Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Catherine Lindsey Satterwhite Date

Chlamydia Surveillance in the United States:
New Analytic Approaches and Alternate Considerations for
Monitoring Trends in Disease Burden

By

Catherine Lindsey Satterwhite, MSPH, MPH
Doctor of Philosophy

Epidemiology

_____

Penelope P. Howards, Ph.D., M.S.
Advisor

_____

David Kleinbaum, Ph.D., M.A.
Advisor

_____

Stuart Berman, M.D., Sc.M.
Committee Member

_____

James Buehler, M.D.
Committee Member

_____

Hillard Weinstock, M.D., M.P.H.
Committee Member

Accepted:

_____

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____

Date

Chlamydia Surveillance in the United States: New Analytic Approaches and Alternate Considerations for Monitoring Trends in Disease Burden

By

Catherine Lindsey Satterwhite
B.A., New York University, 1999
M.P.H., Emory University, 2000
M.S.P.H., Emory University, 2004

Advisors: Penelope P. Howards, Ph.D., M.S.
David Kleinbaum, Ph.D., M.A.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Epidemiology.
2011

Abstract

Chlamydia Surveillance in the United States:
New Analytic Approaches and Alternate Considerations for
Monitoring Trends in Disease Burden
By Catherine Lindsey Satterwhite

Over 1.2 million cases of chlamydia were reported in the U.S. in 2009. Chlamydia may lead to pelvic inflammatory disease (PID), ectopic pregnancy, and infertility. Annual screening is recommended for sexually active women aged <25 years. Reported case rates have increased likely due to increased screening and improved test technology. Other data suggest prevalence has decreased. Studies in this dissertation explored new analytic approaches to enhance the utility of U.S. chlamydia surveillance data.

Using data from the Infertility Prevention Project (IPP), a national chlamydia screening program, chlamydia positivity trends among women aged 15-24 years were assessed in two settings: family planning clinics (2004-2008) and prenatal clinics (2004-2009). Women seeking prenatal care are routinely screened and less influenced by health care seeking behaviors. After evaluating the linearity of year, a multivariate, correlated, longitudinal data analysis with a random intercept was conducted using the clinic as the unit of analysis and treating year as continuous. The odds ratio (OR) associated with a single year change (1.00; 95% confidence interval [CI]: 0.99, 1.00) suggested that positivity did not change from 2004-2008 in family planning clinics. In prenatal clinics, positivity decreased from 2004-2009 (OR: 0.93 per year, 95% CI: 0.92, 0.95). A survey assessing chlamydia screening practices in prenatal clinics was also conducted. Of the 166 clinics completing the survey, 163 (98.2%) had documented chlamydia screening criteria.

Because the purpose of chlamydia screening is to prevent adverse reproductive outcomes, data from 2 large health maintenance organizations, Group Health Cooperative (GH) and Kaiser Permanente Colorado (KPCO), were analyzed to develop a PID case-finding algorithm. A classification and regression tree analysis identified 2 main predictors beyond ICD-9 codes: PID-specific treatment and age 15-25 years. When using ICD-9 codes alone to identify PID cases, the positive predictive value (PPV) was 78.8% in GH and 79.1% in KPCO. Algorithm PPV improved to 86.9% and 84.5%, respectively.

Findings support previous analyses suggesting that chlamydia prevalence is not increasing and may be decreasing. Maximizing the utility of administrative data by examining the privately insured prenatal population and monitoring adverse outcomes can improve surveillance around this important public health issue.

Chlamydia Surveillance in the United States: New Analytic Approaches and Alternate Considerations for Monitoring Trends in Disease Burden

By

Catherine Lindsey Satterwhite
B.A., New York University, 1999
M.P.H., Emory University, 2000
M.S.P.H., Emory University, 2004

Advisors: Penelope P. Howards, Ph.D., M.S.
David Kleinbaum, Ph.D., M.A.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Epidemiology.
2011

Acknowledgement

This work would not have been possible without the support of many people.

First, thanks to my husband, Lewis. You gave me the gift of time, not an easy feat when two small children demand so much of it. Even in the midst of your fellowship, you encouraged me to continue my pursuit of learning. You set an excellent example to our children in demonstrating how a partnership works.

To Ella, my daughter, thank you for your patience with me. Thank you for being full of joy and endless love. Your hugs at the end of the day over the past 3 years lit up my world and made each day better than the next. Still do! You amaze me. Your inquisitive nature, your thirst for learning, even your endless questions… You gave me the gift of the future, of knowing that the work I do today holds promise for tomorrow. I am so excited to watch you grow!

To Lawton, my son, your whole life has literally been spent with me in pursuit of this degree! I found out I was pregnant with you three weeks after I began coursework. You were born 48 hours after I finished my last final and were only five weeks old when I took my qualifying exams. Now, you are almost 2 years old! You gave me the gift of perspective and grounded me, reminding me that my family, my children, are infinitely more important than a homework assignment, a final, or even, yes, a dissertation.

To the rest of my family, thank you for your support. Dad, Laura, and Steven, you guys are awesome; without you, I would not be where I am today. Mary Lynn, H.D., Emily, Phil, Maryn, Lauren, Shamus, and Whit, I am honored to be part of your family.

I would also like to acknowledge the Centers for Disease Control and Prevention (CDC) for their support in obtaining this training and, ultimately, this degree. In particular, thank you to Stuart Berman, who initially suggested I pursue my PhD and offered his support, and to Hillard Weinstock, my supervisor who picked up the slack while I was taking classes and was my most vocal supporter as I developed this dissertation.

Thank you to Penny Howards for being an exemplary advisor, accommodating my sometimes demanding deadlines, and providing valuable feedback on all aspects of my dissertation. Thank you also to David Kleinbaum and Jim Buehler for agreeing to serve on my committee despite very busy schedules and acting as important, contributing members.

And, finally, thank you, Mom. I miss you more than words can say, and I know you are proud of me not only today, but always. You are with me in everything I do. Your strength and guidance helped shape who I am. You know, I will never forget one of the most fundamental things you taught me: that learning should be fun! Every time I take a test, I hear your voice reminding me to treat it like a game, a challenge. So far, so good!

Table of Contents

Tables and Figures

Overview and Rationale

Chlamydia trachomatis infection is the most commonly reported nationally-notifiable disease. Over 1.2 million cases were reported in 2009. Infections may lead to serious adverse outcomes among women, including pelvic inflammatory disease (PID), ectopic pregnancy, tubal-factor infertility, and chronic pelvic pain. Among men, chlamydia may result in urethritis, prostatitis, and epididymitis. The frequency of occurrence, asymptomatic nature of infection, and the possibility of adverse outcomes prompted the development of widespread screening recommendations in 1993. Currently, the Centers for Disease Control and Prevention (CDC) recommends that all sexually-active women under the age of 26 years be screened annually for chlamydia. Given screening recommendations and the high number of reported chlamydia cases, tracking trends in infections is a national priority.

Four nationally-available, recurring data sources are utilized to assess chlamydia disease burden, including temporal trends: morbidity data (national case reports), the National Health and Nutrition Examination Survey (NHANES), data collected through the Infertility Prevention Project (IPP), and data from chlamydia screening performed through the National Job Training Program (NJTP). These data sources cover a variety of populations and vary in representativeness and completeness. Substantial limitations exist when attempting to ascertain chlamydia disease burden and trends, including improved case reporting as data systems are enhanced, changes in test technology to utilize increasingly sensitive tests, and more widespread screening. These limitations result in difficulties comparing national data sources and in determining underlying

temporal trends. In addition, some data are highly dependent upon population healthcare-seeking behaviors and programmatic influences, such as changes in available funding and reporting mechanisms. Morbidity data indicate a steady increase in reported chlamydia cases over the last 10 years; however, other data suggest that the actual burden of chlamydia has remained steady or decreased.

