INFORMING SELECTIVE SCREENING THROUGH MORE ROBUST ESTIMATION OF STI RISK

by

Titilola Oluwaseun Falasinnu

B.A., Hood College, 2003, *magna cum laude*
M.H.S., Johns Hopkins University, 2005

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Healthcare and Epidemiology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

October 2014

Abstract

Background. Due to rising health care costs and increases in client volumes, it is imperative to develop systems that make efficient use of increasingly scarce publicly funded sexual health resources. With this in mind, an internet-based testing program that provides online access to STI testing is in development in British Columbia (BC) to improve the accessibility and limit the burden on health resources. However, much is still unknown about how to implement risk assessment and recommend tests in online settings. Prediction rules have been shown to successfully increase efficiency and cost-effectiveness of STI case finding. The aim of this dissertation was to develop and validate a risk-scoring algorithm for the selective screening of asymptomatic patients at increased risk for chlamydia and gonorrhoea infections.

Methods. The risk-scoring algorithm was derived from a multivariate logistic regression of patient visits at two sexual health clinics in Vancouver between 2000 and 2006 (i.e., derivation population) and validated in a subsequent time period between 2007 and 2012 (i.e., temporal validation population). The model's performance was evaluated using the area under the receiver operating characteristic curve (AUC) and the Hosmer-Lemeshow (H-L) statistic. Geographical validation was performed using seven sexual health clinics outside of Vancouver between 2000 and 2012.

Results. The prevalence of infection was 1.8% (n=10,437), 2.2% (n=14,956), and 5.3% (n=10,425) in the derivation, temporal validation, and geographical validation populations, respectively. The predictors that comprised the algorithm were young age, non-white race/ethnicity, multiple sexual partners, previous chlamydia or gonorrhoea infection. The model discriminative accuracy was good in the derivation population (AUC=0.74, 95% CI: 0.70-0.77) and acceptable in the temporal (AUC=0.64, 95% CI: 0.61-0.67) and geographical (AUC=0.69, 95% CI: 0.67-0.71) validation populations. The model also demonstrated adequate calibration and screening performance in all three populations.

Conclusions. The results from this research will have important implications for scaling up of Internet-based testing in BC. The algorithm could be adapted in an online setting to offer individualized testing recommendations and create educational materials to inform other Web-based content by creating awareness about STI risk factors, which may stimulate health care seeking behaviour among individuals accessing the website.

Preface

All of the work presented henceforth was conducted at the University of British Columbia, Point Grey Campus and at the British Columbia Centre for Disease Control. All projects and associated methods were approved by the University of British Columbia's Research Ethics Board (certificate # H11-02000).

A version of Chapter 2 has been accepted for publication [Falasinnu T, Gustafson P, Hottes TS, Gilbert M, Ogilvie G, Shoveller, J. Predictors identifying those at increased risk for STDs: a theory-guided review of empirical literature and screening recommendations. The International Journal of STD and AIDS. *In press*]. I was the lead investigator, responsible for all major areas of concept formation, search strategy, selection of studies for inclusion, literature synthesis, as well as manuscript composition. Gustafson P, Hottes TS, Gilbert M, and Ogilvie G were involved in the early stages of concept formation and contributed to manuscript edits. Shoveller, J was the supervisory author on this project and was involved throughout the project in concept formation and manuscript composition.

A version of Chapter 3 has been published [Falasinnu T, Gustafson P, Hottes TS, Gilbert M, Ogilvie G, Shoveller, J. A critical appraisal of risk models for predicting sexually transmitted diseases. Sex Transmitted Diseases 2014; May;41(5):321-30]. I was the lead investigator, responsible for all major areas of concept formation, search strategy, selection of studies for inclusion, literature synthesis, as well as manuscript composition. Gustafson P, Hottes TS, Gilbert M, and Ogilvie G were involved in the early stages of concept formation and contributed to manuscript edits. Shoveller, J was the supervisory author on this project and was involved throughout the project in concept formation and manuscript composition.

A version of Chapter 4 has been accepted for publication [Falasinnu T, Gilbert M, Gustafson P, Shoveller, J. Deriving and validating a risk estimation tool for screening asymptomatic chlamydia and gonorrhoea. Sex Transmitted Diseases. *In press*]. I was the lead investigator, responsible for all major areas of concept formation, data analysis, as well as, manuscript composition. Gilbert M and Gustafson P were involved in the early stages of concept formation and contributed to manuscript edits. Gustafson P also contributed data analyses and interpretation. Shoveller, J was the supervisory author on this project and was involved throughout the project in concept formation and manuscript composition.

A version of Chapter 5 has been submitted for publication [Falasinnu T, Gilbert M, Gustafson P, Shoveller, J. Geographical validation of a risk estimation algorithm for optimising chlamydia and gonorrhoea case finding. *Submitted*]. I was the lead investigator for the empirical work presented and responsible for all major areas of concept formation, data analysis, as well as, manuscript composition. Gilbert M and Gustafson P were involved in the early stages of concept formation and contributed to manuscript edits. Gustafson P also contributed data analyses and interpretation. Shoveller, J was the supervisory author on this project and was involved throughout the project in concept formation and manuscript composition.

Table of Contents

List of Tables

List of Figures

Dedication

For Mopelola, who taught me to relentlessly pursue knowledge

Chapter 1: Introduction

1.1 STIs – An important individual and public health problem

The 2012 Chief Public Health Officer's Report on the State of Public Health in Canada drew notable attention to the increasing rates of communicable diseases such as sexually transmitted infections (STI) including HIV and their threat to the health of individuals and populations [1]. STIs are significant sources of morbidity through their impact on reproductive and child health [2]. Undiagnosed and untreated infections often have long-term consequences such as pelvic inflammatory disease, ectopic pregnancy, and infertility [1]. In children born to infected mothers, bacterial STIs (such as chlamydia, gonorrhoea and syphilis) can cause serious conditions, such as low birth weight, pneumonia, and congenital syphilis [1]. STIs also impose a large burden of mortality indirectly through their role in facilitating the sexual transmission of HIV infection [1]. At the population level, untreated STIs can also act as reservoirs for future infection. Furthermore, the indirect costs (in unpaid leave and lost wages) and the intangible costs (e.g., emotional distress, intimate partner violence, shame, stigma, and pain) generate an overall human cost that may be difficult to appraise, but that exacts a significant toll.

These costs disproportionately impact sexually active youth and young adults as this group is well-represented [3] in the case reports for the four nationally reportable STIs in Canada: HIV, chlamydia, gonorrhoea and infectious syphilis – infections that have dissimilar epidemiological profiles [4]. Chlamydia is the most frequently detected and reported bacterial STI and chlamydial incidence has been steadily increasing in British Columbia (BC) since 1997 among all age groups [3]. In 2012, women aged 20-24 years experienced the highest rate of chlamydia in BC (1863.3 cases per 100,000 population), while women aged 15-19 years experienced the second highest chlamydia rate (1433.0 cases per 100,000 population) [3]. Among men, chlamydia rates for young men in BC were 351.9 cases per 100,000 population (men aged 15-19 years) and 810.8 cases per 100,000 population (men aged 20-24 years), compared with the BC population average of 192.3 per 100,000 population among all males in 2012 [3].

Gonorrhoea is the second most frequently diagnosed and reported bacterial STI [3]. There has been gradual rise in gonococcal infection rate since 1997, however, this rate has been relatively stable since 2007. The current incidence levels of this infection have been attributed to a network of

people with high-transmission activities [4]. Males comprise more than two-thirds of reported cases in BC in 2012 [3]. Young men aged 20-29 years and young women aged 15-24 years are the most affected by gonorrhoea [3]. For the last twenty years, public health officials have been concerned that gonorrhoea might become resistant to the last widely available antibiotics used to treat it, a class of drugs called quinolones [3]. Recent trends also indicate that gonorrhoea is now widely resistant to quinolones [4,5]. A study conducted in Ontario found that 28% of *N. gonorrhoeae* isolates were resistant to this class of antibiotics in 2006; in this setting, men who have sex with men (MSM) were significantly more likely to contribute to the quinolones resistance rate [6]. The continual monitoring of antimicrobial resistance trends is essential to ensure high treatment and cure rates for gonorrhoeal infection [7].

Infectious syphilis is the least frequently diagnosed of the nationally reportable bacterial STIs. Syphilis, which was previously rare in Canada, dramatically increased between 2000 and 2012 in males and females, but especially in males [3]. In 2012, 84% of the reported syphilis cases in BC were attributed to MSM transmission [3]. Recently, localised outbreaks of infectious syphilis have been reported in several geographical locations in Canada including Vancouver, Calgary, Montreal, and the Northwest Territories [4]. Most of the outbreaks have been among MSM and some outbreaks have been related to sex work [3].

There has been a steady decline in HIV rates in Canada and BC [8]. However, the HIV epidemic in Canada is complex, with differing infection rates in specific population sub-groups. MSM, injection drug users (IDUs), and adolescent females are high risk populations for infection [8]. MSM comprise the greatest number of new positive HIV tests per year in BC [8]. In 2012, youth (ages 25-29) experienced the highest rate of infection (22.0 cases per 100,000 population). Recent epidemiological trends indicate that HIV prevalence is low among Canadian youth, however, behavioural and STI data surveillance data clearly signify that the potential for HIV transmission remains considerable in this population [8], especially among young Canadians with the following characteristics: street-involvement, MSM, and IDU [9,10].

1.2 Publicly funded sexual health clinics

The determinants of STI epidemics include the three factors that define the rate of increase or basic reproduction rate (R_0) of STIs in populations: infectivity or the probability that an exposed person will acquire the infection (β), rate of exposure between susceptible individuals to infected

individuals (c), and the time that newly infected persons remain infectious (D), and the factors that affect those variables. The higher the value of R_0, the greater the potential for the STI to spread in a population. The three determinants (β, c, and D) are subject to different influences and immense heterogeneity across populations, communities and infection pathogen [11]. In this dissertation, the basic reproduction rate has two overarching functions to guide the understanding of STI risk on an individual (i.e., the cause of the case or the prevalent infection) and population-based (i.e., the cause of the incident cases) level [12,13]. The former function is the basis of the understanding of individualized risk assessment that is the theme of this dissertation. However, a detailed explanation of the population health impact of R_0 is warranted in order to define the broader epidemiological context of this dissertation.

Specifically, the three main strategies for effective population-based control of STIs emerged from an understanding of R_0 [2]. The first strategy is to reduce the risk of transmission during sexual contact through safer sex practices like condom use. The second strategy is to reduce the number of new sexual partners through effective behavioural change initiatives. The third strategy is to reduce the period of infectivity achieved through accessible sexual health clinical services, screening and treatment of individuals at increased risk, and effective contact tracing of STI cases. In many countries, publicly funded sexual health clinics are tasked with the third component of STI control and many of those provide no - or low-cost STI testing and treatment services. However, emerging budgetary limitations appear to be affecting public health infrastructure (even in many high-income countries), which have led to STI clinic closures and/or scale-back of operations in some jurisdictions [14,15]. For example, in BC, five STI clinics were closed in 2011 [16] and to further complicate the situation, in many jurisdictions across BC, STI clinics are operating at maximum capacity amidst rising STI rates [17]. Increased clinic patient volumes contribute to scenarios where clinicians are faced with providing more services with diminishing resources [15,18].

It is also important to appreciate the context of sexual health service provision, which is perhaps unique from other medical disciplines, in several ways. For example, sexual health service provision settings do not typically benefit from high profile public advocacy initiatives (e.g., few citizens organise and engage in letter writing campaigns demanding that government officials provide additional resources to address excessive wait-times or geographical barriers to STI testing) [18]. In fact, a significant proportion of clinic attendees prefer to test anonymously or pseudonymously. This field is one where stigma and popular misconceptions remain significant

barriers to care [18], and where there is little public appreciation of the complexities of providing high quality services. For example, in many jurisdictions, community-based sexual health clinics diagnose a high proportion of STI cases (e.g., as many as 20% of chlamydia cases, 35% of gonorrhoea cases, and 35% of HIV cases) [14]. And, large proportions of patients presenting at these clinics are uninsured or socially marginalised, (e.g., MSM; racial minorities). These clinics also often serve as training sites for new health professionals (e.g., nurses; physicians); and, these clinics are in many cases the focal point for studying the diagnosis, treatment, and prevention of STIs within the healthcare system [14].

1.3 Internet-based sexual health delivery programs

In light of these realities, there is an urgent need for the adoption of innovative service delivery options that provide high quality and comprehensive STI services while limiting the burden on resources (e.g., budgets; staff; infrastructure) [19]. Health agencies are increasingly using the Internet to meet emerging needs for disease control and as adjuncts to existing clinical and public health services [20,21] in order to improve the accessibility (particularly in rural or remote regions), quality, and appropriateness of services [17,22]. For example, internet-based programs for chronic diseases (e.g., asthma, breast cancer, diabetes) have demonstrated effectiveness [23-25]. Many jurisdictions have implemented an array of internet-based sexual health interventions, including testing programs which have been proposed as potentially cost-reducing and patient-centred alternative service delivery models [21,26]. These models are facilitated in part by the increasing adoption of highly sensitive and specific STI testing technologies, including Nucleic Acid Amplification Tests (NAAT) conducted on urine specimens and point of care (rapid) testing, which make the provision of testing services more convenient and accessible in these alternative settings where patients have limited interaction with clinicians [27].

In BC, an Internet-based STI testing service – Get Checked Online (GCO) – is being developed and pilot tested. GCO is intended to complement existing face-to-face clinic-based sexual health services, with the goal of increasing online testing uptake and therefore potentially easing barriers to testing as well as providing an alternative venue to clinic-based services [21,28]. GCO's model of integrated Internet testing, clinic-based sexual health services, and public health and laboratory systems is unique. Clients accessing GCO will complete a risk assessment module, download a test requisition, provide blood and/or urine specimens at designated collection sites, and

retrieve negative results online or positive results in-person or by phone. A more comprehensive description of GCO is available at: [http://bclovebytes.wordpress.com/about/].

Although the feasibility of internet-based testing programs has been established, rigorous empirical research studies are rare and urgently needed to ensure their sustainability, quality, comprehensiveness and flexibility [17]. In the current GCO model, clients who are symptomatic or are sexual contacts of infected individuals are considered to need immediate clinical assessment and will be triaged to receiving care at STI clinics. However, much remains to be learned about how to maximise internet-based testing experiences for those at highest risk for STIs, especially in the absence of clinician interaction as will be the case for eligible GCO clients. In this scenario, universal testing of everyone who presents for testing is not likely to reflect a prudent choice as STI prevalence rates are expected to be low in this population (i.e., asymptomatic and non-sexual contacts) [29]. In order to maximise case finding in this setting, developing criteria for selective screening (based on risk assessment and entails the screening of individuals who meet pre-specified criteria) may be a prudent approach because it minimises the costs associated with testing individuals at low risk [27]. However, unlike traditional (or face-to-face) clinical encounters, STI assessment in internet-based settings is not well articulated.

1.4 STI risk assessment

In traditional clinical encounters, a comprehensive consultation often includes risk assessment and helps clinicians in evaluating a patient's sexual behaviours, and also guides counselling and testing recommendations [30]. During these consultations, estimating the probability of infection in a patient requires an understanding of STI epidemiological trends and the risk factors that facilitate STI transmission. The elements that comprise an STI risk assessment questionnaire include detailed inquiries about the patient's relationship status, sexual risk behaviours, STI history, reproductive health history, substance use and psychosocial history [4,31]. Clinical guidelines recommend inquiring about the patient's relationship status (e.g., regular/casual sex partners) and the nature of the relationship (and timing of the relationship). This is because an individual's risk for STI/HIV may increase with each new sex partner [32,33]. New sexual partners facilitate a potential pool of infection into a sexual relationship [32,33]. In addition, it is often difficult to know the sexual history of new partners or whether they are presently infected with an STI [34]. The increased risk

associated with new sexual partnership has been shown to be potentially problematic among adolescents who are more likely to engage in sexual partnerships of short duration [32-34].

The clinical guidelines also recommend asking patients about any concerns they have relating to their current relationship (e.g., violence, coercion, abuse). The impact of intimate partner violence (IPV) on STI risk factors has been explored extensively, especially among women [35]. Current research indicates that instances of IPV precipitate sexual coercion, which could indirectly facilitate reduced sexual and condom negotiation practices [35]. One survey performed in a publicly-funded sexual health clinic in San Francisco documented a high prevalence of IPV among female patients and identified a significant association between IPV, STI/HIV risk factors, and self-reported STI history [35]. Compared with non-abused female patients, abused women attending this clinic were more likely to report being in high-risk sexual partnerships with IDUs, HIV-infected men, or MSM and sex work [35].

Questions about the patient's sexual risk behaviours such as number of partners, sexual activities (e.g., anal/oral/vaginal), sexual preference or orientation, condom use, and partner recruitment sites (Internet, bathhouse, travelling) are also essential to determine the patient's risk profile. The significance of patterns of sexual partnerships is highlighted in studies of sexual networks. In STI transmission dynamics, the "core" represents the innermost circle of the network with the highest concentration of risky behaviour and infection incidence such as IDUs and commercial sex workers [36]. Individuals in the "periphery" belong to the outermost circle of the network and are considered to be at the lowest risk for infection and have the lowest infection incidence [36]. However, "bridge" members, situated between the core and periphery, have STI incidence that is lower than the core but higher than the periphery and are crucial in the spread of infection from high-risk core groups to the general population [36]. Another emerging facet in the "bridging" of risk groups is the venue of sexual partner recruitment [37]. In a study of MSM recruited in several cities in the United States, the most endorsed venues for meeting sexual partners were found to be bathhouses, bars/clubs, and the Internet [37]. However, this study also found that MSM showing preference for bars/clubs were unlike MSM with preference for bathhouses or the Internet in terms of their demographic profile and sexual behaviour (e.g. age, number of sex partners, condom use) [37]. For example, MSM who recruit sexual partners online have been shown to have higher odds of unprotected anal intercourse [38]. However, it is still unclear as to whether

the Internet facilitates risk behaviours (i.e., promotes risky sex) or is a venue for high-risk sexual behaviours that would happen elsewhere.

The timing of sexual partnerships is another important dimension that influences STI transmission dynamics [39-42]. The literature identifies frequent concurrent partnerships (i.e., having frequent partners overlapping in time) and narrow partnership gap (i.e., having short duration of time between the end of one sexual partnership and the beginning of the next sexual partnership) as features that amplify the spread of STIs [39-42]. In addition, serial monogamy – in which the duration of time elapsed between last sex with one partner and first sex with the next partner is shorter than the duration of infectiousness of any untreated infection is also a known facilitator of STI transmission [32,39,41,43]. Although individuals in these relationships may self-report having lower STI risk, serially monogamous relationships, with known and committed partners, may not provide sufficient protection from STI [39].

Other components of risk assessment are reproductive health history (e.g., contraceptive use, Pap testing, pregnancy) and STI history (e.g., previous STI test, previous STI diagnosis). The use of oral contraception is frequently associated with lower rates of condom use [4]. This is because individuals in relationships often transition from condom use to oral contraception without testing for STI [4]. Consequently, the use of oral contraception is often associated with increased incidence of STI [4]. In addition, previous STI diagnosis is another potential predictor of current STI diagnosis. The endemic rates of chlamydia and gonorrhoea have been attributed more to individuals with repeat infections than those with single infections [44-46]. It has been hypothesised that network level factors may be operating, i.e., an increased prevalence of STI in an individual's social network may be associated with a high risk for new infections as well as frequent infections [44,45]. The assessment of an individual's substance use (e.g., sex under influence, shared equipment for injection) and psychosocial history (e.g., sex trade work, housing) are also essential aspects of sexual health consultations. Alcohol or drug use prior to sexual activity reduces inhibitions and could influence sexual and condom negotiation practices [30,36]. Drug users are important members of the core group [30,36]. The reason for the higher risk of STIs commonly found in drug users is a complex issue and potentially related to their increased number of sexual partners and participation in unprotected sex resulting from sex work (e.g., sex for drugs or money to buy drugs), especially among women [30,36].

1.5 Risk estimation for sexual health decision making

In traditional clinical encounters, whether in STI clinics or general practice settings, the risk factors mentioned in the above section frequently comprise the selective screening criteria that are used as decision aids to provide clinicians with assistance in distilling and applying the scientific literature to recommend specific STI tests and prioritise patient groups. This dissertation differentiates between mass/universal screening, selective screening, and selective testing. Universal screening is a strategy used to identify or predict infections in individuals without signs or symptoms within a defined group, such as a population, family, or workforce [47]. An example of the use of universal screening in sexual health is in the targeting of young women less than 25 years of age for chlamydia screening without regard to risk behaviours [30]. Selective or high risk screening identifies individuals who are known to be at higher risk; for example, commonly cited screening criteria include combinations of risk factors such as multiple sexual partners, inconsistent condom use, and new sexual partners. In contrast, selective testing (for the purposes of this dissertation) refers to the identification of high risk individuals for diagnostic testing when there is a suspicion of infection due either to the presence of symptoms, or high risk exposure (e.g., contact of someone with an STI).

The criteria used for the selective screening of STIs are most frequently derived from clinical guidelines and usually involve a combination of demographic (e.g., age; sex), behavioural (e.g., condom use; number of sexual partners) and clinical risk factors (e.g., previous STI diagnosis). However, recent criticism of screening guidelines highlight the differences that exist between the extent to which evidence drives specific STI screening recommendations and, to a certain extent, query whether in fact guidelines have sufficient evidence supporting them to optimise case finding [48]. Unlike HIV and syphilis, which are predominantly found in high risk or core groups, chlamydia - and gonorrhoea to a lesser extent - is spread among young people in general. However, population-based guidelines (such as screening those under 25 years old for chlamydia or gonorrhoea infection) may not be sufficiently flexible to meet budget constraints for some public health programs.

The generalisability of these screening recommendations is perceived to be uncertain because most recommendations have been developed in geographically and clinically restricted settings without validation in other populations [49,50]. While some factors, such as age, can be expected to contribute significantly to STI burden in virtually any setting, the association of infection with other factors (e.g., race/ethnicity; IDU) may differ considerably across contexts [36].

Other factors cited as contributing to the poor generalisability of screening recommendations include: selection bias due to study samples; differences in risk determinants between study populations; dissimilar populations and settings; and varying operationalisation (or measurements) of risk determinants [51].

There is growing support for more nuanced or personalised risk assessment approaches [52]. The current understanding is that screening recommendations could be improved by tailoring recommendations to the specific circumstances of the patient through the use of *clinical prediction rules (CPRs)*. CPRs use combinations of risk factors that have been statistically demonstrated to be meaningful predictors to calculate a numerical probability for the presence of a specific condition or likelihood of an outcome [53,54]. This allows for more nuanced and precise decision-making than population based guidelines when applied to individual patients [55]. CPRs are used extensively in chronic disease research. One widely used CPR is the Framingham risk score, a tool used for predicting ten-year risk of cardiovascular outcomes like myocardial infarction and stroke [56,57].

CPRs have potential applications in sexual health decision-making. CPRs used to derive selective screening criteria may help prioritise STI testing resources and help clinicians and public health administrators make decisions about where to focus STI screening efforts on the individual and population level. Specifically, CPRs can be used in targeted screening programs, public education initiatives, and risk communication to patients to encourage STI testing or behaviour modification [27,58]. However, the vast majority of these prediction rules remain unevaluated and their accuracy and generalisability are rarely interrogated within the context of STIs, although they have been investigated in chronic disease contexts [59]. For this reason, novel research is required to inform the development of well-performing, evidence-based CPRs that have potential for improving sexual health care service provision. Specifically, this dissertation focuses on deriving a CPR that can be used as a selective screening criterion for chlamydia and gonorrhoea infections.

1.6 Theoretical framework

This dissertation uses a well-established framework for the development of CPRs [60,61]. The framework describes an approach for evaluating CPRs based on the *accuracy* or measures of a CPR's performance and *generalisability* or features that facilitate a CPR's validity [60,62]. Accuracy, generally measured by two metrics (discrimination and calibration), reflects the degree to which the predictions match the outcomes. Discrimination is a measure of how well the CPR distinguishes

between those who do and do not have the disease of interest [62]. Calibration is a measure of how well the CPR's predicted probabilities agree with actual observed risk [62].

No matter how well calibrated or discriminating a CPR may be in development, a CPR that can only predict outcomes accurately in the population in which it is developed does not have huge utility [60]. For a CPR to be most useful it must also be generalisable; and, its accuracy must be both reproducible and transportable. Reproducibility or internal validity is the ability of the CPR to produce accurate predictions among individuals not included in the derivation of the CPR but who are from the same population [60]. Transportability or generalisability refers to the CPR's ability to produce accurate predictions among individuals drawn from a different, but reasonably related population [60]. The generalisability of a CPR is established as it is tested and found to be accurate across increasingly diverse settings (e.g., time periods and geographical locations) [60]. The more numerous and varied the settings in which a CPR is found to have good discriminative and calibration performance, the more likely it will generalise to another untested setting [60]. Consequently, the analyses presented in the current dissertation were informed by the hierarchical and iterative pathways described by the theoretical framework (as shown in Figure 1.1 and expounded upon in the next section).

1.7 Study setting and design

The empirical analyses presented in this dissertation (Chapters 4 and 5) rely on data collected at sexual health clinics located at nine geographical sites in BC. To avoid replication in subsequent chapters, the features of the study setting, population and design are summarised here. Briefly, this study involved a population-based, cross-sectional analysis of electronic records of patients visits collected at publicly funded STI clinics that offer physical examination and treatment for STIs in BC. The electronic records were derived from the STI Information System (STI IS), a database that houses risk assessment information and laboratory results of patient visits at these clinics. Data from each new client consultation between 2000 and 2012 among women and men who have sex with women were included. This analysis was limited to asymptomatic clinic visits that are not sexual contacts of known STI cases. Repeat visits within 30 days of a previous clinical visit were also excluded to avoid including clients receiving confirmatory diagnoses. The outcome measured in this study was chlamydia and/or gonorrhoea infection, which are reportable diseases in the Communicable Disease Regulation of the Public Health Act of British Columbia [63]. Appendix A

shows the geographical distribution of chlamydia and/or gonorrhoea infection rate in 2012 by Health Service Delivery Area (HSDA) in BC. Practitioners at STI clinics may order the following specimens to test for chlamydia or gonorrhoea: urine specimens and swabs (cervical, vaginal, urethral, rectal, oral swabs), which are tested using nucleic acid amplification test (NAAT) or via cultures (gonorrhoea only) [63]. Chlamydia and gonorrhoea infection was measured as a composite outcome because most laboratories use multiplex assays that test for both infections simultaneously [64] and also because the relevant clinical decision is whether to offer this test or not.

The dissertation's methodological framework is outlined in Figure 1.1. Briefly, the prediction rule was created using the data gathered from the *development population* and the generalisability of the criteria was tested in the *validation populations*. The *development population* is comprised of patient visits at the West 12th Avenue and Bute Street clinics in Vancouver *between 2000 and 2006* (n=10,437; where chlamydia and/or gonorrhoea prevalence was 1.8%). Both clinics are low-threshold (i.e., proof of health insurance is not required nor is proof of identity) outpatient clinics operated by the BC Centre for Disease Control (BCCDC). They provide STI assessment and management services, including HIV testing, for clients from throughout the Greater Vancouver area. Chlamydia, gonorrhoea, syphilis, and HIV tests are generally offered to all sexually active clients at each clinic visit.

Both the temporal and geographical generalisability of the prediction rule was examined. The *temporal validation population* included visits at the Vancouver clinics *between 2007 and 2012* (n=14,956; chlamydia and/or gonorrhoea prevalence was 2.2%). The *geographical validation populations* included clinic visits at publicly funded sexual health clinics located in the following geographical locations in BC – Penticton, Kelowna, Kamloops, New Westminster, Boundary, Courtenay, and Prince George (see map in Appendix B). The study analysed computerised records from clients attending these clinics between 2000 and 2012 (n=10,425; chlamydia and/or gonorrhoea prevalence was 5.3%). These clinics use the same electronic charting system as the BCCDC clinics (West 12th and Bute Street); thus, the consistent nature of the data collection methods across the clinics allows for the direct comparison of data between individuals attending the clinics.

1.8 Understanding the roles of proximate and distal determinants

This dissertation uses the proximate determinant framework [65,66] as the theoretical foundation for the selection, operationalisation, and interpretation of explanatory variables (or risk factors) in the prediction rule. Figure 1.2 shows the proximate-determinants framework for infection with STIs, which links social determinants (depicted on the left of Figure 1.2) with biological outcomes (depicted on the right of Figure 1.2). The socio-demographic variables (or distal determinants) operate through proximate determinants that in turn influence the outcome of interest – chlamydia and/or gonorrhoea infection. For example, consider the frequently noted association between involvement in sex work and STIs. Being paid for sex or paying for sex does not, in itself, increase an individual's risk of infection [65,66]. This association exists because those who are paid for sex or those who pay for sex on average have more sexual partners and those partners have a higher probability of being infected [65,66].

