



## Research Paper

# Predicting STI Diagnoses Amongst MSM and Young People Attending Sexual Health Clinics in England: Triage Algorithm Development and Validation Using Routine Clinical Data

Carina King <sup>a,\*</sup>, Gwenda Hughes <sup>a,b</sup>, Martina Furegato <sup>b,c</sup>, Hamish Mohammed <sup>a,b</sup>, John Were <sup>b</sup>, Andrew Copas <sup>a</sup>, Richard Gilson <sup>a</sup>, Maryam Shahmanesh <sup>a</sup>, Catherine H Mercer <sup>a</sup>

<sup>a</sup> Institute for Global Health, University College London, 30 Guilford Street, London, UK

<sup>b</sup> Department of HIV & STIs, Centre for Infectious Disease Surveillance & Control, Public Health England, 61 Colindale Avenue, London, UK

<sup>c</sup> Applied Diagnostic Research and Evaluation Unit, St George's University of London, Cranmer Terrace, London, UK

## ARTICLE INFO

## Article history:

Received 22 June 2018

Received in revised form 16 October 2018

Accepted 15 November 2018

Available online 28 November 2018

## Keywords:

STI

Young people

MSM

Predictive model

Triage

Routine data

## ABSTRACT

**Background:** Sexual health (SH) services increasingly need to prioritise those at greatest risk of sexually transmitted infections (STIs). We used SH surveillance data to develop algorithms to triage individuals attending SH services within two high-risk populations: men who have sex with men (MSM) and young people (YP).

**Methods:** Separate multivariable logistic regression models for MSM and YP were developed using surveillance data on demographics, recent sexual history, prior STI diagnoses and drug/alcohol use from five clinics in 2015–2016 to identify factors associated with new STI diagnoses. The models were prospectively applied in one SH clinic in May 2017 as an external validation.

**Findings:** 9530 YP and 1448 MSM SH episodes informed model development. For YP, factors associated with new STI diagnosis (overall prevalence: 10.6%) were being of black or mixed white/black ethnicity; history of chlamydia diagnosis (previous year); and multiple partners/new partner (previous 3-months). The YPs model had reasonable performance (c-statistic: 0.703), but poor discrimination when externally validated (c-statistic: 0.539). For MSM, being of South Asian ethnicity; being born in Europe (excluding the UK); and condomless anal sex or drug use (both in previous 3-months) were associated with STI diagnosis (overall prevalence: 22.0%). The MSM model had a c-statistic of 0.676, reducing to 0.579 on validation.

**Interpretation:** SH surveillance data, including limited behavioural data, enabled triage algorithms to be developed, but its implementation may be problematic due to poor external performance. This approach may be more suitable to self-triage, including online, ensuring patients are directed towards appropriate services.

**Funding:** NIHR HTA programme (12/191/05).

© 2018 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

Sexually transmitted infections (STIs) remain prevalent in England, with the highest rates seen in young people (YP) and men who have sex with men (MSM) [1,2]. Behavioural interventions can reduce risky sexual behaviours, increase testing and ultimately reduce STI incidence, and attendance at a sexual health clinic provides an opportunity for targeted intervention delivery [3–6]. However, funding for sexual health services in England has decreased since 2016, with reductions of up to 20% occurring alongside overall increases in attendance [7,8].

Data driven triage, developed through predictive statistical models, is relatively common in primary and secondary clinical care [9], such

as the Framingham risk score which supports treatment decisions for cardiovascular disease [10]. In sexual health, triage is common-place [11,12], with clinics stratifying patients according to symptoms, behavioural risks and demographics to receive different services, such as ‘quick checks’ or safe-guarding [13,14]. However, these triage approaches focus primarily on dichotomous pre-defined criteria rather than using statistical, population-level models, which may not necessarily take into account risk behaviours or identify individual patients most in need of interventions [15].

Since 2008, sexual health clinics in England have reported all STI services and diagnoses to a nationally mandated STI surveillance system (GUMCAD) [16,17]. This has allowed monitoring of trends in STI diagnoses, but has lacked information on risk behaviours for more detailed risk stratification. To improve monitoring of the patient case-mix and

\* Corresponding author at: Institute for Global Health, 30 Guilford Street, London WC1N 1EH, UK.

E-mail address: [c.king@ucl.ac.uk](mailto:c.king@ucl.ac.uk) (C. King).

## Research in context

### *Evidence before this study*

A search was conducted in May 2018 of PubMed and Embase databases using the following search terms: (triag\* OR algorithm OR score) AND (STI OR 'sexually transmitted infection' OR 'sexual health'). Papers were considered that included STI risk algorithm development, validation or evaluation of introduction into routine care. Studies which focussed on risks of acquiring a single infection (e.g. HIV), related to the biological diagnosis of an infection, appointment booking systems or were published prior to 2000 were excluded. We found 11 examples of risk algorithms that used demographic, sexual history and behavioural data to develop risk tools, conducted in Australia (n = 2), Canada (n = 1), USA (n = 6), the Netherlands (n = 1) and South Africa (n = 1). Of these only two considered all STIs as the outcome and both were amongst young people, with others focussed specifically on HIV (n = 5) or combinations of chlamydia, gonorrhoea and trichomonas (n = 4). All studies showed reasonable or good model discrimination, with c-statistics in development models ranging from 0.72 to 0.89. Validation models, which used a range of internal and external approaches such as geographical and temporal validation, showed reduced performance (c-statistics: 0.64–0.86). The existing data suggest that this approach could be appropriate for tailoring interventions or types of services based on risk. However, no studies were found that evaluated the impact of these models to direct patients into sexual risk reduction interventions, using all STIs as the outcome in MSM or using electronic patient records in the UK.

### *Added value of this study*

Our study is the first that we know of from the UK that uses electronic surveillance data collected within sexual health clinics, alongside an external prospective validation, to triage service users according to STI risk. Compared to previously published studies, we used a comprehensive STI outcome and included data in the development model from five different sexual health services that are geographically dispersed and offer different levels of services. We developed separate models for young people (aged 16–25 years old) and men who have sex with men, and found they had marginally lower discrimination (i.e. ability to differentiate between those with and without an STI) than those in previously published studies (c-statistic = 0.70 and 0.68, respectively).

### *Implications of all the available evidence*

Our study supports previously published evidence that predictive risk algorithms for STI diagnosis which have reasonable ability to discriminate between people with and without an STI diagnosis based on a combination of demographic, sexual history and behaviour data. However, the evidence-base for scaling-up these triage approaches across diverse clinical settings, geographies and patient profiles, is lacking. Further work is needed on how these models can be further refined with existing electronic surveillance data, and how they can be integrated into sexual health services which transition to offering both online self-testing and clinic-based pathways.

recommended by the British Association for Sexual Health and HIV (BASHH) [18–20]. Electronic patient record (EPR) systems facilitate the capture of these standardised risk behaviours, providing an opportunity to implement and automate triaging within sexual health clinics without expending considerable resources. Examples of triage tools using this approach in sexual health have been developed in other high-income settings, and have shown reasonable ability to predict STI diagnoses [21–23].

