

This version of the article “Falasinnu, T., Gilbert, M., Gustafson, P., & Shoveller, J. (2015). A validation study of a clinical prediction rule for screening asymptomatic chlamydia and gonorrhoea infections among heterosexuals in British Columbia. Sexually Transmitted Infections. doi: 10.1136/sextrans-2014-051992. Epub ahead of print.” is not the final version.

We acknowledge the publisher’s (BMJ) copyright.

Here is a link to the final published version of the article:  
<http://www.ncbi.nlm.nih.gov/pubmed/25933609>

## ORIGINAL ARTICLE

# A validation study of a clinical prediction rule for screening asymptomatic chlamydia and gonorrhoea infections among heterosexuals in British Columbia

Titilola Falasinnu,<sup>1</sup> Mark Gilbert,<sup>2</sup> Paul Gustafson,<sup>3</sup> Jean Shoveller<sup>1</sup>

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/sextrans-2014-051992>).

<sup>1</sup>The School of Population and Public Health, University of British Columbia, Vancouver, British Columbia, Canada

<sup>2</sup>British Columbia Center for Disease Control, Vancouver, British Columbia, Canada

<sup>3</sup>The Department of Statistics, University of British Columbia, Vancouver, British Columbia, Canada

## Correspondence to

Dr Titilola Falasinnu, The School of Population and Public Health, University of British Columbia, 2206 East Mall, Vancouver, British Columbia, Canada BC V6T 1Z3; [lola.falasinnu@ubc.ca](mailto:lola.falasinnu@ubc.ca)

Received 23 December 2014

Revised 25 March 2015

Accepted 11 April 2015

## ABSTRACT

**Background** One component of effective sexually transmitted infections (STIs) control is ensuring those at highest risk of STIs have access to clinical services because terminating transmission in this group will prevent most future cases. Here, we describe the results of a validation study of a clinical prediction rule for identifying individuals at increased risk for chlamydia and gonorrhoea infection derived in Vancouver, British Columbia (BC), against a population of asymptomatic patients attending sexual health clinics in other geographical settings in BC.

**Methods** We examined electronic records (2000–2012) from clinic visits at seven sexual health clinics in geographical locations outside Vancouver. The model's calibration and discrimination were examined by the area under the receiver operating characteristic curve (AUC) and the Hosmer–Lemeshow (H-L) statistic, respectively. We also examined the sensitivity and proportion of patients that would need to be screened at different cut-offs of the risk score.

**Results** The prevalence of infection was 5.3% (n=10 425) in the geographical validation population. The prediction rule showed good performance in this population (AUC, 0.69; H-L p=0.26). Possible risk scores ranged from –2 to 27. We identified a risk score cut-off point of ≥8 that detected cases with a sensitivity of 86% by screening 63% of the geographical validation population.

**Conclusions** The prediction rule showed good generalisability in STI clinics outside of Vancouver with improved discriminative performance compared with temporal validation. The prediction rule has the potential for augmenting triaging services in STI clinics and enhancing targeted testing in population-based screening programmes.

## BACKGROUND

The imperative to provide more efficient sexual health services by public health programmes has led to the development of service models that optimise the use of health human resources, such as internet-based sexually transmitted infections (STIs) testing and triage services.<sup>1–4</sup> The aim of optimising service provision could be facilitated by the use of risk estimation algorithms. In a previous paper,<sup>5</sup> a risk estimation algorithm for optimising asymptomatic chlamydia and gonorrhoea case finding was derived using electronic medical records of patient visits at two sexual health clinics in Vancouver,

British Columbia (BC). This algorithm combines five risk factors: younger age, non-white race/ethnicity, multiple sexual partners, previous chlamydia diagnosis and previous gonorrhoea diagnosis. The prediction rule will eventually be adapted into a tool for facilitating selective screening in Get Checked Online (GCO), a novel internet-based testing programme in BC.<sup>4 6</sup>

In the derivation stage, we specified that the main intended application of the risk score is to help deal with the increasing numbers of people accessing services in the sexual health clinic contexts, particularly the asymptomatic individuals. However, a prediction rule's accuracy in one context may vary from the performance estimates reported in the derivation stage.<sup>7 8</sup> Inconsistent performance may reflect artificial disparities (eg, different study contexts), potentially in combination with genuine disparities (eg, distribution of risk factors).<sup>9</sup> Thus, it is important to consider the prediction rule's accuracy in varied settings, such as general practice clinics or hospitals, different types of hospitals or the same type of clinical setting in different geographical locations.<sup>9</sup>