In this framework, the immediate determinants of infection that comprise R_0 become relevant for individualized risk assessment. Specifically, proximate determinants are also known as biosocial mechanisms. The reason is that proximate determinants have both social/behavioural characteristics (e.g., exposure to susceptible individuals, c) as well as biological characteristics (e.g., efficiency of transmission per act, β) [65,66]. Prior to conducting any statistical analysis, the variables captured in STI IS were mapped to the framework in order to develop a detailed understanding of their contribution to an individual's overall STI risk. The framework acknowledges the 'hierarchy' of STI risk by postulating that structural factors or underlying factors (e.g., economic inequities, barriers to STI prevention and treatment services, law enforcement activities) and socio-demographic determinants (e.g., age, sex, race/ethnicity) must operate through more proximate risk factors (e.g., unprotected sex, number of sexual partners) to cause STIs [31,65-67]. The discussion of the findings of this dissertation will take advantage of the 'upstream to downstream' nature of risk outlined in the framework to understand the associations found between the determinants and STIs and will generate evidence that can be used to inform how GCO can be scaled-up effectively to avoid exacerbating sexual health inequities, particularly amongst vulnerable sub-groups.

1.9 Ethical considerations

Because the empirical parts of this dissertation present a secondary use of healthcare data, there are concerns about patient privacy and confidentiality, an ethical issue particularly salient in

this era of "big data" analytics [68]. The electronic medical records used in this dissertation were accessed in a manner that maintained the security and confidentiality of the data by means such as de-identification and aggregation of data elements, the encryption of electronic files, and the establishment of data breach quality assurance checks. Ethics approval for the dissertation was obtained from the University of British Columbia Research Ethics Board prior to the start of any research activities.

1.10 Rationale for the study

In chronic disease contexts, CPRs are well understood and articulated and in some cases, have become part of routine medical care for diagnostic and prognostic decision-making [56,57]. However, the adoption of prediction rules in sexual health decision-making is in a somewhat nascent stage. There are several essential gaps in knowledge that hinder the implementation of CPRs for use in sexual health. The predictors used in CPRs, especially selective screening criteria, are often derived from clinical guidelines, and few studies have been conducted to determine the degree to which evidence in the published literature supports the predictors included in recommendations. Although empirical studies are available, a critical examination of the research evidence for predictors of STIs has not been explored. Indeed, a majority of the evidence in this area has included symptoms as risk factors, which makes their application to selective screening scenarios problematic. There is a need for a better understanding of the use (or lack) of evidence in this area as it has the potential to inform the ways in which STI test recommendation practices unfold in the future.

The majority of CPR-related research in sexual health has been generated with little attention to whether they are conducted according to methodological recommendations often cited in this field [59]. Overstated and biased results from inadequately designed and reported prediction studies can potentially trigger their premature implementation and lead to erroneous decision-making by STI programs. A rigorous evaluation of prediction rules before their introduction into sexual health practice could not only reduce the number of unwanted and unintended decision-making consequences related to misleading estimates of test accuracy, but, in some cases, could also reduce healthcare costs by precluding unnecessary testing [69]. Thus, a study elucidating the methodological contexts and the predictive performance of published prediction rules used in sexual health services has the potential to be a vital part of this evaluation process.

To concretise selective screening in Internet-based testing contexts, e.g., GCO, STI programs have a choice of adopting a set of previously published screening recommendations, such as those issued by public health organisations, or develop a set of risk estimation algorithms specific to their setting. As no major studies of selective screening have been conducted in BC, public health agencies in this region have little guidance regarding the choice of selective screening criteria. Moreover, Internet-based testing programs have only recently been initiated in Canada. There also has been little discussion of the empirical validity of CPRs, which may be particularly salient for promoting their effective integration with STI service delivery. For these reasons, new research is required to inform the development of a CPR that will be adapted for use in an Internet-based context in BC.

Finally, it is important to acknowledge that the scope of the current dissertation was limited to the examination of chlamydia and gonorrhoea outcomes among asymptomatic women and heterosexual men. This approach was adopted for several reasons. First, this dissertation focused on asymptomatic women and heterosexual men for pragmatic reasons – because sample size restrictions prohibited the examination of other outcomes such as HIV and syphilis, or other populations of interest such as MSM. Second, the targeted population accessing GCO will be limited to those who are asymptomatic, as clients 'presenting' online with symptoms will automatically be referred to an STI clinic. Third, as will be revealed in the literature review section of this dissertation (Chapter 2), there are currently no published prediction rules for screening asymptomatic chlamydia and gonorrhoea. Finally, asymptomatic infections are extremely commonplace (perhaps accounting for 80% to 90% of all infections) [70]. Current research indicates that increasing testing access for high-risk asymptomatic clients has the potential to confer public health benefits as undiagnosed and untreated infections may result in long-term health and health economic impacts that surpass those associated with clinically observable infections [71].

1.11 Study objectives

The current dissertation addresses the abovementioned gaps in knowledge by characterising the attributes and components of prediction rules that encourage their adoption for personalised clinical decision-making in sexual health. The current research also provides one of the first opportunities to advance both the conceptual context of selective screening intervention research in combination with an empirical assessment of the potential contribution of a prediction algorithm in

sexual health contexts. Finally, the work in this dissertation articulates and applies a methodological framework for a hierarchical and iterative approach to prediction rule development for use in sexual health contexts, which may prove useful for additional research in this and other areas. Specifically, the dissertation address the following objectives:

1. **To characterise the scientific evidence examining the associations between risk factors and chlamydia and gonorrhoea infection (Chapter 2).** This dissertation provides the results of a literature review examining whether the risk factors that are often cited in screening recommendations are correlated with their predictive strength in empirical studies using a theory-guided approach to understand the significance of the findings. This review has potential relevance in sexual health practice because an understanding of the evidence in this area may help inform the ways in which existing selective screening criteria continue to be implemented in traditional face-to-face clinical encounters, as well as in novel approaches such as Internet-based initiatives. The findings of this review were also used to identify candidate predictors used in the empirical analyses included in this dissertation.
2. **To critically appraise and generate a knowledge synthesis of prediction rules developed in sexual health settings (Chapter 3).** Motivated by the lack of standards in the reporting of the methodology and findings of prediction rules in sexual health contexts, it was important to examine the design, methods, analysis, and results of other studies that use CPRs. This review determined whether the accuracy and generalisability of prediction rules are correlated with the potential for bias. These findings provided important insights for guiding the methodological approaches used in the derivation and validation of the prediction rule described in this dissertation.
3. **To derive and temporally validate a prediction rule for screening asymptomatic chlamydia and gonorrhoea (Chapter 4).** A prediction rule for identifying those at increased risk for asymptomatic chlamydia and gonorrhoea among women and heterosexual men was derived using multivariate logistic regression of the data from electronic medical records collected at two sexual health clinics in Vancouver collected (2000-2006). The prediction rule was then validated using data from a subsequent time period (2007-2012). Temporal validity is generally considered the first line of generalisability ascertainment for CPRs. This issue is particularly salient because of the shift towards more sensitive diagnostic tests between the time periods represented in this chapter.

4. **To assess the generalisability of the rule across geographic locations in British Columbia (Chapter 5).** The generalisability of the aforementioned prediction rule was examined in seven additional clinics located outside of Vancouver. Because prevalence and social contexts vary between Vancouver and the other smaller cities, it was important to assess whether the prediction rule would show good validity in the geographical validation population.

1.12 Organisation of the dissertation

This dissertation is delineated into six chapters. **Chapter 1** summarises the public health burden of STIs in Canada, described STI risk assessment and selective screening criteria currently used in clinical settings, and provides a rationale for the development of a prediction rule for screening chlamydia and gonorrhoea in BC. The study objectives, settings and the study design also are described in this chapter. **Chapter 2** is a review that aims to provide a theory-informed understanding of the predictive strength of risk factors commonly cited in screening guidelines and empirical studies. **Chapter 3** is a knowledge synthesis and a critical review of CPRs used in sexual health settings. The content of Chapter 3 also informs the analytical approaches used in Chapters 4 and 5. **Chapter 4** describes the derivation, internal validation and temporal validation of a CPR for screening asymptomatic chlamydia and gonorrhoea using data gathered from two Vancouver clinics. **Chapter 5** focuses on the geographical validation of the prediction rule in seven additional clinics across BC. Finally, **Chapter 6** provides a summary of the key findings, practical implications, strengths and limitations of the research, as well as recommendations for future research.

Figure 1.1 The methodological framework for the derivation and validation of the prediction rule

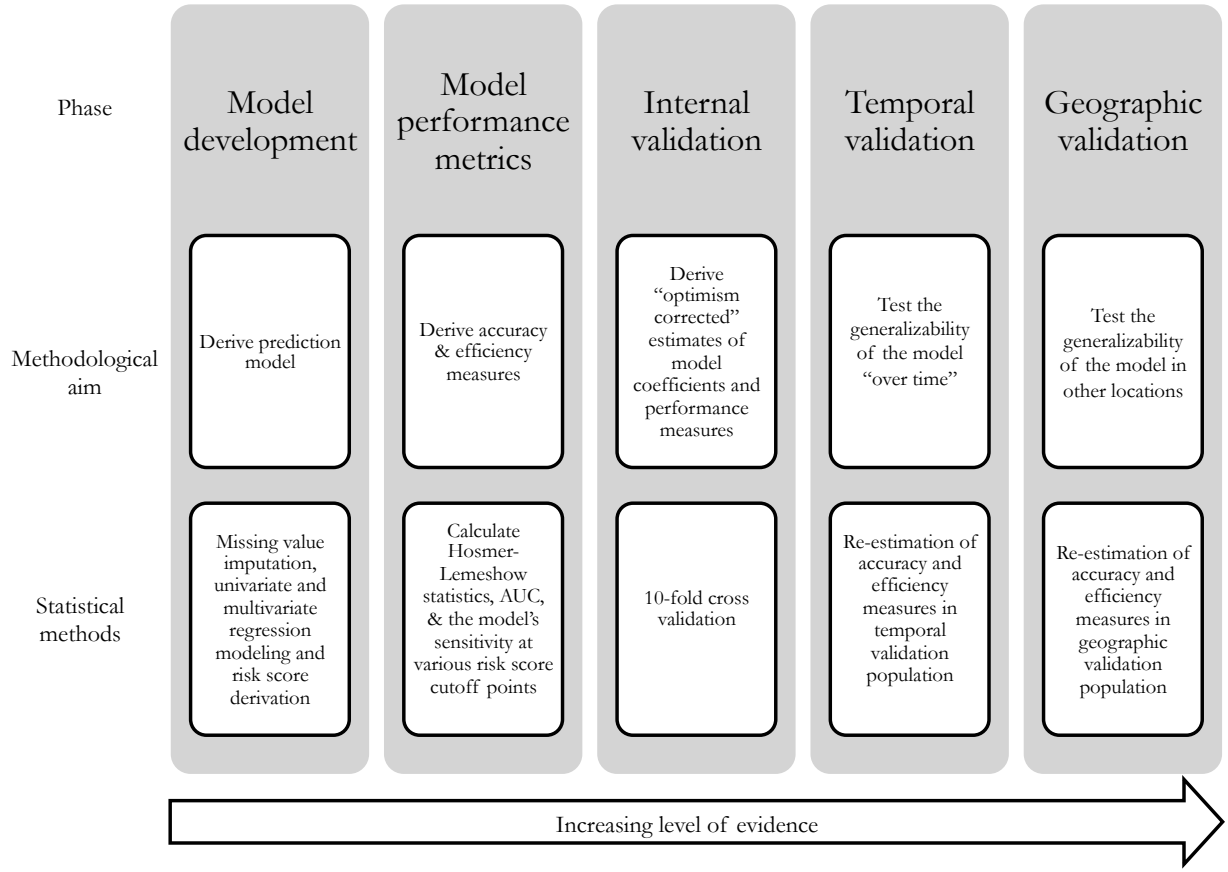
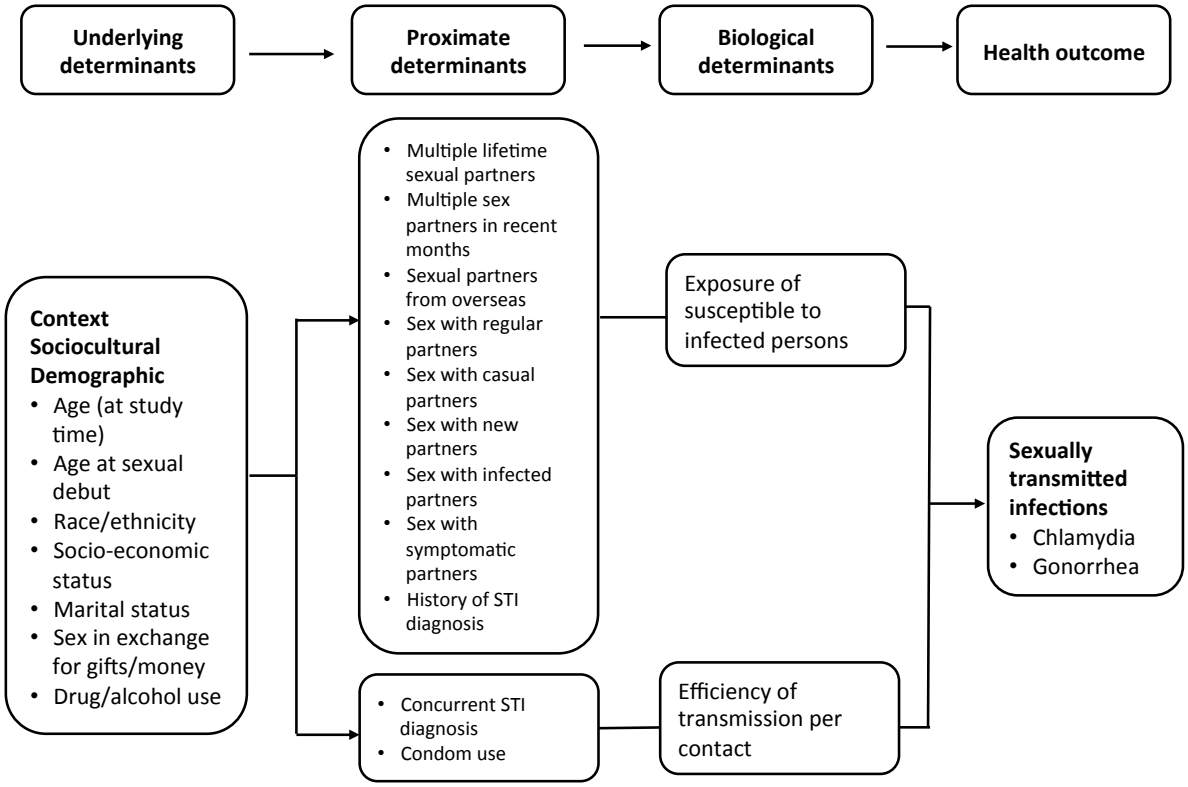


Figure 1.2 Proximate-determinants framework for factors affecting the risk of sexual transmission of chlamydia and gonorrhoea



Chapter 2: Predictors identifying individuals at increased risk for STIs: A theory-guided review of empirical literature and clinical guidelines

Falasinnu T, Gustafson P, Hottes TS, Gilbert M, Ogilvie G, Shoveller, J. Predictors identifying those at increased risk for STDs: a theory-guided review of empirical literature and screening recommendations. The International Journal of STD and AIDS. *In press.*

2.1 Introduction

Sexually transmitted infections such as chlamydia and gonorrhoea are leading causes of substantial morbidity worldwide; left untreated, they are often important exposures in the causal pathway for pelvic inflammatory disease, ectopic pregnancies, and infertility. The fact that most infections do not exhibit symptoms that prompt clinical evaluations often makes the detection of infection challenging [51]. Public health and clinical organisations routinely issue guidelines aimed at helping healthcare providers identify individuals at increased risk for STIs. Plausible candidates for selective screening criteria include known determinants of prevalent infection (e.g., age, multiple partners) [72]. These criteria help sexual health programs minimise costs and increase the probability of detecting infections [51].

The identification of risk factors that comprise an accurate selective screening criterion provides the opportunity to deliver four key benefits in clinical settings: identifying specific patient groups with higher disease prevalence, identifying areas to target prevention services, improving case finding accuracy, and increasing cost-efficiency by limiting the testing of low-risk individuals. However, predictors used for selective testing are often derived from clinical guidelines that are broadly perceived as evidence-based statements; because they are developed by experts, they are assumed to have the same level of certainty and authority as conclusions generated by scientific methods [48,72]. Few studies have been conducted to determine the degree to which evidence in the published literature supports the predictors included in recommendations. A better understanding of the use (or lack) of evidence in this area has the potential to inform the ways in which STI test recommendation practices unfold in the future.

The review presented in this chapter aims to synthesise the evidence supporting predictors of chlamydia and gonorrhoea (the two most common reportable STIs) in high-income jurisdictions in North America, Western Europe, and Australia – regions with comparable STI prevalence and

social determinants of sexual health. The objectives are to identify risk factors cited as being most predictive of chlamydia and gonorrhoea in clinical guidelines and in empirical studies. This review also aims to assess the strength of evidence supporting the associations between these predictors and chlamydia and gonorrhoea outcomes in empirical studies. Finally, this review highlights discrepancies between predictors used in STI clinical guidelines and the relevant empirical evidence.

2.2 Methods

2.2.1 Literature search

This review identified the most recent and commonly cited clinical guidelines for screening chlamydia and gonorrhoea issued by health organisations in high-income countries in North America and Europe, and Australia. This review also identified empirical studies in journals readily accessible to sexual health clinicians or experts in these settings and thus, limited the search to electronic searches of English-language studies published in OVID Medline, as well as manual searches of reference lists of papers identified. MeSH search terms included: 'screening'; and at least one of 'chlamydia', 'gonorrhoea', 'STI', or 'STD'. Articles that measured demographical and behavioural predictors associated with STI diagnoses were included. Studies focusing on psychosocial factors (e.g., knowledge and attitudes, stress, self-esteem) were excluded, as these predictors are not consistently measured across studies. The search was limited to studies published between 2003-2011, an era characterised by the expansion of highly sensitive tests (e.g., Nucleic Acid Amplification Tests) and urine-based testing, in an effort to ensure comparability in specimen collection and STI outcome assessment.

2.2.2 Assessment of methodological quality

Articles selected for inclusion were examined to assess their methodological quality. The following nine criteria, adapted from previously validated methodologies [73,74], were used to score the methodological quality of studies: study design, specimen collection, statistical methods, model selection, study sample description, variable definition, events per variable, outcome assessment, and study generalisability (see Table 2.1). Based on the scoring system, studies with a rating of 'unacceptable' (≤3 points) were removed from further review. Studies rated as 'marginal' (4-5 points), 'acceptable' (6-7 points) and 'commendable' (8-9 points) were kept for the empirical analysis

stage of the review. The studies rated as 'commendable' or 'acceptable' were given greater weight when interpreting consistency or variable relationships.

2.2.3 Data extraction

To help clarify the relative strength or importance of each STI predictor, the predictors extracted from the empirical studies were categorised into two groups based on the proximate-determinants framework [65,75]: (1) distal determinants, which are demographic, social or economic variables distally related to STIs, and (2) proximate determinants, which are directly associated with an individual's probability of exposure to STIs and the efficiency of STI transmission. Also, to determine whether the various factors identified in the included studies were significantly associated with STIs in multivariable analysis (if possible), a significance level of 0.05 was used. Because the studies included in this review were heterogeneous with respect to study design and predictor definition, pooling of data for meta-analyses was not possible. Therefore a qualitative data-synthesis was performed [76,77].

Key predictors mentioned in multiple studies were retained to test the level of empirical support for their association with chlamydia or gonorrhoea infection (STIs). The strengths of the evidence for predictors were assessed as follows [76]: (1) Strong evidence: >75% of 'acceptable' or 'commendable' studies found consistent and statistically significant greater risk of infection; (2) Moderate evidence: 60% - 74% of 'acceptable' or 'commendable' studies found consistent and statistically significant greater risk of infection; (3) Mixed evidence: 45% - 59% of 'acceptable' or 'commendable' studies found consistent and statistically significant greater risk of infection; (4) Weak evidence: <45% of 'acceptable' or 'commendable' studies found consistent and statistically significant greater risk of infection. In addition, predictors were also deemed to have insufficient evidence if few studies (i.e., less than the lowest quartile of the number of studies reviewed) examined the association between the predictor and infection.

2.3 Results

2.3.1 Overview of clinical guidelines

Nine organisations issuing clinical guidelines in high-income contexts were identified (Table 2.2). Six organisations listed clinical guidelines for women only, citing lack of evidence for the

selective screening of men: the US Preventive Services Task Force (USPSTF) [78,79], the American Academy of Family Practice (AAFP) [78,79], the US Centers for Disease Control and Prevention (USCDC) [80], the American College of Obstetrics and Gynaecology (ACOG) [81], the Public Health Agency of Canada (PHAC) [30] , and the Society of Obstetricians and Gynaecologists of Canada (SOGC) [82] (Table 2.2). Three other organisations reported clinical guidelines for males and females combined [83-85].

Almost all the organisations support the universal screening of young sexually active individuals for chlamydia and to some extent gonorrhoea (Table 2.2). Occasionally, clinical guidelines differ across organisations, mainly because of distinctions in goals and target audience, the epidemiology of the various STIs, and differing approaches to evidence review [67]. For example, the USCDC and PHAC focus on public health in the US and Canada, respectively, while other organisations focus on clinical settings [67]. A closer examination of predictors revealed incongruities between the organisations, even among those in the same country (Table 2.2). For example, clinical recommendations issued by both the PHAC and the SOGC agree on screening individuals with the following risk factors for chlamydia or gonorrhoea infection: younger age, sexual contact of STI case, previous STI diagnosis, new sexual partners, multiple sexual partners, use of drugs or alcohol during sex, sex in exchange for money or gifts, and individuals or travelers from endemic regions (Table 2.2). However, the PHAC also identifies the following risk factors: street involvement (or homelessness) and sexual contacts of people with any of the aforementioned risk factors (Table 2.2).

The most consistently listed predictors among clinical recommendations were younger age and multiple sexual partners as all nine organisations recommended screening individuals with these two risk factors (Table 2.2). Approximately 90% of the nine organisations recommended screening individuals who are sexual contacts of STI cases and also those who use alcohol or drugs during sexual relations. The least consistently listed predictors were street involvement (11%), inconsistent condom use (44%), and individuals or travelers from STI endemic regions (56%).

2.3.2 Assessment of included studies

Figure 2.1 provides an overview of the article selection process used in this review. The initial search of OVID Medline identified 841 articles. These were screened for eligibility and to obtain 216 articles for further assessment based on their titles and keywords. After reviewing the

abstracts of these 216 articles, 74 articles were deemed relevant for a closer reading of the full text, leaving 56 articles from the electronic search. An additional 19 articles were identified by hand searching the reference lists of these 56 articles. Thus, 75 studies were rated, most of which included a mix of symptomatic and asymptomatic individuals (Table 2.3).

Approximately 68% of studies examined chlamydia infection only, 8% examined gonorrhoea outcomes only, and 19% examined both outcomes (Table 2.3). Approximately 32% of the studies had female only populations and 20% included male only populations. Nearly 50% were from the United States (Table 2.3). Some studies used population-based participant recruitment (21%); most studies were cross-sectional in design (87%). A minority of the studies examined only asymptomatic populations (8%). Chlamydia prevalence ranged from 1.0% to 14.0% and gonorrhoea prevalence ranged from 0.1% to 33%. Based on the nine criteria of methodological quality, a total of 5 studies (7%) received a score of $\leq$3/9 and thus were rated as 'unacceptable' [86-90]. These studies were omitted from the second stage of review. Studies rated as 'marginal' (24%)[72,91-107], 'acceptable' (41%) [71,108-137], and 'commendable' (28%) [27,51,138-156] were retained in the review. Table 2.4 summarises the 70 studies that were retained for further assessment.

Empirical support for the following seven variables was strong or moderate: age (at time of study), race/ethnicity, multiple lifetime sexual partners, multiple sexual partners in recent months, sex with infected partners, sex with symptomatic partners, and concurrent STI diagnosis. Three predictors had mixed evidence: age at sexual debut, sex with causal partners, and sex with new partners. Eight predictors had weak evidence associating them with STI diagnosis: socio-economic status, marital status, sexual partners from overseas, sex in exchange for gifts/money, drug/alcohol use, condom use, sex with regular partners, and history of STI diagnosis.

Predictors examined by eleven studies or less (i.e., less than the lowest quartile of the number of studies reviewed) were deemed to have insufficient research. There were insufficient empirical studies to determine the nature of the association between the following six variables and infection: sex in exchange for gifts/money, sexual partners from overseas, sex with regular partners, sex with casual partners, concurrent STI diagnosis, and sex with symptomatic partners. Six variables were not included in any clinical guidelines: race/ethnicity, socio-economic status, marital status, sex with regular partners, sex with casual partners, and sex with symptomatic partner.

2.3.3 Assessment of the predictive abilities of underlying determinants

The following underlying determinants of STI diagnosis were identified in this review: age (at study time), age at sexual debut, race/ethnicity, socio-economic status or SES as defined as income or educational attainment, sex work, marital status, and drug/alcohol use (Table 2.4). Younger age was consistently associated with increased risk of STI acquisition (especially chlamydia infection); 85% of 'commendable' studies reported a significant association between age and STI diagnosis in their multivariable analyses (Table 2.4). There was mixed evidence in the current review associating early age at sexual debut to STI diagnosis; only 25% of 'commendable' studies and 56% of 'acceptable' studies that examined this risk factor found a significant relationship (Table 2.4).

The level of empirical evidence in the current review for the association between race/ethnicity and STI diagnosis was moderate – 72% of 'commendable' studies and 68% of 'acceptable' studies measuring this association found significant associations in their multivariable analyses (Table 2.4). Most of these studies found a higher risk of infection among ethnic minority individuals [27,51]. Conversely, the level of empirical evidence for the association between SES and STI diagnosis was rated as weak – only 56% of 'commendable' studies and 18% 'acceptable' studies found a significant relationship (Table 2.4). SES was not associated with infection in multivariable analyses using any measure, including educational level [127,146,153] and employment status [109,118].

This review found weak evidence for the association between STI diagnosis and the following predictors: marital status, drug/alcohol use and sex in exchange for drugs/money (Table 2.4). Only 33% of 'commendable' studies and 33% of 'acceptable' studies found significant associations between drug/alcohol use and STI diagnosis. Only 8 studies (of 'marginal' and 'acceptable' quality) examined the associations between sex in exchange for gifts/money and infection, providing insufficient evidence to comment further here.

2.3.4 Assessment of the predictive abilities of determinants of exposure to infected people

The following proximate determinants of exposure to infected people were identified in this review: multiple lifetime sexual partners, multiple sex partners in recent months, sexual partners from overseas, sex with regular partners, sex with casual partners, sex with new partners, sex with infected partners, history of STI diagnosis, and sex with symptomatic partners (Table 2.4). The

empirical evidence linking multiple lifetime sexual partners (i.e., more than 1 sexual partner) and STIs was strong. However, there was no consistent evidence for the association with higher cut-points (e.g., >2, >5, or >10 partners). All of the 'commendable' studies and 57% of 'acceptable' studies showed significant associations between multiple lifetime sexual partners and STI diagnosis (Table 2.4). This review also found moderate evidence for the association between having multiple sexual partners in recent months and STI diagnosis. Approximately 64% of 'commendable' studies and 57% of 'acceptable' studies that measured this relationship found a significant association between having multiple sexual partners in recent months in multivariable analysis (Table 2.4).

The evidence for the association between STI diagnosis and the following predictors was either mixed, weak and/or explored in few studies: sex with regular partners, sex with casual partners, and sex with partners from overseas (Table 2.4). This review found mixed evidence regarding the association between having a new partner and STI diagnosis (only about 60% of 'commendable' studies and 33% of 'acceptable' studies that measured this variable found significant associations with STI diagnosis). This review found strong evidence between STI diagnosis and sex with infected partners or symptomatic partners. All of the 'commendable' studies measuring these associations found a significant association. This review found weak evidence for the association between having a history of STI diagnosis and infection. Only 33% of 'commendable' studies and 46% of 22 'acceptable' studies that measured this association found a significant relationship.

2.3.5 Assessment of the predictive abilities of determinants of the efficiency of transmission per contact

The following proximate determinants of the efficiency of infection transmission were identified: concurrent STI diagnosis and condom use (Table 2.4). This review found strong evidence for the association between having a concurrent STI diagnosis and infection. Both of the 2 'commendable' studies and 75% of the 4 'acceptable' studies that explored the relationship between having a concurrent STI diagnosis and infection found a significant association, although few studies explored this relationship (n=11). The empirical evidence for the association between the use of condoms and infection was weak across the studies. Only 36% of 'commendable' studies, 26% of 'acceptable' studies and 25% of 'marginal' studies found significant associations between condom use and STI diagnosis. There were no notable trends among the studies examining the relationship

between condom use and STI diagnosis, although self-report of condom use is an enduring measurement problem in this field [157].

2.4 Discussion

The increasing availability and use of non-invasive diagnostic tests for STI diagnosis has re-opened discussions about the feasibility of scaling up selective testing efforts to include the general population. This review was conceptualised as a guide to help decision makers considering undertaking selective testing programs in alternative settings such as the internet. Specifically, the findings may help them augment selective testing by identifying variables to include in risk assessment questionnaires. In the future, clinicians may develop risk estimation or prediction tools by combining the risk factors identified in this review. These prediction tools may be potentially useful for targeted risk communication and STI testing motivation among patients attending sexual and reproductive health clinics [28].