We aimed to develop and validate triage models for young people and MSM using electronic surveillance data collected from sexual health clinics in England, to stratify patients according to their risk of STI diagnosis. This work was conducted as part of a larger pilot feasibility study (Santé Project; ISRCTN: 16738765). The wider aim of this project was to identify and adapt evidence-based brief behavioural risk reduction interventions that could be delivered routinely within sexual health clinics, such as a motivational interviewing session with a health advisor, and use automated triage models to refer “higher risk” service users into this intervention.

## 2. Methods

We developed two predictive models, one for YP and one for MSM, using data collected as part of a pilot of an enhanced specification of GUMCAD, conducted by PHE in 2015–2016. We prospectively implemented the two models in a single sexual health clinic Brighton in May 2017, as part of pilot trial of the triaging process and delivery of a brief motivational interviewing session for “higher risk” patients. The data collected during this pilot was used to externally validate the models.

### 2.1. Enhanced GUMCAD Surveillance Pilot

The dataset used for model development was generated during a pilot of an enhanced GUMCAD specification by PHE from July 2015 to June 2016 [19]. This enhancement has now been approved for routine implementation in England. Five clinics located in southern England, including two in Greater London, took part. Four clinics were comprehensive specialist STI services for all patients and one was specifically a young person's clinic run by Brook (a sexual health charity for young people). The dataset included routinely-collected, standardised information on STI diagnoses and patient demographics [17], and additional data on recent sexual behaviours, drug and alcohol use, and partner notifications (Supplementary Table 1). The behavioural variables were based on those recommended for sexual history taking by BASHH and are well-supported in the literature as being risks for STI acquisition [20].

### 2.2. Definitions

Young people were defined as all women, and men who self-reported as heterosexual and reported no sex with men in the previous 3-months, aged 16–25 years old. MSM were defined as men who reported any sex with men in the previous 3-months, or self-reported as bisexual or homosexual, aged 16 and over. Within the GUMCAD data collected during the study period, gender was coded as a binary variable based on patient's self-reported gender. This means that we were unable to differentiate between or include non-binary and trans-gender identities in the models.

The outcome for both models was a new specified STI diagnosis at the current clinic attendance, including: HIV; primary, secondary or early latent syphilis; gonorrhoea; chlamydia; hepatitis A, B and C; lymphogranuloma venereum (LGV); trichomonas; or herpes. Recurrent herpes, genital warts and non-specific genital infections were excluded.

Demographic predictors included: categorised age; quintiles of socio-economic status using the Index of Multiple Deprivation score, derived from the patient's area of residence [24]; self-reported ethnicity, country of birth, sexual orientation, and gender; HIV status; and STI diagnosis in the previous 12 months based on clinical records. Previous STI diagnosis included only those diagnosed within the same clinic

interpretation of STI trends, Public Health England (PHE) proposed an enhanced GUMCAD dataset including additional behavioural and sexual history variables collected during routine patient consultations, as

as patient ID codes are unique to each clinic. Behavioural predictors included in both models were: number of sexual partners in the last 3-months; problematic alcohol use, assessed using a standardised tool according to local clinical guidelines; and any recreational drug use in the last 3-months. Young person specific variables were defined as: any new partners in the last 3-months and condom use at last sex. MSM specific variables were defined as: condomless anal intercourse (CAI) and sex with known HIV positive partners in the last 3-months. These predictors were decided a priori based on published literature and guidelines.

### 2.3. Sample Size

As we used secondary data for model development, we did not conduct an a priori sample size calculation. However, as a general rule to prevent over-fitting, 10 outcome events (i.e. STI diagnoses) per degree of freedom in the model (i.e. predictor variable) are recommended [25].

### 2.4. Model Development

All pre-defined variables were included in the primary models, except if missing data were too common, and after re-categorisation to reduce degrees of freedom and merge empty or low frequency cells. Continuous variables were investigated for linearity with the log odds of STI diagnosis and modelled as splines, continuous and categorical; interaction terms for age, ethnicity and gender were also investigated. We opted for the simplest model, without interaction terms and using categorical variables, comparing models using the Bayesian Information Criteria (BIC) to check whether they were statistically supported. Regarding missing data, we assumed data may not be missing at random (i.e. we assumed data may be missing depending on its value), and therefore standard multiple imputation may not be valid and was not used as our primary approach. In addition, if a variable has high levels of missing data, then its inclusion in a triage tool may be impractical [26]. We chose a pragmatic primary approach of including variables with less than one third missing data and coded missingness as a distinct category retaining all observations in the model. Variables with more than one third missing data were excluded.

Multivariable logistic regression, using a maximum likelihood approach, was used to generate the models in Stata SE14, with default settings. Global p-values were calculated using Wald tests with Stata's post-estimation command *-test-*. Model calibration was assessed with the Hosmer-Lemeshow goodness of fit test [27,28], and model discrimination was based on the c-statistic [26,29]. A c-statistic of >0.8 is considered strong discrimination and >0.7 as reasonable for a clinical tool; 0.5 indicates performance the same as chance at predicting the outcome [30]. Sensitivity and specificities were calculated for different model prediction thresholds. The range of the predicted probabilities of the outcome from across the participants was described. Internal validation was conducted using k-fold cross-validation, with 10 folds.

The following sensitivity analyses were conducted to vary model assumptions and methodology: 1. A forward stepwise approach (with a p-value threshold of <0.2) to model selection to explore the performance of a potentially more parsimonious model; 2. A model including only demographic data to evaluate the added-value of behavioural variables; 3. Imputing missing values (10 imputations with chained equations) in variables with less than one third missing. All analyses were conducted in Stata SE14.

### 2.5. External Validation

External validation is recommended as an independent assessment of model performance to assess the extent of over-fitting and the resulting optimism of its performance in practice [31]. Both models were externally validated during a prospective implementation of the triaging process at a large sexual health clinic in May 2017. We

compared the model's predictions of STI diagnosis with each patient's actual STI diagnosis, calculating the c-statistic and sensitivity and specificity for different model thresholds.

Patients attending the clinic during specific time blocks over a one-month period were asked to self-complete the demographic and behavioural questions on tablets in the waiting room by a member of study staff. On completing the questionnaire, a ticket was printed containing a random study ID number which clinic staff were asked to enter into the clinic's EPR system. The data from the tablet-based system were linked to clinic records using the study ID number, and if not available deterministic matching was undertaken using: date of attendance, time of appointment, patient age and gender.

## 3. Results

A total of 28,514 episodes of care were recorded in the enhanced GUMCAD pilot, including 9530 young people and 1448 MSM (Table 1). The STI diagnosis prevalence in YP was 10.6% (n = 1005; 95% CI: 9.9%, 11.1%) and 22.0% (n = 318; 95% CI: 19.9%, 24.2%) in MSM. This was similar to the national average during the same time period for young people (10.6% vs. 10.8%), but higher in MSM (22.0% vs. 14.9%). Young people and MSM differed from each other in terms of behavioural data, with MSM twice as likely to report any recreational drug use compared to young people (14% versus 7%), and 15% of MSM reporting ≥5 partners in the previous 3-months compared to only 2% in young people. MSM had less missing data (Table 1).