Before making recommendations regarding the widespread use of the risk estimation algorithm derived through our previous work,<sup>5</sup> we conducted an external validation study in an independent setting to further test the parameters of the tool. It is essential to confirm that the algorithm is generalisable to a plausibly related setting (in addition to the previous comparisons conducted with the derivation population) that reflects the level of heterogeneity that will be encountered in real-life applications of the algorithm.<sup>10</sup> As was demonstrated previously, the algorithm showed reasonable discrimination and calibration upon validation in two different time periods (ie, temporal validation).<sup>11</sup> While temporal validation is often cited as the first step in demonstrating the transferability of a prediction algorithm,<sup>12</sup> it cannot assess the utility of the algorithm to other clinics or cities.<sup>10</sup> Geographical validation provides a more rigorous proof of validation than temporal validation owing to the hypothesised differences in patient mix, risk factor definitions and disease prevalence.<sup>10</sup>

It should be appreciated that a prediction rule's performance is often lower upon external validation.<sup>13</sup> In these settings, decision-makers have the option to adopt a previously derived prediction rule if it is found to perform adequately or they can derive or remodel a new prediction rule using



**To cite:** Falasinnu T, Gilbert M, Gustafson P, et al. *Sex Transm Infect* Published Online First: [please include Day Month Year] doi:10.1136/sextrans-2014-051992

their own population. Here, we describe the results of a test to assess the generalisability of an algorithm derived in Vancouver against a population of patients attending sexual health clinics in seven other geographical settings in BC. In addition, we examine the implications of deriving a new prediction rule using the data from the geographical validation population (ie, remodelled prediction rule) and compare its performance to the 'Vancouver' prediction rule.

## METHODS

The geographical validation dataset was derived from electronic medical records of clients attending publicly funded sexual health clinics located in seven locations in BC between 2000 and 2012: Penticton, Kelowna, Kamloops, New Westminster, Boundary, Courtenay and Prince George (see map in [online supplementary figure 1](#)). This analysis was limited to clinic visits among asymptomatic women and heterosexual men who are not sexual contacts of STI cases and not receiving confirmatory positive testing. The aim was to estimate the risk of chlamydia and/or gonorrhoea infection. In the current paper, the original model, regression coefficients and the simplified risk scores derived from the Vancouver clinic data are applied to a *geographical validation population*.<sup>10</sup> A detailed description of the study protocol, analysis plan and predictor definitions has been previously published.<sup>14</sup>

The 'Vancouver' risk estimation algorithm uses a logistic regression formula to relate its five predictors to chlamydia or gonorrhoea risk. The regression coefficients and their associated scoring points are listed in [table 1](#). Multiple imputation methods (five rounds) were used to impute missing values in the geographical validation population.<sup>15–17</sup> This analysis imputed missing values using IVEware, a software application that performs multiple imputations of missing values using the Sequential Regression Imputation Method.<sup>17</sup> All predictors and the outcome variables were included in the imputation model and the results of the five imputed datasets were combined to

obtain final estimates.<sup>15–16</sup> To assess the performance of the model in the geographical validation population, the model's discrimination was estimated by calculating the area under the receiver operating characteristic curve (AUC). The AUC gives the likelihood that a randomly selected infected individual would have a higher model predicted probability of chlamydia or gonorrhoea infection than a randomly selected non-infected individual.<sup>18</sup> The closer an AUC is to 100%, the better the model.<sup>18</sup>

Calibration was assessed with the Hosmer–Lemeshow goodness-of-fit statistic, which investigates (under the null hypothesis that there is no difference) the difference between the model predictions and the actual observations using deciles of predicted probabilities to categorise patients.<sup>18</sup> A p value >0.05 indicates a good fit.<sup>18</sup> The model's calibration was also examined by graphically plotting the prevalence of chlamydia and/or gonorrhoea infection in groups of the simplified risk scores. To aid population-based screening decision-making, a risk score was derived for each clinic visit in the geographical validation population by adding up the scoring points derived from [table 1](#). An evaluation of the sensitivity (or the fraction of infected cases identified) and the proportion of the population that would be screened at different risk score cut points was also performed. A well-performing screening tool detects >90% of cases, while screening <60% of the population.<sup>19</sup>

## RESULTS

During the years 2000–2012, there were 10 425 patient visits that met the inclusion criteria at sexual health clinics at the following geographical sites: Penticton, Kelowna, Kamloops, New Westminster, Boundary, Courtenay and Prince George. [Online supplementary figure 2](#) is a flow chart showing the selection of clinic visits whose data comprised this validation study. The prevalence of chlamydia and/or gonorrhoea infection was 5.3% (higher than the derivation population). [Table 2](#) shows the distribution of the baseline characteristics of patient visits.