This review summarises the literature on the most consistent predictors of chlamydia and gonorrhoea infection, which were identified after a wide range of patient characteristics was assessed. Age, number of sexual partners, sex with a symptomatic partner, sex with an infected partner and concurrent STI diagnosis are perhaps the most intuitive predictors of STI risk. The current review has reinforced the importance of these variables. The current review uses a theoretical framework –the proximate-determinants framework – which hypothesises that relationships between the underlying or socio-demographic characteristics and STI diagnosis should be non-significant, after adjustment for the proximate determinants or socio-behavioural predictors [65,75]. In this review, several underlying predictors (e.g., age and race/ethnicity) were found to be significantly associated with STI diagnosis – including in several studies that adjusted for proximate determinants [65,75]. This finding suggests that some proximate determinants may potentially be measured with error or that some important proximate determinants were not included in the studies [65,75]. It also highlights the complexity inherent in the self-reported nature of proximate determinants such as condom use, coupled with measurement issues in underlying predictors (e.g., race/ethnicity) [158].

There was reasonably good concordance between risk factors *consistently* listed in the recommendations (e.g., younger age, multiple sexual partners, and sexual contact of STD case) and predictors found to have *strong* empirical support in the literature (i.e., if a risk factor is included in

more guidelines, then there is more evidence supporting its predictive capacity). However, when the guidelines are considered individually or in isolation, there are large inconsistencies between the STD risk factors cited by the organisations and this may reflect the lack of clarity, transparency, and consensus about how high risk behaviours influence infection. Although the social determinants of health are comparable at the national level for these guidelines, however, another possible explanation for the differences in recommendations may be due to regional differences (e.g., between cities) in the prevalence of risk factors due to varying social determinants of sexual health. These inconsistencies could also reflect the overall priorities of organisations issuing recommendation statements (e.g., minimising harm from screening and cost-effectiveness versus concerns about missed cases).

In considering the limitations of the current review, there was a paucity of empirical studies examining the strength of predictors in asymptomatic populations. Because most STIs are asymptomatic, it is important to elucidate the predictive ability of risk factors in the absence of symptoms. However, most studies combined symptomatic and asymptomatic individuals in multivariable analyses likely masking important differences between these populations. This review was unable find to identify clear trends in associations between predictors and STIs among the six studies that explicitly examined asymptomatic populations. This limitation may have considerable impact on some findings. It is likely that the effect sizes of other predictors were underestimated in multivariable regression models that included symptoms as covariates because of the strong association between symptoms and STIs. Furthermore, the significant associations between some predictors and STIs may be a function of selection bias in clinics settings because of differential health seeking behaviours [159,160]. It is hypothesised that when health-seeking behaviour of certain population sub-groups (e.g., older adults, racial/ethnic minorities) is low, its percentage of symptomatic clinic visits will be high because these individuals are less likely to visit sexual health clinics for routine screening [159,160]. Thus, the associations found may not actually reflect a higher concentration of risk in certain sub-groups but is a reflection of the over-representation of these sub-groups in symptomatic populations presenting to clinics. Further studies assessing asymptomatic populations are, thus, warranted.

Also, the use of a single reviewer to screen articles and synthesise evidence may have introduced some error/bias [161], although, Buscemi and colleagues showed that this concern may not be valid [162]. The studies included in this review were identified by searching OVID Medline

only. Searching OVID Medline but not other databases (such as Embase or CINAHL) may have led to biased findings as studies archived in OVID Medline only have been shown to produce larger estimates in meta-analyses [163]. However, one review found diminishing marginal returns with increased database searching, with the authors recommending that hand searching of relevant reference lists (as was done in this review) may be more effective than exhaustive database searching [163].

Finally, this review was unable to synthesise the evidence separately for special populations (e.g., men who have sex with men, men who have sex with women) as clinical guidelines and empirical analyses for these groups were inconsistently reported. It is recommended that periodic reviews of the evolving evidence base in this substantive area have good potential to contribute to the planning and evaluation of selective testing intervention practices as they unfold in the future. There is, however, a need to continue to build the evidence base, particularly for special populations, and as STI testing intervention practices expand to non-traditional settings (e.g., the internet; the home).

2.5 Summary

This review identified several key gaps in the literature. There was a lack of empirical evidence exploring risk factors in asymptomatic contexts. Although this review found that predictors that were most consistently mentioned across clinical guidelines were also found to have high predictive strength in the empirical data, the differences in the risk factors specified in each of these screening recommendations indicate that, individually, they may not be broadly applicable. Specifically, these screening guidelines are unable to account for differences in patient populations and clinical contexts. Thus, the empirical portions of this dissertation (Chapters 4 and 5) will address these gaps in knowledge by exploring the development of tailored criterion for screening asymptomatic chlamydia and gonorrhoea.

Table 2.1 A summary of the criteria to be used to evaluate and rate the quality of methodology for each study*

Element	*Quality of methodology criteria*
A. Study design	The study design was clearly evident (e.g., cross-sectional, cohort or case – control).
B. Specimen Collection	The specimen collection methodology was clearly defined. If the same specimen collection methodology was not used consistently among study participants, methods to adjust for this difference were included in the analyses.
C. Statistical methods	Multivariate analyses were used to examine the relationship between predictors and STD outcome. Methods used to examine subgroups and interactions were stated.
D. Model selection	A clear attempt at model selection was made. Step-wise procedures or an a priori specification of p-value cutoff points for variable inclusion were considered acceptable for this criterion.
E. Study sample description	The characteristics of the sample and the selection criteria for the study were clearly stated such as: location, time period, sampling method, sample size, entry criteria and exclusions. The study also included a description of non-respondents. The extent and nature of missing data were assessed to estimate potential bias.
F. Variable definition	Sources of data and details of methods of assessment were given for each variable.
G. Events per variable	The final sample size that was actually used in the analysis comprised of ≥10 study outcomes per variable.
H. Outcome assessment	STD outcomes were measured by the most sensitive and specific tests available at the time of data collection.
I. Study generalisability	STD screening was evaluated in an appropriate spectrum of subjects and sampling was based upon a multi-center study or involved a single study with a relatively large sample size of consecutive visits to a clinic.

*Each quality criterion was awarded 1 point whenever present. Overall ratings were calculated by totaling the points and designated as Commendable (8-9 points), Acceptable (6-7 points), Marginal (4-5 points), and Unacceptable (≤3 points)

Figure 2.1 The process of selecting articles for the inclusion in the literature review of predictors of chlamydia and gonorrhoea infection

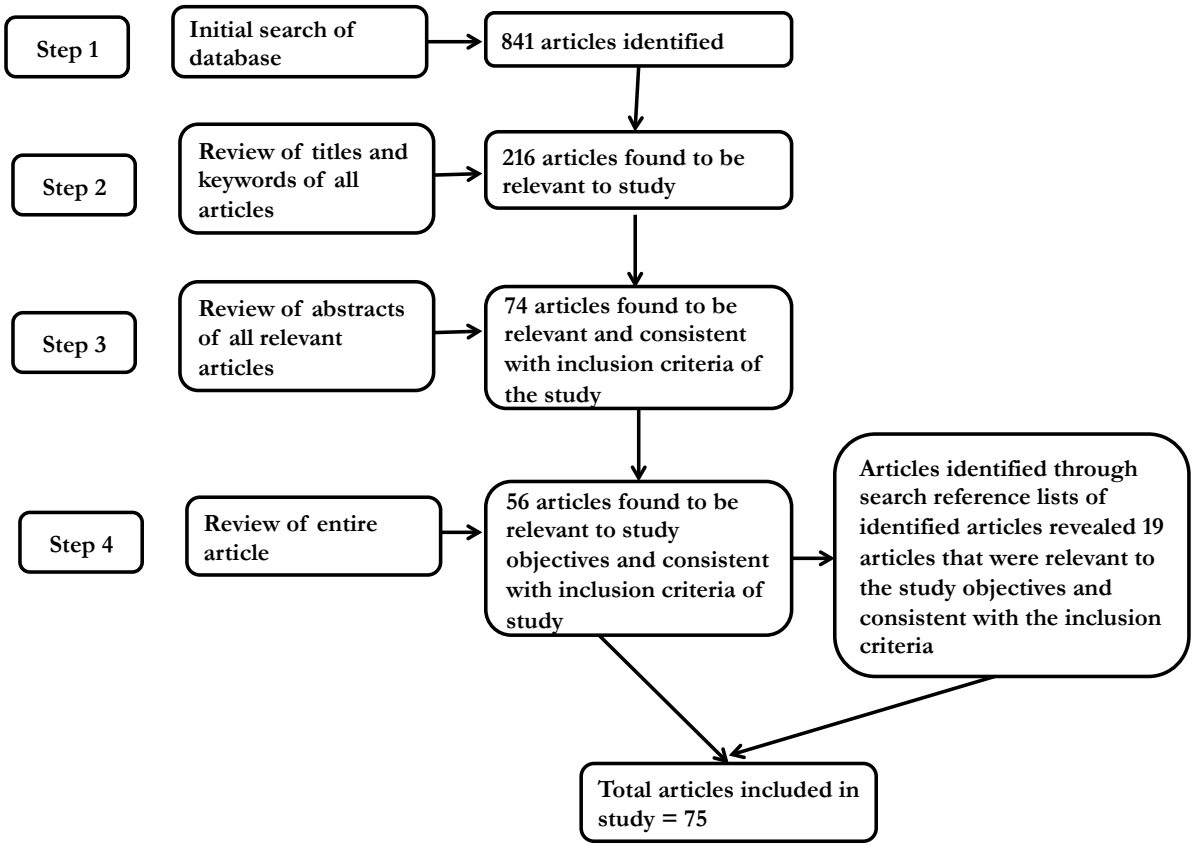


Table 2.2 Comparison of risk factors listed in chlamydia and gonorrhoea screening recommendations*

	Risk factors										
Organisation**	Younger age	Sexual contact of STD case	Previous STD	New sexual partners	Multiple sexual partners	Drug or alcohol use§	Sex in exchange for money or gifts	Inconsistent condom use	Individuals or travelers from endemic regions	Street involvement	Sex with anyone with listed risk factors
US Preventive Services Task Force (USPSTF)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes			
American Academy of Family Physicians (AAFP)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes			
US Centers for Disease Control and Prevention (USCDC)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes			
American Congress of Obstetricians and Gynecologists (ACOG)#	Yes	Yes	Yes		Yes	Yes	Yes		Yes		Yes

	Risk factors										
Organisation**	Younger age	Sexual contact of STD case	Previous STD	New sexual partners	Multiple sexual partners	Drug or alcohol use§	Sex in exchange for money or gifts	Inconsistent condom use	Individuals or travelers from endemic regions	Street involvement	Sex with anyone with listed risk factors
Society of Obstetricians and Gynecologists of Canada (SOGC)	Yes	Yes	Yes	Yes	Yes	Yes	Yes		Yes		
Public Health Agency of Canada (PHAC)	Yes	Yes	Yes	Yes	Yes	Yes	Yes		Yes	Yes	Yes
European STD Guidelines (IUSTI Europe)	Yes	Yes		Yes	Yes			Yes			
British Association for Sexual Health and HIV (BASHH)#	Yes	Yes	Yes		Yes	Yes			Yes		
Australasian Society for HIV Medicine (ASHM)#	Yes			Yes	Yes	Yes			Yes		

*The following are American organisations: USPSTF, AAFP, USCDC, ACOG; the following are Canadian organisations: SOGC, PHAC; the following are European organisations: IUSTI, BASHH; and ASHM is an Australian organisation

**The following organisations listed risk factors for women only: USPSTF, AAFP, USCDC, ACOG, PHAC, and SOGC. The IUSTI Europe ASHM and BASHH reported screening recommendations for males and females combined

#Risk factors listed are for STDs in general, not just for chlamydia and gonorrhoea

§Drug and alcohol use during sexual relations

Table 2.3 Percentage distribution of studies included in the literature review, by level of empirical support, according to selected characteristics (n=75 studies)

Characteristic	N	Unacceptable (%)	Marginal (%)	Acceptable (%)	Commendable (%)
STI					
Chlamydia only	51	9.8	25.5	39.2	25.5
Gonorrhoea only	6	0.0	33.3	66.7	0.0
Chlamydia and GC	14	0.0	14.3	35.7	50.0
Other**	4	0.0	25.0	50.0	25.0
Gender					
Females only	24	4.2	25.0	29.2	41.7
Males only	15	20.0	26.7	33.3	20.0
Both (with stratification*)	20	0.0	15.0	55.0	30.0
Both (without stratification)	16	6.3	31.3	50.0	12.5
Country					
Australia	7	14.3	28.6	28.6	28.6
Italy	3	0.0	33.3	0.0	66.7
Canada	2	0.0	50.0	50.0	0.0
Netherlands	3	0.0	0.0	0.0	100.0
Switzerland	2	0.0	0.0	100.0	0.0
UK	11	9.1	54.5	27.3	9.1
USA	36	2.8	13.9	50.0	33.3
Other	12	16.7	25.0	50.0	8.3
Setting					
Population/registry	16	0.0	25.0	43.8	31.3
Emergency room	3	0.0	0.0	0.0	100.0
Sexual health clinic	11	0.0	27.3	45.5	27.3
Family planning clinic	5	0.0	40.0	40.0	20.0
GP clinic	4	0.0	75.0	0.0	25.0
HIV study	3	0.0	33.3	0.0	66.7
Military	4	25.0	0.0	50.0	25.0
Detention facility	11	9.1	9.1	63.6	18.2
Other	18	16.7	22.2	44.4	16.7
Study design					
Case control	5	0.0	20.0	60.0	20.0
Cross sectional	65	7.7	24.6	43.1	24.6
Prospective	5	0.0	20.0	0.0	80.0
Presence of symptoms					
Asymptomatic only	6	0.0	66.7	16.7	16.7
Symptomatic and asymptomatic	69	7.2	20.2	43.5	29.0

* Reported separate stratified models or an interaction term

**Other includes combinations of Chlamydia, gonorrhoea, syphilis and HIV outcomes

Table 2.4 A summary of the level of empirical support for variables associated with STD outcomes according to the ratings of methodological quality (n=70 studies)

	*Proportion of studies (n) with significant findings according to ratings**								
	Marginal		Acceptable		Commendable		Total		*Level of empirical support*
	N	%	N	%	N	%	N	%	
Underlying determinants									
Age (at time of study)	20	55.0	44	59.1	20	85.0	84	64.3	Strong evidence
Age (at sexual debut)	1	0.0	9	55.6	4	25.0	14	42.9	Mixed evidence
Race/Ethnicity	10	70.0	37	67.6	18	72.2	65	69.2	Moderate evidence
Socio-economic status	1	0.0	11	18.2	9	55.6	21	33.3	Weak evidence
Marital status	3	100.0	10	20.0	4	25.0	17	35.3	Weak evidence
Drug/alcohol use	7	0.0	12	33.3	3	33.3	22	22.7	Weak evidence
Sex in exchange for gifts/money	2	0.0	6	16.7	--	--	8	12.5	Weak evidence‡
Proximate determinants									
Exposure of susceptible to infected people									
Multiple lifetime sexual partners	1	100.0	7	57.1	9	100.0	17	82.4	Strong evidence
Multiple sex partners in recent months	12	83.3	29	55.2	14	64.3	55	63.6	Moderate evidence
Sexual partners from overseas	3	0.0	4	25.0	1	100.0	8	25.0	Weak evidence‡
Sex with regular partners	4	25.0	3	33.3	3	0.0	10	20.0	Weak evidence‡
Sex with casual partners	3	33.3	6	66.7	2	0.0	11	45.5	Mixed evidence‡
Sex with new partners	5	100.0	9	22.2	10	60.0	24	54.2	Mixed evidence
Sex with infected partner	4	100.0	12	75.0	6	100.0	22	86.4	Strong evidence
Sex with symptomatic partner	1	100.0	5	80.0	5	100.0	11	90.9	Strong evidence‡
History of STI diagnosis	7	57.1	22	45.5	12	33.3	41	43.9	Weak evidence
Efficiency of transmission per contact									
Concurrent STI diagnosis	5	100.0	4	75.0	2	100.0	11	90.9	Strong evidence‡
Condom use	12	25.0	19	26.3	11	36.4	42	28.6	Weak evidence

*Some articles were counted more than once if they conducted separate analysis by gender or STI

‡Insufficient research

Chapter 3: A critical appraisal of risk models for predicting sexually transmitted infections

Falasinnu T, Gustafson P, Hottes TS, Gilbert M, Ogilvie G, Shoveller, J. A critical appraisal of risk models for predicting sexually transmitted diseases. Sex Transmitted Diseases *2014; May;41(5):321-30*

3.1 Introduction

Internet-based screening programs have been proposed as potentially cost-reducing alternative service delivery models in recent years [26]. These new models are facilitated in part by the increasing adoption of highly sensitive and specific STI testing technologies, including Nucleic Acid Amplification Tests (NAAT) conducted on urine specimens and rapid testing, which make the provision of testing services more convenient and accessible in these alternative settings where patients have limited interaction with clinicians [27]. However, strategies for identifying those at increased risk for STIs are urgently needed in alternative service delivery models. Current screening recommendations issued by public health and professional organisations have deficiencies (e.g., over-simplification, lack of generalisability) that limit their adoption for individualised decision-making to the specific patient, and there is growing support for more nuanced risk assessment considerations that involve more than the body of evidence concerning screening recommendations alone [52].

Risk prediction approaches that reflect a continuous risk spectrum have been widely adopted in chronic disease decision-making (e.g., the Framingham risk score) and have been proposed as alternatives to screening recommendations in various contexts [122]. Clinical prediction rules (CPRs) are tools that provide estimates of absolute risk based on the combination of several patient characteristics, thus allowing for more nuanced and precise decision-making than screening recommendations when applied to individual patients [55]. To become routinely incorporated into sexually health service delivery, CPRs must demonstrate good performance and generalisability. To date, there has been no review undertaken providing a critical appraisal of the methodological quality and performance measures of CPRs in sexual health contexts. In the review presented in this chapter, the research methods in the CPRs used for sexual health services are investigated to identify key variables, measures, and methods of analysis. The aim is to identify key methodological strengths and weaknesses in this area of research and examine how research practices differ across STIs, time periods or settings. Specifically, the objectives of this review were to: (1) identify and characterise

prediction rules developed and validated for STI screening; (2) describe and critically appraise the methodological issues essential to the suitability of derived models for clinical or public health application; and (3) synthesise the literature on the performance of these models.

3.2 Methods

3.2.1 Search strategy

A search of OVID Medline was conducted to identify potentially relevant studies published between 2003 and 2012, an era characterised by the expansion of highly sensitive tests (e.g., Nucleic Acid Amplification Tests) and convenient specimen collection (e.g., urine) in an effort to ensure comparability in specimen collection and STI outcome assessment. Because of the lack of MeSH terms denoting CPRs, the following terms were used: a combination of 'chlamydia', 'gonorrhoea, 'HIV', 'syphilis', 'STI', or 'STD'; and 'screening'. This review focused on these four STI outcomes because they are of particular interest to public health diagnosis and treatment. This review was also limited to studies published in English language journals with populations derived from North America, Western Europe, and Australia – regions with comparable STI prevalence and social determinants of sexual health. A search of the grey literature was not performed in order to focus on studies that are published in journals read by sexual health clinicians and decision-makers.

The inclusion of articles was assessed through a three-step process. First, the title and abstract of each article were evaluated after being identified through the OVID Medline search for relevance. Second, the complete manuscript of articles identified as relevant through the title and abstract review were manually reviewed. Eligible publications had to report a scoring or assessment tool for risk prediction or stratification. Finally, the references of identified articles were reviewed to find additional articles that may have been missed in the initial electronic search. Only publications that provided measures of association between variables and STI outcomes in the final multivariable prediction model were included. Reporting of at least one empirical performance metric of the prediction model (e.g., discriminative and/or calibration measures) was preferable but not crucial for inclusion.

3.2.2 Assessment of methodological quality

For each article concerned with CPR derivation and validation, study characteristics such as STI outcome, objectives, study setting, predictors, and the presentation format were extracted. Next, the assessment of methodological quality was limited to only studies that derived prediction models. Two reviewers (T.F. and J.S.) performed data extraction and quality assessment independently. Table 3.1 shows the 16-item quality assessment checklist derived from assessment tools previously applied in this area [73,164]. The presence or absence of each item was recorded as a score of 1 or 0, respectively; the maximum total score was 16. The reviewers resolved scoring discrepancies through deliberation and arrived at consensus.

3.2.3 Assessment of CPR performance measures

To evaluate the discriminative performance of both derived and validated models, the area under the receiver operating characteristic curve (AUC) or C-statistic, which indicates the ability of CPRs to discriminate between patients with or without STI outcomes were extracted [27]. For model calibration, an assessment of whether the publication reported on the difference between the observed and predicted rates of the STI outcome, as well as the corresponding test statistic or p-value was conducted [58]. Calibration metrics assess the ability of a CPR to accurately predict the level of observed risk [58]. To define the performance of the prediction models, data on their sensitivity and efficiency were extracted. Sensitivity was defined as the percentage of cases detected and efficiency was the percentage of patients that would have been tested based on the predictive criteria. As initially defined by La Montagne and colleagues, thresholds of 60% efficiency and 90% sensitivity were used as ideal benchmarks for CPR performance [51,96]. Lower percentage values of efficiency at higher sensitivity values are deemed to be better indicators of performance. Thus, performance was considered acceptable if >90% of cases were detected while testing <60% of patients [51,96].

3.3 Results

Figure 3.1 provides an overview of the article selection process used in this review. The review identified 216 potentially relevant articles after screening the title and abstract of the 841 articles initially identified from the OVID Medline search. Of these, 78 full-text articles examining

predictors of STIs were evaluated for inclusion; however, only 16 studies evaluated derived and/or validated prediction rules in sexual health contexts [27,51,58,72,96,99,118,119,122,136,145,153,154,165-167].

3.3.1 Characteristics of studies included

Table 3.2 summarises data from studies that derived and validated STI prediction rules. Twelve studies evaluated chlamydia outcomes [27,51,96,99,118,119,122,136,145,153,154,167], three studies evaluated HIV outcomes [58,165,166], and four studies evaluated gonorrhoea outcomes [72,136,153,154]. Three of the included studies were population-based studies [27,119,145]. Half of the studies were conducted in STI (n=5) [51,58,96,122,165] and family planning (n=3) [51,72,99] clinics. There was a significant amount of variation in the risk factors included in the prediction models; however, there were some predictors (e.g., age, clinical symptoms, and number of sexual partners) that were common to all CPRs.

All studies except for one developed or validated CPRs using cross-sectional data. Trick and colleagues developed a CPR for screening chlamydia and gonorrhoea infection in a detention facility in the United States using case-control data [136]. Logistic regression was used to derive all the CPRs in this review. The review identified two distinct CPR presentation formats: point-based scoring systems (n=6) [27,58,118,122,145,153] and predictions by predictor combinations (n=10) [51,72,96,99,119,136,154,165-167]. Point-based scoring systems were derived from the beta-coefficients of predictors in multivariable regression models; for example, a selective screening criteria based on this presentation format may categorise individuals into risk groups based on risk scores derived from a logistic regression equation, with higher risk scores correlating with higher probability of infection. In contrast, predictions made by predictor combinations were derived from unweighted checklists of variables; for example, a selective screening criteria based on this presentation format may categorise individuals into risk groups based on the presence or absence of number of risk factors such as younger age and multiple sexual partners.

3.3.2 Evaluation of methodological quality

Table 3.3 summarises the methodological quality of the CPR derivation studies included in this review. Studies met between 9 and 15 quality items. All studies had adequate study design, study

sample description, statistical methods, model selection methods, and ease of use. The most poorly addressed CPR quality items were missing values (n=3) [58,96,165], calibration measures (n=5) [27,58,122,166,167], and variable definition (n=8) [58,96,99,118,119,153,165,166]. The highest quality studies were two studies that developed CPRs for HIV screening, with each meeting 15 of 16 quality items [166]. In general, studies that developed scoring systems met more quality items (median, 13; range, 12-15) than studies whose CPRs were based on predictor combinations (median, 11.5; range, 9-15). Studies published in the past 5 years met more quality items (median, 13; range, 12-15) than older studies (median, 11; range, 9-15).

3.3.3 Assessment of rule performance

Table 3.4 shows the performance of derived STI prediction rules. STI prevalence ranged from 0.02% (for acute HIV) [166] to 29.2% (for chlamydia/gonorrhoea) [154]. Six studies were not internally validated (i.e., apparent validation) [51,72,96,99,136,167]. Five prediction rules were internally validated using split-sample validation [119,154,165,166], three studies used bootstrapping validation [27,118,153], and one study used cross-validation [58]. Eleven studies reported AUC ranging from 0.64 to 0.88, indicating a modest-to-good discriminatory performance [27,58,118,119,122,136,145,153,165-167]. Three studies reported adequate calibration statistics in one of two ways by: (1) calculating the Hosmer-Lemeshow goodness-of-fit test; or (2) graphically comparing predicted STI prevalence with observed STI prevalence, drawing a linear regression line through the points, and calculating its R^2 and slope [27,58,122].

A comparison of the efficiency of CPRs was conducted by examining the proportion of the population tested when the model's sensitivity is greater than 90% (Table 3.4). Efficiencies at this sensitivity cutoff varied between 42% and 100%. Eight studies attained or were close to attaining the performance benchmark of testing less than 60% of the target population or sub-populations [27,51,72,96,119,165-167]. The review found good discrimination was correlated with higher efficiencies. The most efficient study developed a strategy for the selective testing of women for chlamydia in general practice in Belgium [167]. The authors identified the following testing algorithm: test women aged <35 years with >1 partner in the past year; or test women with any two of the following characteristics: age 18-27 years; frequent postcoital bleeding; no contraception; partner(s) with urinary complaints [167]. This algorithm detected 92% of infections and only 38% of the population was tested. The AUC was 0.88, indicating excellent discrimination accuracy [167].

3.3.4 Validation of STI prediction rules

Table 3.5 shows the results of the external validation of STI prediction rules. Two prediction rules were externally validated. Gotz and colleagues developed a predictive rule for screening chlamydia in the general population in the Netherlands and the following predictors comprised the final model: age; level of residential urbanisation; ethnicity; urogenital symptoms; lifetime number of sexual partners; new partner in previous 2 months; and no condom at last sexual encounter [27]. The Dutch chlamydia risk score was validated in two additional settings: (1) a population-based study in Amsterdam; and (2) an outreach screening project among high-risk youth in Rotterdam [145]. Also, the risk score developed by Haukoos and colleagues helped identify patients with increased probability of diagnosed HIV infection in an STI clinic in Denver, Colorado was validated in an emergency department in Cincinnati, Ohio [58]. The Denver HIV risk score comprised the following predictors: age, gender, race/ethnicity, sexual practices, injection drug use, and past HIV testing [58].

Overall, the AUC in validation studies (0.66 to 0.75) was lower than that of derivation populations. The difference in the AUC ranged from -0.13 to -0.10 (Table 3.5) indicating poorer discrimination in validation populations [58,145]. Calibration and efficiency measures were also worse in the validation populations [58,145].

3.4 Discussion

This review identified 16 publications involving the development and validation of 15 CPRs for STI testing. Using a framework for evaluating the quality of the CPRs, the identified studies were scored based on the presence or absence of 16 items commonly deemed to be of importance for high quality prediction rules. The review evaluated the discrimination and calibration performance of the CPRs identified. The review also assessed the rules based on their ability to test <60% of the population while detecting >90% of infections. Here, the key findings of this review are contextualised.

The considerable variability in the presentation format makes comparing the existing CPRs for STI testing challenging. In this review, 10 studies made predictions about infection using predictor combinations while the rest issued scoring systems. All CPRs derived from scoring systems included a discussion of a clinical scenario in which decision-making may be influenced by

the predicted risk, provided a decision aid to stratify patients by risk, and also explicitly detailed a risk threshold for testing based on risk scores. It is recommended that future CPR derivation studies adopt scoring systems based on several considerations. First, scoring systems mirror the probability of infection, thus, allowing for personalised predictions at the individual level. Second, although the scoring systems are generally used to identify those at increased risk, they may also be used to identify those at low risk, thus limiting the number of tests performed and reducing the likelihood of false-positive tests in scenarios such as low prevalence settings where the yield of new diagnoses is low [58]. Third, scoring systems may have the additional advantage of informing clinicians and patients about personalised risk allowing both to share in STI testing decision-making [58].

The quality of the methodological approaches used to derive and validate CPRs determines their performance and future applicability. Although multiple areas of strength were noted, a number of methodological weaknesses were common in the reviewed CPR studies. These were inadequate handling of missing values, absence of calibration measures, and inadequate variable definition. There is considerable evidence that CPRs that fall short of these quality criteria are likely biased [168]. First, missing values in clinical data rarely occur completely at random [169]. Commonly missing values are often correlated with predictor and outcome measures. Most studies included in this review excluded participants with missing values, which may have led to not only loss of statistical power, but also selection bias of subjects [169]. Second, calibration measures are essential in model validation studies to ascertain whether the predicted probabilities equal observed probability of the outcome in consideration [169]. For example, consider a scenario where an individual presenting for testing in an online setting or triaging at a clinic may be interested in knowing their STI risk. The AUC, reported twice as often as calibration measures in studies included in this review and a measure of discriminatory accuracy, may not be optimal in stratifying individuals into risk categories [170]. In this situation, the agreement between predicted risk and actual risk is paramount and the absence of calibration metrics makes it impossible to determine if the CPR gives an accurate assessment of risk. Third, without a clear description of how CPR variables were assessed or defined, it is difficult for clinicians to reproduce these predictors in risk assessment, thus affecting the real-world application of the CPR [168].