### 3.1. Young Person's Model

Amongst young people, we included prior chlamydia diagnosis rather than all STIs as exploratory analyses showed that it improved model fit. Drug and problematic alcohol use were excluded due to missing data; sexual orientation was excluded as it was one of the eligibility criteria for men. The primary model included 34 degrees of freedom, meeting the required 10 outcomes per degree of freedom.

Being female and older were associated with lower odds of STI diagnosis, while being of black or mixed white and black ethnicity had higher odds of STI diagnosis, compared to being white British (Table 2). Significant behavioural risks included: chlamydia diagnosis in the previous year, having one or more partners compared to none and having a new partner in the prior 3-months. Condom use at last sex was protective. The model had reasonable performance, with a c-statistic of 0.703 (Fig. 1a), and the Hosmer-Lemeshow test did not raise concerns about model fit (p-value = 0.160). Following internal validation, the c-statistic reduced slightly to 0.687. The predicted probabilities of STI diagnosis ranged from 1 to 75%, with a mean of 12%. Fig. 2a plots the model's predicted risk of STI diagnosis against the population risk percentile, demonstrating the range in predicted risks, and suggesting that the risk of STI diagnosis is concentrated in a small percentage of the population. Taking an example, where we prioritise a balance between sensitivity and specificity, a predicted risk of STI diagnosis of ≥10% would result in a sensitivity of 67%, specificity of 63%, and would classify 40% of patients as "high risk" (Figs. 1a and 2a).

Sensitivity analysis using a forward stepwise approach retained all variables. A model fitted using multiple imputation for missing values showed similar performance (c-statistic: 0.688), and a slightly narrower range of predicted risks (1–68%). The magnitude and direction of associations were similar to the primary model. Using demographic data only showed considerably worse model performance (c-statistic: 0.590) and a limited range of predicted risks (2–24%).

### 3.2. MSM Model

Amongst MSM, a range of prior STI diagnoses were reported (including HIV, syphilis, chlamydia and gonorrhoea), but contained too few observations as individual predictors, therefore a single binary variable

**Table 1**  
Description of GUMCADv3 enhanced surveillance data.

Variables	Total (N, %) N = 23,103	Young people (N, %) N = 9530	MSM (N, %) N = 1448
<b>Demographic variables</b>			
Gender			
Male	9491 (41%)	2983 (31%)	1448 (100%)
Female	13,612 (59%)	6547 (69%)	–
Age			
16–20 years	2938 (13%)	2628 (18%)	77 (5%)
20–24 years	6052 (26%)	6902 (82%)	297 (21%)
25–34 years	8664 (38%)	–	562 (39%)
35–44 years	3291 (14%)	–	262 (18%)
45–64 years	2007 (9%)	–	213 (15%)
≥65 years	151 (1%)	–	37 (3%)
Sexual orientation <sup>a</sup>			
Heterosexual	17,758 (77%)	7809 (82%)	51 (4%)
Bisexual	540 (2%)	120 (1%)	299 (21%)
Homosexual	1034 (4%)	27 (0%)	963 (67%)
Missing	3771 (16%)	1574 (17%)	135 (9%)
Country/continent of birth			
UK	15,682 (68%)	6813 (71%)	1049 (72%)
Europe	2095 (9%)	643 (7%)	153 (11%)
Africa	1134 (5%)	309 (3%)	40 (3%)
Americas	821 (4%)	217 (2%)	43 (3%)
Asia	289 (1%)	51 (1%)	19 (1%)
Other	618 (3%)	190 (2%)	54 (4%)
Missing	2464 (11%)	1307 (14%)	90 (6%)
Ethnicity			
White	13,639 (59%)	6072 (64%)	1003 (69%)
British			
Other White	2554 (11%)	785 (8%)	185 (13%)
South Asian	661 (3%)	201 (2%)	33 (2%)
Other Asian	463 (2%)	165 (2%)	36 (2%)
Black	1353 (6%)	448 (5%)	28 (2%)
Caribbean			
Other Black	2379 (10%)	991 (10%)	54 (4%)
White & Black mixed	826 (4%)	418 (4%)	31 (2%)
Other Mixed	548 (2%)	249 (3%)	38 (3%)
Any Other	233 (1%)	79 (1%)	17 (1%)
Missing	447 (2%)	122 (1%)	23 (2%)
Index of Multiple Deprivation quintiles			
Lowest	4731 (20%)	1744 (18%)	294 (20%)
2nd quintile	6019 (26%)	2364 (25%)	354 (24%)
3rd quintile	4257 (18%)	1768 (19%)	259 (18%)
4th quintile	4291 (19%)	1937 (20%)	273 (19%)
Highest	2917 (13%)	1363 (14%)	213 (15%)
Missing	888 (4%)	354 (4%)	55 (4%)
Previous STI diagnosis <sup>b</sup>			
No	21,526 (93%)	8795 (92%)	1329 (92%)
Yes	1577 (7%)	735 (8%)	119 (8%)
<b>Behavioural variables</b>			
Number of partners <sup>c</sup>			
None	1068 (5%)	365 (4%)	73 (5%)
1 partner	10,893 (47%)	4336 (46%)	410 (28%)
2–4 partners	4037 (17%)	1660 (17%)	506 (35%)
≥5 partners	649 (3%)	206 (2%)	215 (15%)
Missing	6456 (28%)	2963 (31%)	244 (17%)
New partners <sup>c</sup>			
No	–	2658 (28%)	–
Yes	–	2663 (28%)	–
Missing	–	4209 (44%)	–
Condom use at last sex			
No	–	3881 (41%)	–
Yes	–	2014 (21%)	–
Missing	–	3635 (38%)	–
Anal sex with known HIV +ve <sup>c</sup>			
No	–	–	786 (54%)
Yes	–	–	124 (9%)
Missing	–	–	538 (37%)
Condomless anal sex <sup>c</sup>			
No	–	–	419 (29%)
Yes	–	–	535 (37%)
Missing	–	–	494 (34%)
Receptive condomless anal sex <sup>c</sup>			
No	–	–	138 (10%)
Yes	–	–	350 (24%)
Missing	–	–	960 (66%)
Problematic alcohol use <sup>d</sup>			
No	4558 (20%)	1890 (20%)	192 (13%)
Yes	203 (1%)	102 (1%)	22 (2%)
Missing	18,342 (79%)	7538 (79%)	1234 (85%)
Drug use <sup>c</sup>			
No	10,212 (44%)	3860 (41%)	795 (55%)
Yes	1537 (7%)	686 (7%)	199 (14%)
Missing	11,354 (49%)	4984 (52%)	454 (31%)

<sup>a</sup> These relate to females only in the young people, and in the MSM, self-reported heterosexuals who reported same sex male partners were included in the MSM group.