**Table 1** Prediction rule for quantifying the probability of asymptomatic chlamydia and/or gonorrhoea infection among heterosexuals in Vancouver, British Columbia

Variable	aOR (95% CI)	Beta	Scoring points
Intercept	–	–6.0381	–
Age (years)			
14–19	5.25 (2.82 to 9.79)	1.6589	8
20–24	1.92 (1.19 to 3.10)	0.6544	3
25–29	1.20 (0.74 to 1.94)	0.1794	1
30–39	0.71 (0.42 to 1.18)	–0.3471	–2
≥40	Ref	Ref	Ref
Race/ethnicity			
White	Ref	Ref	Ref
Non-white	2.74 (2.03 to 3.70)	1.0093	5
No. of sexual partners in previous six months			
0	Ref	Ref	Ref
1–2	2.90 (0.97 to 8.75)	1.0565	5
≥3	3.03 (0.99 to 9.30)	1.1000	6
Previous chlamydia diagnosis			
Yes	4.34 (3.19 to 5.89)	1.4574	7
No	Ref	Ref	Ref
Previous gonorrhoea diagnosis			
Yes	1.27 (0.70 to 2.34)	0.2426	1
No	Ref	Ref	Ref

aOR, adjusted OR.

**Table 2** Population characteristics of heterosexual visits at sexual health clinics in the derivation and geographical validation populations

Variable	Derivation population, 2000–2006		Geographical validation population, 2000–2012	
	N	%	N	%
Chlamydia or gonorrhoea case				
Yes	184	1.8	556	5.3
No	10 253	98.2	9869	94.7
Gender				
Female	3496	33.5	4431	42.5
Male	6941	66.5	5994	57.5
Age (years)				
14–19	257	2.5	1748	16.8
20–24	1962	18.8	2922	28.0
25–29	2651	25.4	1969	18.9
30–39	3181	30.5	1850	17.7
≥40	2386	22.9	1935	18.6
Race/ethnicity				
White	7732	74.1	7741	74.3
Non-white	2705	25.9	2684	25.7
No. of sexual partners in previous six months				
0	644	6.2	590	5.7
1–2	6857	65.7	7004	67.2
≥3	2936	28.1	2830	27.1
Condom use				
Never	2362	22.6	2650	25.4
Always	5269	50.5	4489	43.1
Sometimes	2806	26.9	3286	31.5
Sex with partners recruited online				
Yes	417	4.0	341	3.3
No	10 020	96.0	10 083	96.7
Injection drug use				
Yes	211	2.0	351	3.4
No	10 226	98.0	10 073	96.6
Sex with injection drug user				
Yes	455	4.4	407	3.9
No	9983	95.6	10 018	96.1
Sex with commercial sex worker				
Yes	1381	13.2	826	7.9
No	9057	86.8	9599	92.1
Previous chlamydia diagnosis				
Yes	1518	14.5	1660	15.9
No	8919	85.5	8765	84.1
Previous gonorrhoea diagnosis				
Yes	619	5.9	310	3.0
No	9819	94.1	10 115	97.0
Total	10 437	100.0	10 425	100.0

\*Imputed values.

The derivation population (Vancouver) is included for comparison. The majority of patient visits in the geographical validation population had the following demographic characteristics: male gender (57.5%), aged between 20 and 24 years (28.0%), and white ethnicity (74.3%). More than two-thirds of patient visits reported having 1–2 sexual partners in the previous six months and approximately 43% reported consistent condom use. Approximately 3% of patient visits reported injection drug use and the same proportion reported having sex with partners recruited online. Previous chlamydia diagnosis was reported among nearly 16% of patient visits.

The geographical validation population differed from the derivation population by having a higher proportion of the

following characteristics: women, younger individuals, inconsistent condom use and injection drug use (table 2). There were also some differences between the derivation and geographical validation populations in terms of the unadjusted ORs examining the associations between the predictors and the outcome. Gender and condom use were significantly associated with infection in the geographical validation population—associations that were not significant in the derivation population (table 3). Race/ethnicity was not significantly associated with the outcome in the geographical validation population unlike the derivation population (table 3).