To supplement chlamydia surveillance, PID is often assessed; PID is the most immediate adverse outcome of chlamydial infection. However, assessment of PID has its own challenges, including the lack of an easily applied case definition. PID diagnoses are based on clinical impression. To identify possible PID cases from administrative datasets, ICD-9 codes are commonly used. However, cases identified on the basis of ICD-9 codes may not be clinically diagnosed PID; thus, case identification using ICD-9 codes is problematic.

The purpose of this dissertation was to explore new analytic approaches to enhance the utility of U.S. chlamydia surveillance data. In addition, a PID case-finding algorithm was developed using additional data elements available in administrative data (beyond ICD-9 codes). The new algorithm was then evaluated to assess whether PID case-finding using the algorithm improved upon methods using ICD-9 codes alone.

Study Questions and Methods

1. *How can family planning data from the Infertility Prevention Project, a national chlamydia screening program, be better used to monitor trends in chlamydial infections?*

   Previous analyses of IPP data have focused on using individual test-based data (positive/negative results) to ascertain positivity trends. Analyzing the proportion of positive tests at the clinic level may improve on test-based analyses by indirectly adjusting for unmeasured factors, such as clinic-based screening practices and general population characteristics. Using the clinic as the unit of analysis, a correlated, longitudinal data analysis with a random intercept was conducted using data from 2004-2008. Clinics reporting at least 3 years of data (at least 25 tests conducted among women aged 15 to 24 years of age) to IPP were included. Positivity was calculated among women aged 15 to 24 who were tested in eligible family planning clinics.

2. *Is the prenatal population an alternate population for monitoring chlamydia trends?*

   With the exception of estimates generated from NHANES, chlamydia prevalence surveillance is generally limited to higher risk populations, such as women attending family planning clinics (a population that may be seeking healthcare because of high risk behaviors) and socioeconomically disadvantaged young men and women entering NJTP. NHANES cannot be easily or economically reproduced at state and local levels, nor is it a reliable long-term national source for chlamydia surveillance; as prevalence becomes low and less stable and declines over time, standard errors

increase, limiting point estimate precision. The biases associated with service-based clinic prevalence are likely to be minimized when looking at the prenatal population compared to other populations captured in national data, because this population is likely to be a more stable population that is less impacted by general healthcare-seeking behaviors. Chlamydia trends from 2004-2009 among participating IPP prenatal clinics were examined utilizing a correlated, longitudinal data analysis with a random intercept. The clinic was the unit of analysis (cluster). Results will be compared to other national sources, particularly NHANES, to evaluate the role of prenatal clinics in terms of future surveillance utility. In addition, a survey was conducted among a random sample of prenatal clinics reporting data through IPP in 2008 to assess chlamydia screening practices and policies. Data were also collected to estimate screening coverage among sampled prenatal clinics.

3. *How can administrative data be better used to detect PID cases, an important adverse outcome of untreated chlamydial infections?*

While monitoring PID is a critical component to evaluating the impact of national chlamydia screening efforts, PID surveillance is challenging. PID is diagnosed by a clinician based on non-specific signs and symptoms, such as lower abdominal pain and cervical motion tenderness. In the absence of a laboratory-based case definition, ICD-9 codes are generally relied upon when identifying PID cases from automated data for surveillance purposes. Several ICD-9 codes may be used to indicate a diagnosis of PID, but use varies. The most commonly referenced ICD-9 code, 614.9 (female pelvic inflammatory disease not otherwise specified) has a positive predictive

value (PPV) of only 18.1% for the CDC PID surveillance case definition. A PID case-finding algorithm that moves beyond exclusive reliance on ICD-9 codes may represent an improvement in the methodology used to identify PID cases and allow for more accuracy in burden and trend ascertainment. An automated PID case-finding algorithm was developed and evaulated using data from Group Health Cooperative (GH), located in Seattle, WA) and Kaiser Permanente Colorado (KPCO). To develop the algorithm, a classification and regression tree (CART) analysis was performed to evaluate the predictive value of ICD-9 codes and other available administrative data (procedure codes, treatment information, demographics) in identifying confirmed PID cases (by medical record review). To evaluate the performance of the new algorithm, the positive predictive value (PPV) was calculated as the proportion of algorithm-classified PID cases that were confirmed to be clinically-diagnosed PID by medical record review. The algorithm PPV was then compared to the PPV associated with using ICD-9 codes alone.

Chapter 1

Background

Chlamydia trachomatis infection is the most commonly reported nationally-notifiable disease (1). Over 1.2 million cases were reported in 2009 (2).

Chlamydia trachomatis is an obligate intracellular pathogen. First visualized from orangutan conjunctival specimens in the early 1900s, *C. trachomatis* was initially thought to be a protozoan organism, due to characteristic cellular inclusions (3). Numerous *C. trachomatis* serovars have been identified and may infect several anatomic sites in men and women. Serovars A, B, Ba, and C are associated with trachoma, an infection of the eye that may lead to blinding. Serovars B and D-K are associated with genital transmission, infecting columnar epithelial cells located in the endocervix, urethra, upper reproductive tract (fallopian tubes, endometrium, uterus, epididymis), rectum, pharynx, vagina (prepubertal), eye (neonatal ophthalmia or adult conjunctivitis), lung (infant pneumonia), and joints (reactive arthritis). This group of serovars causes chlamydial infections commonly classified as sexually transmitted infections (STIs). Lymphogranuloma venereum (serovars L1, L2, L3), a sexually transmitted disease with slightly different characteristics, is also caused by *C. trachomatis* infection. The focus of this dissertation is genital chlamydia (serovars B and D-K).

C. trachomatis was not isolated from the genital tract until 1959 (3). Since that time, chlamydia has been well-described. Literature based on cross-sectional studies suggests that chlamydial infection is largely asymptomatic among both men and women (4). In the National Longitudinal Survey of Adolescent Health (AdHealth) Wave III cycle, more than 95% of the young men and women surveyed and found to be positive for chlamydia

reported no symptoms (5). Similarly, in other screened populations not explicitly seeking healthcare, 86-98% had asymptomatic chlamydial infections (6, 7, 8). When present, chlamydia symptoms may include discharge, dysuria, mucopurulent discharge, and cervical friability (9). Chlamydia is effectively treated with a single dose of azithromycin (1 g orally) or 7 days of doxycycline (100 mg twice a day) (10, 11).

Untreated chlamydial infections, whether symptomatic or not, may lead to pelvic inflammatory disease (PID) (9). PID is associated with further reproductive adverse outcomes, including ectopic pregnancy, tubal factor infertility, and chronic pelvic pain (12).

Chlamydia Test Technology

Chlamydia test technology has substantially changed over time. Depending upon the test technology utilized, chlamydia testing can be performed on a number of specimen types including swabs collected from the cervix, vagina (self-collected or clinician-collected), urethra (primarily men), rectum, or throat. The current optimal test technology utilized to detect genital *C. trachomatis* infections are nucleic acid amplification tests (NAATs) (Table 1.1) (11). NAATs may also be used on urine specimens, which allows for non-invasive testing.