<sup>b</sup> Within the previous 12 months.

<sup>c</sup> Within the previous 3 months.

<sup>d</sup> As assessed and defined by local clinical guidelines.

indicating STI diagnosis in the previous year was used. Problematic alcohol use and receptive CAI were excluded due to missingness. The primary model included 36 degrees of freedom, and may therefore be over-fitted (Table 3).

Overall none of the demographic variables were statistically significant, although being of South Asian ethnicity and being born outside of the UK in Europe had increased odds of STI diagnosis compared to being White British and born in the UK, respectively. Significant behavioural risks included CAI and recreational drug use in the prior 3-months. The model had reasonable performance, with a c-statistic of 0.676 and the Hosmer-Lemeshow test did not raise concerns about model fit (p-value = 0.224) (Fig. 1b). Following internal cross-validation, the c-statistic was 0.612, showing poorer discrimination. The predicted probability of STI diagnosis ranged from 3 to 71%, with a mean predicted risk of 16%; Fig. 2b shows a higher level of predicted risk across the development population, compared to young people, with a gradual increase in predicted risk of STI diagnosis. As an illustration, a threshold of ≥20% would give a sensitivity of 66% and specificity of 60%, but classify 46% of patients as “high risk”.

The forward stepwise model excluded age, deprivation quintile, number of partners in the prior 3-months and ethnicity. This model was favoured by the BIC as more parsimonious, but it had poorer discrimination (c-statistic: 0.658). The multiple imputation model had comparable performance to the primary model (c-statistic: 0.676). The predicted risks ranged from 4 to 71%, also showing very similar discrimination, and similar direction and magnitude of relationships with the outcome. A demographic only model showed poor performance and discrimination (c-statistic: 0.553), with predicted risks ranging from 7 to 23%.

### 3.3. External Validation

A total of 246 MSM and 306 YP successfully self-completed the triage questionnaire during the one-month pilot period, representing 19.8% (N = 2697) of all patients who attended the clinic during this period. We were able to link 46.2% (n = 246) of these patient episodes to the clinic's EPR system for the external validation. There were 136 young people, of whom 23.5% (n = 32; 95% CI: 16.7%, 31.6%) had an STI diagnosis, and 110 MSM with an STI prevalence of 28.1% (n = 31; 95% CI: 20.0%, 37.6%) included in the validation analysis.

Data on sexual history and behaviour had higher completion than the enhanced GUMCAD dataset, but postcode for calculating deprivation score had lower completion. This pattern was seen in both MSM and young people (Supplementary Tables 2 and 3). This may reflect patients' willingness to provide information in different formats such as using self-completed form, or the level of staff compliance in collecting or documenting information in routine care.

The young person's primary model in the external validation dataset had a c-statistic of 0.539. Using a predicted risk of ≥10%, which has the best balance between sensitivity and specificity in the development model, resulted in a specificity of <5%, but 100% sensitivity. This shows poor discrimination. For the MSM model the c-statistic was 0.579, and using a predicted risk of STI diagnosis of ≥30% gave a sensitivity of 48% and specificity of 73%. This showed a similar balance between sensitivity and specificity to the development dataset, but is a higher risk cut-off and model discrimination decreased.

## 4. Discussion

We were able to develop triage models for young people and MSM using data routinely collected within sexual health clinics in England, with both models meeting the threshold for being a clinically reasonable diagnostic tool in development. The inclusion of STI history and recent behaviours was crucial to meeting the threshold of reasonable performance and improvements in data quality may provide opportunities for further refinement. However, balancing sensitivity, specificity

**Table 2**  
Multivariable regression model for STI diagnosis in young people (development).

Variable		Adjusted OR <sup>a</sup>	Coefficient	p-Value	95% Confidence Interval	
Gender	Male	1.00		<0.001	0.62	0.83
	Female	0.71	−0.34			
Ethnicity	White	1.00		<0.001	0.95	1.86
	British					
	White, other	1.33	0.28			
	South Asian	0.73	−0.32			
	Asian, other	0.94	−0.06			
	Black	2.65	0.98			
	Caribbean					
	Black, other	1.57	0.45			
	White &	1.85	0.61			
	Black mixed					
	Mixed, other	0.88	−0.13			
	Other	0.69	−0.37			
Missing	0.85	−0.16		0.42	1.73	
Country/continent of birth	UK	1.00		0.106	0.72	1.47
	Europe	1.03	0.03			
	Africa	0.66	−0.42			
	Americas	0.77	−0.26			
	Asia	0.42	−0.86			
	Other	0.89	−0.12			
	Missing	0.78	−0.24			
Age	16–17 years	1.00		0.003	0.59	1.00
	18–19 years	0.77	−0.26			
	20–21 years	0.81	−0.21			
	22–23 years	0.70	−0.36			
	24–25 years	0.62	−0.48			
	Missing	0.85	−0.16			
Index of Multiple Deprivation	Quintile 1 (highest)	1.00		0.237	0.74	1.10
	Quintile 2 (high)	0.91	−0.10			
	Quintile 3 (medium)	1.00	<−0.01			
	Quintile 4 (low)	0.84	−0.18			
	Quintile 5 (lowest)	0.80	−0.23			
	Missing	1.14	0.13			
Previous chlamydia <sup>b</sup>	No	1.00		<0.001	2.88	4.65
	Yes	3.66	1.30			
Number of partners <sup>c</sup>	0 partners	1.00		0.023	1.19	3.91
	1 partner	2.16	0.77			
	2–4 partners	2.51	0.92			
	≥5 partners	2.58	0.95			
	Missing	1.49	0.40			
New partners <sup>c</sup>	No	1.00		<0.001	1.19	1.77
	Yes	1.45	0.37			
	Missing	1.89	0.64			
Condom use at last sex	No	1.00		<0.001	0.41	0.62
	Yes	0.50	−0.69			
	Missing	0.35	−1.04			

<sup>a</sup> All variables in the table included in the multivariable model.

<sup>b</sup> Within the previous 12 months.

<sup>c</sup> Within the previous 3 months. Model coefficient = −2.34; McFadden's pseudo R<sup>2</sup> = 7.8%. OR = odds ratio.

and the proportion of patients classified as “high risk”, these models may be limited in their practical use in a clinical setting with limited resources for providing risk reduction interventions.

The young person's model identified several significant predictors of STI diagnosis, which correspond to previously established associations. Multiple partners and a prior chlamydia diagnosis are established risks for STIs amongst young people [32–34], and being of black (including Black Caribbean) or mixed white and black ethnicity agrees with previous findings from the UK [35,36]. Possible explanations for these associations are different levels of sexual health knowledge, and therefore behaviours, amongst younger and black ethnic minorities [37]. Being female, older than 17 years and reporting condom use at last sex were associated with lower odds of STI diagnosis, again echoing studies from other settings [32,38,39].