The Vancouver risk model demonstrated good discrimination in the geographical validation population. The AUC in the

**Table 3** Chlamydia and/or gonorrhoea prevalence and unadjusted ORs (derivation and geographic validation populations)\*

Variable	Derivation population (N=10 437)		Geographical validation population (N=10 425)	
	%	OR (95% CI)	%	OR (95% CI)
Gender				
Female	2.1	1.31 (0.97 to 1.77)	6.8	1.67 (1.40 to 1.98)
Male	1.6	Ref	4.2	Ref
Age (years)				
14–19	7.4	6.49 (3.58 to 11.75)	9.7	9.47 (6.00 to 14.94)
20–24	2.8	2.30 (1.46 to 3.63)	7.7	7.40 (4.72 to 11.58)
25–29	1.8	1.50 (0.94 to 2.38)	4.9	4.55 (2.82 to 7.34)
30–39	1.1	0.88 (0.53 to 1.45)	2.4	2.15 (1.27 to 3.62)
≥40	1.2	Ref	1.1	Ref
Race/ethnicity				
White	1.2	Ref	5.2	Ref
Non-white	3.4	2.90 (2.16 to 3.89)	5.7	1.10 (0.78 to 1.55)
No. of sexual partners in previous six months				
0	0.5	Ref	2.0	Ref
1–2	1.8	3.43 (1.10 to 10.73)	4.9	2.50 (1.26 to 4.98)
≥3	2.0	3.92 (1.23 to 12.45)	7.1	3.71 (1.83 to 7.55)
Condom use				
Never	1.6	Ref	4.5	Ref
Sometimes	2.0	1.22 (0.84 to 1.77)	6.0	1.36 (1.07 to 1.72)
Always	1.4	0.88 (0.56 to 1.39)	5.0	1.11 (0.85 to 1.46)
Sex with partners recruited online				
Yes	1.2	0.67 (0.27 to 1.63)	2.7	0.48 (0.25 to 0.94)
No	1.8	Ref	5.4	Ref
Injection drug use				
Yes	1.6	0.89 (0.26 to 3.06)	7.5	1.45 (0.81 to 2.58)
No	1.8	Ref	5.3	Ref
Sex with injection drug user				
Yes	1.0	0.55 (0.20 to 1.54)	6.6	1.27 (0.78 to 2.07)
No	1.8	Ref	5.3	Ref
Sex with commercial sex worker				
Yes	1.2	0.65 (0.35 to 1.20)	2.7	0.48 (0.29 to 0.79)
No	1.8	Ref	5.6	Ref
Previous chlamydia diagnosis				
Yes	5.1	4.40 (3.27 to 5.93)	11.7	3.10 (2.58 to 3.72)
No	1.2	Ref	4.1	Ref
Previous gonorrhoea diagnosis				
Yes	2.1	1.21 (0.69 to 2.14)	4.2	0.77 (0.44 to 1.36)
No	1.7	Ref	5.4	Ref
Total	1.8	–	5.3	–

\*Missing data among the predictor variables were handled using a multiple imputation procedure with five resampling replications, which generated an augmented database with (5 × 10 437) 52 185 and (5 × 10 425) 52 125 observations with complete data in the derivation and temporal validation populations, respectively. With the imputed sample, we estimated baseline characteristics and developed prediction models. The average of all five imputed samples is shown in this table.

geographical validation population was 0.69, 95% CI 0.67 to 0.71, while the AUC in the derivation population was 0.74, 95% CI 0.70 to 0.77 (online supplementary figure 3). A *p* value of 0.26 for the Hosmer–Lemeshow goodness-of-fit test also indicated good calibration. Online supplementary figure 4 shows the calibration in the geographic validation population was good as the prevalence of chlamydia and/or gonorrhoea infection increased with increasing risk score, which ranged from 0.2% in the lowest risk score category to 23.7% in the highest risk category. We also explored the use of the risk score for selective screening (table 4). This analysis identified a risk score cut-off level of ≥6 points that would identify approximately 95% of infections while screening 78% of the geographical validation population. In the derivation population, the same risk score cut-off of ≥6 points identified 91% of cases and the fraction screened was 68% of the population.

## CONCLUSIONS

Validation studies aim to provide evidence that a risk scoring algorithm can be generalised to new populations. The ‘Vancouver’ risk estimation tool showed slightly better discrimination in the geographical validation population (AUC=0.69) than in the temporal validation population (AUC=0.64). The risk estimation tool performed well in the geographical validation population despite the fact that the geographical validation population differed from the derivation population regarding some predictors (eg, age, condom use and previous infection). The geographic validation regions have higher rates of chlamydia and gonorrhoea infection and dissimilar STI epidemiology and social determinants of sexual health;<sup>21–23</sup> the temporal validation population also was less heterogeneous than the derivation and geographical validation populations. Further