Clinical usefulness of CPRs in this review was measured using sensitivity and efficiency. The ideal scenario was to identify the most cases while testing the fewest number of people [51]. There is a natural tension between finding the most cases and limiting the number of people unnecessarily

tested. In scoring systems, this relationship is closely tied to the identification of the most efficient cutoff level and depends on costs and priorities. The cutoff level also depends on the context; for example, in systematic population-based screening programs where the screening of high-risk individuals is priority, missed infections are unavoidable and the 90% sensitivity and 60% efficiency benchmark may be appropriate. If identifying the most cases is priority, then the cutoff point will be lower. Indeed, proving cost-effectiveness of CPRs is highly relevant in the current era in which HIV testing is being expanded to include individuals in settings (e.g., general practice) that would be expected to be at lower HIV risk than individuals that are currently offered in high-risk settings (e.g., STI clinics) [58].

Evidence showing performance of a CPR in new populations is an important consideration before recommending its widespread adoption. The review identified only two CPRs that have been validated in different populations [58,145]. External validation (e.g., using populations from a different time period or a different geographical setting) is essential to provide good estimates of these rules' performances in new settings [58,145]. It is recommended that more validation studies of existing CPRs should be performed, ideally conducted by independent investigators, to ensure their generalisability [171]. Also, decision-makers may consider incorporating or updating features of existing rules instead of developing new models in their own setting [171]. This approach minimises unnecessary derivation of new rules in addition to providing estimates of the performance of existing CPRs in different settings [171].

This review has several limitations. The overall findings may have been impacted by the inconsistent quality of the included publications and a small overall sample size. The electronic OVID Medline search may have not identified all CPRs used for STI testing. However, the search process was conducted on several occasions to guarantee that all studies were identified. Also, a review of the bibliographies of all identified CPR publications revealed no additional articles[164]. The equal weights assigned to the items on methodological quality checklist ignored the possibility that some items are more important than others and also failed to acknowledged the subjectivity of the items; however, two reviewers reached consensus on presence or absence of each quality item for each study [164].

Finally, CPRs provide more information about an individual patient than even complex recommendations can accommodate. Accurate CPRs have the capability to limit resource use but are infrequently developed and uncommonly used in sexual health services. Validation and

translational efforts are required to convert existing CPRs into powerful decision aids that can improve their uptake in sexual health contexts. Future developments in the use of CPRs in sexual health practice should address their clinical consequence and comparative usefulness, external validity, and implementation impact.

3.5 Summary

With the increase in prediction research for sexual health, there has been an increased interest in the methodology of this research because poorly done or poor reported prediction research is likely to have limited reliability and applicability and will, therefore, be of little use in decision-making. It appears that research in this field – despite its proliferation in the past 15 years – holds room for improvement, and its quality could substantially be enhanced if researchers paid attention to: (1) the quality of the data they use (especially missing data); (2) the accuracy metrics used (e.g., discrimination and calibration); and (3) the internal and external validity of their findings. In this dissertation, a methodological framework (Figure 1.1) was developed to address these four issues in the empirical parts of this dissertation (Chapters 4 and 5).

Figure 3.1 The process of selecting articles for inclusion in the literature review of derived and/or validated prediction rules in sexual health settings

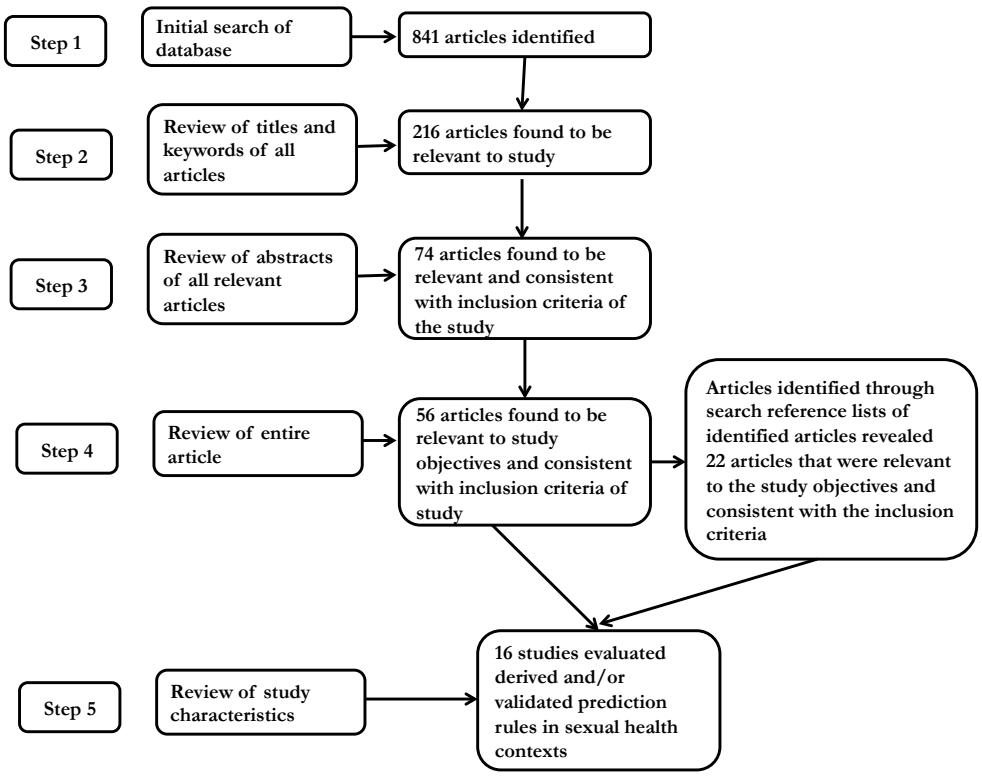


Table 3.1 A summary of the criteria used to evaluate and rate the quality of methodology for each study

Element	*Quality of methodology criteria*	*Scores Yes = 1 No = 0*
A. Study design	The study design was clearly evident (e.g., cross-sectional, cohort or case – control).	Yes/No
B. Specimen Collection	The specimen collection methodology (e.g., urine, blood, urethral/vaginal swab, etc) was clearly defined. If the same specimen collection methodology was not used consistently among study participants, methods to adjust for this difference were included in the analyses.	Yes/No
C. Study sample description	The characteristics of the sample and the selection criteria for the study were clearly stated. The characteristics included all of the following: location, time period, sampling method, sample size, entry criteria and exclusions. The study also included a description of non-respondents (e.g., those who refused to participate or did not respond to invitations).	Yes/No
D. Variable definition	Sources of data and details of methods of assessment (measurement) were given for each variable of interest. Each variable of interest was measured using the source regarded as most valid for that particular variable (e.g., medical records for STI outcomes; self-report/diary for sexual behaviours). The validity of the variables measured in the study reflects the true situation. For particular variables that are difficult to directly assess (e.g., type of sexual behaviour, condom use frequency or selectivity), the information was sought from the most objective sources (e.g., diaries).	Yes/No
E. Outcome assessment	STI outcomes were measured by the most sensitive and specific tests available at the time of data collection. Chlamydia and gonorrhoea outcomes were elicited using NAAT instead of culture or self-report. Syphilis positivity was determined using a combination of RPR and TPPA or NAAT; and HIV positivity is determined using Western Blot/ELISA.	Yes/No
F. Missing values	The authors reported missing values per predictor, or number or percentage of participants with missing values and also specified procedures for dealing with missing values	Yes/No
G. Sample size	The final sample size that was actually used in the analysis (i.e., to predict an STI outcomes) comprised of ≥10 study outcomes per variable.	Yes/No
H. Statistical methods	Multivariate analyses were used to examine the relationship between predictors or indicators and STI outcome rather than, or in addition to, univariate statistics. Methods used to examine subgroups and interactions were stated.	Yes/No
I. Selection method	A clear attempt at model selection was made. Step-wise procedures (e.g., backward selection) or an a priori specification of p-value cutoff points for variable inclusion were considered acceptable for this criterion.	Yes/No
J. Calibration measures	The authors reported the rule's agreement between predicted outcome and observed outcome (e.g., calibration plot, calibration intercept and slope, or the Hosmer-Lemeshow statistic)	Yes/No
K. Discrimination measures	The authors reported measures that estimate the rule's ability to discriminate between those with and without the STI outcome of interest (e.g., C-statistic/AUC-ROC)	Yes/No

Element	*Quality of methodology criteria*	*Scores* *Yes = 1* *No = 0*
L. Classification measures	Classification measures such as the rule's sensitivity and specificity were reported	Yes/No
M. Validity assessment	The predictive performance of the prediction model was assessed using validation (e.g., split sample, cross-validation, or bootstrapping) techniques	Yes/No
N. 95% CI of rule properties	The authors reported the degree of uncertainty around the calibration, discrimination, and classification measures	Yes/No
O. Clinical sensibility	Predictors, outcome, and predictive values were clinically meaningful	Yes/No
P. Ease of use	Ease of use in immediate setting	Yes/No

Table 3.2 Study characteristics of CPRs derived and/or validated for sexual health services

Article	STI outcome	Country	Objectives	Setting	Risk factors included in CPR	Presentation format of the CPR
Al-Tayyib et al, 2008 [153]	Chlamydia, Gonorrhoea	USA	To develop and evaluate screening algorithms to predict current chlamydial and gononococcal infections in an emergency department (ED)	Emergency department	Age, marital status, education, new sex partner within past 2 years, non-ED as primary source for healthcare, non antibiotic use past month, dysuria or discharge past 3 months	Point-based scoring system
Andersen et al, 2003 [119]	Chlamydia	Denmark	To evaluate self-reportable predictive criteria based on sexual behaviour, lifestyle, and sociodemographic characteristics to limit the number of individuals needed to be screened in a population-based screening program for young men and women	Population-based register	Age, more than 3 sex partners in lifetime, sex partner known to have urogenital complaints, previous positive chlamydia test result, no condom use during most recent intercourse, more than 6 months since most recent intercourse	Predictions based on predictor combinations
Facente et al, 2011 [165]	HIV	USA	To develop and evaluate formal criteria for targeting individuals at high risk of acute HIV infection among patients presenting for HIV testing at a municipal STI clinic	STI clinic	Unprotected receptive anal intercourse, HIV+ partner, injection drug use, MSM, recent STI	Predictions based on predictor combinations
Gotz et al, 2005 [27]	Chlamydia	Netherlands	To develop a prediction rule for estimating the risk of chlamydial infection as a basis for prediction rule	Population-based register	Age, residence, ethnicity, education, urinogenital symptoms, lifetime number of partners, new partners in previous 2 months, no condom last sexual contact	Point-based scoring system

Article	STI outcome	Country	Objectives	Setting	Risk factors included in CPR	Presentation format of the CPR
Gotz et al, 2006 [145]	Chlamydia	Netherlands	To assess the validity of the previously developed prediction rule	Population based screening program and a community outreach testing project	Age, residence, ethnicity, education, urinogenital symptoms, lifetime number of partners, new partners in previous 2 months, no condom last sexual contact	Point-based scoring system
Haukoos et al, 2012 [58]	HIV	USA	To derive and validate an instrument to accurately identify patients at risk for HIV infection	STI clinic	Age, gender, ethnicity, sex with a male, vaginal intercourse, receptive anal intercourse, injection drug use, past HIV testing	Point-based scoring system
Hocking et al, 2005 [96]	Chlamydia	Australia	To estimate chlamydia prevalence and risk factors for infection and to assess the performance of chlamydia-screening criteria among clients attending a large STI clinic	STI clinic	Age, condom use <50% of time in last 3 months, 2+ sexual partners in last 12 months, contact with infection, injection drug use, present with symptoms of NSU	Predictions based on predictor combinations
La Montagne et al, 2004 [51]	Chlamydia	USA	To assess: (1) the performance of screening criteria in the North-western USA, (2) predictors of CT infection, and (3) optimization of these criteria	STD clinics, Family Planning clinics, and other clinics	Age, race, specimen type, test type, abnormal appearance of cervix, mucopurulent cervicitis, friability, ectopy, PID, exposure to CT, exposure to GC, exposure to NGU, sex with symptomatic partner, CT last 12 months, new sex partner, multiple sex partners	Predictions based on predictor combinations
Manhart et al, 2007 [72]	Gonorrhoea	USA	To develop selective testing criteria for gonorrhoea among young women screened for chlamydial infection	Family planning clinics, women's clinics, and school based clinics	Exposed to an STD other than chlamydia, symptomatic partner, dysuria, presumptive treatment for CT infection, abnormal vaginal discharge, pregnancy related visit, new partner during the preceding 60 days, black/Native American, clinic in top quartile of male urethral GC infection rates	Predictions based on predictor combinations

Article	STI outcome	Country	Objectives	Setting	Risk factors included in CPR	Presentation format of the CPR
Merchant et al, 2010 [154]	Chlamydia, Gonorrhoea	USA	To create a system for predicting which male emergency department (ED) patients with suspected CT/GC would have laboratory confirmed infections based on clinical factors	Emergency department	Age, race, no insurance status, any genital discharge, any Chlamydia and/or gonorrhoea contact	Predictions based on predictor combinations
Miller et al, 2009 [166]	HIV	USA	To develop risk assessment model for acute HIV infection	HIV counseling and testing sites	Race, MSM, test site, sex with HIV+ person, local HIV prevalence	Predictions based on predictor combinations
Paukku et al, 2003 [99]	Chlamydia	Finland	To find screening criteria for a population with low prevalence of genital Chlamydia infection	Family planning clinic and Student health clinic	Contraception, marital status, casual sex, 3+ sexual partners in past year, 3+ lifetime number of partners	Predictions based on predictor combinations
Stein et al, 2008 [118]	Chlamydia	USA	To develop predictive guidelines appropriate for chlamydia appropriate for community settings	High schools	Race/ethnicity, partner's race/ethnicity, number of partners in previous year, perceived risk of prevalent STI, student status, age, marital status, pregnancy history, hormonal contraception	Point-based scoring system
Trick et al, 2006 [136]	Chlamydia, Gonorrhoea	USA	To determine risk factors for chlamydia or gonorrhoea urethral infection among adult male detainees	Jail	Age, number of partners in past year, most recent sexual encounter, drug use within previous months, symptoms, STD in previous 5 years, condom use	Predictions based on predictor combinations
Verhoeven et al, 2003 [167]	Chlamydia	Belgium	To estimate the prevalence of CT in women in general practice and to assess risk factors associated with infection	General practices	Age, 1+ partner in past year, no contraception, dysuria/frequent urination, partner having urinary complaints, frequent postcoital bleeding, location of practice (inner city)	Predictions based on predictor combinations

Article	STI outcome	Country	Objectives	Setting	Risk factors included in CPR	Presentation format of the CPR
Wand et al, 2011 [122]	Chlamydia	Australia	To develop and validate a risk scoring tool to identify those who are at increased risk of chlamydia infection	STI clinic	Age, marital status, residence, culturally and linguistically diverse, backpackers, current sex worker, current smoker, genital or anal symptoms, STI contact, STI test, contraception use, inconsistent condom use, past chlamydia infection, 1+ sexual partners in previous 3 months, knowledge of HIV status	Point-based scoring system

Table 3.3 Assessment of the methodological quality of CPR derivation studies*

Authors	Study design	Specimen Collection	Study sample description	Variable definition	Outcome assessment	Missing values	Events per variable	Statistical methods	Selection method
Al-Tayyib et al, 2008 [153]	X		X	X	X		X	X	X
Andersen et al, 2003 [119]	X	X	X	X	X		X	X	X
Facente et al, 2011 [165]	X	X	X	X	X	X	X	X	X
Gotz et al, 2005 [27]	X	X	X		X		X	X	X
Haukoos et al, 2012 [58]	X	X	X	X	X	X	X	X	X
Hocking et al, 2005[96]	X	X	X	X	X	X		X	X
La Montagne et al, 2004 [51]	X		X		X		X	X	X
Manhart et al, 2007 [72]	X		X		X		X	X	X
Merchant et al, 2010 [154]	X	X	X		X		X	X	X
Miller et al, 2009 [166]	X	X	X	X	X		X	X	X
Paukku et al, 2003 [99]	X	X	X	X	X			X	X
Stein et al, 2008 [118]	X	X	X	X	X		X	X	X
Trick et al, 2006 [136]	X	X	X		X		X	X	X
Verhoeven et al, 2003 [167]	X	X	X		X		X	X	X
Wand et al, 2011 [122]	X		X		X		X	X	X
Total	15	11	15	8	15	3	13	15	15

*X=Quality item is present

Table 3.3 continued

Authors	Calibration measures	Discrimination measures	Classification measures	Validity assessment	95% CI of rule properties	Clinical sensibility	Ease of use	Total
Al-Tayyib et al, 2008 [153]		X	X	X	X	X	X	13
Andersen et al, 2003 [119]		X	X	X	X	X	X	14
Facente et al, 2011 [165]		X	X		X	X	X	14
Gotz et al, 2005 [27]	X	X	X	X	X	X	X	14
Haukoos et al, 2012 [58]	X	X		X	X	X	X	15
Hocking et al, 2005 [96]			X			X	X	11
La Montagne et al, 2004 [51]			X			X	X	9
Manhart et al, 2007 [72]			X			X	X	9
Merchant et al, 2010 [154]		X	X	X	X	X	X	13
Miller et al, 2009 [166]	X	X	X	X	X	X	X	15
Paukku et al, 2003 [99]			X			X	X	10
Stein et al, 2008 [118]		X	X	X	X		X	13
Trick et al, 2006 [136]		X	X			X	X	11
Verhoeven et al, 2003 [167]	X	X	X			X	X	12
Wand et al, 2011 [122]	X	X	X	X		X	X	12
Total	5	11	14	8	8	14	15	

*X=Quality item is present

Table 3.4 Assessment of the performance measures of CPRs derived for sexual health services

Article	STI**	STI prevalence	Method of internal validation	Predictive criteria	AUC	Calibration#	% needed to test to achieve 90% sensitivity
Al-Tayyib et al, 2008 [153]							
CT and/or GC in females	CT, GC	346/3,080 (11.2%)	Bootstrap	Risk score ≥2 out of 7 points	0.64	-	<83%
CT only in females		230/3,080 (7.5%)		Risk score ≥3 out of 7 points	0.69	-	<79%
CT and/or GC in males		188/2,457 (7.7%)		Risk score ≥3 out of 7 points	0.72	-	<73%
CT only in males		161/2,457 (6.6%)		Risk score ≥3 out of 7 points	0.71	-	<72%
Andersen et al, 2003 [119]	CT	127/1,920 (6.6%)	Split-sample	≥2 of 3 risk factors	0.68	-	>63%[A]*
Facente et al, 2011 [165]	HIV	137/12,353 (1.1%)	Split-sample	Engaged in receptive anal intercourse	0.67	-	>55%[B]*
Gotz et al, 2005 [27]	CT	160/7,005 (2.3%)	Bootstrap	Risk score ≥7 out of 14 points	0.79	H-L test = 0.5	54%[C]*
Haukoos et al, 2012 [58]	HIV	504/92,635 (0.54%)	Cross-validation	Risk score ≥20 out of 81 points	0.85	Slope=0.95, R^2=0.94	>50%[D]*
Hocking et al, 2005 [96]							
Women: Non-sex workers	CT	82/2,084 (4.0%)	Apparent	Any 2 of 5 risk factors	-	-	>56%[E]*
Men: MSM		56/614 (9.1%)		Screen all	-	-	100%
Men: Non-MSM		138/2,028 (6.8%)		Any 2 of 5 risk factors	-	-	>62%[F]*
La Montagne et al, 2004 [51]							
STD clinics	CT	2,499/34,228 (7.3%)	Apparent	Age ≤24 or age>24 if any clinical signs	-	-	64%*
Family planning clinics		12,471/304,183 (4.9%)		Age ≤24 or age>24 if any clinical signs or exposure to STI	-	-	75%
Other clinics		2,716/71,471 (3.8%)		Age ≤24 or age>24 if any clinical signs or behavioural risk	-	-	72%
Manhart et al, 2007 [72]	GC	173/3,278 (0.3%)	Apparent	Any 1 of 8 risk factors	-	-	>52%[G]*
Merchant et al, 2010 [154]	CT, GC	240/822 (29.2%)	Split-sample	≥3 of 5 risk factors	-	-	-
Miller et al, 2009 [166]	HIV	44/222,975 (0.02%)	Split-sample	Any 1 of 5 risk factors	0.86	-	<42%[H]*
Paukku et al, 2003 [99]	CT	42/1,198 (3.5%)	Apparent	Age ≥30	-	-	-
Stein et al, 2008 [118]							
Women	CT	316/5,854 (5.1%)	Bootstrap	Predicted probability cutoff at 2%	0.70	-	<70%
Men		236/5,074 (3.9%)		Predicted probability cutoff at 1%	0.69	-	<70%
Trick et al, 2006 [136]	CT, GC	348/5,634 (6.2%)	Apparent	Test all	0.65	-	100%
Verhoeven et al, 2003 [167]	CT	39/774 (5.0%)	Apparent	Age<35 with >1 partner or age 18-27 with >1 risk factors	0.88	-	<38%[I]*
Wand et al, 2011 [122]							
MSM	CT	609/10,154 (6.0%)	Split-sample	Risk score ≥15 out of 40 points	0.71	H-L test<0.21	<93%
Heterosexual males		1,167/16,667 (7.0%)		Risk score ≥20 out of 40 points	0.74	H-L test<0.21	>88%
Heterosexual females		954/19,081 (5.0%)		Risk score ≥15 out of 40 points	0.72	H-L test<0.21	89%

*CPR attained or was close to attaining the performance benchmark of testing less than 60% of the target population or sub-populations

A: Proportion tested is actually 63% to detect 86% of infections; B: Proportion tested is actually 55% to detect 89% of infections; C: value interpolated between 45.0% needed to achieve 86.8% sensitivity and 62.4% needed to achieve 93.1% sensitivity; D: Proportion tested is actually 50% to detect 89% of infections; E: Proportion tested is actually 56% to detect 86% of infections; F: Proportion tested is actually 62% to detect 88% of infections; G: Proportion tested is actually 52% to detect 89% of infections; H: Proportion tested is actually 42% to detect >90% of infections; I: Proportion tested is actually 56% to detect 86% of infections; I: Proportion tested is actually 38% to detect 92% of infections

**CT=Chlamydia; GC=Gonorrhoea

#Calibration can be measured using (1) the Hosmer-Lemeshow (H-L) goodness-of-fit test where p-values greater than 0.05 indicate poor goodness of fit, or (2) assessing whether the calibration slope of the plotted estimated and observed probabilities is 1

Table 3.5 External validation of STI prediction rules

Article	Original model	STI*	STI prevalence	Predictive criteria	AUC	Change from original AUC	Calibration	% needed to test to achieve 90% sensitivity
Gotz et al, 2006 [145]								
Population-based program	Gotz et al, 2005	CT	52/1,413 (3.7%)	Risk score ≥7 out of 14 points	0.66	-0.13	H-L test=0.02	<77%
Community outreach project			19/152 (12.5%)	Risk score ≥9 out of 14 points	0.68	-0.11	H-L test=0.20	<76%
Haukoos et al, 2012 [58]	Haukoos et al, 2012	HIV	168/22,983 (0.7%)	Screen all	0.75	-0.10	Slope=1.07, R^2=0.98	100%

*CT=Chlamydia

Chapter 4: Deriving and validating a risk estimation tool for screening asymptomatic chlamydia and gonorrhoea

Falasinnu T, Gilbert M, Gustafson P, Shoveller, J. Deriving and validating a risk estimation tool for screening asymptomatic chlamydia and gonorrhoea. Sex Transmitted Diseases. *In press*

4.1 Introduction

Sexual healthcare is facing a critical moment as healthcare costs escalate in almost every high-income country and the recent economic crisis exacerbates the financial challenges already facing publicly funded healthcare provision [15]. In this context, improving sexual healthcare delivery is arguably of paramount interest. Moreover, inadequate access to sexual healthcare has potentially detrimental effects on individuals and the community [172]. Thus, it is imperative to adopt systems that optimise the delivery of comprehensive and high quality STI testing services while minimising public health budget demands [14]. One important feature of efficient STI control programs, especially in novel health service delivery models, such as internet-based testing programs, is making certain that those at increased risk of STIs have access to screening services because identifying and treating infections in this group can effectively terminate onward transmission and therefore prevent new cases of disease.

However, much is still unknown about how best to maximise access for those at highest risk, particularly in contexts where STI clinics are over-burdened with symptomatic clients [172]. As well, increasing access to high-risk asymptomatic clients may also confer public health benefits as undiagnosed and untreated infections can frequently progress to long-term complications such as pelvic inflammatory disease and infertility (in women) and epididymitis (in men) [71]. In order to maximise case finding in asymptomatic visits, selective screening (based on risk assessment and entails the screening of individuals who meet pre-specified criteria) may be a prudent approach because it minimises the costs associated with testing low risk individuals [27]. The current body of knowledge indicates that risk prediction rules that capture a continuous risk spectrum are excellent tools for targeted screening. However, CPRs in sexual health contexts have not been fully explored in a variety of screening scenarios, including those where STI clinic clients present in the absence of symptoms.

This chapter examines the performance of a selective screening strategy (derived from a clinical prediction rule) in identifying persons at increased risk of *Chlamydia trachomatis* and/or gonococcal infections. Specifically, this chapter aimed to derive a risk scoring tool for screening for asymptomatic chlamydia and/or gonorrhoea infection among patients seen at sexual health clinics between 2000 and 2006 (derivation population) and test the generalisability of the algorithm in a more recent time period among patients seen between 2007 and 2012 (temporal validation population).

4.2 Methods

4.2.1 Study population

This analysis used electronic medical records from asymptomatic patients tested for chlamydia or gonorrhoea between 2000 and 2012 at two sexual health clinics in Vancouver, BC. All statistical analyses were performed using SAS version 9.3. This analysis included visits among all women and heterosexual men, a population often targeted by screening recommendations issued by public health organisations. This analysis also excluded visits among persons who were sexual contacts of STI cases and follow-up visits for positive results. For this analysis, a range of demographic and behavioural information from patient visits were extracted. These included age, gender, race/ethnicity, number of sexual partners in the previous six months, condom use, injection drug use (IDU), sex with partners recruited online, sex with IDU, sex with commercial sex workers (CSW), and previous diagnosis with chlamydia and/or gonorrhoea infection.

The presence of missing data is a frequently encountered problem in the derivation and validation of prediction rules [173]. The default strategy is to delete all incomplete observation from the analysis; however, this is often a precarious and wasteful approach as variables are rarely missing at random. In this analysis, variables such as race/ethnicity, condom use, and number of sexual partners in previous 6 months have rates of missingness ranging from 8.9% to 42.2%. Imputation techniques, especially multiple imputation, have been increasingly advocated to address the issue of missing values [174-176]. This analysis imputed missing values using IVEware, a software application that performs multiple imputations of missing values [177]. In this method, imputations for each missing variable are produced based on a regression model using other variables as predictors in a cyclic manner [178]. Missing data among predictors were imputed five times using the

Sequential Regression Imputation Method (SRIM) and this resulted in estimates that were averaged using Rubin's rules [175].

In prediction modeling, the effective sample size is dependent on the number of individuals who experience the outcome of interest. To reduce the risk of false positive findings (i.e., type I errors), some authors have recommended that at least 10 individuals having the outcome of interest are needed per variable to allow for accurate prediction modeling (i.e, events per variable, or EPV) [54,179]. In this analysis, the derivation and validation populations were sufficiently powered, having 15 and 28 EPV, respectively. All associated methods were approved by the University of British Columbia's Research Ethics Board (certificate # H11-02000).

4.2.2 Derivation of risk estimation tool

The outcome measured was diagnosis with chlamydia and/or gonorrhoea infection. This analysis examined a composite of both infections because most diagnostic tests use multiplex assays that test for both infections simultaneously. This analysis used chi-square tests to analyse categorical variables and used Student's *t*-test to analyse continuous variables. Univariate logistic regression was used in the derivation population to identify potential STI risk factors. To simplify risk score generation and facilitate application in clinical and population-based settings, continuous variables (e.g., age and number of sexual partners) were categorised. Tests for interaction between gender and other risk factors were performed. Predictors found to be significant in the univariate analyses were included in the final logistic regression model using backward elimination; predictors that were included in the model had p-values<0.20. To be conservative, the final regression model included only variables with $p<0.05$ in at least one of the imputed datasets [175]. The risk factors in the final model were used to construct the equations used for the clinical prediction rule. To aid use in screening decision-making, simplified risk scores were derived by multiplying the regression coefficients (betas) by 5 and rounded to the nearest integer. Sum scores for each visit were then derived by adding up the risk scores. These sum scores are direct reflection of the probability of infection [27].

4.2.3 Performance measures

This analysis estimated the model's ability to discriminate between participants with or without infection as measured by the area under the receiver operating characteristic curve (AUC). An AUC value closer to a 100% shows that the model has excellent discriminative ability, while a value close to 50% indicates no value [55,60]. Over-fitting occurs when the same data used to create the model is also used to evaluate its performance leading to an optimistic assessment of the model's performance measures [180]. Ten-fold cross validation techniques were performed to estimate how the model will generalise to an independent population and correct for this optimism bias [180]. In addition, this analysis estimated 95% confidence intervals around the optimism-corrected AUC from the calculation of the approximate standard error of the cross-validated datasets [55].