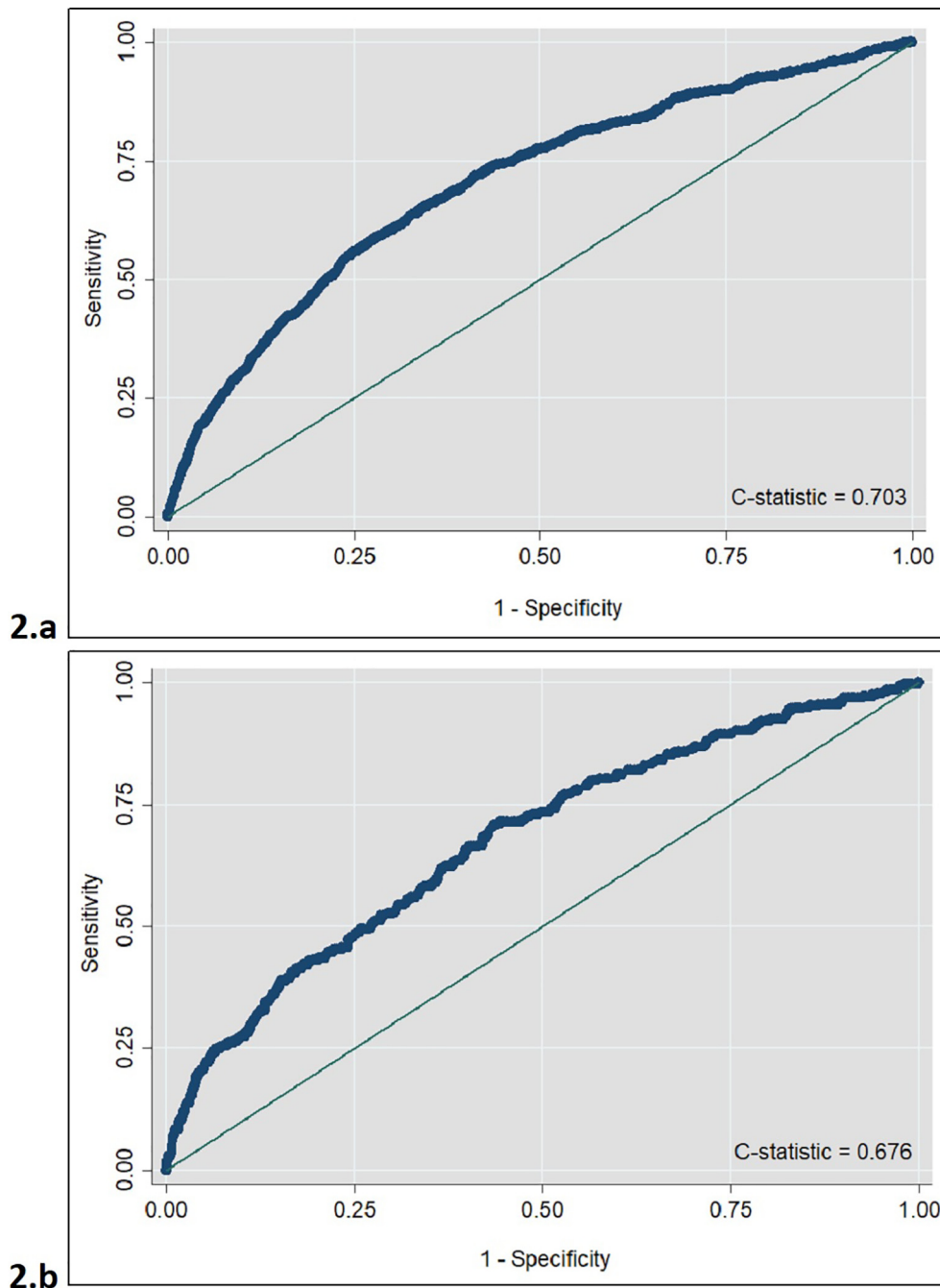
The MSM model identified two significant predictors of STI diagnosis: having had CAI and drug use in the prior 3-months. The use of recreational drugs amongst MSM as a risk for STI diagnosis corresponds to findings from multiple studies [40–42]. The lack of association with number of recent partners contradicted published literature, although it should be noted that there was a trend in the adjusted odds ratios [41,43,44]. When we used a forward stepwise approach, number of partners was not retained in the model, along with age, deprivation and ethnicity. This may be the result of correlation between drug use and partner numbers, or that MSM attending sexual health clinics have higher numbers of partners than the general population.

External validation showed poor ability for the models to predict STI diagnosis in a different clinical setting than the development data. Validation, both internal and external, generally shows worse model performance indicators than development models, so to some extent this was expected. However, additional contributing factors may have been the difference in STI prevalence between datasets, which was more pronounced in YP where performance declined more (YP: 11% versus 24%; MSM: 22% versus 28%). The demographic distribution of patients in the validation dataset differed considerably from the development dataset, and represented one clinical service (Supplementary Tables 2 and 3). As the risk of STI diagnosis at the population level varies according to demographics, it is likely that in different regions these demographic factors will vary in significance according to the local epidemiology. This raises a more general challenge with using standardised prediction models derived from national data, to increase power and generate robust models. However, if the magnitude or direction of association between predictors and the outcome vary between clinics or regions, then this approach may be unreasonable. More sophisticated machine learning methods for triage tool development could lead to improved model performance, but need to balance statistical performance with interpretability and real-world application in a clinical setting. Previous research has shown that transparency in algorithms and their assumptions are important to service providers, so a simplistic approach to development could be a strength during implementation [45].

Another explanation for the reduced discrimination in the validation could be the way in which the information was collected, with patients self-completing the questions instead of answering them in consultation with a healthcare provider. A previous study in sexual health clinics found that patients reported higher levels of sexual risk behaviours when self-completing, compared to face-to-face consultations [46]. This is likely supported by our findings, with no missing data recorded on the number of partners in the prior 3-months for MSM or YP in the validation, compared to 17% and 31% in the development dataset, respectively. This suggests patients may be more willing to report these outside of a clinical consultation (e.g. drug use in MSM), and therefore the associations between these questions and STI diagnosis would likely differ from the development dataset. Alternatively, it may be that this data is collected during routine consultations, but healthcare providers do not comply with documentation requirements, reducing the quality of routinely collected data.

While there were similarities in demographics between young people and MSM and the wider clinic populations from the five pilot sites, these sub-populations may not be representative of attendances at sexual health services nationally. For MSM specifically, the STI diagnosis rate in the development population was higher than the nationally-reported rate for the same time period (22.0% versus 14.9%), however, none of the large London clinics which have higher numbers of MSM patients were included in our analyses [47].