**Table 4** Sensitivity and specificity of cut-off scores in the derivation and geographical validation populations

Score cut-off	Derivation population (%)				Geographical validation population (%)			
	Sensitivity	Specificity	Fraction screened	PPV	Sensitivity	Specificity	Fraction screened	PPV
≥-2	100.0	0.0	100.0	1.8	100.0	0.0	100.0	5.3
≥-1	100.0	1.2	98.8	1.8	100.0	0.7	99.3	5.4
≥0	100.0	1.3	98.7	1.8	100.0	0.7	99.3	5.4
≥1	100.0	2.6	97.4	1.8	99.9	1.7	98.4	5.4
≥2	99.5	3.7	96.4	1.8	99.9	2.5	97.7	5.5
≥3	99.5	3.7	96.4	1.8	99.9	2.5	97.7	5.5
≥4	96.7	16.7	83.5	2.0	97.1	10.8	89.7	5.8
≥5	95.8	22.2	78.1	2.2	96.2	13.7	86.9	5.9
≥6	91.2	32.7	67.8	2.4	94.5	22.5	78.4	6.4
≥7	84.9	47.7	52.8	2.8	90.1	34.1	67.2	7.2
≥8	82.2	53.0	47.7	3.0	86.0	37.9	63.3	7.2
≥9	72.2	65.2	35.4	3.6	71.6	52.9	48.4	7.9
≥10	67.3	70.6	30.1	3.9	64.0	58.9	42.3	8.1
≥11	62.5	74.8	25.8	4.3	61.8	63.0	38.4	8.6
≥12	52.6	81.4	19.2	4.8	59.3	67.4	34.1	9.3
≥13	47.7	84.5	16.1	5.2	57.6	69.5	31.9	9.6
≥14	35.1	91.1	9.3	6.6	41.5	83.1	18.2	12.1
≥15	32.4	94.2	6.3	9.1	32.4	88.2	12.9	13.4
≥16	25.9	95.9	4.5	10.1	25.6	91.1	9.8	13.9
≥17	21.2	96.9	3.4	10.9	22.4	92.4	8.4	14.3
≥18	19.9	97.2	3.1	11.4	21.8	92.8	8.0	14.5
≥19	14.3	98.5	1.8	14.4	17.4	95.9	4.9	19.1
≥20	11.6	99.0	1.2	16.9	14.8	96.9	3.7	21.2
≥21	7.3	99.5	0.6	20.9	8.1	98.5	1.8	23.7
≥22	2.8	99.9	0.2	28.0	3.5	99.4	0.7	26.1
≥23	2.8	99.9	0.1	33.8	3.4	99.5	0.7	26.3
≥24	2.8	99.9	0.1	33.8	3.4	99.5	0.7	26.3
≥25	2.8	99.9	0.1	33.8	3.4	99.5	0.7	26.3
≥26	0.7	100.0	0.0	30.0	1.5	99.8	0.3	28.1
≥27					0.1	100.0	0.0	18.2

analysis also revealed that a remodelled algorithm using data from the geographical validation population performed no better than the 'Vancouver' prediction algorithm as indicated by the non-significant  $\chi^2$  statistic testing the difference between the AUCs. These findings provide strong evidence that the risk score is robust and valid and likely has generalisable discrimination and calibration in varied settings.

In the geographical validation population, choosing the 'Vancouver' cut-off point of  $\geq 6$  would require screening 78% of the population to find 95% of the cases and equates to a reduction of 22% in the number of individuals that would need to be screened. However, using the  $\geq 6$  cut-off point would fail to meet the efficiency benchmark of screening <60% of the population. Increasing the cut-off point to  $\geq 7$  would require screening 67% of the population to achieve a sensitivity of 90%, which would be closer to the efficiency benchmark. In applying the prediction rule to a population with a higher prevalence of infection and risk behaviours (ie, more severe case mix) such as the geographical validation population, it is expected that choosing a higher cut-off point will increase the efficiency of screening decision-making.<sup>24 25</sup>

Alternatively in this setting, applying the age-based screening criterion (ie, age <25 years) to the geographical validation population would require screening of 45% of the population but would only detect 71% of the cases, a performance that falls short of the screening benchmark. However, increasing the

cut-off to age <30 years would require screening 64% of the population with a sensitivity of 88%, a performance that is close to the benchmark and also similar to using the cut-off point of  $\geq 7$ . This finding suggests that using age alone could be a viable option as a screening criterion in the geographical validation population, but not in the derivation population. This finding was not surprising because the distribution of age in the geographical validation population was more heterogeneous than in the derivation population—a situation that often leads to better discrimination and optimum screening performance.<sup>25</sup> However, age-based screening criterion may be contraindicated in settings where a majority of the population presenting for screening is comprised of younger individuals (eg, youth clinics). Also, if the prediction rule is found to perform less than adequately in these settings, universal screening may be a more suitable alternative.