No true gold standard test for chlamydia exists; NAAT performance is superior to the traditional gold standard, C. *trachomatis* culture (13), with sensitivities of greater than 90% and specificity levels of approximately 99% (14, 15, 16). First introduced in the late

1990s, use of NAAT technology was initially cost-prohibitive (17). However, as costs decreased and additional studies demonstrated clear advancements over prior generation tests, usage has increased. In 2000, 24.5% of all chlamydia tests conducted in surveyed public health laboratories were NAATs; by 2004, the proportion had increased to 64.4% (18, 19). In 2007, the most recent survey of public health laboratories showed that 81.6% of all tests performed were NAATs (20). A survey conducted among U.S. Army laboratories in 2007 showed similarly high NAAT usage rates: 78.6% of reported chlamydia tests were conducted using NAAT technology (21). Before the introduction of NAATs, the advent of non-culture tests for *C. trachomatis* detection, such as direct fluorescent antibody (DFA) tests, enzyme immuno-assays (EIA), and DNA probes made it possible to test large numbers of women and men with high accuracy and relative rapidity when compared to culture (22, 23, 24). Chlamydia culture is difficult to perform and was not suitable for use in screening large populations (24).

Table 1.1 FDA-approved chlamydia tests, 2008

Technology (abbreviation)	Test Type*	Manufacturer	Brand Name
Transcription mediated amplification (TMA)	NAAT	Gen-Probe, Inc. (San Diego, CA)	Aptima Combo 2
Strand displacement assay (SDA)	NAAT	Becton Dickinson (Sparks, MD)	BDProbeTec
Polymerase chain reaction (PCR)	NAAT	Roche Molecular Diagnostics (Indianapolis, IN)	Amplicor
Nucleic acid hybridization ("DNA-probe")	non-NAAT	Gen-Probe, Inc. (San Diego, CA)	PACE 2
Signal amplification	non-NAAT	Digene (Valencia, CA)	Hybrid Capture
Enzyme immunoassay (EIA)	non-NAAT	Various	Various
Direct fluorescent assay (DFA)	non-NAAT	Various	Various
Culture	non-NAAT	NA	NA

*NAAT: Nucleic acid amplification test

While improvements in test technology have facilitated diagnosis, significant challenges persist in determining and interpreting the impact of these changing technologies on chlamydia burden and trends. The increased sensitivity of available chlamydia tests has resulted in better detection of prevalent infections; older test technologies with limited sensitivities likely missed infections (25). If test type is not known, increases in chlamydia burden due to better test technology may appear to represent increases in actual disease burden, and not just enhanced case detection. Studies have demonstrated how changes in chlamydia test technology can effect estimates of chlamydia prevalence. Dicker et al. found that chlamydia positivity in Philadelphia, Pennsylvania, family planning clinics increased by 46% when NAATs replaced DNA probes (4.1% to 6.0%) (25). Likewise, when family planning clinics in Oregon and Washington began to utilize EIA technology instead of DFA, positivity increased by 21% (3.3% to 4.0%). An analysis of data from the National Job Training Program (NJTP) revealed a one-year increase (2005 to 2006) in prevalence from 9.1% to 13.9% (53% increase) associated with a dramatic shift in test technology: from 2005 to 2006, NAAT usage went from 21% to 88% of all tests (8). When chlamydia trends were assessed, prevalence in the NJTP increased from 2003-2007, but after adjustment for test technology and other confounding factors, a statistically significant decrease was reported. Thus, test technology usage is a critical consideration in chlamydia surveillance.

Chlamydia Screening Recommendations

The prevalence of asymptomatic infection combined with the possibility of adverse outcomes prompted the development of widespread screening recommendations in 1993 (24). Results from a randomized controlled trial demonstrated that chlamydia screening reduced the incidence of PID (26). Initial guidelines suggested that sexually-active women under the age of 20 years, women with mucopurulent cervicitis (MPC), and women aged 20 to 24 years who either used barrier contraception inconsistently or who had greater than one or a new sex partner(s) be screened for chlamydia. Screening for women older than 24 years was also recommended if multiple risk factors were present. These screening guidelines were subsequently broadened. Currently, the Centers for Disease Control and Prevention (CDC) recommends that all sexually-active women under the age of 26 years be screened annually for chlamydia (11). In addition, since 2001, the United States Preventive Services Task Force (USPSTF) has also recommended screening of young, sexually-active women (27, 28). In 2007, USPSTF updated their chlamydia screening recommendations to change the upper age bound from under 26 years to under 25 years of age, a change from the CDC-recommended upper age range (29). The age change was made only to be consistent with nationally-reported surveillance data age groupings (30). Both CDC and USPSTF recommend consideration of chlamydia screening for older women with risk factors. Possible risk factors identified include: a history of chlamydia or another STD, new or multiple sex partners, inconsistent condom use, and exchanging sex for money or drugs. In sum, both CDC and USPSTF, as well as most major medical organizations, recommend that sexually-active women under the age of 25 years be screened annually for chlamydia (31). CDC also

suggests rescreening of women who test positive for chlamydia 3 months after the initial infection (11).

Chlamydia screening is also recommended for pregnant women. CDC recommends that all pregnant women, regardless of age, be screened for chlamydia at their first prenatal visit (11). Additional screening, in the third trimester, is recommended for all women under the age of 25 years and older women at increased risk. Likewise, USPSTF recommends screening of pregnant women (29). However, USPSTF only recommends screening of pregnant women aged 24 years or younger or older women at increased risk. Pregnant women over the age of 24 years, without any risk factors, are not recommended for routine chlamydia screening (29). USPSTF recommends first trimester screening, followed by subsequent third trimester screening if indicated by sustained or new risk behaviors.

Chlamydia screening has been recognized by the National Commission on Prevention Priorities (NCPP) as one of the most beneficial and cost-effective preventive services among those recommended by USPSTF or the Advisory Committee on Immunization Practices, based on a standardized analysis evaluating and comparing these services (32). NCPP also identified chlamydia screening as one of the most under-utilized preventive services. The National Committee for Quality Assurance added chlamydia screening as a measure in the Healthcare Effectiveness Data and Information Set (HEDIS) in 1999 (33). Both commercial healthcare plans and Medicaid managed care plans report on the measure. HEDIS is used by greater than 90% of U.S. health plans as a quality indicator

to assess plan performance (34). Although chlamydia screening coverage, as measured using HEDIS, is still low, it has increased steadily over time. From 2000 to 2007, screening coverage among young women aged 16-25 years (16-26 years for 2000 to 2002) in commercial and Medicaid populations increased from 25.3% to 41.6% (33). Coverage was consistently higher among Medicaid populations than commercial populations. In 2008, women aged 16-20 years in commercial plans had a chlamydia screening coverage rate of 40.1%; women aged 21-24 years had a rate of 43.5% (35). Among Medicaid populations in 2008, women aged 16-20 years had a chlamydia screening coverage rate of 52.7%; women aged 21-24 years had a rate of 59.4%.

No national chlamydia screening recommendations exist for men. In 2006, CDC held a consultation to address male chlamydia screening (36). The purpose of the consultation was to provide guidance for programs currently already screening men, or planning to do so in the future, not whether or not male screening should be recommended (37). Experts emphasized that the primary screening focus should be young women, per existing chlamydia screening recommendations. Consultation attendees identified some priority venues where male chlamydia screening might be appropriate, such as adult corrections facilities and STD clinics, if programs chose to screen males (38). USPSTF also reviewed the literature on male chlamydia screening, but found insufficient evidence to make a recommendation (29). While the burden of disease in men is not trivial, for male screening to be cost-effective, it must result in a decrease in adverse outcomes among women due to chlamydial infection (39, 40). These adverse outcomes among women, such as PID and tubal-factor infertility, are associated with the majority of chlamydia-

related costs (41). One recent study by Peterman et al. found that intensive venue-based chlamydia screening of men in Philadelphia adult corrections facilities did not have an impact on the community burden of disease among women, as measured by chlamydia positivity in family planning clinics (42).