This study had several strengths and limitations, the main strength being the implementation of an external prospective validation, a robust method of model validation [31]. Despite being best practice, external validation is rarely conducted alongside model development, limiting conclusions about real-world applicability. However, we acknowledge that the sample of patients whose self-completed electronic data could be linked to clinic EPR data was small, and those included in

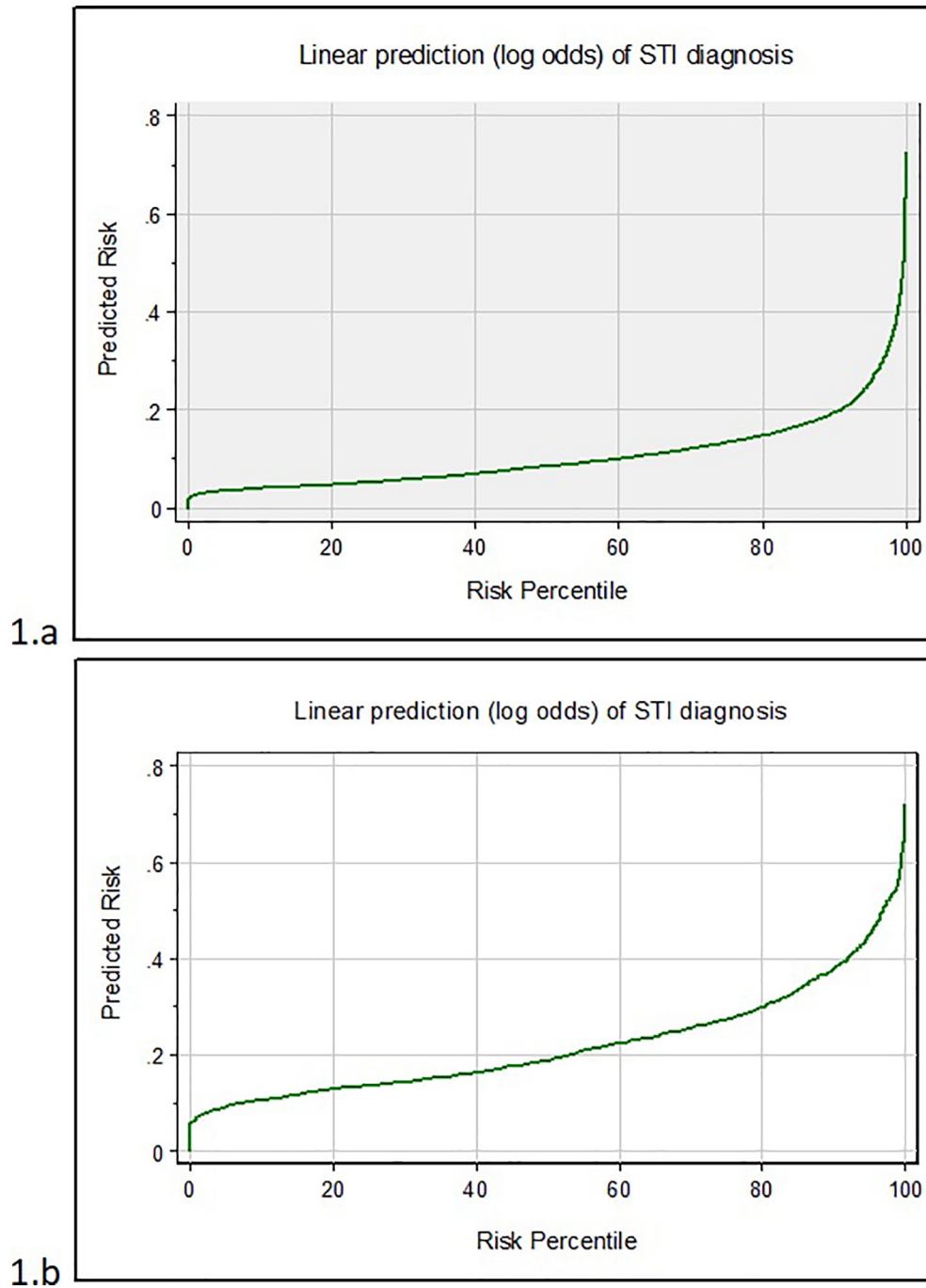


**Fig. 1.** Receive Operating Characteristic (ROC) curves for STI diagnosis in young people and MSM (development models). a: Young people; b: MSM.

analysis had some differences in demographics to the clinic population (e.g. more heterosexual men and fewer young white British participants). This may have biased the performance of the triage model during validation, and limits the strength of our conclusions about external performance. The prospective study provided valuable lessons about algorithm implementation, with challenges, including differences in risk profiles between clinical settings, as well as opportunities to for further investigation, such as patient self-triage.

A key limitation was the high levels of missingness in the development dataset, and restrictions that using routine data posed (e.g. binary gender classification, or the lack of presenting symptom information). This was a dataset developed as part of a pilot implementation, where part of the purpose was to determine the feasibility of implementation and data quality. BASHH guidelines recommend the collection of these key sexual history and behavioural variables, which are deemed feasible

to collect in routine care, especially in sexual health clinics [20]. However, it may be that in practice this is not routinely done, or as clinics were piloting a new system, not all staff were aware that these needed to be documented. As missingness differed between young people and MSM, it suggests these risks are not consistently investigated with all patients. Rather, clinical staff are selective in who they ask and record data for, based on assumptions relating to personal characteristics or local clinic policies, which could result in ascertainment bias. For example, MSM were more likely to have drug use recorded than young people (69% vs. 48%), potentially reflecting provider's awareness of chemsex being more common in this group and a sexual risk factor thus impacting on their clinical decision-making [48]. More complete behavioural data may have improved model performance, and allowed for additional variables to be included (e.g. problematic alcohol use). This is likely to have somewhat limited the predictive performance,



**Fig. 2.** Risk predictiveness curves for STI diagnosis in young people and MSM (development models). a: Young people; b: MSM.

especially for key risk factors such as sexualized drug use. The MSM primary model was fit with 36 degrees of freedom for 318 outcomes; in order to avoid over-fitting a minimum of 360 STI diagnoses amongst MSM would have been required.

Triaging patients according to their risk of STI diagnosis using routinely collected data within sexual health clinics was feasible; however, at a minimum, basic behavioural data are needed to improve the discrimination of these models. The ability to include additional, or more complete, behavioural data would likely improve performance further. However, the feasibility of using these models at a national scale in a standardised way may not be achievable, as patient populations may vary in their risk associations across different clinics and geographical

regions. Further model refinement, potentially exploring machine learning methodologies, and subsequent external validation are still needed. Additionally, further investigation on whether this approach may be more suitable to online patient pathways, which are becoming increasingly common for sexual health, would be valuable [49]. An online algorithm could be developed and applied to a patient's self-completed questionnaire, to ensure those at higher risk of STI diagnosis are directed towards appropriate services, such as a clinical appointment versus self-testing kit sent to their home. While this work demonstrated that developing such a tool was feasible using routine data, further refinement, validation and assessments of the real-world applicability are still needed.