This analysis assessed calibration performance by calculating the Hosmer-Lemeshow goodness of fit statistic, which measures whether the predicted probability of infection corresponds with the observed probability. A well-calibrated model gives a corresponding *p*-value >0.05 [145]. This analysis also assessed the calibration of the simplified risk scores by visually examining the prevalence of chlamydia and/or gonorrhoea infection in groups of the risk scores [58]. This analysis also examined the sensitivity (i.e., proportion of all cases identified) and fraction of patients that would need to be screened at different cut-offs of the risk scores. The benchmark set for a well performing tool is one that identifies more than 90% of cases while screening 60% or less of the population [51].

4.3 Results

4.3.1 Derivation population

Figure 4.1 is a flow chart showing the selection of clinic visits whose data were used in this analysis. The chlamydia and/or gonorrhoea infection rate was 1.8% in the derivation population (n=10,437). Table 4.1 shows the baseline distribution of candidate predictors. The following were the demographic characteristics of the majority of patient visits: male (67%), individuals between 30 and 39 years old (31%) and white race (74%). Individuals who reported consistent condom use during sexual contact comprised approximately 27% of clinic visits. Sexual contact with a CSW was documented in 13% of patient visits (Table 4.1).

Univariate predictors of chlamydia and/or gonorrhoea infection are shown in Table 4.2. In the derivation population, the following predictors were not significantly associated with infection and were subsequently not included in the final logistic regression model: gender, condom use, sex with partners recruited online, IDU, sex with IDU, and sex with CSW. There were no significant differences between the risk factors and the outcome by gender (i.e., no evidence of interaction effects). Table 4.3 shows the results of the final multivariable regression model used for developing the prediction rule. The model included age in years (categorised as 14-19, 20-14, 25-29, 30-39, ≥40); race/ethnicity (white or non-white); number of sexual partners (0, 1-2, ≥3); previous chlamydia diagnosis (yes or no); and previous gonorrhoea diagnosis (yes or no).

Figure 4.2 shows the receiver operating curves (ROCs) for the chlamydia and/or gonorrhoea risk estimation model in the derivation and temporal validation populations. The model demonstrated good discrimination in the derivation population (AUC=0.75, 95% CI: 0.72-0.80). As internal validation indicates an upper limit of the expected performance in new settings, the ten fold cross-validation indicated the lack of evidence for overfitting (AUC=0.74, 95% CI: 0.70-0.77). The model demonstrated strong calibration in the derivation population, indicating good fit; the Hosmer-Lemeshow χ^2 statistic was 3.4 (8 df, p=0.91). The coefficients yielded risk scores, with a minimum sum score of -2 and a maximum sum score of 26. To visualise the calibration of the prediction rule, the total sample was divided into six groups as shown in Figure 4.3 which illustrates the observed proportion of chlamydia and/or gonorrhoea infection as a function of the sum score derived from the final model. Higher sum scores were correlated with higher prevalence, further bolstering the good calibration indicated by the Hosmer-Lemeshow statistic.

The simplified risk scores can be applied for selective screening decision-making. Table 4.4 shows the screening performance estimates at different cutoff levels of the sum scores. In order to identify all cases (i.e., 100% sensitivity), approximately 97% of the population would need to be screened at a sum score cutoff point of ≥1. However, by reducing the cut off point of the risk score to ≥6, only 68% of the population would need to be screened to identify 91% of the cases, making this close to the benchmark of screening <=60% while identifying more than 90% of cases.

4.3.2 Temporal validation population

The validation sample consisted of 14,956 clinic visits, of which 2.2% were diagnosed with chlamydia and/or gonorrhoea infection. There were slight differences between the derivation and

validation populations; for example, the temporal validation population had lower overall prevalence of injection drug use, previous chlamydia diagnosis and previous gonorrhoea diagnosis. This population also had higher overall prevalence of patient visits reporting sex with partners recruited from the Internet (Table 4.1). There were no major differences between the derivation and validation populations in terms of the unadjusted odds ratios examining the associations between the predictors and the outcome, except for the number of sexual partners in previous 6 months (Table 4.2). The model demonstrated acceptable discrimination in the temporal validation population (AUC=0.64, 95% CI: 0.61, 0.67) (Figure 4.2). The model also showed good calibration upon validation (Hosmer-Lemeshow χ^2=8.8, 8 df, p=0.36). When categorised into the same six risk categories as the derivation population, chlamydia and/or gonorrhoea infection prevalence ranged from 0.1% in the lowest risk category to 16.1% in the highest risk category (Figure 4.3). When the simplified risk scores were considered in the temporal validation population, choosing the risk cutoff point of ≥6 would identify 83% of cases while screening 68% of the population (Table 4.4).

4.4 Discussion

This chapter derived and validated a risk-scoring tool for assessing the risk of chlamydia and/or gonorrhoea infection among asymptomatic women and heterosexual men accessing sexual health clinics in Vancouver, BC using predictors that are relatively easy to assess in a clinical encounter. By using two populations, an earlier time frame (2000-2006) for the development of the tool and a later time frame (2007-2012) for the validation, this chapter provides insight into the value of the tool in sexual health practice in a more recent time period. As expected, the discrimination and calibration performance of the tool were more impressive in the derivation population because the logistic regression model was built on the derivation population [60,179,181]. The model revealed no evidence for lack of fit in calibration performance in the temporal validation population. Good calibration is important for health services delivery. Applying the risk stratification system (derived from the simplified risk score) indicated that the risk of infection increased with higher risk scores. The predicted risk in each risk score category may also facilitate shared decision-making between patients and their clinicians and could be used to counsel patients.

The discriminatory performance in the temporal validation group was lower relative to the derivation population. This issue draws attention to the likelihood of over-fitting, which if present, could impact the algorithm's reliability in additional settings [60]. However, over-fitting in this

analysis was unlikely as judged by the cross-validated AUC of the derivation population, which was not different from the naïve AUC. The waning in discriminatory performance of the model could be due to the difference in case-mix between the two time periods [181]. Although the later time frame had a slightly higher prevalence of infection, individuals comprising this population reported lower proportions of the risk factors included in the final model. As a result, discrimination between cases and non-cases in the more homogeneous temporal validation population was more difficult than in the more heterogeneous derivation population [181]. It is hypothesised that the rise of less invasive and more sensitive diagnostic tests (e.g., urine-based) may account for the increased testing uptake among lower risk individuals. In addition, there is also the possibility that missed risk factors could have impacted the discriminatory performance of the model in the temporal validation population. Future studies are needed to evaluate whether additional predictors (e.g., highly detailed assessments of behavioural risk) might further improve STI prediction. In these future scenarios, the practical value and clinical usefulness of the updated risk score should be evaluated by assessing the ability of newly proposed models to improve screening decision-making processes.

The lower discriminatory performance also reduced the case detection ability (sensitivity) at each consecutive risk cutoff point of the model in the temporal validation population. Although the sensitivity value was not as high in the temporal validation population as the derivation population at each risk cutoff point, these results are an improvement to the current clinical situations in some jurisdictions that entails the universal screening of all asymptomatic patients in STI clinical contexts. The choice of the ideal risk cutoff points in decision-making depends on programmatic goals. The consequences of the cutoff points must be evaluated by weighing the potential public health and fiscal costs of missing an infection against the burden on current health human resources through extensive cost effectiveness and cost benefit analyses [27].

The analysis presented in this chapter has several strengths and limitations. The large sample sizes relative to the number of variables in the derivation and temporal validation populations are important strengths. This is the first study in sexual health contexts that conducted temporal validation of a derived prediction rule, an issue particularly salient in this field because of the shift to more sensitive diagnostic tests over this time period. However, the data from which the model was derived came from STI clinics and this could limit generalisability to other settings such as reproductive health clinics or general practice settings. As well, the algorithm is based on the assumption of a fixed epidemic; however, this assumption should be re-evaluated periodically over

time to reflect the evolution of patient risk profiles. This process could be facilitated by the adoption of electronic medical records whereby updated prediction rules can be easily re-programmed and integrated into routine clinical care. Despite these limitations, a pragmatic risk-scoring tool was derived from a model that included a diverse patient population. As funding available for sexual health services decrease (or in some cases, remain stagnant) and STI rates increase, public health programs are in need of novel strategies to maximise service efficiency. Future research is needed to determine whether the adoption of risk estimation tools, such as the one developed here, will result in economic savings or have long-term impacts on resource allocation (including health human services and clinic operations).

4.5 Summary

In this chapter, the accuracy and temporal generalisability of a prediction rule for screening asymptomatic chlamydia and/or gonorrhoea was determined. The good performance of the prediction rule "over time" indicates that sexual health decision-makers may cautiously use the prediction rule in future patients who are similar to the development and validation population. However, validation in varied study sites is still necessary before the rule can be implemented in clinical practice. Chapter 5 of this dissertation addresses this issue.

Figure 4.1 Study population selection

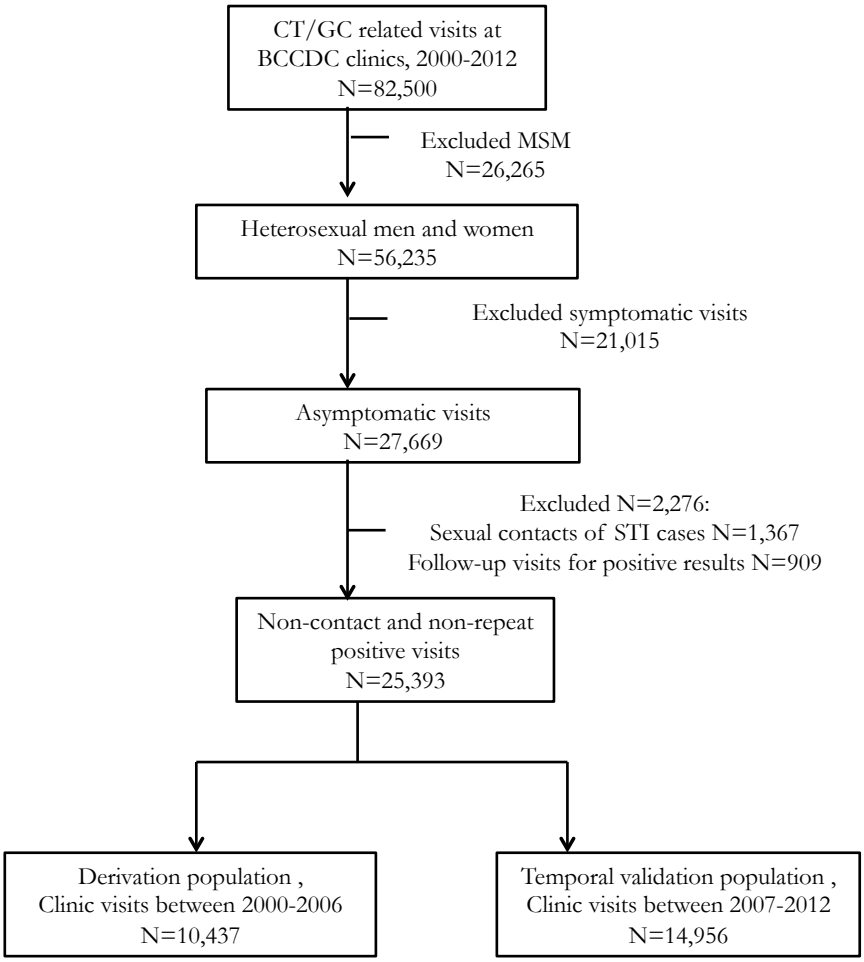


Table 4.1 Population characteristics of the derivation and temporal validation populations*

	Derivation population		Temporal validation population	
Variable	N	(%)	N	(%)
Chlamydia and/or gonorrhoea case				
Yes	184	1.8%	331	2.2%
No	10,253	98.2%	14,625	97.8%
Gender				
Female	3,496	33.5%	5,341	35.7%
Male	6,941	66.5%	9,615	64.3%
Age (yrs)				
14-19	257	2.5%	249	1.7%
20-24	1,962	18.8%	2,142	14.3%
25-29	2,651	25.4%	4,347	29.1%
30-39	3,181	30.5%	4,892	32.7%
≥40	2,386	22.9%	3,327	22.2%
Race/ethnicity				
White	7,732	74.1%	10,402	69.6%
Non-white	2,705	25.9%	4,554	30.4%
No. of sexual partners in previous 6 months				
0	644	6.2%	734	4.9%
1-2	6,857	65.7%	9,252	61.9%
≥3	2,936	28.1%	4,971	33.2%
Condom use*				
Never	2,362	22.6%	3,178	21.2%
Sometimes	5,269	50.5%	7,630	51.0%
Always	2,806	26.9%	4,148	27.7%
Sex with partners recruited online				
Yes	417	4.0%	1,375	9.2%
No	10,020	96.0%	13,581	90.8%
Injection drug use				
Yes	211	2.0%	134	0.9%
No	10,226	98.0%	14,822	99.1%
Sex with injection drug user				
Yes	455	4.4%	405	2.7%
No	9,983	95.6%	14,551	97.3%
Sex with commercial sex worker				
Yes	1,381	13.2%	1,840	12.3%
No	9,057	86.8%	13,116	87.7%
Previous chlamydia diagnosis				
Yes	1,518	14.5%	1,728	11.6%
No	8,919	85.5%	13,228	88.4%
Previous gonorrhoea diagnosis				
Yes	619	5.9%	441	2.9%
No	9,819	94.1%	14,515	97.1%

Variable	Derivation population		Temporal validation population	
	N	(%)	N	(%)
Total	10,437	100.0%	14,956	100.0%

*Missing data among the predictor variables was handled using a multiple imputation procedure with 5 resampling replications, which generated an augmented databases with (5*10,437) 52,185 and (5*14,956) 74,780 observations with complete data in the derivation and temporal validation populations, respectively. With imputed sample, we estimated baseline characteristics and developed prediction models. The average of all 5 imputed samples are shown in this table.

Table 4.2 Chlamydia and/or gonorrhoea prevalence and unadjusted ORs in the derivation and temporal validation populations*

	Vancouver clinics, 2000-2006 (N=10,437)		Vancouver clinics, 2007-2012 (N=14,956)	
Variable	%	OR (95% CI)	%	OR (95% CI)
Gender				
Female	2.1%	1.31 (0.97-1.77)	2.3%	1.08 (0.86-1.35)
Male	1.6%	Ref	2.2%	Ref
Age (yrs)				
14-19	7.4%	6.49 (3.58-11.75)	6.8%	6.28 (3.49-11.30)
20-24	2.8%	2.30 (1.46-3.63)	3.3%	2.94 (1.97-4.37)
25-29	1.8%	1.50 (0.94-2.38)	2.6%	2.28 (1.57-3.31)
30-39	1.1%	0.88 (0.53-1.45)	1.9%	1.65 (1.12-2.41)
≥40	1.2%	Ref	1.2%	Ref
Race/ethnicity				
White	1.2%	Ref	2.0%	Ref
Non-white	3.4%	2.90 (2.16-3.89)	2.8%	1.44 (1.10-1.88)
No. of sexual partners in previous 6 months				
0	0.5%	Ref	0.4%	Ref
1-2	1.8%	3.43 (1.10-10.73)	2.0%	6.48 (0.93-45.43)
≥3	2.0%	3.92 (1.23-12.45)	2.8%	8.97 (1.37-58.65)
Condom use				
Never	1.6%	Ref	2.2%	Ref
Sometimes	2.0%	1.22 (0.84-1.77)	2.4%	1.12 (0.83-1.50)
Always	1.4%	0.88 (0.56-1.39)	1.9%	0.87 (0.62-1.21)
Sex with partners recruited online				
Yes	1.2%	0.67 (0.27-1.63)	1.4%	0.61 (0.38-0.97)
No	1.8%	Ref	2.3%	Ref
Injection drug use				
Yes	1.6%	0.89 (0.26-3.06)	2.4%	1.02 (0.23-4.54)
No	1.8%	Ref	2.2%	Ref
Sex with injection drug user				
Yes	1.0%	0.55 (0.20-1.54)	1.3%	0.51 (0.09-2.78)
No	1.8%	Ref	2.2%	Ref
Sex with commercial sex worker				
Yes	1.2%	0.65 (0.35-1.20)	1.2%	0.50 (0.30-0.84)
No	1.8%	Ref	2.4%	Ref
Previous chlamydia diagnosis				
Yes	5.1%	4.40 (3.27-5.93)	4.7%	2.52 (1.95-3.26)
No	1.2%	Ref	1.9%	Ref
Previous gonorrhoea diagnosis				
Yes	2.1%	1.21 (0.69-2.14)	2.1%	0.96 (0.49-1.87)
No	1.7%	Ref	2.2%	Ref
Total	1.8%	--	2.2%	--

*Missing data among the predictor variables were handled using a multiple imputation procedure with 5 resampling replications, which generated an augmented databases with (5*10,437) 52,185 and (5*14,956) 74,780 observations with complete data in the derivation and temporal validation populations, respectively. With the imputed sample, I estimated baseline characteristics and developed prediction models. The average of all 5 imputed samples are shown in this table.

Table 4.3 Prediction rule for quantifying the probability of chlamydia and/or gonorrhoea infection using the derivation population

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

Figure 4.2 Receiver operating characteristic curves for the derivation and temporal validation populations

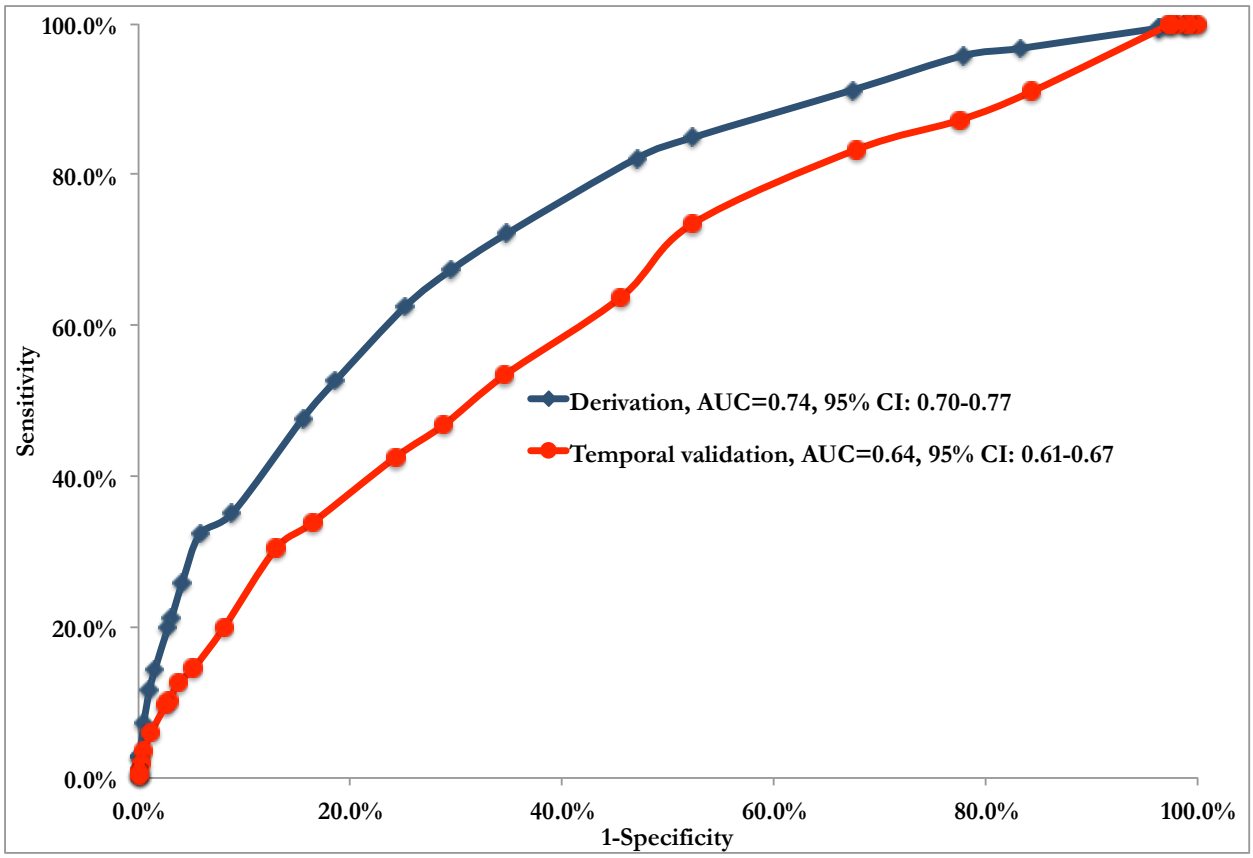
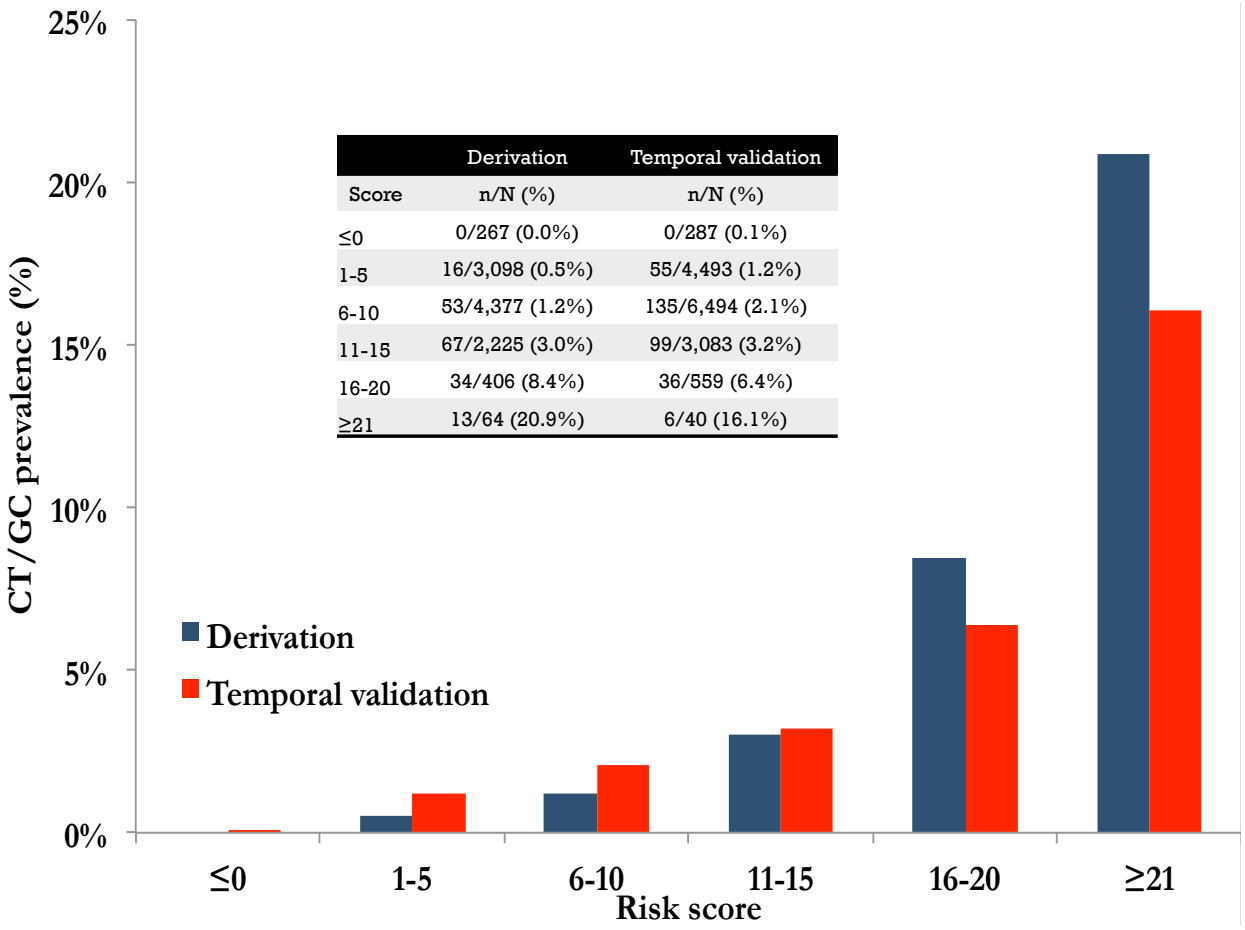


Figure 4.3 Prevalence of chlamydia and/or gonorrhoea infection within risk score categories in the derivation and temporal validation populations



Score	Derivation n/N (%)	Temporal validation n/N (%)
≤0	0/267 (0.0%)	0/287 (0.1%)
1-5	16/3,098 (0.5%)	55/4,493 (1.2%)
6-10	53/4,377 (1.2%)	135/6,494 (2.1%)
11-15	67/2,225 (3.0%)	99/3,083 (3.2%)
16-20	34/406 (8.4%)	36/559 (6.4%)
≥21	13/64 (20.9%)	6/40 (16.1%)

Table 4.4 Sensitivity and specificity of cutoff scores in the derivation and temporal validation populations

	Derivation population				Temporal validation population			
Score Cutoff	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%	2.2%
≥-1	100.0%	1.2%	98.8%	1.8%	100.0%	0.9%	99.1%	2.2%
≥0	100.0%	1.3%	98.7%	1.8%	100.0%	0.9%	99.1%	2.2%
≥1	100.0%	2.6%	97.4%	1.8%	99.9%	2.0%	98.1%	2.3%
≥2	99.5%	3.7%	96.4%	1.8%	99.9%	2.7%	97.3%	2.3%
≥3	99.5%	3.7%	96.4%	1.8%	99.9%	2.7%	97.3%	2.3%
≥4	96.7%	16.7%	83.5%	2.0%	91.0%	15.7%	84.4%	2.4%
≥5	95.8%	22.2%	78.1%	2.2%	87.2%	22.6%	77.7%	2.5%
≥6	91.2%	32.7%	67.8%	2.4%	83.3%	32.3%	68.0%	2.7%
≥7	84.9%	47.7%	52.8%	2.8%	73.4%	47.8%	52.7%	3.1%
≥8	82.2%	53.0%	47.7%	3.0%	63.8%	54.4%	46.0%	3.1%
≥9	72.2%	65.2%	35.4%	3.6%	53.5%	65.4%	35.0%	3.4%
≥10	67.3%	70.6%	30.1%	3.9%	46.8%	71.2%	29.2%	3.6%
≥11	62.5%	74.8%	25.8%	4.3%	42.5%	75.8%	24.6%	3.8%
≥12	52.6%	81.4%	19.2%	4.8%	33.9%	83.6%	16.8%	4.5%
≥13	47.7%	84.5%	16.1%	5.2%	30.4%	87.0%	13.4%	5.0%
≥14	35.1%	91.1%	9.3%	6.6%	19.9%	92.0%	8.3%	5.3%
≥15	32.4%	94.2%	6.3%	9.1%	14.5%	94.8%	5.4%	6.0%
≥16	25.9%	95.9%	4.5%	10.1%	12.7%	96.2%	4.0%	7.0%
≥17	21.2%	96.9%	3.4%	10.9%	10.3%	97.1%	3.0%	7.5%
≥18	19.9%	97.2%	3.1%	11.4%	9.6%	97.4%	2.7%	7.9%
≥19	14.3%	98.5%	1.8%	14.4%	6.1%	98.8%	1.3%	10.2%
≥20	11.6%	99.0%	1.2%	16.9%	3.6%	99.5%	0.6%	14.1%
≥21	7.3%	99.5%	0.6%	20.9%	1.9%	99.8%	0.3%	16.1%
≥22	2.8%	99.9%	0.2%	28.0%	0.6%	100.0%	0.1%	23.8%
≥23	2.8%	99.9%	0.1%	33.8%	0.6%	100.0%	0.1%	24.4%
≥24	2.8%	99.9%	0.1%	33.8%	0.6%	100.0%	0.1%	24.4%
≥25	2.8%	99.9%	0.1%	33.8%	0.6%	100.0%	0.1%	24.4%
≥26	0.7%	100.0%	0.0%	30.0%	0.3%	100.0%	0.0%	33.3%

Chapter 5: Geographical validation of a risk estimation algorithm for optimising chlamydia and gonorrhoea case finding

Falasinnu T, Gilbert M, Gustafson P, Shoveller, J. Geographical validation of a risk estimation algorithm for optimising chlamydia and gonorrhoea case finding. *Submitted*

5.1 Introduction

The imperative to adopt more efficient sexual health service provision by public health programs has led to the development of service models that minimise the use of health human resources, such as internet-based sexually transmitted infections (STI) testing and triage services [21,22,29,182]. The aim of optimising service provision could be facilitated by the use of risk estimation algorithms. In Chapter 4, 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, BC. This algorithm combines five risk factors: younger age, non-white race/ethnicity, multiple sexual partners, previous chlamydia diagnosis, and previous gonorrhoea diagnosis.

The measures of a prediction rule's performance are sometimes thought of as immutable features that can be determined theoretically and subsequently applied in practice [69]. However, a prediction rule's accuracy in one context may vary from the performance estimates reported in the derivation stage. This issue has been explored extensively in chronic disease research; for example, the Framingham risk score has shown tremendous variability in discriminative and calibration abilities in varying settings [56,183]. This variability is never random or due to chance. There may be genuine disparities that account for the prediction rule's accuracy in varied settings, such as general practice clinics or hospitals, different types of hospital, or the same type of clinical setting in different geographical locations [69]. Whether the measures of a prediction rule's accuracy are transferable to additional contexts depends on recognising the possible reasons for the variability in discrimination and calibration metrics across settings [69]. This inconsistent performance may be due to artifactual disparities (e.g., different study design features in different contexts) or genuine disparities (such as distribution of risk factors) [69].