Several studies have established the validity of prediction rules as screening criteria, especially where chlamydia and/or gonorrhoea prevalence is low (ie, <2%), as is the case in the derivation population used in this study.<sup>26</sup> Although it has been suggested that universal screening or using criteria based on age would be cost effective in settings with prevalence of 2% or more, publicly funded sexual health services in these settings are constrained by available funding and limited resources. In considering such practical constraints within the BC, we suggest a cautious approach to such global screening approaches.<sup>26</sup>

Specifically, screening women <25 years old could prove to be cost prohibitive in settings where individuals in this age group comprise the highest proportion of clinic visits.<sup>27</sup> Caution should also be exercised before using age as a criterion in internet-based testing scenarios such as GCO where good calibration (and not just discrimination) would be essential. In these scenarios, the risk score categories and their associated prevalence prove more useful than age-based criterion for patients trying to decide whether to take the STI test. This is because the process of moving from screening criteria that focus on specific risk factors (eg, age) to prediction rules acknowledges a more comprehensive risk spectrum.<sup>28</sup>

The findings of this analysis also were compared with other external validation studies that examined the validity of previously derived clinical prediction rules (CPRs) in new geographic settings.<sup>24–29</sup> Gotz and colleagues derived a prediction rule for chlamydia infection for the selective screening of high-risk individuals in Rotterdam, the Netherlands.<sup>24</sup> The prediction rule showed fair external validity in two independent settings: a population-based study in Amsterdam and an outreach screening project among high-risk youth in Rotterdam. The AUC was 0.79 (95% CI 0.76 to 0.84) in the derivation sample, 0.66 (95% CI 0.58 to 0.74) in the Amsterdam sample and 0.68 (95% CI 0.58 to 0.79) in the Rotterdam sample.<sup>24</sup> A second study by Haukoos *et al*<sup>29</sup> derived and validated an algorithm to accurately identify patients at risk for HIV infection, using patient data from an STI clinic in Denver, Colorado (1996–2008). Validation was performed using an independent population from an urban emergency department in Cincinnati, Ohio. The results of the study showed that the risk score showed reasonable generalisability; the AUC was 0.85 (95% CI 0.83 to 0.88) in the derivation sample and 0.75 (95% CI 0.70 to 0.78) in the validation sample.<sup>29</sup>

The AUC of the 'Vancouver' risk estimation in the derivation population was lower compared with the AUCs of the two aforementioned studies.<sup>24–29</sup> This can be explained by the omission of symptoms in the 'Vancouver' risk estimation tool, which have been shown to be significantly associated with infection.<sup>30</sup> Specifically, unlike previous risk estimation tools in sexual health settings, the 'Vancouver' risk estimation tool was limited to asymptomatic patients, an important improvement as most STIs infrequently present with symptoms. The loss in discriminative ability between the derivation and validation populations in the other studies ranged (in absolute percentage points) from 10% points to 13% points compared with a loss of 10% points and 5% points in the temporal and geographical validation populations, respectively, in this analysis.

There were several strengths to the geographic validation process undertaken here, including the large overall population size, the independence of the clinicians in the geographical validation population from the derivation population, and the systematic analysis of its discrimination and calibration. The current study was the first to derive and validate a locally specific risk assessment tool to quantify STI risk in a Canadian setting. Risk assessment tools ideally should be derived from large representative samples.<sup>31</sup> This study included 13 years of electronic health records comprising 40 000 patient visits to publicly funded STI clinics in BC, representing a high percentage of the population of individuals using this service in the province. As with most administrative datasets, the dataset was not deliberately built for the derivation of risk algorithms, resulting in some missing information for several predictors, which we have attempted to mitigate through the use of imputation (something rarely done in prediction modelling studies).

The use of imputation techniques yielded discrimination and calibration performance measures similar to those of complete case analyses in which individuals with missing values on any of the considered variables were excluded and baseline analyses in which individuals with missing values on a variable were assumed to be in the lowest category (data not shown).<sup>15</sup> This finding suggests that the algorithm was valid despite the consequential risk factor misclassification associated with the data imputation process. Overall, however, the use of imputation techniques offers improved study power and limited bias in the estimated regression coefficients.<sup>15</sup>

Caution should be exercised in generalising the findings of this analysis to even more diverse geographic settings. Several additional analyses are recommended before the widespread implementation of the risk estimation algorithm. For example, the algorithm's screening performance could be prospectively verified in internet-based STI testing contexts. Furthermore, while the derivation and validation populations are an unbiased representation of STI clinic clients in BC, the current results might or might not be valid for other settings in BC (or other Canadian provinces, or even other global settings). It would be reasonable to argue that *STI clinic clients* also may vary significantly from *patients seeking care in primary care settings or online contexts*; and, therefore, the results of the current CPR should not be directly extrapolated to other settings without additional validation studies that could provide stronger evidence for the generalisability of the risk estimation algorithm.