**Table 3**  
Multivariable regression model for STI diagnosis in MSM (development).

Variable		Adjusted OR <sup>a</sup>	Coefficient	p-Value	95% Confidence Interval
Ethnicity	White British	1.00		0.351	
	White, other	0.67	−0.40		0.35 1.30
	South Asian	2.53	0.93		1.05 6.10
	Asian, other	1.43	0.36		0.48 4.21
	Black	0.57	−0.56		0.20 1.67
	Caribbean				
	Black, other	0.98	−0.02		0.47 2.03
	White & Black mixed	0.76	−0.28		0.29 1.97
	Mixed, other	1.19	0.17		0.53 2.70
	Other	1.05	0.04		0.28 3.90
	Missing	1.99	0.69		0.76 5.20
County/continent of birth	UK	1.00		0.110	
	Europe	2.46	0.90		1.26 4.78
	Africa	1.00	0.002		0.42 2.42
	Americas	1.43	0.36		0.60 3.40
	Asia	1.23	0.21		0.37 4.16
	Other	0.88	−0.13		0.33 2.32
	Missing	0.65	−0.43		0.35 1.23
	Age	16–19 years	1.00		0.297
20–24 years	0.75	−0.28		0.41 1.39	
25–34 years	0.79	−0.24		0.44 1.39	
35–44 years	0.63	−0.47		0.34 1.17	
45–64 years	0.55	−0.59		0.29 1.06	
≥65 years	0.41	−0.89		0.13 1.25	
Index of Multiple Deprivation	Quintile 1 (highest)	1.00		0.444	
	Quintile 2 (high)	0.93	−0.07		0.63 1.36
	Quintile 3 (medium)	0.87	−0.14		0.57 1.32
	Quintile 4 (low)	1.08	0.08		0.72 1.63
	Quintile 5 (lowest)	0.66	−0.41		0.41 1.07
	Missing	1.15	0.14		0.56 2.35
Previous STI <sup>b</sup>	No	1.00		0.150	
	Yes	1.40	0.33		0.89 2.20
Number of partners <sup>c</sup>	0 partners	1.00		0.442	
	1 partner	1.24	0.21		0.55 2.76
	2–4 partners	1.30	0.26		0.58 2.93
	≥ 5 partners	1.70	0.53		0.73 3.97
	Missing	1.01	0.01		0.44 2.36
Condomless anal intercourse <sup>c</sup>	No	1.00		<0.001	
	Yes	1.95	0.67		1.39 2.73
	Missing	0.89	−0.12		0.43 1.86
Known HIV positive partner <sup>c</sup>	No	1.00		0.181	
	Yes	1.52	0.42		0.98 2.37
	Missing	1.15	0.14		0.59 2.22
Any recreational drug use <sup>c</sup>	No	1.00		0.003	
	Yes	1.89	0.64		1.31 2.74
	Missing	1.29	0.25		0.87 1.91

<sup>a</sup> All variables in the table included in the multivariable model.

<sup>b</sup> Within the previous 12 months.

<sup>c</sup> Within the previous 3 months. Model coefficient = −1.73; McFadden's pseudo R<sup>2</sup> = 7.0%. OR = odds ratio.

## Acknowledgments

We would like to thank the staff at all the clinics who took part in the enhanced GUMCAD phase 2 pilot (Bedford, Bristol, Barnet, Southend and Croydon) and all the staff at Claude Nicol for taking part in the external validation pilot (specifically Dr. Daniel Richardson, Ms. Laura Clark and Ms. Catherine Hendrickx). We would like to acknowledge Dr. Alex Pollard and Mr. Alex Langrish (Brighton and Sussex Medical School) for participant recruitment during the pilot, and Mrs. Sarika Desai (Public Health England) for input into the analysis methods. We would like to acknowledge the remaining investigators on the SANTE project: Dr. Carrie Llewellyn (Brighton and Sussex Medical School), Dr. Fiona Burn, Dr. Alison Rodger, Dr. Julia Bailey, Dr. Alison Howarth

(University College London), Prof. Charles Abraham (University of Exeter) and Dr. Anupama Roy.

## Funding

Funding was provided by the National Institute for Health Research Health Technology Assessment programme (12/191/05). The funders did not have any role in the study design, data collection, analysis, interpretation or writing of the manuscript.

## Conflict of Interest

Authors declare no conflicts of interest.

## Author Contributions

The manuscript was conceived by CK, MS, GH, CM, AC and RG. The model development was conducted by CK, with input from MS, CM, GH, AC, RG, MF and HM. The enhanced GUMCAD pilot was overseen by GH, HM, JW and MF. The design of the prospective external validation was done by CK, MS, RG, GH, CM and AC and overseen by CK, MS and RG. Data management and analysis for algorithm development and validation was conducted by CK. The manuscript was drafted by CK, with significant input from CM and GH. All authors read and commented on the final manuscript.

## Appendix A. Supplementary Data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.eclinm.2018.11.002>.

## References

- [1] PHE. Sexually Transmitted Infections and Chlamydia Screening in England, 2016. PHE; 2017.
- [2] HPR. Sexually Transmitted Infections and Chlamydia Screening in England, 2012; 2013.
- [3] Johnson BT, Carey MP, Marsh KL, Levin KD, Scott-Sheldon LA. Interventions to reduce sexual risk for the human immunodeficiency virus in adolescents, 1985–2000: a research synthesis. *Arch Pediatr Adolesc Med* 2003;157(4):381–8.
- [4] Kalichman SC, Carey MP, Johnson BT. Prevention of sexually transmitted HIV infection: a meta-analytic review of the behavioral outcome literature. *Ann Behav Med* 1996;18(1):6–15.
- [5] Havinghurst RJ. Human Development and Education. New York: Longmans Green and Co.; 1952.
- [6] Long L, Abraham C, Paquette R, et al. Brief interventions to prevent sexually transmitted infections suitable for in-service use: a systematic review. *Prev Med* 2016; 91:364–82.
- [7] The King's Fund. Understanding NHS Financial Pressures: How Are They Affecting Patient Care?; 2017.
- [8] Iacobucci G, Torjesen I. Cuts to sexual health services are putting patients at risk, says King's Fund. *BMJ* 2017;356.
- [9] Moons KGM, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ* 2009;338.
- [10] Kannel WB, McGee D, Gordon T. A general cardiovascular risk profile: the Framingham study. *Am J Cardiol* 1976;38(1):46–51.
- [11] Dombrowski JC, Golden MR. Modernizing operations to improve efficiency and refine the role and mission of STI clinics. *Sex Transm Dis* 2013;40(1):81–4.
- [12] Robinson AJ, Rogstad K. Modernization in GUM/HIV services: what does it mean? *Int J STD AIDS* 2003;14(2):89–98.
- [13] Jones R, Menon-Johansson A, Waters AM, Sullivan AK. eTriage – a novel, web-based triage and booking service: enabling timely access to sexual health clinics. *Int J STD AIDS* 2010;21(1):30–3.
- [14] Handy P, Pattman R. Triage up front. *Sex Transm Infect* 2005;81(1):59–62.
- [15] Fairley CK, Williams H, Lee DM, Cummings R. A plea for more research on access to sexual health services. *Int J STD AIDS* 2007;18(2):75–6.
- [16] PHE. Genitourinary medicine clinic activity dataset (GUMCADv2). Available from: <https://www.gov.uk/guidance/genitourinary-medicine-clinic-activity-dataset-gumcadv2>; 2013.
- [17] Savage EJ, Mohammed H, Leong G, Duffell S, Hughes G. Improving surveillance of sexually transmitted infections using mandatory electronic clinical reporting: the genitourinary medicine clinic activity dataset, England, 2009 to 2013. *Eurosurveillance* 2014;19(48).
- [18] Mohammed H, Nardone A, Gilbert V, Desai S, Hughes G. Monitoring sti risk behaviour and partner notification outcomes through routine national surveillance: a pilot study in England. *Sex Transm Dis* 2014;41(Suppl. 1).