Epidemiology of Chlamydial Infections in the United States

An estimated 2.8 million chlamydia cases occur annually in the U.S (43). However, many of these cases are not detected and diagnosed. Despite this, *Chlamydia trachomatis* infections are still the most commonly reported nationally-notifiable disease (1). Chlamydia was made a nationally notifiable disease in 1995 and by 2000 was reported by all states. Over 1.2 million cases were reported in 2009; three times more chlamydia cases were reported than gonorrhea cases, the next most frequently reported notifiable disease (2). Rates of reported chlamydia are highest among young women, reflective of screening recommendations. Among women aged 14 to 19 years, in 2009, there were 3,329.3 reported cases per 100,000 population; among women aged 20 to 24, the rate was 3,273.9. Reported case rates among men are substantially lower (in 2009, 1,120.6 cases per 100,000 men aged 20 to 24 years). Lower reported rates in men are likely due to limited testing and detection of chlamydial infections in this population, when compared to broad screening among women. Racial disparities exist in reported chlamydia rates; in 2009, rates among black men and women were eight times greater than among white men and women (2).

For the past 20 years, reported overall chlamydia case rates (all ages, both sexes) have steadily increased, from 160.2 cases per 100,000 population in 1990 to 409.2 cases per 100,000 in 2009 (2). Given the estimated burden of infection, more widespread screening, data system enhancements, and use of increasingly sensitive tests, increases in the number of chlamydia cases reported are expected. Case report data do not represent trends in disease burden, only trends in case detection (44). The majority of chlamydia cases are reported from non-STD clinic settings, such as family planning clinics and private providers (2). Among women, only 10% of cases are reported from STD clinics, suggesting that many reported cases may be due to screening activities; women seeking care at STD clinics are likely symptomatic or a partner of an infected man. Therefore, chlamydia burden may not be well assessed by reported case rates and may be better assessed by population-based prevalence data. Chlamydia case report data currently are highly reflective of programmatic activities, such as increased chlamydia screening among young women; therefore, as screening coverage increases, case reports would be expected to also increase, at least until actual reported cases (1.2 million) more closely mirror estimated cases (2.8 million) (43, 45). In contrast, prevalence data reflect actual disease burden among a defined population uniformly tested for infection.

An analysis of chlamydia data from the National Health and Nutrition Examination Survey (NHANES), a continuous, population-based survey conducted annually, showed that overall chlamydia prevalence among U.S. men and women aged 14 to 39 years was 2.0% (95% confidence interval [CI], 1.6% to 2.5%) (46). NHANES consists of annual data on approximately 5,000 U.S., non-institutionalized men and women, selected using

complex sampling methodology. Stratified by age group, chlamydia prevalence was highest among young men and women aged 20 to 29 years (3.2%). Non-Hispanic blacks bore a disproportionate burden of disease with a prevalence of 5.3%, compared to a prevalence of 1.5% among white men and women. The burden of disease was similar among men (2.0%) and women (2.5%). When NHANES results were examined over time, chlamydia prevalence decreased from 1999-2006 in the overall population of men and women aged 14 to 39 years of age (47). NHANES is an important source of chlamydia prevalence data, but the stability of point estimates may be questionable as prevalence decreases, thus limiting the ability to detect changes in prevalence over time as standard errors increase.

The Infertility Prevention Program (IPP) is a national program, administered primarily through family planning clinics, targeting young women for chlamydia screening. Test-based data reported through IPP are used to calculate chlamydia positivity (positive tests/total tests). Positivity approximates prevalence closely, based on a comparison of the two measures in family planning clinics (48). In two regions of the country (Regions III and VIII), the mean absolute difference between positivity and prevalence was 0.1% (4.8% vs. 4.7% and 3.4% vs. 3.3%, respectively; in Region X, mean positivity and prevalence were the same (5.5%). IPP data suggest that, among women attending family planning clinics aged 15 to 24 years, the median state-specific positivity has steadily increased over time, from 1997 to 2009 (2). State IPP positivity varies substantially. Positivity is highest in the Southeast, consistent with case rates. The highest positivity in 2009, 15.5%, was reported in the U.S. Virgin Islands; positivity was lowest in Vermont

(3.5%). NAAT usage has increased in IPP, so crude positivity is impacted by improvements in test technology.

Data from the National Job Training Program (NJTP), a program serving young, socio-economically disadvantaged men and women aged 16-24 years, are not subject to some of the primary limitations present in case report data and IPP data. Nearly all NJTP participants are screened for chlamydia at program entrance, and the test technology utilized is known and consistent (8); the population is defined and tested in a standardized way. In addition, the NJTP population is screened opportunistically, at a relatively low cost, as compared to the costly NHANES. Chlamydia prevalence is high; in 2009, the median state-specific prevalence among women was 11.3% (2). Among men, prevalence was 7.0%. While NJTP data represent a high-risk population not broadly generalizable, this relatively consistent population (stable demographics and social characteristics) provides important insight into chlamydia burden. Significant decreases in chlamydia prevalence were detected in the young, at-risk men and women entering the NJTP from 2003 to 2007 (8). Prior studies show that chlamydia prevalence has been declining in this population since 1990 (49, 50).

Reinfection

Chlamydial reinfection is common. Approximately 14% of women with an initial chlamydial infection will be reinfected within the year following diagnosis (range: 0% to 32%) (51). Modeling results based upon a meta-analysis of studies from active cohorts suggested that previously-infected women had a peak reinfection rate of 20.9% by 13

months after the initial infection (51). These data strongly support recommendations to rescreen women testing positive for chlamydia. Men also have high rates of repeat chlamydial infection, but data are limited (52).

Pregnant Women

Among pregnant women, untreated chlamydia can lead to adverse birth outcomes and infection of the infant (neonatal ophthalmia) (3, 53, 54, 55). From NHANES, 1999 to 2002, among those women with a positive pregnancy test at the time the survey was administered, 2.0% also had chlamydia (46). Positivity among women aged 15 to 24 years who attended publicly-funded prenatal clinics and were screened for chlamydia is substantially higher than general population estimates. In 2009, the median state-specific positivity rate was 7.7% (range: 3.6% to 20.4%) (2). However, this estimate includes women who may have sought a pregnancy test only, who were not necessarily already aware they were pregnant. This population may be different than women in standard prenatal care (already aware of and committed to their pregnancy). Women seeking a pregnancy test only may perceive themselves to be at risk for pregnancy; therefore, an STD risk is also likely present. In a recent study, Geisler et al. reported a chlamydia prevalence of 12% among women aged 16-45 years who were seeking a pregnancy test only (South Carolina family planning clinics) (56). Prevalence did not vary based upon whether or not the women tested positive for pregnancy. Similar risk has been reported among women seeking emergency contraception and induced abortions, groups where pregnancy was not desired and the risk behaviors (e.g., lack of barrier contraception) leading to that pregnancy also favor STD acquisition (57, 58).

Chlamydia Surveillance, United States

Currently, four nationally-available, recurring data sources are used to monitor chlamydia disease burden, including temporal trends. The first, and most commonly cited, source is case report data, including chlamydia case rates. Such data are collected annually from all 50 states, the District of Columbia, Guam, Puerto Rico, and the Virgin Islands (2). Chlamydia was made a nationally notifiable disease in 1995 and was reported by all states by 2000. A second source of chlamydia data is NHANES, a nationally-representative probability sample of U.S. civilian population (non-institutionalized), combining interviews and physical examinations. NHANES, while extremely valuable due to its generalizability, is costly to conduct; to reproduce NHANES at a community level is cost-prohibitive. Therefore, general population chlamydia prevalence estimates, like those available from NHANES, are rarely available at the state and local levels. Moreover, the national sample size limits sub-group analyses; funding does not permit larger samples. An additional important source of chlamydia data is from IPP. IPP data are primarily from young women attending family planning clinics, a health-care seeking population, and are subject to programmatic decisions that may reflect fuding decreases in an environment of limited resources. Changes in population, screening policies, test technology, and reporting practices over time make chlamydia trends difficult to interpret. While IPP data do not allow for broad generalizability, the population captured encompasses a variety of risk levels. Finally, data from NJTP provide chlamydia prevalence on a consistent population universally screened using a single test technology at program entrance; however, this high-risk population is not broadly representative.