In conclusion, a new era in evidence-based decision-making regarding STI testing and progress in relation to the adoption of prediction rules may be at hand. The advent of online approaches to risk estimation, the emergence of new statistical methods, as well as increasingly sophisticated theory, all reflect the potential to continue to make advanced in improving testing and treatment of STIs. To date, however, few prediction rules have been validated and, hence, the dissemination and usage of prediction rules in STI service provision remains in the nascent stages. The well-performing prediction rule derived and broadly validated here provides evidence that risk estimation tools have a place in sexual health service provision. New investments in research and practice are required to facilitate the effective integration of prediction rules into routine sexual health service provision and more attention should be paid to their scaling up and to the scientific evaluation of their effects over time.

### Key messages

- ▶ This article highlights the geographical validation of the 'Vancouver' risk estimation tool for screening asymptomatic chlamydia and gonorrhoea.
- ▶ The prediction tool showed adequate discrimination and calibration upon validation in seven clinics outside of Vancouver.
- ▶ These findings are encouraging and bolster confidence in recommending this tool for use in sexual health services and programmes.
- ▶ The risk score could be easily implemented and is accurate enough to convey important screening considerations.

**Handling editor** Jackie A Cassell

**Twitter** Follow Mark Gilbert at @mpjgilbert

769 **Contributors** TF was the lead investigator for the empirical work presented and  
770 responsible for all major areas of concept formation, data analysis, as well as  
771 manuscript composition. MG and PG were involved in the early stages of concept  
772 formation and contributed to manuscript edits. PG also contributed data analyses  
773 and interpretation. JS was the supervisory author on this project and was involved  
774 throughout the project in concept formation and manuscript composition.

775 **Funding** TF was supported by the Canadian Institutes of Health Research (CIHR)  
776 Doctoral Research Award. PG was supported by a grant from the Natural Sciences  
777 and Engineering Research Council of Canada.

778 **Competing interests** None declared.

779 **Ethics approval** University of British Columbia's Research Ethics Board (certificate  
780 # H11-02000).

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

## 782 REFERENCES

- 783 1 Chai SJ, Aumakhan B, Barnes M, et al. Internet-based screening for sexually  
784 transmitted infections to reach nonclinic populations in the community: Risk factors  
785 for infection in men. *Sex Transm Dis* 2010;37:756–63.
- 786 2 Gaydos CA, Dwyer K, Barnes M, et al. Internet-based screening for chlamydia  
787 trachomatis to reach non-clinic populations with mailed self-administered vaginal  
788 swabs. *Sex Transm Dis* 2006;33:451–7.
- 789 3 Shamos SJ, Mettenbrink CJ, Subiadur JA, et al. Evaluation of a testing-only  
790 "express" visit option to enhance efficiency in a busy STI clinic. *Sex Transm Dis*  
791 2008;35:336–40.
- 792 4 Hottes TS, Farrell J, Bondyra M, et al. Internet-based HIV and sexually transmitted  
793 infection testing in british columbia, canada: opinions and expectations of  
794 prospective clients. *J Med Internet Res* 2012;14:e41.
- 795 5 Falasinnu T, Gilbert M, Gustafson P, et al. Deriving and validating a risk estimation tool  
796 for screening asymptomatic chlamydia and gonorrhoea. *Sex Transm Dis* 2014;41:706.
- 797 6 Gilbert M. Impact of an integrated internet-based testing program for HIV and  
798 sexually transmitted infections in british columbia. *CIHR's Partnership for Health  
799 System Improvement Grant*. 2010(350,000 CAD).
- 800 7 Knottnerus J, Buntinx F, eds. *The evidence base of clinical diagnosis: theory and  
801 methods of diagnostic research*. 2009.
- 802 8 D'Agostino RB Sr, Grundy S, Sullivan LM, et al. CHD Risk Prediction Group.  
803 Validation of the framingham coronary heart disease prediction scores: Results of a  
804 multiple ethnic groups investigation. *JAMA* 2001;286:180–7.
- 805 9 Siontis GC, Tzoulaki I, Siontis KC, et al. Comparisons of established risk prediction  
806 models for cardiovascular disease: Systematic review. *BMJ* 2012;344:e3318.
- 807 10 Moons KG, Kengne AP, Grobbee DE, et al. Risk prediction models: II. external  
808 validation, model updating, and impact assessment. *Heart* 2012;98:691–8.
- 809 11 Toll DB, Janssen KJ, Vergouwe Y, et al. Validation, updating and impact of clinical  
810 prediction rules: a review. *J Clin Epidemiol* 2008;61:1085–94.
- 811 12 Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic  
812 information. *Ann Intern Med* 1999;130:515–24.
- 813 13 Ivanov J, Tu JV, Naylor CD. Ready-made, recalibrated, or remodeled? issues in the  
814 use of risk indexes for assessing mortality after coronary artery bypass graft surgery.  
815 *Circulation* 1999;99:2098–104.