- [19] PHE. Genitourinary medicine clinic activity dataset (GUMCADv3) pilot. Available from: <https://www.gov.uk/guidance/genitourinary-medicine-clinic-activity-dataset-gumcadv3-pilot>; 2015.
- [20] Brook G, Bacon L, Evans C, et al. UK national guideline for consultations requiring sexual history taking. *Int J STD AIDS* 2013;2013.
- [21] Chambers LC, Manhart LE, Katz DA, Golden MR, Barbee LA, Dombrowski JC. Comparison of algorithms to triage patients to express care in a sexually transmitted disease clinic. *Sex Transm Dis* 2018;45(10):696–702.
- [22] Wand H, Bryant J, Pitts M, et al. Development of a risk algorithm to better target STI testing and treatment among Australian Aboriginal and Torres Strait Islander people. *Arch Sex Behav* 2017;46(7):2145–56.
- [23] Li Z, Liu H, Tu W. A sexually transmitted infection screening algorithm based on semiparametric regression models. *Stat Med* 2015;34(20):2844–57.
- [24] Ministry of Housing CaLG. The English Indices of Deprivation 2010; 2011.
- [25] Steyerberg EW, Harrell FE, Borsboom GJJM, Eijkemans MJC, Vergouwe Y, Habbema JDF. Internal validation of predictive models. *J Clin Epidemiol* 2001;54(8):774–81.
- [26] Moons KGM, Kengne AP, Woodward M, et al. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. *Heart* 2012;98(9):683–90.
- [27] Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S. A comparison of goodness-of-fit tests for the logistic regression model. *Stat Med* 1997;16(9):965–80.
- [28] Kramer A, Zimmerman J. Assessing the calibration of mortality benchmarks in critical care: the Hosmer-Lemeshow test revisited. *Crit Care Med* 2007;35(9):2052–6.
- [29] Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* 2010;21(1):128–38.
- [30] Aylin P, Bottle A, Majeed A. Use of administrative data or clinical databases as predictors of risk of death in hospital: comparison of models. *BMJ* 2007;334(7602):1044.
- [31] Moons KGM, Kengne AP, Grobbee DE, et al. Risk prediction models: II. External validation, model updating, and impact assessment. *Heart* 2012;98(9):691–8.
- [32] Kang M, Rochford A, Skinner SR, et al. Sexual behaviour, sexually transmitted infections and attitudes to chlamydia testing among a unique national sample of young Australians: baseline data from a randomised controlled trial. *BMC Public Health* 2014;14.
- [33] Fairbairn AP, Tyler H, Su JY, Tilley EL. Risk factors and associations for the diagnosis of sexually transmitted infections in Aboriginal women presenting to the Alice Springs Hospital emergency department. *Emerg Med Australas* 2010;22(3):216–23.
- [34] Hwang LY, Ma YF, Moscicki AB. Biological and behavioral risks for incident *Chlamydia trachomatis* infection in a prospective cohort. *Obstet Gynecol* 2014;124(5):954–60.
- [35] Fenton KA, Mercer CH, McManus S, et al. Ethnic variations in sexual behaviour in Great Britain and risk of sexually transmitted infections: a probability survey. *Lancet* 2005;365(9466):1246–55.
- [36] 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(1):41–6.
- [37] Coleman L, Testa A. Sexual health knowledge, attitudes and behaviours among an ethnically diverse sample of young people in the UK. *Health Educ J* 2007;66(1):68–81.
- [38] Miller GC, McDermott R, McCulloch B, Fairley CK, Muller R. Predictors of the prevalence of bacterial STI among young disadvantaged Indigenous people in north Queensland, Australia. *Sex Transm Infect* 2003;79(4):332–5.
- [39] Hillis SD, Nakashima A, Marchbanks PA, Addiss DG, Davis JP. Risk-factors for recurrent *Chlamydia trachomatis* infections in women. *Am J Obstet Gynecol* 1994;170(3):801–6.
- [40] Kim AA, Kent CK, Klausner JD. Risk factors for rectal gonococcal infection amidst resurgence in HIV transmission. *Sex Transm Dis* 2003;30(11):813–7.
- [41] Hirschfield S, Remien RH, Walavalkar I, Chiasson MA. Crystal methamphetamine use predicts incident STD infection among men who have sex with men recruited online: a nested case-control study. *J Med Internet Res* 2004;6(4):42–9.
- [42] Mayer KH, O’Cleirigh C, Skeer M, et al. Which HIV-infected men who have sex with men in care are engaging in risky sex and acquiring sexually transmitted infections: findings from a Boston Community Health Centre. *Sex Transm Infect* 2010;86(1):66–70.
- [43] Imrie J, Lambert N, Mercer CH, et al. Refocusing health promotion for syphilis prevention: results of a case-control study of men who have sex with men on England’s south coast. *Sex Transm Infect* 2006;82(1):80–3.
- [44] Wilkinson A, El-Hayek C, Fairley CK, et al. Incidence and risk factors associated with chlamydia in men who have sex with men: a cohort analysis of Victorian Primary Care Network for Sentinel Surveillance data. *Sex Transm Infect* 2012;88(5):319–24.
- [45] Aicken CRH, Armstrong NT, Cassell JA, et al. Barriers and opportunities for evidence-based health service planning: the example of developing a Decision Analytic Model to plan services for sexually transmitted infections in the UK. *BMC Health Serv Res* 2012;12:202.
- [46] Richens J, Copas A, Sadiq ST, et al. A randomised controlled trial of computer-assisted interviewing in sexual health clinics. *Sex Transm Infect* 2010;86(4):310–4.
- [47] Hughes G, Field N. The epidemiology of sexually transmitted infections in the UK: impact of behavior, services and interventions. *Future Microbiol* 2015;10(1):35–51.
- [48] Bourne A, Weatherburn P. Substance use among men who have sex with men: patterns, motivations, impacts and intervention development need. *Sex Transm Infect* 2017;93(5):342–6.
- [49] Cleary M, O’Sullivan J. P145 London sexual health transformation programme. *Sex Transm Infect* 2017;93(Suppl. 1):A65.