-y
- Whittington WL, Kent C, Kissinger P, Oh MK, Fortenberry JD, Hillier SE, et al. Determinants of persistent and recurrent *Chlamydia trachomatis* infection in young women: results of a multicenter cohort study. *Sex Transm Dis.* (2001) 28:117–23. doi: 10.1097/00007435-200102000-00011
- Russell AN, Zheng X, O'Connell CM, Taylor BD, Wiesenfeld HC, Hillier SL, et al. Analysis of factors driving incident and ascending infection and the role of serum antibody in chlamydia trachomatis genital tract infection. *J Infect Dis.* (2016) 213:523–31. doi: 10.1093/infdis/jiv438
- Visser M, van Aar F, Koedijk FD, Kampman CJ, Heijne J. Repeat Chlamydia trachomatis testing among heterosexual STI outpatient clinic visitors in the Netherlands: a longitudinal study. *BMC Infect Dis.* (2017) 17:782. doi: 10.1186/s12879-017-2871-1
- Guy R, Wand H, Franklin N, Fairley CK, Chen MY, O'Connor CC, et al. Retesting for chlamydia at sexual health services in Australia, 2004–08. *Sexual Health.* (2011) 8:242–7. doi: 10.1071/SH10086
- Wijers J, Dukers-Muijers N, Hoebe C, Wolffs PFG, van Liere G. The characteristics of patients frequently tested and repeatedly infected with *Chlamydia trachomatis* in Southwest Limburg, the Netherlands. *BMC Public Health.* (2020) 20:1239. doi: 10.1186/s12889-020-09334-9
- Vajdic CM, Middleton M, Bowden FJ, Fairley CK, Kaldor JM. The prevalence of genital Chlamydia trachomatis in Australia 1997/2004: a systematic review. *Sexual Health.* (2005) 2:169–83. doi: 10.1071/S05018

41. Jasumback CL, Perry SH, Ness TE, Matsenjwa M, Masangane ZT, Mavimbela M, et al. Point-of-care testing to guide treatment and estimate risk factors for sexually transmitted infections in adolescents and young people with human immunodeficiency virus in Eswatini. *Open Forum Infect Dis.* (2020) 7:ofaa052. doi: 10.1093/ofid/ofaa052

42. Allan-Blitz LT, Leon SR, Bristow CC, Konda KA, Vargas SK, Flores JA, et al. High prevalence of extra-genital chlamydial or gonococcal infections among men who have sex with men and transgender women in Lima, Peru. *Int J STD AIDS.* (2017) 28:138–44. doi: 10.1177/0956462416630909

43. Fairley CK, Sze JK, Vodstrcil LA, Chen MY. Computer-assisted self interviewing in sexual health clinics. *Sex Transm Dis.* (2010) 37:665–8. doi: 10.1097/OLQ.0b013e3181f7d505

44. Chow EPF, Hocking JS, Ong JJ, Phillips TR, Fairley CK. Sexually transmitted infection diagnoses and access to a sexual health service before and after the national lockdown for COVID-19 in Melbourne, Australia. *Open Forum Infect Dis.* (2020) 8:ofaa536. doi: 10.1093/ofid/ofaa536