Approximately 120 training centers nation-wide screen about 30,000 men and 15,000 women entering NJTP annually(8). Universal screening of women began in 1990; male screening began in 2003.

Each of these data sources is a component of the U.S. national surveillance portfolio utilized to monitor chlamydia disease burden and trends. Using these sources, CDC attempts to estimate disease burden and determine if prevention programs are working.

Importance and Challenges of Monitoring PID

PID is the most immediate adverse outcome of chlamydial infection. Untreated chlamydia leads to PID in approximately 10-15% of cases (59, 60). PID may lead to further adverse outcomes, including tubal-factor infertility, ectopic pregnancy, and chronic pelvic pain (12). The specific contribution of chlamydia to each of these adverse outcomes is unknown (61). However, approximately 10-20% of all infertility is estimated to be due to tubal and pelvic pathology (59). Given that the objective of screening is to reduce adverse outcomes, monitoring PID is important in assessing chlamydia prevention and control efforts, despite the fact that chlamydia is only one of many factors contributing to PID occurrence.

In the absence of a laboratory-based case definition, PID diagnosis is made based upon clinical signs and symptoms (11). This diagnosis lacks specificity and is not easily standardized. The "gold standard" for diagnosing tubal involvement is laparoscopy, which can detect both symptomatic (clinically diagnosed PID) and asymptomatic

infections (62). However, laparoscopy is an invasive procedure not commonly used in clinical practice. Several ICD-9 codes may be used to indicate a diagnosis of PID, but use is inconsistent. Nevertheless, ICD-9 codes have been used to attempt to identify PID cases for analysis and surveillance purposes. However, the most commonly referenced ICD-9 code, 614.9 (female pelvic inflammatory disease not otherwise specified) has a positive predictive value (PPV) of only 18.1% for the CDC PID case definition (63). When coupled with a positive chlamydia test, the PPV increases to 56%.

Trends in PID diagnoses are difficult to ascertain. At the national level, data for monitoring PID trends are routinely obtained from complex sample surveys, a health services survey, and a survey of hospital admissions. While each of these data sources has limitations, all suggest a downward trend in PID diagnoses (2). From 1985 to 2001, PID diagnosed in hospital and ambulatory settings decreased (64). While national surveys provide important information, a better data source may be administrative claims data. A recent analysis of a national insurance claims database also revealed decreases (65). Overall, challenges in monitoring PID, both in case ascertainment and in determining the best data source, persist.

Contribution of Dissertation

The purpose of this dissertation is to explore new analytic approaches to enhance the utility of U.S. chlamydia surveillance portfolio data. In order to address some limitations of analyzing IPP individual test-based data, such as a lack of data on clinic practices, data will be analyzed at the clinic level. Analyzing the proportion of positive tests at this level

may improve on using test-based analyses by indirectly adjusting for unmeasured factors, such as clinic-based screening practices and general population characteristics. To explore enhanced usage of existing national IPP data, data from women attending prenatal clinics will be assessed and analyzed. Trends in chlamydia positivity will be evaluated in this population, using a clinic-based analysis approach. The biases associated with service-based clinic prevalence are likely to be minimized when looking at the prenatal population compared to other populations captured in national data, because this population is likely to be a more stable population that is less impacted by general healthcare-seeking behaviors. Results of the prenatal analysis will be compared to other national sources, particularly NHANES, to evaluate the role of prenatal clinics in future surveillance.

PID is an important adverse outcome of chlamydial infection. Therefore, in addition to the chlamydia trends analyses, a new PID case-finding algorithm will be developed and evaluated, as compared to the current case-finding approach based on ICD-9 codes alone. The new algorithm incorporates other key data elements, prescribed treatment and patient age.

Chapter 2

How can family planning data from the Infertility Prevention Project, a national chlamydia screening program, be better used to monitor trends in chlamydial infections?

Chlamydia Positivity Trends Among Women Attending Family Planning Clinics: United States, 2004-2008

Catherine Lindsey Satterwhite[1], MSPH, MPH; LaZetta Grier[1], BS; Rachel Patzer[2], MPH; Hillard Weinstock[1], MD, MPH; Penelope Howards[2], PhD, MS; David Kleinbaum[2], PhD, MA

[1]Division of STD Prevention, CDC, Atlanta, Georgia
[2]Department of Epidemiology, Emory University, Atlanta, Georgia

ABSTRACT

Background: Annual chlamydia screening is recommended for all sexually active women aged <25 years. Substantial limitations exist in ascertaining chlamydia trends. Reported case rates have increased likely due to increased screening and improved test technology. Other data suggest that prevalence has decreased.

Methods: Data from the Infertility Prevention Project (IPP), a national chlamydia screening program, were used to assess trends in chlamydia positivity from 2004-2008 among women aged 15-24 years who were tested in family planning clinics reporting data to IPP. Using the clinic as the unit of analysis, a correlated, longitudinal data analysis with a random intercept was conducted among clinics reporting ≥3 years of data during the analysis timeframe. Sensitivity analyses were performed to address the impact of various clinic participation levels in addition to assessment of various correlation structures.

Results: Over 5 million chlamydia tests were reported to IPP family planning clinics from 2004-2008. The majority of tests were conducted among white women (clinic-specific mean: 63.2%, interquartile range [IQR]: 37.6%-91.5%); the clinic-specific mean percent of tests conducted among black women was 17.9% (IQR: 0.8%-25.7%). Overall chlamydia positivity from 2004-2008 was 7.0%. The odds ratio associated with a single year change (1.00; 95% confidence interval: 0.99, 1.00) suggested that chlamydia positivity did not change from 2004-2008, after controlling for clinic-specific population factors (age, race, test usage, and geography).

Conclusions: Findings support previous analyses suggesting that chlamydia prevalence is not increasing despite apparent increasing rates based on case reports.

BACKGROUND

Chlamydia trachomatis infection, a sexually transmitted disease associated with serious adverse outcomes among women, including pelvic inflammatory disease, ectopic pregnancy, tubal-factor infertility, and chronic pelvic pain, is the most commonly reported nationally notifiable disease in the United States (1, 9). Over 1.2 million cases were reported to the Centers for Disease Control and Prevention (CDC) from state and local health departments in 2008 (45). However, it is estimated that 2.8 million chlamydia cases occurred annually, suggesting that under-detection of cases is substantial (43). Chlamydia screening recommendations were first made in 1993 and expanded in 2001 (24, 27). Currently, the U.S. Preventive Services Task Force recommends that all sexually active women under the age of 25 years be screened annually for chlamydia (29). Given the national effort to prevent chlamydia and its complications, efforts to monitor trends in infections are critical.

Data sources available to assess chlamydia disease burden and temporal trends on a national scale are limited. Trends in U.S. chlamydia case report data, collected routinely from state and local health departments, show case rates increasing over the last two decades (45). However, increasing trends are likely due to better case detection through improvements in test technology and more widespread screening (44). Contrary to national case report data, a recent analysis of data from the National Health and Nutrition Examination Survey (NHANES) suggested that chlamydia prevalence from 1999 to 2006 was stable or decreasing among a nationally-representative sample of men and women aged 14-39 years (47). While a valuable data source, NHANES allows for only analyses

at the national level, not smaller geographic areas, and is costly to reproduce at the local level. Moreover, a limited sample size restricts the ability to track chlamydia trends over time in subgroups.