- 816 14 Falasinnu T, Gustafson P, Gilbert M, et al. Risk prediction in sexual health contexts:  
817 Protocol. *JMIR Res Protoc* 2013;2:e57.
- 818 15 Vergouwe Y, Royston P, Moons KG, et al. Development and validation of a  
819 prediction model with missing predictor data: A practical approach. *J Clin Epidemiol*  
820 2010;63:205–14.
- 821 16 Janssen KJ, Vergouwe Y, Donders AR, et al. Dealing with missing  
822 predictor values when applying clinical prediction models. *Clin Chem*  
823 2009;55:994–1001.
- 824 17 Survey Research Center, Institute for Social Research, University of Michigan.  
825 IVEware: Imputation and Variance Estimation software. Updated 2013. <http://www.isr.umich.edu/src/smp/ive/> (accessed 8 jul, 2013).
- 826 18 Sanagou M, Wolfe R, Leder K, et al. External validation and updating of a  
827 prediction model for nosocomial pneumonia after coronary artery bypass graft  
828 surgery. *Epidemiol Infect* 2014;142:540–4.
- 829 19 La Montagne DS, Patrick LE, Fine DN, et al. Infertility Prevention Project.  
830 Re-evaluating selective screening criteria for chlamydial infection among women in  
831 the U S pacific northwest. *Sex Transm Dis* 2004;31:283–9.
- 832 20 DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or  
833 more correlated receiver operating characteristic curves: A nonparametric approach.  
834 *Biometrics* 1988;44:837–45.
- 835 21 Goldenberg S, Shoveller J, Koehoorn M, et al. Barriers to STI testing among youth  
836 in a canadian oil and gas community. *Health Place* 2008;14:718–29.
- 837 22 Goldenberg S, Shoveller J, Ostry A, et al. Youth sexual behaviour in a boomtown:  
838 Implications for the control of sexually transmitted infections. *Sex Transm Infect*  
839 2008;84:220–3.
- 840 23 Shoveller J, Johnson J, Rosenberg M, et al. Youth's experiences with STI testing  
841 in four communities in british columbia, canada. *Sex Transm Infect*  
842 2009;85:397–401.
- 843 24 Gotz HM, Veldhuijzen IK, Habbema JD, et al. Prediction of chlamydia trachomatis  
844 infection: Application of a scoring rule to other populations. *Sex Transm Dis*  
845 2006;33:374–80.
- 846 25 Ewout W. Steyerberg. *Clinical prediction models: a practical approach to  
847 development, validation, and updating*. New York: Springer, 2009.
- 848 26 Han Y, Coles FB, Hipp S. Screening criteria for chlamydia trachomatis in family  
849 planning clinics: accounting for prevalence and clients' characteristics. *Fam Plann  
850 Perspect* 1997;29:163–6.
- 851 27 Miller WC, Hoffman IF, Owen-O'Dowd J, et al. Selective screening for chlamydial  
852 infection: which criteria to use? *Am J Prev Med* 2000;18:115–22.
- 853 28 Eddy DM, Adler J, Patterson B, et al. Individualized guidelines: the potential for  
854 increasing quality and reducing costs. *Ann Intern Med* 2011;154:627–34.
- 855 29 Haukoos JS, Lyons MS, Lindsell CJ, et al. Derivation and validation of the denver  
856 human immunodeficiency virus (HIV) risk score for targeted HIV screening. *Am J  
857 Epidemiol* 2012;175:838–46.
- 858 30 Xu F, Stoner BP, Taylor SN, et al. "Testing-only" visits: an assessment of missed  
859 diagnoses in clients attending sexually transmitted disease clinics. *Sex Transm Dis*  
860 2013;40:64–9.
- 861 31 Wand H, Guy R, Donovan B, et al. Developing and validating a risk scoring tool for  
862 chlamydia infection among sexual health clinic attendees in australia: A simple  
863 algorithm to identify those at high risk of chlamydia infection. *BMJ Open* 2011;1:  
864 e000005.