Therefore, before the use of the risk estimation algorithm derived through the work described in Chapter 4 may be recommended in clinical or population health decision-making, external validation in an independent setting is required. This is essential to confirm that the algorithm is generalisable to a plausibly related setting (as compared with the derivation population) that reflects the level of heterogeneity that will be encountered in real-life applications of the algorithm [179]. As was demonstrated in Chapter 4, the

algorithm showed reasonable discrimination and calibration upon validation in a later time period (i.e., temporal validation) [181]. Temporal validation is often cited as the first step in demonstrating the transferability of a prediction algorithm [60]. However, temporal validation cannot assess the transportability of the algorithm to other clinics or cities [179]. 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 [179]. Thus, this chapter sought to test the generalisability of the Vancouver-derived algorithm in a population of patients attending sexual health clinics in seven other geographical settings in BC.

5.2 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: Penticton, Kelowna, Kamloops, New Westminster, Boundary, Courtenay, and Prince George (see map in Appendix B). This analysis was limited to clinic visits among asymptomatic females and heterosexual males 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. Here, in Chapter 5, the original model, regression coefficients, and the simplified risk scores derived from the Vancouver clinic data are applied to a *geographical validation population* [179].

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 4.3. Multiple imputation methods (5 rounds) were used to impute missing values in the geographical validation population [174,175]. This analysis imputed missing values using IVEware, a software application that performs multiple imputations of missing values using the Sequential Regression Imputation Method (SRIM) [177]. All predictors and the outcome variables were included in the imputation model and the results of the 5 imputed datasets were combined to obtain final estimates [174,175]. 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 [184]. The closer an AUC is to 100%, the better the model [184].

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 [184]. A p-value >0.05 indicates a good fit [184]. 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 4.3. 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 [51].

In addition, a sub-analysis was performed to determine whether the 'Vancouver' prediction rule performed better than deriving a new prediction model altogether using the geographical validation population. The remodelled algorithm was derived with the same methodology as the 'Vancouver' prediction rule developed in Chapter 4 using logistic regression modelling. The final model consisted of predictors that were used to score the geographical validation population. Discrimination and calibration measures were also calculated for the remodelled algorithm and compared to the estimates derived using the 'Vancouver' prediction rule. The Delong method for comparing two correlated ROC curves was used to test whether that the AUC derived from the re-modeled algorithm was different from the one derived from the 'Vancouver' algorithm [185]. The Delong chi-square test was used to assess whether the difference between the AUCs is different from zero; a non-significant chi-square statistic is interpreted as evidence that there is no difference between the AUCs [185]. All associated methods were approved by the University of British Columbia's Research Ethics Board (certificate # H11-02000).

5.3 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. The prevalence of chlamydia and/or gonorrhoea infection was 5.3% (higher than the derivation population). Table 5.1 shows the distribution of the baseline characteristics of patient visits. 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 20-24 years (28.0%), and white ethnicity (74.3%). More than two-thirds of patient visits reported having 1-2 sexual partners in the previous 6 months and approximately 43% reported always using condoms. Approximately 3% of patient visits reported injection

drug use and the same proportion reported having sex with partners recruited online. Nearly 16% of patient visits reported a previous chlamydia diagnosis.

The geographical validation population differed from the derivation population by having a higher proportion of the following characteristics: females, younger individuals, inconsistent condom use, and injection drug use (Table 5.1). There were also some differences between the derivation and geographical validation populations in terms of the unadjusted odds ratios 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 5.2). Race/ethnicity was not significantly associated with the outcome in the geographical validation population unlike the derivation population (Table 5.2).

The Vancouver risk model demonstrated good discrimination in the geographical validation population. The AUC in the geographical validation population was 0.69, 95% CI: 0.67-0.71, while the AUC in the derivation population was 0.74, 95% CI: 0.70-0.77 (Figure 5.1). The *P*-value of 0.26 for the Hosmer-Lemeshow goodness-of-fit test also indicated good calibration. Figure 5.2 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. The risk model may be adapted for selective screening of individuals with relatively high scores (Table 5.3). This analysis identified a risk score cutoff 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 cutoff of ≥6 points identified 91% of cases and the fraction screened was 68% of the population.

A remodelled risk algorithm was derived by performing a logistic regression analysis on the geographical validation population. The following predictors were included in this remodelled algorithm: gender, age, number of sexual partners in previous 6 months, injection drug use, sex with IDU, previous chlamydia diagnosis, and previous gonorrhoea diagnosis (Table 5.4). The AUC of the remodelled algorithm was 0.72 (95% CI: 0.70-0.74) and the Hosmer-Lemeshow goodness of fit *P*-value was 0.47. The *p*-value of the Delong chi-square testing the difference between the AUCs of the remodeled algorithm and the 'Vancouver' prediction rule in geographical validation population was 0.236.

5.4 Discussion

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). Remarkably, 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 (e.g., 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 [186-188] (see map in Appendix A). A sub-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-square 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.

The findings of this analysis were compared to other external validation studies in sexual health contexts. Two published studies that examined the validity of previously derived CPRs in new geographic settings [58,145] were identified. Gotz and colleagues derived a prediction rule for chlamydia infection for the selective screening of high-risk individuals in Rotterdam, The Netherlands [145]. 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: 0.76-0.84) in the derivation sample, 0.66 (95% CI: 0.58-0.74) in the Amsterdam sample, and 0.68 (95% CI: 0.58- 0.79) in the Rotterdam sample [145]. A second study by Haukoos and colleagues (2011) 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) [58]. 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: 0.83-0.88) in the derivation sample and 0.75 (95% CI: 0.70-0.78) in the validation sample [58].

The AUC of the 'Vancouver' risk estimation in the derivation population was lower compared to the AUCs of the two aforementioned studies (Gotz et al., 2006; Haukoos et al., 2011). 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 [189]. 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 to a loss of 10% points and 5% points in the temporal and geographical validation populations, respectively in this analysis.

Overall, these findings are encouraging and bolster confidence in recommending this risk estimation algorithm for use in sexual health services and programs. The risk score could be easily implemented and is accurate enough to convey important screening considerations. The risk score is based on information easily obtained from clinical encounters and does not require intensive data collection resources. However, 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*; 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.

5.5 Strengths and limitations

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. One limitation of this study is that it was not possible to test the accuracy of the risk score at individual geographical sites due to sample size limitations because the insufficient events per variable at individual clinics precluded this analysis. Differences in prevalence, methods of patient selection, and data collection between the derivation population and other sexual health service contexts are more likely to be detected when multiple independent geographical sites have applied the tool [60]. A site-specific analysis would have tested the accuracy of the risk estimation tool in the hands of diverse clinicians at diverse geographical sites and, consequently, could have provided more thorough evidence of the generalisability of the tool than the single validation conducted at the multiple sites. Unfortunately, this particular approach was beyond the feasibility reach of the current study.

5.6 Summary

This chapter highlighted the geographical validation of the 'Vancouver' risk estimation tool - the final step in the methodological framework (Figure 1.1). The prediction rule showed adequate discrimination and calibration upon validation in seven clinics outside of Vancouver. However, despite the good performance of the risk estimation tool in the geographical validation, the integration of this tool into routine sexual health services may face a number of challenges, e.g., the availability of electronic medical records, patient and clinician acceptability. The concluding section of this dissertation (i.e., Chapter 6) reflects upon many of those challenges. Chapter 6 also provides an opportunity to contextualise the overall findings of this dissertation as well as provide directions for future research.

Table 5.1 Population characteristics of visits at sexual health clinics in the derivation and geographical validation populations

	Derivation population, 2000-2006		Geographical validation population, 2000-2012	
Variable	N	%	N	%
Chlamydia or gonorrhoea case				
Yes	184	1.8%	556	5.3%
No	10,253	98.2%	9,869	94.7%
Gender				
Female	3,496	33.5%	4,431	42.5%
Male	6,941	66.5%	5,994	57.5%
Age (yrs)				
14-19	257	2.5%	1,748	16.8%
20-24	1,962	18.8%	2,922	28.0%
25-29	2,651	25.4%	1,969	18.9%
30-39	3,181	30.5%	1,850	17.7%
≥40	2,386	22.9%	1,935	18.6%
Race/ethnicity				
White	7,732	74.1%	7,741	74.3%
Non-white	2,705	25.9%	2,684	25.7%
No. of sexual partners in previous 6 months				
0	644	6.2%	590	5.7%
1-2	6,857	65.7%	7,004	67.2%
≥3	2,936	28.1%	2,830	27.1%
Condom use				
Never	2,362	22.6%	2,650	25.4%
Always	5,269	50.5%	4,489	43.1%
Sometimes	2,806	26.9%	3,286	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	9,983	95.6%	10,018	96.1%
Sex with commercial sex worker				
Yes	1,381	13.2%	826	7.9%
No	9,057	86.8%	9,599	92.1%
Previous chlamydia diagnosis				
Yes	1,518	14.5%	1,660	15.9%
No	8,919	85.5%	8,765	84.1%
Previous gonorrhoea diagnosis				
Yes	619	5.9%	310	3.0%

Variable	Derivation population, 2000-2006		Geographical validation population, 2000-2012	
	N	%	N	%
No	9,819	94.1%	10,115	97.0%
Total	10,437	100.0%	10,425	100.0%

*Imputed values

Table 5.2 Chlamydia and/or gonorrhoea prevalence and unadjusted ORs (derivation and geographic validation populations)*

	Derivation population (N=10,437)		Geographical validation population (N=10,425)	
Variable	%	OR (95% CI)	%	OR (95% CI)
Gender				
Female	2.1%	1.31 (0.97-1.77)	6.8%	1.67 (1.40-1.98)
Male	1.6%	Ref	4.2%	Ref
Age (yrs)				
14-19	7.4%	6.49 (3.58-11.75)	9.7%	9.47 (6.00-14.94)
20-24	2.8%	2.30 (1.46-3.63)	7.7%	7.40 (4.72-11.58)
25-29	1.8%	1.50 (0.94-2.38)	4.9%	4.55 (2.82-7.34)
30-39	1.1%	0.88 (0.53-1.45)	2.4%	2.15 (1.27-3.62)
≥40	1.2%	Ref	1.1%	Ref
Race/ethnicity				
White	1.2%	Ref	5.2%	Ref
Non-white	3.4%	2.90 (2.16-3.89)	5.7%	1.10 (0.78-1.55)
No. of sexual partners in previous 6 months				
0	0.5%	Ref	2.0%	Ref
1-2	1.8%	3.43 (1.10-10.73)	4.9%	2.50 (1.26-4.98)
≥3	2.0%	3.92 (1.23-12.45)	7.1%	3.71 (1.83-7.55)
Condom use				
Never	1.6%	Ref	4.5%	Ref
Sometimes	2.0%	1.22 (0.84-1.77)	6.0%	1.36 (1.07-1.72)
Always	1.4%	0.88 (0.56-1.39)	5.0%	1.11 (0.85-1.46)
Sex with partners recruited online				
Yes	1.2%	0.67 (0.27-1.63)	2.7%	0.48 (0.25-0.94)
No	1.8%	Ref	5.4%	Ref
Injection drug use				
Yes	1.6%	0.89 (0.26-3.06)	7.5%	1.45 (0.81-2.58)
No	1.8%	Ref	5.3%	Ref
Sex with injection drug user				
Yes	1.0%	0.55 (0.20-1.54)	6.6%	1.27 (0.78-2.07)
No	1.8%	Ref	5.3%	Ref
Sex with commercial sex worker				
Yes	1.2%	0.65 (0.35-1.20)	2.7%	0.48 (0.29-0.79)
No	1.8%	Ref	5.6%	Ref
Previous chlamydia diagnosis				
Yes	5.1%	4.40 (3.27-5.93)	11.7%	3.10 (2.58-3.72)
No	1.2%	Ref	4.1%	Ref
Previous gonorrhoea diagnosis				
Yes	2.1%	1.21 (0.69-2.14)	4.2%	0.77 (0.44-1.36)
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 5 resampling replications, which generated an augmented databases 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, I estimated baseline characteristics and developed prediction models. The average of all 5 imputed samples are shown in this table.

Table 5.3 Sensitivity and specificity of cutoff scores in the derivation and geographical validation populations

	Derivation population				Geographical validation population			
Score Cutoff	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%

Table 5.4 Locally derived prediction algorithm using the geographical validation population

Variable	aOR (95% CI)	Beta
Intercept	--	-5.5634
Gender		
Female	1.18 (0.98-1.42)	0.1658
Male	Ref	Ref
Age (yrs)		
14-19	8.52 (5.35-13.56)	2.1418
20-24	6.67 (4.24-10.49)	1.8968
25-29	4.11 (2.55-6.62)	1.4117
30-39	1.95 (1.16-3.29)	0.6673
≥40	Ref	Ref
No. of sexual partners in previous 6 months		
0	Ref	Ref
1-2	2.09 (1.15-3.80)	0.7280
≥3	2.93 (1.60-5.37)	1.0635
Injection drug use		
Yes	1.55 (1.00-2.42)	0.4323
No	Ref	Ref
Sex with injection drug user		
Yes	1.33 (0.86-2.05)	0.2750
No	Ref	Ref
Previous chlamydia diagnosis		
Yes	2.84 (2.35-3.43)	1.0434
No	Ref	Ref
Previous gonorrhoea diagnosis		
Yes	1.23 (0.68-2.23)	0.2085
No	Ref	Ref

Figure 5.1 Receiver operating characteristic curves for the derivation and geographical validation populations

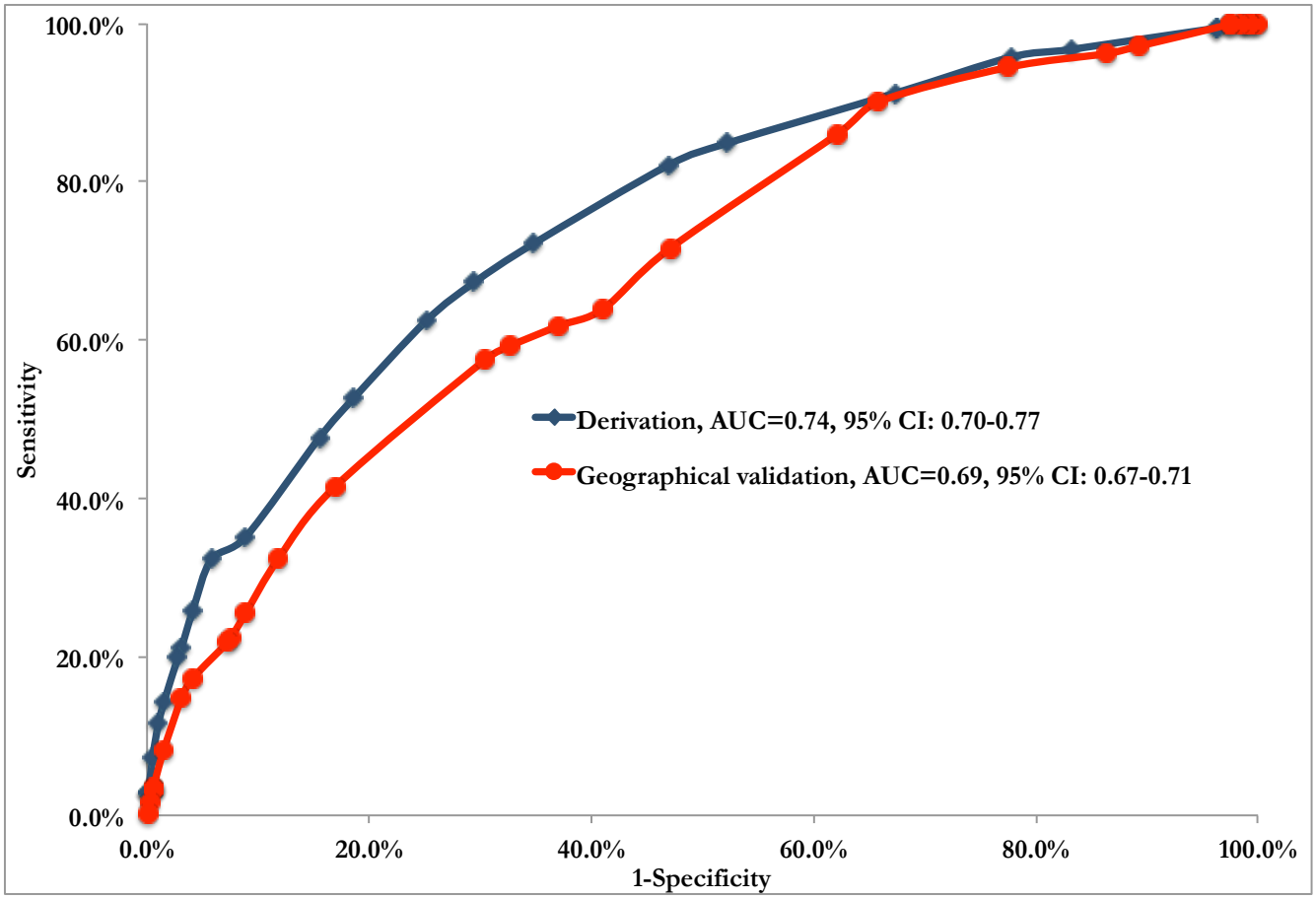
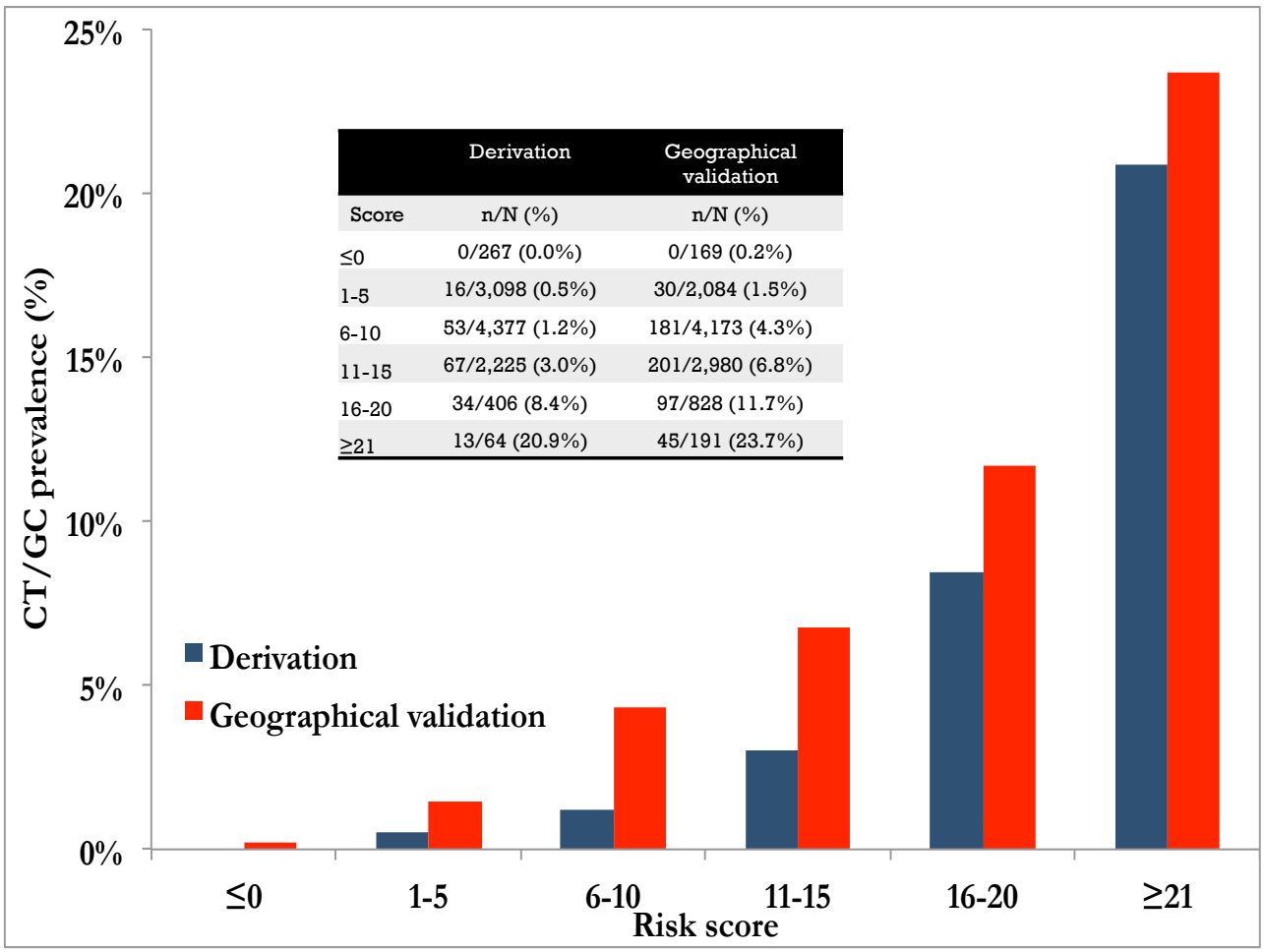


Figure 5.2 Prevalence of chlamydia and/or gonorrhoea infection within risk score categories in the derivation and geographical validation populations



	Derivation	Geographical validation
Score	n/N (%)	n/N (%)
≤0	0/267 (0.0%)	0/169 (0.2%)
1-5	16/3,098 (0.5%)	30/2,084 (1.5%)
6-10	53/4,377 (1.2%)	181/4,173 (4.3%)
11-15	67/2,225 (3.0%)	201/2,980 (6.8%)
16-20	34/406 (8.4%)	97/828 (11.7%)
≥21	13/64 (20.9%)	45/191 (23.7%)

Chapter 6: Discussion and future directions

6.1 Overview of study findings

The purpose of this dissertation was to conduct an integrated series of inquiries examining the development and validation of a risk estimation tool for screening asymptomatic chlamydia and/or gonorrhoea. Here, in the final chapter, I integrate and synthesise the empirical findings, identify the theoretical and policy implications with respect to the sexual health service provision, highlight the overall study strengths and limitations, and provide suggestions for future research. Specifically, this final chapter highlights insights regarding strategies that could be used to further bridge the gap between evidence from research and routine sexual health service provision.

In this first section of Chapter 6, the reviews presented in Chapters 2 and 3 are used to contextualise the findings of the empirical chapters (Chapters 4 and 5). Chapter 2 presented a theory-guided review of risk factors for chlamydia and gonorrhoea infection identified in empirical studies and in screening guidelines issued by health organisations. The intention of that review was to provide an overall assessment of the predictive value of these risk factors and in this way, guide an understanding of the predictors that eventually comprise the prediction rule presented in the empirical chapters. In Chapter 3, a methodological review of prediction rules used in sexual health services was presented. Drawing on the findings of the methodological review, a methodological framework was developed highlighting distinct stages required for the development and validation of a clinical prediction rule that was then used in the subsequent empirical chapters.

In Chapter 4, a prediction rule which uses routinely collected risk factors was derived for screening asymptomatic chlamydia and/or gonorrhoea in an earlier time period (2000-2006) and validated in a subsequent time period (2007-2012) using electronic medical records collected at two STI clinics in Vancouver. The predictors of chlamydia and/or gonorrhoea infection identified were young age, non-white race/ethnicity, multiple sexual partners, previous chlamydia, and previous gonorrhoea infection. These risk factors were also identified as strong predictors in the empirical literature and screening guidelines (see Chapter 2). The observed statistical significance of race/ethnicity should be interpreted in light of other social and structural factors (e.g., socioeconomic status). Research shows that the acquisition of STI is related to background STI prevalence rates within particular sexual networks, which are clearly different in various socioeconomic and ethnic groups [72]. In this study, the only behavioural risk factor or proximate

determinant found to be independently associated with infection was the number of sexual partners; this is consistent with the findings of another study examining selective screening criteria in the United States [51], where behavioural risk factors were not found to add efficiency as selective factors. As noted in Chapter 2, the weak association between proximate determinants and infection can be explained by the difficulty in measuring behavioural risk factors in day-to-day clinical risk assessment. To be feasible in clinical settings, risk assessment tools must be as parsimonious as possible, and are as such not necessarily reflective of comprehensive assessments of risk behaviour. Although the review presented in Chapter 2 found weak evidence for the association between having a history of infection and STI diagnosis, this dissertation found that previous chlamydia or gonorrhoea infection was predictive of current infection. This finding was potentially due to the use of medical records to elicit this information as opposed to relying on patient recollection.

The prediction rule demonstrated adequate discriminative and calibration ability for chlamydia and/or gonorrhoea infection in the derivation and temporal validation populations. Each risk factor in the final model received a certain number of points that were added into a risk score, which mirrors the probability of infection. By choosing a specific risk score as cut-off level for screening, the number of persons to be screened will be reduced, and the prevalence in the screened population will be higher than in the general population. The aim of selective screening is to increase sensitivity (percentage of cases detected) and to increase efficiency (decrease the percentage of the population to be screened). In the derivation population, if screening were advised for people with a score of 6 or above, 91% of the cases would be detected by screening only 68% of the population; the optimal threshold of the benchmark of 90% sensitivity by screening 60% would be close to being reached (Table 4.4). In the screened population, chlamydia and/or gonorrhoea prevalence would be 2.4% compared to 1.8% overall. When only considering the screened population, 97.6% of the screened population would test negative. Of all non-infected people, 33% would avoid being unnecessarily screened. However, the 32% reduction in the number of individuals that need to be screened using the ≥6 risk score cutoff compared to screening the whole population shows that the efficiency of screening in population-based programmes may be improved by targeting screening in this way [145]. As noted in Chapter 2, age <25 years is a screening criterion used in the US and Canada. Using this criterion in the derivation population would require screening 21% of the population while identifying only 40% of cases, a performance that falls short of the screening

benchmark even after increasing the age cutoff to <30 years, indicating that the prediction rule could be useful in decision making in this setting.

In Chapter 5, the performance of the 'Vancouver' prediction rule was evaluated in seven additional geographical settings in BC. The discriminative performance of the prediction rule in the geographical validation population was moderate (AUC=0.69), but was better than the temporal validation population (AUC=0.64), indicating that the temporal validation population was less heterogeneous than the derivation and geographical validation populations. In this population, choosing the cutoff point of ≥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 ≥6 cutoff point would fail to meet the efficiency benchmark of screening <60% of the population. Increasing the cutoff point to ≥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 (i.e., more severe case mix) such as the geographical validation population, it is expected that a choosing a higher cutoff point will increase the efficiency of screening decision-making [55,145].

Alternatively in this setting, applying the age-based screening criterion (i.e., 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 cutoff to age <30 years would require screening 64% of the population with a sensitivity of 88%, a performance that's close to the benchmark and also similar to using the cutoff point of ≥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 [55].

Several studies have established the validity of prediction rules as screening criteria, especially where chlamydia and/or gonorrhoea prevalence is low (i.e., <2%) such as found in the derivation population [190]. Although it has been suggested that universal screening or using age as a viable discriminant factor would be cost effective in settings with prevalence of 2% or more, however,

publicly funded sexual health clinics are constrained by available funding and limited resources must continue to be applied where they will be able to detect and treat as much infection as possible [190]. Specifically, screening women less than 25 years old could prove to be cost prohibitive in settings where individuals in this age group comprise the highest proportion of clinic visits [50]. 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 an essential attribute of risk assessment. In these scenarios, the risk score categories and their associated prevalence presented in chapters 4 and 5 would 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 particular risk factors (e.g., age) to prediction rules acknowledges a more comprehensive spectrum and the continuous nature of risk factors [191].

6.2 Study strengths and limitations

Collectively, these research findings offer novel and important insights into risk estimation in sexual health contexts. Although specific strengths and limitations are presented in the discussion sections of Chapters 3-5, several overarching points need to be considered when interpreting the overall findings and implications of the current dissertation. For example, in Chapter 2, there were few high-quality studies that specifically investigated the risk factors associated with infection amongst males. Most of the studies included in the review were conducted in contexts where women's bodies are the primary foci of surveillance (e.g., routine screening of sexually active females during annual physical exams). Also, it should be noted that many organisations issuing screening recommendations highlight the lack of evidence and the potential effects of under-screening (and the resultant service gap and neglected reservoir of infection) amongst males, especially young men [192]. In order to be consistent with the guidelines in the literature, gender was not treated as a predictor but rather males and females were treated as different populations in articles that reported multivariate analyses stratified by gender.

The empirical work conducted in Chapters 4-5 attempted to address at least a portion of the identified gap by deriving and validating a screening tool in a population of women and heterosexual men. The apriori analysis plan was to develop prediction rules separately for males and females. However, the predictors and beta coefficients included in the final models did not differ significantly by gender. Clearly, testing women for STIs remains important from an individual and a public health

perspective; but, to do so while not adequately accounting for men within screening interventions could have unintended consequences (e.g., stigmatization of women's bodies; reinforcing stereotypes about men's 'freedom' from sexual health responsibilities) [193] . While developing screening interventions that engage women and men presents a complex and enduring public health challenge [194], the prediction rule derived in this dissertation is a promising approach for fostering gender equity in this area [195].

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 [122]. 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. A methodological strength of the current analysis was the use of imputation to replace those missing data, which is rarely done in prediction modeling 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) [175]. 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 [175].

The analyses presented in this dissertation combined chlamydia and gonorrhea outcomes because of the single test offered for both. However, as noted in Chapter 1, there are differences in the epidemiology of these two infections that affect the findings of the literature review and the empirical analyses presented. Chlamydia infections contribute the highest STI incidence rate in the BC. It is characterized by a more generalized distribution of case reports and affects mostly young people. In contrast, gonorrheal infections have a lower incidence rates and are driven by core group activities. The apriori expectation was that the discrimination of the model would be higher for gonorrhoea outcomes only because of the concentration of this infection in individuals with higher risk profiles. A sub-analysis found that there was no difference in the performance of the prediction rule for detecting gonorrhoeal infection (AUC = 0.73) and chlamydia infection (AUC = 0.75) only.