There are additional data sources that supplement case report data and NHANES and allow for assessment of national chlamydia trends. The National Job Training Program (NJTP) is a program serving young, socioeconomically disadvantaged men and women aged 16-24 years. All participants are screened for chlamydia at program entrance. In this high-risk population, chlamydia prevalence declined from 2003 to 2007 (8). In addition, data from young women screened for chlamydia are available through the Infertility Prevention Project (IPP). IPP is a national program targeting young, sexually active women for chlamydia and gonorrhea screening to prevent sequelae leading to infertility.

Previous analyses of data reported through IPP have focused on using the individual test-based data to ascertain positivity trends (45, 66). These analyses have generally suggested an increase in chlamydia positivity over time among young women attending family planning clinics. However, there are substantial limitations when using this approach, primarily, the lack of covariate availability and subsequent inability to adequately assess and control for confounding. Analyzing the proportion of positive tests at the clinic level, rather than the individual encounter level, may help minimize some of these limitations. In such an analysis, the clinic itself may be considered a proxy for possible confounders, such as screening practices, demographic and behavioral

population characteristics, and health care access. Treating the clinic as a confounder in individual test-based analyses is not possible due to the large number of participating clinics; regression models fail when adding clinic as a covariate (i.e., a large number of parameters are required). Thus, analyzing data at the clinic level may improve upon individual test-based analyses. Developing reproducible methodology to better utilize IPP family planning data is an important step in assessing chlamydia prevalence trends in the United States overall and at state and local levels.

The objective of this analysis was to describe trends in chlamydia positivity from 2004 to 2008 among women aged 15-24 years who were tested at publicly-funded family planning clinics participating in IPP. Given the limitations of using individual test result data reported through IPP, we applied an analytic approach utilizing clinic-level data to determine if a linear trend existed in positivity.

MATERIALS AND METHODS

Data Source and Study Population

Administered primarily through family planning clinics, data on chlamydia test results have been reported through IPP since 1997. Data are routinely collected from facilities participating in IPP and reported to CDC on a quarterly basis. IPP data are test-based (i.e., each observation is one test conducted, and an individual may have multiple tests), with no personal or unique identifiers included. Available variables in IPP data include demographics (age, sex, race/ethnicity, geography) and information specific to the test

performed (technology, specimen type, and test results). Variables describing the facility conducting the test (state, region) and the type of facility (family planning, prenatal, etc.) are also available from a facility reference file. Information found on the facility reference file is combined with the test-based data using a unique facility identifier variable.

The study population consisted of family planning clinics reporting data to IPP from 2004 to 2008. Sensitivity analyses were performed using various levels of clinic participation (clinics reporting ≥3, ≥4, or 5 years of data). Since findings were consistent, regardless of levels of participation, family planning clinics reporting ≥3 years of data to IPP from 2004 to 2008 were included to utilize the maximum available data. Family planning clinics were defined as either a stand-alone family planning clinic, or as a designated component of an integrated clinic, in which family planning visits may be distinguished. Only individual test results from women aged 15-24 years were included. In order to contribute data for a single calendar year, a clinic must have reported at least 25 total tests (positive and negative) conducted among women aged 15-24 years during that year. Exclusion of data from low-volume clinics was intended to reduce the influence of outlier chlamydia positivity values and increase analytic stability.

Analysis

A correlated, longitudinal analysis was conducted to assess trends over a 5-year time span. The unit of analysis was the individual clinic performing chlamydia tests (clinic-

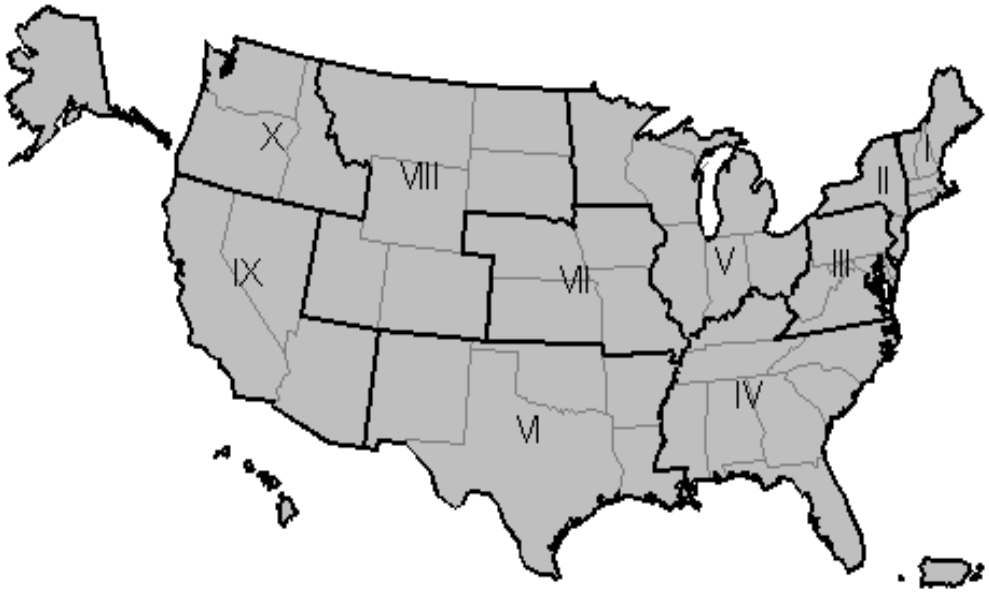
based analysis), as opposed to an analysis based only on individual test results that would assume independence of test results within and across clinics.

The outcome of interest was chlamydia positivity within a clinic, defined as the proportion of positive tests out of all positive and negative tests reported by that clinic (events/trials model). This approach allowed for the incorporation of clinic size (denominator) into the modeled outcome. Chlamydia positivity has been found to be a reasonable approximation of prevalence (48). To assess for trends, the primary independent exposure of interest was defined as calendar year, from 2004 to 2008. In order to determine if year should be treated as ordinal or categorical, a linearity assessment was performed using two methods. First, estimated logit plots were created, which suggested a linear pattern. Linearity was then assessed using a generalized estimating equation model, to account for the correlated data. Similar to the logit plot results, modeling showed a general linear relationship between estimates associated with dummy variables representing year as categorical; therefore, year was treated as ordinal for the purposes of this analysis.

Four covariates were considered. The proportion of tests conducted using NAAT technology performed by a given clinic, was likely important because these tests demonstrate substantially increased sensitivity over prior-generation tests (16). Two demographic covariates were also considered: the proportion of tests occurring among young women aged 15 to 19 years, the group generally considered to be at the highest risk for chlamydial infection, and the proportion of tests occurring among black women,

a group disproportionally affected by chlamydia (45). The fourth covariate identified the region (Figure 2.1) where the clinic was geographically located. All two-way interaction terms (product of year with each of the 4 covariates) were assessed, in order to detect differences in trends over time in subgroups.

Figure 2.1. U.S. Department of Health and Human Services regions



Using SAS (version 9.2, GLIMMIX procedure), a random effects regression model with an events/trials approach for the outcome (chlamydia positivity, logit [probability [total positive chlamydia tests/total positive and negative chlamydia tests in each clinic for each analysis year]) was applied to assess linear trends from 2004 to 2008. The exposure of interest (year) and other explanatory variables were fixed, but the intercept was treated as random, given the substantial variation in the underlying chlamydia positivity across clinics (e.g, in 2004, median: 5.8%, range: 0.0% to 29.5%). A Wald test was conducted to test for the significance of the random intercept. A random slope was also considered.

After reviewing respective logit plots and confirming appropriateness, the three proportional covariates (age, race, test technology) were treated as continuous in the model.

Prior to assessing correlation structures, all covariates and interaction terms were entered into an initial model to assess for covariate collinearity. Using a logistic model that accounted for correlated data but assumed a fixed effect rather than a random effect for the intercept (SAS version 9.2, GENMOD procedure), condition indices (>30) were evaluated first, followed by variance decomposition proportions (>0.5), both produced using the inverse of the information matrix (collingenmodv9c.sas macro, Emory University, Atlanta, GA, modified).

A variety of possible correlation structures were considered. The default G-matrix for this analysis was defined by σ_0^2, a scalar parameter given to the variance component associated with the random intercept. There was no obvious R-matrix correlation structure that described the correlations between two different outcomes (i.e., observed proportions) within the same clinic that could be identified on the basis of clinical or biologic rationale. However, a $\sigma_1^2 I_5$ R correlation matrix may be appropriate if clinic populations are similar over time. This matrix, with an empirical option and a random intercept, produces an overall covariance structure that is approximately exchangeable (compound symmetric). Unstructured, auto-regressive, and toeplitz structures were also tested. Results were compared by evaluating coefficient estimates and standard errors to identify if results were consistent or discrepant when using different correlation

structures. In addition, results of type III tests of fixed effects (F-statistics and associated significance tests) were obtained and compared.

To assess for confounding and precision, all possible model combinations involving subsets of the four covariates were considered by comparing the odds ratio (OR) point estimates associated with year to the OR point estimate from the full analytic model containing all four covariates. Sixteen possible models, including the full model, were considered. If the OR estimate was within 10% of the OR estimate for the full model, the model was determined to be eligible for further consideration by assessing precision, based upon the 95% confidence intervals (CI) around the OR point estimate. The final model was selected based upon precision and appropriateness of including covariates based on demonstrated prior associations with the outcome (chlamydia positivity).

All statistical analyses were performed using SAS version 9.2 (SAS Institute Inc., Cary, NC).

RESULTS

From 2004 to 2008, more than 5 million tests conducted among young women aged 15-24 years were reported to IPP. Approximately 1 million tests were reported annually. Of the 4,253 clinics reporting data to IPP for at least 1 year during the analysis time frame, 58.2% (2,475) reported ≥3 years of data, representing 88% of all tests reported. This group of clinics was located in 49 states. About 30% of clinics were located in Region IV (Figure 2.1).

The clinic-specific mean percent of tests conducted among black women aged 15-24 years was 17.9% (Table 2.1). Slightly less than half of the population tested was comprised of young women aged 15-19 years. Mean usage of NAAT technology increased over time, from 69.7% in 2004 to 87.6% in 2008; however, median usage was consistently 100.0%. When stratified by year, race/ethnicity and age did not vary substantially (data not shown).

Table 2.1. Clinic-specific population characteristics of family planning clinics reporting ≥3 years of data to the Infertility Prevention Project from chlamydia tests conducted among women aged 15-24 years, 2004-2008

	Clinic Percent or Count*		
	Mean	**Median**	**(Interquartile Range)**
Race/Ethnicity+			
Hispanic	11.7	2.6	(0.4-11.8)
Non-Hispanic Black	17.9	5.4	(0.8-25.7)
Non-Hispanic White	63.2	73.2	(37.6-91.5)
Other	3.3	1.0	(0-2.8)
Unknown/Missing	3.9	0.2	(0-1.9)
Age Group+			
15-19 Years	45.4	45.3	(38.2-52.1)
20-24 Years	54.5	54.7	(47.9-61.8)
NAAT+ Usage			
2004	69.7	100.0	(2.9-100.0)
2005	77.5	100.0	(99.5-100.0)
2006	83.0	100.0	(100.0-100.0)
2007	86.1	100.0	(97.9-100.0)
2008	87.6	100.0	(100.0-100.0)
Number of Tests	396	250	(123-512)

NAAT=nucleic acid amplification test
*Mean and median of all clinic values
+Proportions

The overall mean clinic-specific chlamydia positivity from 2004 to 2008 was 7.0%. In 2004, positivity was 6.9%, remaining fairly stable through 2008, when positivity was 7.2%. When compared to clinics reporting only 1 or 2 years of data to IPP not included

in the analysis, clinics reporting ≥3 years of data had a slightly lower mean proportion of women tested who were black (17.9% vs. 20.4%); the distribution of age and NAAT usage was nearly identical. The overall mean positivity among clinics not included in the analysis was 7.5%.

Correlated Analysis of Continuously-reporting Clinics

Only the $\sigma_1^2 I_5$ R-matrix model converged consistently. The random intercept was statistically significant and retained in the model. Inclusion of the random slope did not alter findings, and the random slope was not significant. All 4 product terms (involving year with each covariate) were sequentially dropped because of collinearity; consequently, the resulting model assessed was a no-interaction model.

After adjusting for all covariates in the no-interaction model, the estimated effect of year, the independent predictor of interest, was null (OR: 1.00; CI: 0.99, 1.00; P=0.69), suggesting that chlamydia positivity did not change from 2004 to 2008 (Table 2.2). OR values for the continuous variables for the proportion of tests conducted among 15-19 year-old women (OR: 1.00; CI: 1.00, 1.00; P<0.001), among black women (OR: 1.02; CI: 1.01, 1.02; P<0.001), and using NAAT technology (OR: 1.00; CI: 1.00, 1.00; P<0.001) were extremely close to 1 (with narrow confidence intervals). The estimated effect of region varied.

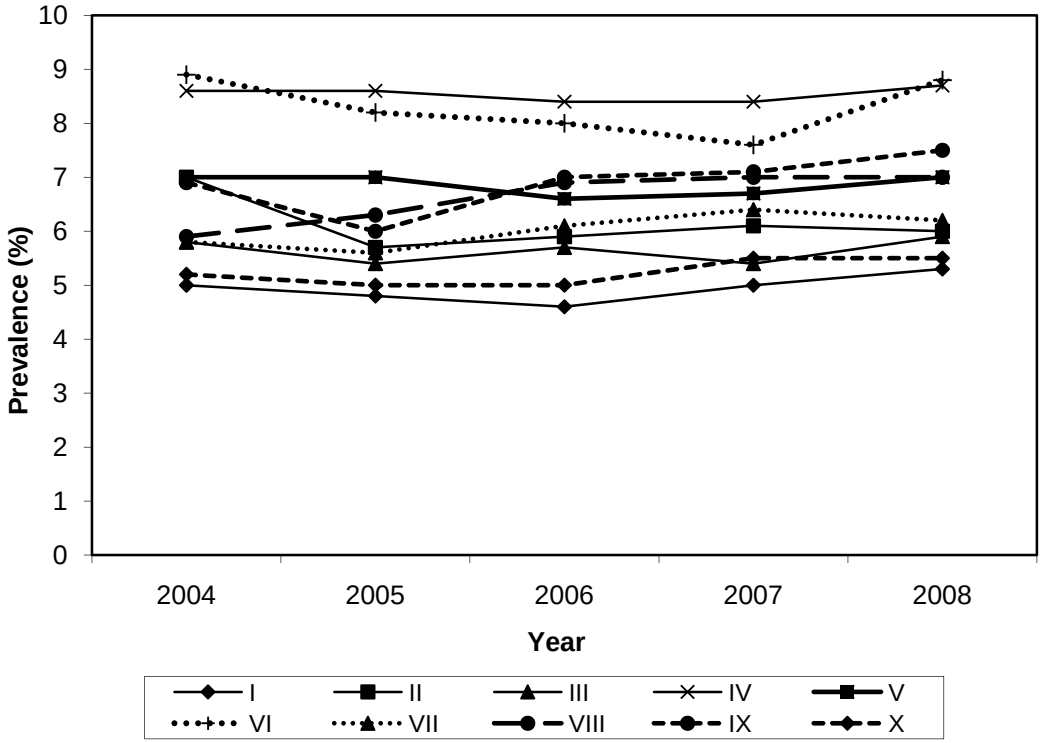
Table 2.2. Clinic-based model output assessing chlamydia positivity in family planning clinics reporting data to the Infertility Prevention Project from chlamydia tests conducted among women aged 15-24 years, 2004-2008