This finding gives more confidence in the decision to combine both outcomes. However, this dissertation did not combine previous diagnosis with chlamydia and gonorrhoea infection into a single variable in the multivariate analysis because of the reduction in discrimination and model fit measures that resulted from this combination.

Also, this was the first study that adopted a conceptual framework to guide the understanding of the predictors that comprise a prediction rule. This is an important strength of the study because it is essential to understand the reasoning behind a variable's statistical significance. The proximate-determinants framework was adopted because it is the only framework that has been validated in STI settings –specifically, in a seminal HIV study that evaluated the contribution of proximate and underlying determinants for predicting individual level HIV risk [65]. The framework hypothesizes that if the proximate determinants are measured well, the underlying determinants should be non-significant because the underlying determinants predict the proximate determinants. But this was obviously not the case in this study. The most significant variables are sociodemographic determinants. Only one behavioural variable (number of sexual partners in previous six months) was significant. This finding highlights the difficulty in validating a framework in contexts where the data was not collected with the framework specified apriori.

In addition, it was also impossible to evaluate determinants that measure important concepts such as the force of infection, sexual network and partnership structures of the population, and partnership level characteristics. Specifically, the decrease in the value of the performance metrics of the prediction rule may have resulted from the absence of these determinants of infection [145]. Currently, the variables captured in the STI databases do not include many of these partnership level characteristics. STI risk is determined by the characteristics of sexual partnerships and sexual networks, as these are components of transmission dynamics that hinder or facilitate safer sex practices [145,196] . For example, an emerging area of consideration in STI epidemiology is age mixing, which is defined as sexual partnerships between younger women and older men, a situation that is associated with lower condom negotiation skills and increased rates of infection [196]. The lack of these sexual partnership characteristics and missing network level determinants such as concurrency may have contributed to the sub-optimal performance of the prediction rule upon validation [145]. Further validation and updating of the prediction rule is, thus, warranted. For example, when using the prediction rule in online contexts, it is essential that partnership

characteristics be included and the performance of this updated prediction rule should be assessed [145].

In terms of limitations, the validity of the risk factors comprising the prediction rule depends on the accuracy of the self-reported health behaviors. There is a risk of recall and social desirability biases because of the self-reported nature of stigmatised activities and behaviours. There is also the risk of over-reporting of perceived normative behaviours, such as condom use. Such reporting biases would artificially inflate the relationships between infection and risk factors. However, because the clinical risk assessment interviews are confidential and are conducted by clinicians who are typically not acquainted with the respondents, strong motivations to self-present are unlikely. Moreover, the *outcome* variables do not rely on self-report and thus are not subject to recall or social desirability biases.

Another important limitation of the current work is that the inclusion of race/ethnicity in a prediction rule might give rise to concerns about stigma and discrimination in the absence of careful attention and explanation of network effects. Race/ethnicity, an underlying determinant, was evidenced as a major risk factor for STIs. It has been posited that race/ethnicity is a marker of a core group with high prevalence of infection because of poor access to care [11]. The findings of this research suggest that additional research is needed to unravel how the evidence associating race/ethnicity with infection can be more effectively understood, particularly in terms of how that knowledge could be used to better inform the scaling-up of screening interventions, without exacerbating existing sexual health inequities, particularly amongst vulnerable sub-groups (e.g., people who are racialised and/or stereotyped by virtue of their ethnic identity). This is a pressing concern that demands sophisticated theoretical and methodological work to sort out how sexual health inequities become concentrated amongst particular population sub-groups, who also are more likely than the general population to suffer from both an accumulation of negative exposures over the life course as well as a concentration of multiple risk factors because of shared 'fundamental causes' associated with their social positions[197]. Without this kind of work, existing interventions are at risk of failing to address the needs of these subgroups, potentially exacerbating the negative health and social sequalae associated with infection.

Finally, the conceptual challenges of adopting prediction rules for sexual health contexts in comparison to chronic disease contexts must be acknowledged. As acknowledged in Chapter 3, prediction rules are new in sexual health context, while these tools are well articulated in chronic

disease contexts with an abundance of guidelines statements that effectively determine the methodological standards that govern how they are developed and validated e.g., the STARD statement [59]. There are also differences in terms of the clinical utility of the tools; they are used in sexual health settings for selective screening and used for prognostic or diagnostic purposes in chronic disease settings. Specifically, in converting the prediction rules in sexual health setting to decision-making instruments, the aim is to maximize sensitivity of case finding and limit the number of people unnecessarily tested. However, in chronic disease contexts, the aim is often is to rule out or rule in a disease state. On a population level, the intervention in sexual health settings is a decision to screen, while oftentimes the intervention in question chronic disease research is an invasive or therapeutic one, e.g., using mammograms or prescribing statins in breast cancer risk models and the Framingham risk score. In this situation, the choice of risk cutoff point is one that balances the harms and benefits of the intervention. However, the only harm in sexual health context is missed infections. Also, the variables that comprise sexual health risk estimation tools are usually demographic and behavioural in nature, while those used for chronic disease decision-making are usually biological in nature with causal underpinnings and are usually more objectively measured (e.g., blood pressure, cholesterol level). Altogether, these differences account for the difficulty in achieving the high performance values in sexual health contexts that often found in chronic disease contexts.

6.3 Practical implications of research findings

The findings of the empirical chapters are situated at the nexus of at least two potentially competing sets of priorities in public health: one set of priorities focuses on maximising or improving the health of the individual patient as compared with another set of priorities that are focused on achieving the most cost-effective decision [69]. A tension emerges when policy makers must make choices about the acceptability and feasibility of various approaches to STI testing, which might range from highly efficient approaches from a cost perspective to very nuanced and patient-centred approaches from a clinical practice standpoint. The algorithm developed in this dissertation may prove useful to public health programs that lack the resources for universal or age-based chlamydia/gonorrhoea screening programs. The independent validation of the algorithm suggests that it has potential to inform screening decisions, especially in low STI prevalence settings. For example, to aid the scaling-up of GCO, BC's internet-based STI testing program, the algorithm

could be adapted into a self-selection tool for filtering GCO participants based on risk profile. Only participants with sufficient risk score would be recommended to receive testing through GCO [198]. However, in this scenario, it would be important to demonstrate whether the risk score has beneficial effects on GCO uptake and acceptability among participants.

It is also anticipated that the prediction rule could potentially facilitate decision-making in traditional clinical encounters where clinicians could enter basic demographic and behavioural data directly into the client's computerised medical record during the consultation. The prediction rule could be used to display an alert on the computer screen to prompt clinicians to offer specific STI tests to those at increased risk of infection. In addition, the prediction rule could be used to inform ongoing clinical recommendations related to selective screening of STI clients and standardise STI testing at the clinics in BC, potentially enabling testing that is targeted to higher risk individuals, thereby reducing the unnecessary testing of those without the infection and saving costs. However, the use of the prediction rule depends on numerous contextual sources of evidence, including clinical knowledge and experience as well as other features of the clinical practice landscape. For example, to accommodate new epidemiological trends, such as Quinolone and Cephalosporin resistance in gonorrhoea treatment, the cutoff point for an existing CPR for infection may need to be lowered (which reduces the number of false negatives) to highlight the importance of sensitivity (failure to provide the appropriate intervention is highly undesirable in such instances).

Even within contexts where universally screening approaches remain the norm, the algorithm could be helpful in prioritising or triaging patients. For example, triage services are becoming routine in publicly funded sexual health clinics [29,172]. These triage services require patients to fill out risk assessment questionnaires in the waiting room prior to receiving services at the sexual health clinics [29,172]. The risk estimation algorithm developed in the current dissertation could be adapted into a decision-making tool that prioritises patients by risk scores. This may enable clinicians to effectively determine whether triaging those with low risk scores to receive "express" services (e.g., providing self-collected specimens) or referring others to more comprehensive services (e.g., complete physical examinations) [199] . This process may help decrease wait times, improve clinical workflow, and reduce unnecessary clinician face time.

The algorithm developed through the current dissertation also might be used to inform the development of evidence-based risk assessment questions for use in internet-based settings or mobile applications, other than GCO (e.g., a smart-phone 'app'). For example, the risk assessment

tool could be incorporated into other online websites (e.g., chat rooms; dating sites), where individuals accessing those websites could also estimate their STI risk. The algorithm also could be programmed with suggested sexual lifestyle modification messages specific to calculated risk estimates. If programmed into an app, the algorithm also could provide a readily accessible screening guideline reference for clinicians and patients – particularly if it was designed to encourage two-way communication and shared decision-making during clinic visits. The risk scores developed from the prediction rule also have important implications for risk communication and testing motivation because they can increase risk perception by creating tailored risk messages to different groups.

6.4 Recommendations for future research directions

When considering adopting prediction rules for their settings, decision makers may want to consider incorporating or updating features of existing rules instead of developing new prediction rules in their own setting. Using appropriate statistical methods for adjusting for differences between the derivation and validation populations that are well-articulated in chronic disease research (e.g., case mix differences, prevalence) may have potential applicability in sexual health contexts [200]. Those approaches are promising because they minimize the unnecessary derivation of new rules. If the prediction rule derived in this dissertation performs inadequately or poorly in another population (e.g., among a pilot sample of potential GCO clients), the prediction model can be adjusted or updated using the data from the new setting to improve its performance in that population [200]. In sexual health contexts, for example, the prevalence of disease can affect the generalisability of a CPR from one setting to another (e.g., when a CPR developed in a *high* STI prevalence setting, such as STI clinics, is validated in a *low* STI prevalence setting, such as general practice settings). The predictive strength of the CPR in the general practice settings may be modest resulting from systematically inflated predicted probabilities [200]. By adjusting only the intercept of the original prediction rule, the predicted probability estimates for the new setting can be improved [200].

After successful validation in multiple settings, the next step is to assess how CPRs impact current clinical and population-based sexual health practice after they are implemented [54]. At present, only one implementation study of STI prediction rules has been conducted [198]. Future impact analyses should be completed to evaluate the effects of the implementation of the CPRs on numerous implementation and scale-up concerns, including the effects on overall client volume, number of tests ordered, costs, trends in STI prevalence, wait times, and number of clients turned

away. Before-after comparisons, experimental and well-designed observational study designs could yield this important information [29]. Furthermore, detailed understandings of the implementation context of proposed CPRs (e.g., health human resources; 'fit' with current clinical practice; patient factors) and clear examination of features of clinical realities that might pose potential barriers (e.g., current workflow; availability of electronic health records) are necessary before these innovations are recommended in clinical practice. For example, future research could examine: (1) clinicians' understanding and willingness to use CPRs; and (2) patients' trust, preference, participation, and acceptance in adopting CPRs [29,198].

6.5 Final conclusions

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 in this dissertation provides evidence that risk estimation tools have a place in sexual health service provision and public health. New investments in research 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.

Bibliography

1. Public Health Agency of Canada. **The Chief Public Health Officer's Report on the State of Public Health in Canada**. 2012.

2. Mayaud P, McCormick D. Interventions against sexually transmitted infections (STI) to prevent HIV infection. Br Med Bull 2001;58:129-153.

3. BC Centre for Disease Control. STI in British Columbia: Annual Surveillance Report 2012. 2014.

4. Public Health Agency of Canada. Canadian Guidelines on Sexually Transmitted Infections, 2009. 2009.

5. Tapsall J. Multidrug-resistant Neisseria gonorrhoeae. CMAJ 2009 Feb 3;180(3):268-269.

6. Ota KV, Jamieson F, Fisman DN, Jones KE, Tamari IE, Ng LK, et al. Prevalence of and risk factors for quinolone-resistant Neisseria gonorrhoeae infection in Ontario. CMAJ 2009 Feb 3;180(3):287-290.

7. MacDonald NE, Stanbrook MB, Flegel K, Hebert PC, Rosenfield D. Gonorrhea: what goes around comes around. CMAJ 2011 Oct 4;183(14):1567.

8. BC Centre for Disease Control. HIV in British Columbia: Annual Surveillance Report 2012. 2013.

9. Marshall BD, Shannon K, Kerr T, Zhang R, Wood E. Survival sex work and increased HIV risk among sexual minority street-involved youth. J Acquir Immune Defic Syndr 2010 Apr;53(5):661-664.

10. Marshall BD, Kerr T, Shoveller JA, Patterson TL, Buxton JA, Wood E. Homelessness and unstable housing associated with an increased risk of HIV and STI transmission among street-involved youth. Health Place 2009 Sep;15(3):753-760.

11. Aral SO. Determinants of STD epidemics: implications for phase appropriate intervention strategies. Sex Transm Infect 2002 Apr;78 Suppl 1:i3-13.

12. Rockhill B. The privatization of risk. Am J Public Health 2001 Mar;91(3):365-368.

13. Rockhill B, Kawachi I, Colditz GA. Individual risk prediction and population-wide disease prevention. Epidemiol Rev 2000;22(1):176-180.

14. Golden MR, Kerndt PR. Improving Clinical Operations: Can We and Should We Save Our STD Clinics? Sex Transm Dis 2010 Apr;37(4):264-265.

15. Rietmeijer CA, Mettenbrink C. Why we should save our STD clinics. Sex Transm Dis 2010 Sep;37(9):591.

16. Perkins E. STI testing clinic closure unacceptable for Grand Forks. The Boundary Sentinel 2011 December 30 2011.

17. Gilbert M. Impact of an Integrated Internet-based Testing Program for HIV and Sexually Transmitted Infections in British Columbia. CIHR's Partnership for Health System Improvement Grant 2010 2014(350,000 CAD).

18. Fairley CK, Vodstrcil LA, Read T. The importance of striving for greater efficiency. Sex Health 2011 Mar;8(1):3-4.

19. Masaro CL, Johnson J, Chabot C, Shoveller J. STI service delivery in British Columbia, Canada; providers' views of their services to youth. BMC Health Serv Res 2012 Aug 6;12:240-6963-12-240.

20. Koekenbier RH, Davidovich U, van Leent EJ, Thiesbrummel HF, Fennema HS. Online-mediated syphilis testing: feasibility, efficacy, and usage. Sex Transm Dis 2008 Aug;35(8):764-769.

21. Hottes TS, Farrell J, Bondyra M, Haag D, Shoveller J, Gilbert M. Internet-based HIV and sexually transmitted infection testing in British Columbia, Canada: opinions and expectations of prospective clients. J Med Internet Res 2012 Mar 6;14(2):e41.

22. Chai SJ, Aumakhan B, Barnes M, Jett-Goheen M, Quinn N, Agreda P, et al. Internet-based screening for sexually transmitted infections to reach nonclinic populations in the community: risk factors for infection in men. Sex Transm Dis 2010 Dec;37(12):756-763.

23. Araujo L, Jacinto T, Moreira A, Castel-Branco MG, Delgado L, Costa-Pereira A, et al. Clinical efficacy of web-based versus standard asthma self-management. J Investig Allergol Clin Immunol 2012;22(1):28-34.

24. van den Berg SW, Peters EJ, Kraaijeveld JF, Gielissen MF, Prins JB. Usage of a generic web-based self-management intervention for breast cancer survivors: substudy analysis of the BREATH trial. J Med Internet Res 2013 Aug 19;15(8):e170.

25. van Bastelaar KM, Pouwer F, Cuijpers P, Riper H, Snoek FJ. Web-based depression treatment for type 1 and type 2 diabetic patients: a randomized, controlled trial. Diabetes Care 2011 Feb;34(2):320-325.

26. Greenland KE, Op de Coul EL, van Bergen JE, Brouwers EE, Fennema HJ, Gotz HM, et al. Acceptability of the internet-based Chlamydia screening implementation in the Netherlands and insights into nonresponse. Sex Transm Dis 2011 Jun;38(6):467-474.

27. Gotz HM, van Bergen JE, Veldhuijzen IK, Broer J, Hoebe CJ, Steyerberg EW, et al. A prediction rule for selective screening of Chlamydia trachomatis infection. Sex Transm Infect 2005 Feb;81(1):24-30.

28. Falasinnu T, Gustafson P, Gilbert M, Shoveller J. Risk prediction in sexual health contexts: protocol. JMIR Res Protoc 2013 Dec 3;2(2):e57.

29. Shamos SJ, Mettenbrink CJ, Subiadur JA, Mitchell BL, Rietmeijer CA. Evaluation of a testing-only "express" visit option to enhance efficiency in a busy STI clinic. Sex Transm Dis 2008 Apr;35(4):336-340.

30. Public Health Agency of Canada. **Canadian Guidelines on Sexually Transmitted Infections - Updated January 2010**. 2011; Available at: http://www.phac-aspc.gc.ca/std-mts/sti-its/cgsti-ldcits/index-eng.php. Accessed 07/22, 2013.

31. Aral SO. Sexual risk behaviour and infection: epidemiological considerations. Sex Transm Infect 2004 Dec;80 Suppl 2:ii8-12.

32. Niccolai LM, Ethier KA, Kershaw TS, Lewis JB, Meade CS, Ickovics JR. New sex partner acquisition and sexually transmitted disease risk among adolescent females. J Adolesc Health 2004 Mar;34(3):216-223.

33. Green J, Fulop N, Kocsis A. Determinants of unsafe sex in women. Int J STD AIDS 2000 Dec;11(12):777-783.

34. Ott MA, Katschke A, Tu W, Fortenberry JD. Longitudinal associations among relationship factors, partner change, and sexually transmitted infection acquisition in adolescent women. Sex Transm Dis 2011 Mar;38(3):153-157.

35. Bauer HM, Gibson P, Hernandez M, Kent C, Klausner J, Bolan G. Intimate partner violence and high-risk sexual behaviors among female patients with sexually transmitted diseases. Sex Transm Dis 2002 Jul;29(7):411-416.

36. Aral SO. Behavioral aspects of sexually transmitted diseases: core groups and bridge populations. Sex Transm Dis 2000 Jul;27(6):327-328.

37. Chiasson MA, Hirshfield S, Remien RH, Humberstone M, Wong T, Wolitski RJ. A comparison of on-line and off-line sexual risk in men who have sex with men: an event-based on-line survey. J Acquir Immune Defic Syndr 2007 Feb 1;44(2):235-243.

38. Lewnard JA, Berrang-Ford L. Internet-based partner selection and risk for unprotected anal intercourse in sexual encounters among men who have sex with men: a meta-analysis of observational studies. Sex Transm Infect 2014 Jun;90(4):290-296.

39. Foxman B, Newman M, Percha B, Holmes KK, Aral SO. Measures of sexual partnerships: lengths, gaps, overlaps, and sexually transmitted infection. Sex Transm Dis 2006 Apr;33(4):209-214.

40. Drumright LN, Gorbach PM, Holmes KK. Do people really know their sex partners? Concurrency, knowledge of partner behavior, and sexually transmitted infections within partnerships. Sex Transm Dis 2004 Jul;31(7):437-442.

41. Gorbach PM, Drumright LN, Holmes KK. Discord, discordance, and concurrency: comparing individual and partnership-level analyses of new partnerships of young adults at risk of sexually transmitted infections. Sex Transm Dis 2005 Jan;32(1):7-12.

42. Morris M, Kurth AE, Hamilton DT, Moody J, Wakefield S. Concurrent partnerships and HIV prevalence disparities by race: linking science and public health practice. Am J Public Health 2009 Jun;99(6):1023-1031.

43. Kraut-Becher JR, Aral SO. Gap length: an important factor in sexually transmitted disease transmission. Sex Transm Dis 2003 Mar;30(3):221-225.

44. Ellen JM, Gaydos C, Chung SE, Willard N, Lloyd LV, Rietmeijer CA. Sex partner selection, social networks, and repeat sexually transmitted infections in young men: a preliminary report. Sex Transm Dis 2006 Jan;33(1):18-21.

45. Hughes G, Brady AR, Catchpole MA, Fenton KA, Rogers PA, Kinghorn GR, et al. Characteristics of those who repeatedly acquire sexually transmitted infections: a retrospective cohort study of attendees at three urban sexually transmitted disease clinics in England. Sex Transm Dis 2001 Jul;28(7):379-386.

46. Phipps W, Kent CK, Kohn R, Klausner JD. Risk factors for repeat syphilis in men who have sex with men, San Francisco. Sex Transm Dis 2009 Jun;36(6):331-335.

47. Wilson JM, Jungner YG. Principles and practice of mass screening for disease. Bol Oficina Sanit Panam 1968 Oct;65(4):281-393.

48. Sniderman AD, Furberg CD. Why guideline-making requires reform. JAMA 2009 Jan 28;301(4):429-431.

49. van Valkengoed IG, Boeke AJ, Morre SA, van den Brule AJ, Meijer CJ, Deville W, et al. Disappointing performance of literature-derived selective screening criteria for asymptomatic Chlamydia trachomatis infection in an inner-city population. Sex Transm Dis 2000 Oct;27(9):504-507.

50. Miller WC, Hoffman IF, Owen-O'Dowd J, McPherson JT, Privette A, Schmitz JL, et al. Selective screening for chlamydial infection: which criteria to use? Am J Prev Med 2000 Feb;18(2):115-122.

51. La Montagne DS, Patrick LE, Fine DN, Marrazzo JM, Region X Infertility Prevention Project. Re-evaluating selective screening criteria for chlamydial infection among women in the U S Pacific Northwest. Sex Transm Dis 2004 May;31(5):283-289.

52. Owens DK. Improving practice guidelines with patient-specific recommendations. Ann Intern Med 2011 May 3;154(9):638-639.

53. Beattie P, Nelson R. Clinical prediction rules: what are they and what do they tell us? Aust J Physiother 2006;52(3):157-163.

54. Childs JD, Cleland JA. Development and application of clinical prediction rules to improve decision making in physical therapist practice. Phys Ther 2006 Jan;86(1):122-131.

55. Ewout W. Steyerberg. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. New York: Springer; 2009.

56. Siontis GC, Tzoulaki I, Siontis KC, Ioannidis JP. Comparisons of established risk prediction models for cardiovascular disease: systematic review. BMJ 2012 May 24;344:e3318.

57. D'Agostino RB S, Grundy S, Sullivan LM, Wilson P, CHD Risk Prediction Group. Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. JAMA 2001 Jul 11;286(2):180-187.

58. Haukoos JS, Lyons MS, Lindsell CJ, Hopkins E, Bender B, Rothman RE, et al. Derivation and validation of the Denver Human Immunodeficiency Virus (HIV) risk score for targeted HIV screening. Am J Epidemiol 2012 Apr 15;175(8):838-846.

59. Bossuyt PM, Reitsma JB, Standards for Reporting of Diagnostic Accuracy. The STARD initiative. Lancet 2003 Jan 4;361(9351):71.

60. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Ann Intern Med 1999 Mar 16;130(6):515-524.

61. Wallace E, Smith SM, Perera-Salazar R, Vaucher P, McCowan C, Collins G, et al. Framework for the impact analysis and implementation of Clinical Prediction Rules (CPRs). BMC Med Inform Decis Mak 2011 Oct 14;11:62-6947-11-62.

62. Cook N. **Use and Misuse of the Receiver Operating Characteristic Curve in Risk Prediction**. Circulation 2007;115:928.

63. BC Centre for Disease Control. British Columbia Annual Summary of Reportable Diseases, 2009. 2010.

64. Mahony JB, Luinstra KE, Tyndall M, Sellors JW, Krepel J, Chernesky M. Multiplex PCR for detection of Chlamydia trachomatis and Neisseria gonorrhoeae in Genitourinary specimens. J Clin Microbiol 1995 Nov;33(11):3049-3053.

65. Lewis JJ, Donnelly CA, Mare P, Mupambireyi Z, Garnett GP, Gregson S. Evaluating the proximate determinants framework for HIV infection in rural Zimbabwe. Sex Transm Infect 2007 Aug;83 Suppl 1:i61-69.

66. Lopman B, Nyamukapa C, Mushati P, Mupambireyi Z, Mason P, Garnett GP, et al. HIV incidence in 3 years of follow-up of a Zimbabwe cohort--1998-2000 to 2001-03: contributions of proximate and underlying determinants to transmission. Int J Epidemiol 2008 Feb;37(1):88-105.

67. Meyers D, Wolff T, Gregory K, Marion L, Moyer V, Nelson H, et al. USPSTF recommendations for STI screening. Am Fam Physician 2008 Mar 15;77(6):819-824.

68. Murdoch TB, Detsky AS. The inevitable application of big data to health care. JAMA 2013 Apr 3;309(13):1351-1352.

69. Knottnerus J, Buntinx F editors. **The Evidence Base of Clinical Diagnosis:** Theory and Methods of Diagnostic Research. ; 2009.

70. Fisman DN, Laupland KB. Sexually transmitted infections in Canada: A sticky situation. Can J Infect Dis Med Microbiol 2011 Fall;22(3):80-82.

71. Geisler WM, Chow JM, Schachter J, McCormack WM. Pelvic examination findings and Chlamydia trachomatis infection in asymptomatic young women screened with a nucleic acid amplification test. Sex Transm Dis 2007 Jun;34(6):335-338.

72. Manhart LE, Marrazzo JM, Fine DN, Kerani RP, Golden MR. Selective testing criteria for gonorrhea among young women screened for Chlamydial infection: contribution of race and geographic prevalence. J Infect Dis 2007 Sep 1;196(5):731-737.

73. Ownsworth T, McKenna K. Investigation of factors related to employment outcome following traumatic brain injury: a critical review and conceptual model. Disabil Rehabil 2004 Jul 8;26(13):765-783.

74. Turner BJ, Fleming JM, Ownsworth TL, Cornwell PL. The transition from hospital to home for individuals with acquired brain injury: a literature review and research recommendations. Disabil Rehabil 2008;30(16):1153-1176.

75. Boerma JT, Weir SS. Integrating demographic and epidemiological approaches to research on HIV/AIDS: the proximate-determinants framework. J Infect Dis 2005 Feb 1;191 Suppl 1:S61-7.

76. de Rooij A, Roorda LD, Otten RH, van der Leeden M, Dekker J, Steultjens MP. Predictors of multidisciplinary treatment outcome in fibromyalgia:a systematic review. Disabil Rehabil 2013 Mar;35(6):437-449.

77. Ariens GA, van Mechelen W, Bongers PM, Bouter LM, van der Wal G. Psychosocial risk factors for neck pain: a systematic review. Am J Ind Med 2001 Feb;39(2):180-193.

78. U.S. Preventive Services Task Force. *Screening for Gonorrhea.* 2005; Available at: http://www.uspreventiveservicestaskforce.org/uspstf/uspsgono.htm. Accessed 07/22, 2013.

79. U.S. Preventive Services Task Force. Screening for chlamydial infection: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2007 Jul 17;147(2):128-134.

80. Workowski KA, Berman S, Centers for Disease Control and Prevention (CDC). Sexually transmitted diseases treatment guidelines, 2010. MMWR Recomm Rep 2010 Dec 17;59(RR-12):1-110.

81. American Congress of Obstetricians and Gynecologists. Annual Women's Healthcare: High-risk factors. 2013; Available at: http://www.acog.org/About_ACOG/ACOG_Departments/Annual_Womens_Health_Care/High-Risk_Factors. Accessed 07/22, 2013.

82. The Society of Obstetricians and Gynaecologists of Canada. **Screening recommendations for asymptomatic individuals**. 2012; Available at: http://www.sexualityandu.ca/health-care-professionals/sexually-transmitted-infections/screening-recommendations-for-asymptomatic-individuals. Accessed 07/22, 2013.

83. British Association for Sexual Health and HIV. Sexually Transmitted Infections: UK National Screening and Testing Guidelines. 2006; Available at: http://www.bashh.org/documents/59/59.pdf. Accessed 07/22, 2013.

84. Australasian Society for HIV Medicine. Guidelines for Managing Sexually Transmitted Infections: Screening of Asymptomatic Men and Women. 2013; Available at: http://silverbook.health.wa.gov.au/Default.asp?PublicationID=1&SectionID=236. Accessed 07/22, 2013.

85. Radcliffe KW, European Branch of the International Union against Sexually Transmitted Infection and the European Office of the World Health Organization. Introduction. European STD guidelines. Int J STD AIDS 2001 Oct;12 Suppl 3:2-3.

86. Kong FY, Hocking JS, Link CK, Chen MY, Hellard ME. Sex and sport: chlamydia screening in rural sporting clubs. BMC Infect Dis 2009;9:73.

87. Greene JP, Stafford E. Prevalence of Chlamydia trachomatis among active duty male soldiers reporting to a troop medical clinic for routine health care. South Med J 2007 May;100(5):478-481.

88. Mak RP, Van Renterghem L, Traen A. Chlamydia trachomatis in female sex workers in Belgium: 1998-2003. Sex Transm Infect 2005 Feb;81(1):89-90.

89. Menon-Johansson AS, Winston A, Matthews G, Portsmouth S, Daniels D. The first point prevalence study of genital Chlamydia trachomatis infection in young male inmates in the UK. Int J STD AIDS 2005 Dec;16(12):799-801.

90. Powell J, O'Connor C, O'hlarlaithe M, Saunders J, De Freitas J. Chlamydia trachomatis prevalence in men in the mid-west of Ireland. Sex Transm Infect 2004 Oct;80(5):349-353.

91. Annan NT, Sullivan AK, Nori A, Naydenova P, Alexander S, McKenna A, et al. Rectal chlamydia--a reservoir of undiagnosed infection in men who have sex with men. Sex Transm Infect 2009 Jun;85(3):176-179.

92. Dicker LW, Mosure DJ, Kay RS, Shelby L, Cheek JE, Region VIII Infertility Prevention P. An ongoing burden: chlamydial infections among young American Indian women. Matern Child Health J 2008 Jul;12(Suppl 1):25-29.

93. Hince D, Brett T, Mak D, Bulsara M, Moorhead R, Arnold-Reed D. Opportunistic screening for chlamydia in young men. Aust Fam Physician 2009 Sep;38(9):734-737.

94. Spasovski MS, Simjanovska LJ, Taleski V, Petrova N, Lazetic L, Popeska Z, et al. Screening of Chlamydia trachomatis urogenital infections among the male and female population of the Republic of Macedonia. J Eur Acad Dermatol Venereol 2005 Jul;19(4):427-430.

95. Simms I, Talebi A, Rhia J, Horner P, French RS, Sarah R, et al. The English National Chlamydia Screening Programme: variations in positivity in 2007/2008. Sex Transm Dis 2009 Aug;36(8):522-527.

96. Hocking J, Fairley CK. Do the characteristics of sexual health centre clients predict chlamydia infection sufficiently strongly to allow selective screening? Sex Health 2005;2(3):185-192.

97. Tipple C, Hill SC, Smith A. Is screening for pharyngeal Chlamydia trachomatis warranted in high-risk groups?. Int J STD AIDS 2010 Nov;21(11):770-771.

98. Mossong J, Muller M, Majery N, Mardaga C, Decruyenaere F, Schneider F. Screening for Chlamydia trachomatis in secondary schools, family planning and occupational health centres in Luxembourg. Sex Transm Infect 2009 Oct;85(6):455-458.

99. Paukku M, Kilpikari R, Puolakkainen M, Oksanen H, Apter D, Paavonen J. Criteria for selective screening for Chlamydia trachomatis. Sex Transm Dis 2003 Feb;30(2):120-123.

100. Ivens D, Macdonald K, Bansi L, Nori A. Screening for rectal chlamydia infection in a genitourinary medicine clinic. Int J STD AIDS 2007 Jun;18(6):404-406.

101. Landes M, Thorne C, Barlow P, Fiore S, Malyuta R, Martinelli P, et al. Prevalence of sexually transmitted infections in HIV-1 infected pregnant women in Europe. Eur J Epidemiol 2007;22(12):925-936.

102. Barry PM, Kent CK, Klausner JD. Risk factors for gonorrhea among heterosexuals--San Francisco, 2006. Sex Transm Dis 2009 Feb;36(2 Suppl):S62-6.

103. Hirst H, Dinsmore WW. Chlamydia screening programme, HMP Hydebank Wood Young Offender Centre, Belfast, Northern Ireland, UK. Int J STD AIDS 2009 May;20(5):360-361.

104. Schillinger JA, Dunne EF, Chapin JB, Ellen JM, Gaydos CA, Willard NJ, et al. Prevalence of Chlamydia trachomatis infection among men screened in 4 U.S. cities. Sex Transm Dis 2005 Feb;32(2):74-77.

105. Matteelli A, Beltrame A, Carvalho AC, Casalini C, Forleo MA, Gulletta M, et al. Chlamydia trachomatis genital infection in migrant female sex workers in Italy. Int J STD AIDS 2003 Sep;14(9):591-595.

106. Steenbeek A, Tyndall M, Sheps S, Rothenberg R. An epidemiological survey of chlamydial and gonococcal infections in a Canadian arctic community. Sex Transm Dis 2009 Feb;36(2):79-83.

107. Geisler WM, James AB. Chlamydial and gonococcal infections in women seeking pregnancy testing at family-planning clinics. Am J Obstet Gynecol 2008 May;198(5):502.e1-502.e4.

108. Bakken IJ, Nordbo SA, Skjeldestad FE. Chlamydia trachomatis testing patterns and prevalence of genital chlamydial infection among young men and women in central Norway 1990-2003: a population-based registry study. Sex Transm Dis 2006 Jan;33(1):26-30.

109. Bozicevic I, Grgic I, Zidovec-Lepej S, Cakalo JI, Belak-Kovacevic S, Stulhofer A, et al. Urine-based testing for Chlamydia trachomatis among young adults in a population-based survey in Croatia: feasibility and prevalence. BMC Public Health 2011;11:230.

110. Gaydos CA, Kent CK, Rietmeijer CA, Willard NJ, Marrazzo JM, Chapin JB, et al. Prevalence of Neisseria Gonorrhoeae among men screened for Chlamydia Trachomatis in four United States cities, 1999-2003. Sex Transm Dis 2006 May;33(5):314-319.

111. Jackson Y, Sebo P, Aeby G, Bovier P, Ninet B, Schrenzel J, et al. Prevalence and associated factors for Chlamydia trachomatis infection among undocumented immigrants in a primary care facility in Geneva, Switzerland: a cross-sectional study. J Immigr Minor Health 2010 Dec;12(6):909-914.

112. Joesoef MR, Mosure DJ. Prevalence of Chlamydia in young men in the United States from newly implemented universal screening in a national job training program. Sex Transm Dis 2006 Oct;33(10):636-639.

113. Richardson E, Sellors JW, Mackinnon S, Woodcox V, Howard M, Jang D, et al. Prevalence of Chlamydia trachomatis infections and specimen collection preference among women, using self-collected vaginal swabs in community settings. Sex Transm Dis 2003 Dec;30(12):880-885.

114. Singh D, Fine DN, Marrazzo JM. Chlamydia trachomatis infection among women reporting sexual activity with women screened in Family Planning Clinics in the Pacific Northwest, 1997 to 2005. Am J Public Health 2011 Jul;101(7):1284-1290.

115. Steiner AS, Haller DM, Elger BS, Sebo P, Gaspoz JM, Wolff H. Chlamydia trachomatis infection in a Swiss prison: a cross sectional study. Swiss Med Wkly 2010;140:w13126.

116. Sutton TL, Martinko T, Hale S, Fairchok MP. Prevalence and high rate of asymptomatic infection of Chlamydia trachomatis in male college Reserve Officer Training Corps cadets. Sex Transm Dis 2003 Dec;30(12):901-904.

117. Uuskula A, Kals M, Denks K, Nurm U, Kasesalu L, Dehovitz J, et al. The prevalence of chlamydial infection in Estonia: a population-based survey. Int J STD AIDS 2008 Jul;19(7):455-458.

118. Stein CR, Kaufman JS, Ford CA, Leone PA, Feldblum PJ, Miller WC. Screening young adults for prevalent chlamydial infection in community settings. Ann Epidemiol 2008 Jul;18(7):560-571.

119. Andersen B, van Valkengoed I, Olesen F, Moller JK, Ostergaard L. Value of self-reportable screening criteria to identify asymptomatic individuals in the general population for urogential Chlamydia trachomatis infection screening. Clin Infect Dis 2003 Apr 1;36(7):837-844.

120. Bernstein KT, Chow JM, Ruiz J, Schachter J, Horowitz E, Bunnell R, et al. Chlamydia trachomatis and Neisseria gonorrhoeae infections among men and women entering California prisons. Am J Public Health 2006 Oct;96(10):1862-1866.

121. Goulet V, de Barbeyrac B, Raherison S, Prudhomme M, Semaille C, Warszawski J, et al. Prevalence of Chlamydia trachomatis: results from the first national population-based survey in France. Sex Transm Infect 2010 Aug;86(4):263-270.

122. Wand H, Guy R, Donovan B, McNulty A. Developing and validating a risk scoring tool for chlamydia infection among sexual health clinic attendees in Australia: a simple algorithm to identify those at high risk of chlamydia infection. BMJ Open 2011 Jan 1;1(1):e000005.

123. Belenko S, Dembo R, Weiland D, Rollie M, Salvatore C, Hanlon A, et al. Recently arrested adolescents are at high risk for sexually transmitted diseases. Sex Transm Dis 2008 Aug;35(8):758-763.

124. Asbel LE, Newbern EC, Salmon M, Spain CV, Goldberg M. School-based screening for Chlamydia trachomatis and Neisseria gonorrhoeae among Philadelphia public high school students. Sex Transm Dis 2006 Oct;33(10):614-620.

125. Robertson AA, Thomas CB, St Lawrence JS, Pack R. Predictors of infection with Chlamydia or gonorrhea in incarcerated adolescents. Sex Transm Dis 2005 Feb;32(2):115-122.

126. Kahn RH, Moseley KE, Thilges JN, Johnson G, Farley TA. Community-based screening and treatment for STDs: results from a mobile clinic initiative. Sex Transm Dis 2003 Aug;30(8):654-658.

127. Doherty IA, Adimora AA, Schoenbach VJ, Aral SO. Correlates of gonorrhoea among African Americans in North Carolina. Int J STD AIDS 2007 Feb;18(2):114-117.

128. McDonagh P, Ryder N, McNulty AM, Freedman E. Neisseria gonorrhoeae infection in urban Sydney women: prevalence and predictors. Sex Health 2009 Sep;6(3):241-244.

129. Rao GG, Bacon L, Evans J, Dejahang Y, Michalczyk P, Donaldson N, et al. Prevalence of Neisseria gonorrhoeae infection in young subjects attending community clinics in South London. Sex Transm Infect 2008 Apr;84(2):117-121.

130. Bozicevic I, Fenton KA, Martin IM, Rudd EA, Ison CA, Nanchahal K, et al. Epidemiological correlates of asymptomatic gonorrhea. Sex Transm Dis 2006 May;33(5):289-295.

131. Gorgos L, Fine D, Marrazzo J. Chlamydia positivity in American Indian/Alaska Native women screened in family planning clinics, 1997-2004. Sex Transm Dis 2008 Aug;35(8):753-757.

132. LaMontagne DS, Fenton KA, Randall S, Anderson S, Carter P. Establishing the National Chlamydia Screening Programme in England: results from the first full year of screening. Sex Transm Infect 2004 Oct;80(5):335-341.

133. Lofy KH, Hofmann J, Mosure DJ, Fine DN, Marrazzo JM. Chlamydial infections among female adolescents screened in juvenile detention centers in Washington State, 1998-2002. Sex Transm Dis 2006 Feb;33(2):63-67.

134. Satterwhite CL, Tian LH, Braxton J, Weinstock H. Chlamydia prevalence among women and men entering the National Job Training Program: United States, 2003-2007. Sex Transm Dis 2010 Feb;37(2):63-67.

135. Arcari CM, Gaydos JC, Howell MR, McKee KT, Gaydos CA. Feasibility and short-term impact of linked education and urine screening interventions for Chlamydia and gonorrhea in male army recruits. Sex Transm Dis 2004 Jul;31(7):443-447.

136. Trick WE, Kee R, Murphy-Swallow D, Mansour M, Mennella C, Raba JM. Detection of chlamydial and gonococcal urethral infection during jail intake: development of a screening algorithm. Sex Transm Dis 2006 Oct;33(10):599-603.

137. Javanbakht M, Murphy R, Harawa NT, Smith LV, Hayes M, Chien M, et al. Sexually transmitted infections and HIV prevalence among incarcerated men who have sex with men, 2000-2005. Sex Transm Dis 2009 Feb;36(2 Suppl):S17-21.

138. Gaydos CA, Howell MR, Quinn TC, McKee KT,Jr, Gaydos JC. Sustained high prevalence of Chlamydia trachomatis infections in female army recruits. Sex Transm Dis 2003 Jul;30(7):539-544.

139. Geisler WM, Chyu L, Kusunoki Y, Upchurch DM, Hook EW,3rd. Health insurance coverage, health care-seeking behaviors, and genital chlamydial infection prevalence in sexually active young adults. Sex Transm Dis 2006 Jun;33(6):389-396.

140. Joffe A, Rietmeijer CA, Chung SE, Willard N, Chapin JB, Lloyd LV, et al. Screening asymptomatic adolescent men for Chlamydia trachomatis in school-based health centers using urine-based nucleic acid amplification tests. Sex Transm Dis 2008 Nov;35(11 Suppl):S19-23.

141. McDonnell DD, Levy V, Morton TJ. Risk factors for Chlamydia among young women in a northern california juvenile detention facility: implications for community intervention. Sex Transm Dis 2009 Feb;36(2 Suppl):S29-33.

142. Scott Lamontagne D, Baster K, Emmett L, Nichols T, Randall S, McLean L, et al. Incidence and reinfection rates of genital chlamydial infection among women aged 16-24 years attending general practice, family planning and genitourinary medicine clinics in England: a prospective cohort study by the Chlamydia Recall Study Advisory Group. Sex Transm Infect 2007 Jul;83(4):292-303.

143. Skjeldestad FE, Marsico MA, Sings HL, Nordbo SA, Storvold G. Incidence and risk factors for genital Chlamydia trachomatis infection: a 4-year prospective cohort study. Sex Transm Dis 2009 May;36(5):273-279.

144. van Bergen J, Gotz HM, Richardus JH, Hoebe CJ, Broer J, Coenen AJ, et al. Prevalence of urogenital Chlamydia trachomatis increases significantly with level of urbanisation and suggests targeted screening approaches: results from the first national population based study in the Netherlands. Sex Transm Infect 2005 Feb;81(1):17-23.

145. Gotz HM, Veldhuijzen IK, Habbema JD, Boeke AJ, Richardus JH, Steyerberg EW. Prediction of Chlamydia trachomatis infection: application of a scoring rule to other populations. Sex Transm Dis 2006 Jun;33(6):374-380.

146. Latino MA, Caneparo A, Rosso C, De Maria D, De Intinis G, Intorcia P, et al. Prevalence and risk factors for Chlamydia trachomatis infection in young women in north-west of Italy. Minerva Ginecol 2008 Feb;60(1):29-37.

147. Salfa MC, Latino MA, Regine V, De Maria D, De Intinis G, Camoni L, et al. Prevalence and determinants of Chlamydia trachomatis infection among sexually active women in Turin, Italy. ITALIAN JOURNAL OF PUBLIC HEALTH 2011;8(3):295-301.

148. Berry SA, Ghanem KG, Page KR, Gange SJ, Thio CL, Moore RD, et al. Increased gonorrhoea and chlamydia testing did not increase case detection in an HIV clinical cohort 1999-2007. Sex Transm Infect 2011 Oct;87(6):469-475.

149. Datta SD, Sternberg M, Johnson RE, Berman S, Papp JR, McQuillan G, et al. Gonorrhea and chlamydia in the United States among persons 14 to 39 years of age, 1999 to 2002. Ann Intern Med 2007 Jul 17;147(2):89-96.

150. Einwalter LA, Ritchie JM, Ault KA, Smith EM. Gonorrhea and chlamydia infection among women visiting family planning clinics: racial variation in prevalence and predictors. Perspect Sex Reprod Health 2005 Sep;37(3):135-140.

151. Hardick J, Hsieh YH, Tulloch S, Kus J, Tawes J, Gaydos CA. Surveillance of Chlamydia trachomatis and Neisseria gonorrhoeae infections in women in detention in Baltimore, Maryland. Sex Transm Dis 2003 Jan;30(1):64-70.

152. Jin F, Prestage GP, Mao L, Kippax SC, Pell CM, Donovan B, et al. Incidence and risk factors for urethral and anal gonorrhoea and chlamydia in a cohort of HIV-negative homosexual men: the Health in Men Study. Sex Transm Infect 2007 Apr;83(2):113-119.

153. Al-Tayyib AA, Miller WC, Rogers SM, Leone PA, Law DC, Ford CA, et al. Evaluation of risk score algorithms for detection of chlamydial and gonococcal infections in an emergency department setting. Acad Emerg Med 2008 Feb;15(2):126-135.

154. Merchant RC, DePalo DM, Liu T, Rich JD, Stein MD. Developing a system to predict laboratory-confirmed chlamydial and/or gonococcal urethritis in adult male emergency department patients. Postgrad Med 2010 Jan;122(1):52-60.

155. Mehta SD, Hall J, Lyss SB, Skolnik PR, Pealer LN, Kharasch S. Adult and pediatric emergency department sexually transmitted disease and HIV screening: programmatic overview and outcomes. Acad Emerg Med 2007 Mar;14(3):250-258.

156. Chen MY, Rohrsheim R, Donovan B. Chlamydia trachomatis infection in Sydney women. Aust N Z J Obstet Gynaecol 2005 Oct;45(5):410-413.

157. Vanable PA, Carey MP, Brown JL, DiClemente RJ, Salazar LF, Brown LK, et al. Test-retest reliability of self-reported HIV/STD-related measures among African-American adolescents in four U.S. cities. J Adolesc Health 2009 Mar;44(3):214-221.

158. Adimora AA, Schoenbach VJ, Martinson FE, Donaldson KH, Fullilove RE, Aral SO. Social context of sexual relationships among rural African Americans. Sex Transm Dis 2001 Feb;28(2):69-76.

159. Guan J, Wu Z, Li L, Lin C, Rotheram-Borus MJ, Detels R, et al. Self-reported sexually transmitted disease symptoms and treatment-seeking behaviors in China. AIDS Patient Care STDS 2009 Jun;23(6):443-448.

160. Matser A, Luu N, Geskus R, Heijman T, Heiligenberg M, van Veen M, et al. Higher Chlamydia trachomatis Prevalence in Ethnic Minorities Does Not Always Reflect Higher Sexual Risk Behaviour. PLoS One 2013 Jun 14;8(6):e67287.

161. Khangura S, Konnyu K, Cushman R, Grimshaw J, Moher D. Evidence summaries: the evolution of a rapid review approach. Syst Rev 2012 Feb 10;1(1):10-4053-1-10.

162. Buscemi N, Hartling L, Vandermeer B, Tjosvold L, Klassen TP. Single data extraction generated more errors than double data extraction in systematic reviews. J Clin Epidemiol 2006 Jul;59(7):697-703.

163. Ganann R, Ciliska D, Thomas H. Expediting systematic reviews: methods and implications of rapid reviews. Implement Sci 2010 Jul 19;5:56-5908-5-56.

164. Maguire JL, Kulik DM, Laupacis A, Kuppermann N, Uleryk EM, Parkin PC. Clinical prediction rules for children: a systematic review. Pediatrics 2011 Sep;128(3):e666-77.

165. Facente SN, Pilcher CD, Hartogensis WE, Klausner JD, Philip SS, Louie B, et al. Performance of risk-based criteria for targeting acute HIV screening in San Francisco. PLoS ONE 2011;6(7):e21813.

166. Miller WC, Leone PA, McCoy S, Nguyen TQ, Williams DE, Pilcher CD. Targeted testing for acute HIV infection in North Carolina. AIDS 2009 Apr 27;23(7):835-843.

167. Verhoeven V, Avonts D, Meheus A, Goossens H, Ieven M, Chapelle S, et al. Chlamydial infection: an accurate model for opportunistic screening in general practice. Sex Transm Infect 2003 Aug;79(4):313-317.

168. Kulik DM, Uleryk EM, Maguire JL. Does this child have appendicitis? A systematic review of clinical prediction rules for children with acute abdominal pain. J Clin Epidemiol 2013 Jan;66(1):95-104.

169. Bouwmeester W, Zuithoff NP, Mallett S, Geerlings MI, Vergouwe Y, Steyerberg EW, et al. Reporting and methods in clinical prediction research: a systematic review. PLoS Med 2012;9(5):1-12.

170. Cook NR. Statistical evaluation of prognostic versus diagnostic models: beyond the ROC curve. Clin Chem 2008 Jan;54(1):17-23.

171. Echouffo-Tcheugui JB, Kengne AP. Risk models to predict chronic kidney disease and its progression: a systematic review. PLoS Med 2012;9(11):e1001344.

172. Fairley CK, Williams H, Lee DM, Cummings R. A plea for more research on access to sexual health services. Int J STD AIDS 2007 Feb;18(2):75-76.

173. Heymans MW, van Buuren S, Knol DL, van Mechelen W, de Vet HC. Variable selection under multiple imputation using the bootstrap in a prognostic study. BMC Med Res Methodol 2007 Jul 13;7:33.

174. Janssen KJ, Vergouwe Y, Donders AR, Harrell FE,Jr, Chen Q, Grobbee DE, et al. Dealing with missing predictor values when applying clinical prediction models. Clin Chem 2009 May;55(5):994-1001.

175. Vergouwe Y, Royston P, Moons KG, Altman DG. Development and validation of a prediction model with missing predictor data: a practical approach. J Clin Epidemiol 2010 Feb;63(2):205-214.

176. Janssen KJ, Donders AR, Harrell FE,Jr, Vergouwe Y, Chen Q, Grobbee DE, et al. Missing covariate data in medical research: to impute is better than to ignore. J Clin Epidemiol 2010 Jul;63(7):721-727.

177. Survey Research Center, Institute for Social Research, University of Michigan. **IVEware: Imputation and Variance Estimation Software**. 2013; Available at: http://www.isr.umich.edu/src/smp/ive/. Accessed 08/07, 2013.

178. He Y, Raghunathan T. **On the Performance of Sequential Regression Multiple Imputation Methods with Non Normal Error Distributions**. Communications in Statistics - Simulation and Computation 2009;38(4).

179. Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart 2012 May;98(9):691-698.

180. Vickers AJ, Cronin AM. Traditional statistical methods for evaluating prediction models are uninformative as to clinical value: towards a decision analytic framework. Semin Oncol 2010 Feb;37(1):31-38.

181. Toll DB, Janssen KJ, Vergouwe Y, Moons KG. Validation, updating and impact of clinical prediction rules: a review. J Clin Epidemiol 2008 Nov;61(11):1085-1094.

182. Gaydos CA, Dwyer K, Barnes M, Rizzo-Price PA, Wood BJ, Flemming T, et al. Internet-based screening for Chlamydia trachomatis to reach non-clinic populations with mailed self-administered vaginal swabs. Sex Transm Dis 2006 Jul;33(7):451-457.

183. D'Agostino RB S, Grundy S, Sullivan LM, Wilson P, CHD Risk Prediction Group. Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. JAMA 2001 Jul 11;286(2):180-187.

184. Sanagou M, Wolfe R, Leder K, Reid CM. External validation and updating of a prediction model for nosocomial pneumonia after coronary artery bypass graft surgery. Epidemiol Infect 2014 Mar;142(3):540-544.

185. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988 Sep;44(3):837-845.

186. Goldenberg S, Shoveller J, Koehoorn M, Ostry A. Barriers to STI testing among youth in a Canadian oil and gas community. Health Place 2008 Dec;14(4):718-729.

187. Goldenberg S, Shoveller J, Ostry A, Koehoorn M. Youth sexual behaviour in a boomtown: implications for the control of sexually transmitted infections. Sex Transm Infect 2008 Jun;84(3):220-223.

188. Shoveller J, Johnson J, Rosenberg M, Greaves L, Patrick DM, Oliffe JL, et al. Youth's experiences with STI testing in four communities in British Columbia, Canada. Sex Transm Infect 2009 Sep;85(5):397-401.

189. Xu F, Stoner BP, Taylor SN, Mena L, Martin DH, Powell S, et al. "Testing-only" visits: an assessment of missed diagnoses in clients attending sexually transmitted disease clinics. Sex Transm Dis 2013 Jan;40(1):64-69.

190. Han Y, Coles FB, Hipp S. Screening criteria for Chlamydia trachomatis in family planning clinics: accounting for prevalence and clients' characteristics. Fam Plann Perspect 1997 Jul-Aug;29(4):163-166.

191. Eddy DM, Adler J, Patterson B, Lucas D, Smith KA, Morris M. Individualized guidelines: the potential for increasing quality and reducing costs. Ann Intern Med 2011 May 3;154(9):627-634.

192. Chen MY, Donovan B. Screening for genital Chlamydia trachomatis infection: are men the forgotten reservoir? Med J Aust 2003 Aug 4;179(3):124-125.

193. Oliffe JL, Chabot C, Knight R, Davis W, Bungay V, Shoveller JA. Women on men's sexual health and sexually transmitted infection testing: a gender relations analysis. Sociol Health Illn 2013 Jan;35(1):1-16.

194. Shoveller JA, Knight R, Johnson J, Oliffe JL, Goldenberg S. 'Not the swab!' Young men's experiences with STI testing. Sociol Health Illn 2010 Jan;32(1):57-73.

195. Shoveller J, Knight R, Davis W, Gilbert M, Ogilvie G. Online sexual health services: examining youth's perspectives. Can J Public Health 2012 Jan-Feb;103(1):14-18.

196. Mercer CH, Copas AJ, Sonnenberg P, Johnson AM, McManus S, Erens B, et al. Who has sex with whom? Characteristics of heterosexual partnerships reported in a national probability survey and implications for STI risk. Int J Epidemiol 2009 Feb;38(1):206-214.

197. Phelan JC, Link BG, Tehranifar P. Social conditions as fundamental causes of health inequalities: theory, evidence, and policy implications. J Health Soc Behav 2010;51 Suppl:S28-40.

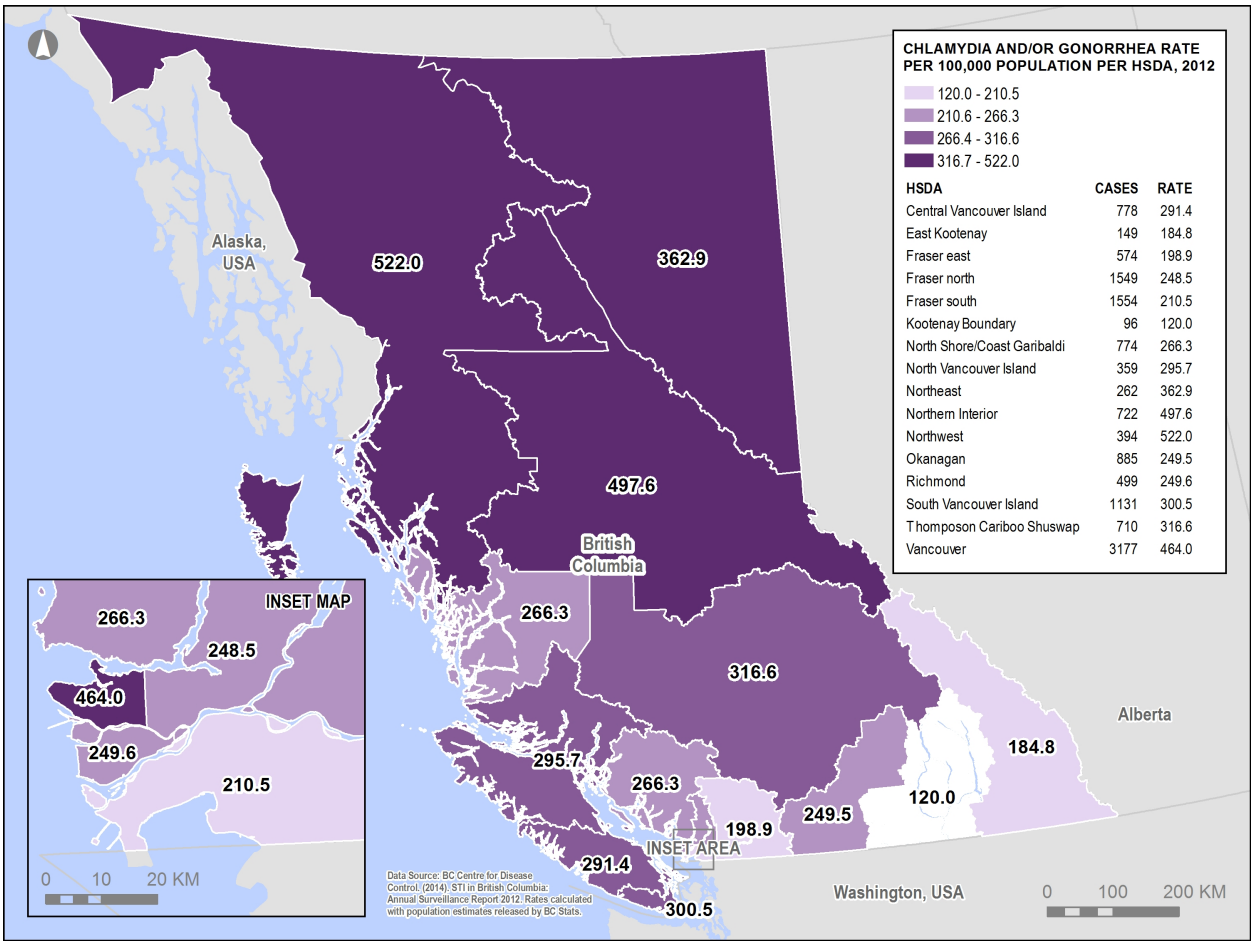
198. van den Broek IV, Brouwers EE, Gotz HM, van Bergen JE, Op de Coul EL, Fennema JS, et al. Systematic selection of screening participants by risk score in a Chlamydia screening programme is feasible and effective. Sex Transm Infect 2012 Apr;88(3):205-211.

199. Martin L, Knight V, Ryder N, Lu H, Read PJ, McNulty A. Client feedback and satisfaction with an express sexually transmissible infection screening service at an inner-city sexual health center. Sex Transm Dis 2013 Jan;40(1):70-74.

200. Janssen KJ, Vergouwe Y, Kalkman CJ, Grobbee DE, Moons KG. A simple method to adjust clinical prediction models to local circumstances. Can J Anaesth 2009 Mar;56(3):194-201.

Appendices

Appendix A : Map showing the distribution of chlamydia and/gonorrhoea rates by Health Service Delivery Area (HSDA), 2012



Appendix B : Map showing the geographical distribution of STI clinics used to derive and validate the prediction rule