	Effect	Odds Ratio (95% CI)
Full Model	Year	1.00 (0.99, 1.00)
	Proportion 15-19	1.00 (1.00, 1.00)
	Proportion Black	1.02 (1.01, 1.02)
	Proportion NAAT	1.00 (1.00, 1.00)
	Region I	0.73 (0,67, 0.79)
	Region II	0.81 (0.72, 0.90)
	Region III	0.65 (0.61, 0.69)
	Region IV	0.90 (0.85, 0.96)
	Region V	0.97 (0.92, 1.03)
	Region VI	1.14 (1.07, 1.22)
	Region VII	0.93 (0.87, 0.99)
	Region VIII	1.16 (1.05, 1.27)
	Region IX	1.19 (1.08, 1.30)
	Region X	ref
Reduced Model*	Year	1.02 (1.01, 1.03)

CI=Confidence interval

When examining crude chlamydia positivity by region, the mean clinic-specific positivity appeared to fluctuate (Figure 2.2). However, when trends were modeled for each region, adjusting for race, test technology, and age, some minor variation in trends was noted, but no meaningful, substantial changes in chlamydia positivity from 2004 to 2008 were seen. In regions I, II, IX, and X, no changes in positivity were evident. In regions IV (OR: 0.95; CI: 0.95, 0.96), V (OR: 0.98; CI: 0.97, 1.00), and VI (OR: 0.97; CI: 0.95, 0.99), a suggestion of a decrease was detected, although the magnitude of change was negligible. Similarly, in regions VII (OR: 1.04; CI: 1.02, 1.05) and VIII (OR: 1.04; CI: 1.01, 1.07), a suggestion of an increase was detected, but changes were effectively null. The model for region III did not converge due to data limitations.

Figure 2.2 Mean clinic-specific chlamydia positivity among women aged 15-24 years who attended family planning clinics reporting data to the Infertility Prevention Project, by region, 2004-2008



DISCUSSION

Our analysis suggests that chlamydia positivity did not change among the population of women aged 15-24 years who were screened in family planning clinics reporting data to IPP from 2004 to 2008. When combined with other findings from earlier prevalence studies examining chlamydia trends, such as recently conducted analyses using data from NHANES and the NJTP, this analysis adds more evidence that chlamydia prevalence is not increasing (8, 47).

This conclusion runs counter to common misinterpretations using case report data, which is often incorrectly cited as evidence that the burden of chlamydia is increasing. From

2004 to 2008, reported chlamydia case rates increased more than 20% among women aged 15-24 years (3). During this same time period, NAAT usage increased. In 2004, 64.4% of all chlamydia tests conducted in surveyed public health laboratories were NAATs; by 2007, 81.6% of all tests performed utilized NAAT technology (19, 20). Expanded use of more sensitive test technology likely resulted in increased case detection, as has been previously reported (25). Concurrently, chlamydia screening coverage increased. Among sexually active women aged 16-20 years with commercial health insurance who were seeking health care, coverage increased from 32.6% in 2004 to 40.1% in 2008 (35). Similarly, coverage among young women in Medicaid managed care increased from 45.9% to 52.7%. Both of these factors, increasing NAAT usage and increasing coverage, affect case report trends.

Findings presented in this analysis differ from other analyses of IPP data. IPP data from women aged 15-24 years who were tested for chlamydia in family planning clinics suggested an upward trend in median state-specific positivity, which increased from 6.3% in 2004 to 7.4% in 2008 (45, 67); no adjustments for clinic or other possible confounders were made. This possible increase, similar to increases in chlamydia case reports, is at least partially explained by increasing NAAT usage; for instance, when NAAT usage among women screened in NJTP increased from about 20% to 88%, positivity increased from 9.1% to 13.9% in the absence of any other measured population changes (8). When overall crude chlamydia positivity was stratified by region, slight increases were noted in most regions (2). Conversely, findings reported here showed that chlamydia prevalence changed little by region. In this analysis, test technology, represented in the model as the

proportion of tests performed in a clinic that utilized NAATs, was treated as a confounder. However, because use different test technologies have varying sensitivity and specificity values, even within different types of NAATs, measurement error is a concern. Adjusting for test technology in the model presented in this analysis does not adequately address this type of error; rather, including test technology represents broader clinic practices.

Analyzing IPP data at the clinic-level may be preferable to individual test-based analyses. Few covariates are available in IPP data. However, clinic populations are likely similar over time, and use of a correlated analysis approach allows for clinics to act as a proxy for unmeasured confounding. Inclusion of a random intercept accounted for the natural heterogeneity among clinics on the prediction of chlamydia positivity trends. Because of the initial uncertainty in identifying a single appropriate correlation matrix, several were tested. While only the $\sigma_1^2 I_5$ R-matrix model consistently converged, this correlation matrix approximates a compound symmetric structure, supporting the original posit of a compound symmetric matrix being a possible matrix, if clinic populations maintain some consistency over time. Failure of models using other R-matrix structures should not have adverse implications. Overall, this analysis likely better characterizes national trends in chlamydia prevalence and positivity than unadjusted test-based analyses commonly reported.

As a result of sensitivity analyses assessing the importance of clinic participation in examining trends, IPP data were maximized by including all clinics reporting at least 3

years of data. This important finding allowed for inclusion of 58% of all clinics reporting any data from 2004 to 2008, and 88% of all tests reported. If analyses had revealed that stability and estimates were compromised by allowing incompletely-participating clinic data, only 42% of available clinics would have been included in the analysis (those clinics reporting 5 years of data, 2004 to 2008). Programmatic and funding decisions frequently affect annual clinic participation. Thus, a more inclusive approach should allow for maximum data usage, enhancing IPP data utility and allowing for broader generalizability within the IPP family planning clinic population. In addition, the developed approach is easily reproducible for future surveillance usage.

Monitoring trends in chlamydia prevalence is a critical component to assessing the impact of prevention efforts. While prevention strategies may be having some effect more remains to be done. Despite steady improvements in screening eligible women seeking healthcare, chlamydia trends remain stable. The burden of disease is substantial; this analysis shows regional variation in positivity of 5% to 9% among young women aged 15 to 24 years tested in family planning clinics analyzed. In addition, other reports have demonstrated substantial racial/ethnic disparities (45, 46, 47).

This analysis has several limitations. Although use of data summarized at the clinic-level likely minimized the influence of some unmeasured confounders, only a limited number of covariates were available. In particular, sexual behaviors were not measured and thus could not be accounted for in the analysis. Changes in clinic characteristics, such as screening policies, or uncaptured differences in clinic population, such as minor changes

within a demographic stratum that resulted in a lower-risk population being screened may be missed using the modeling approach presented. In addition, as with all analyses of national IPP data, lack of personal identifiers meant that some individuals may have been tested more than once, thus contributing more than one test result to the analysis. Another limitation was the inability to assess for effect measure modification. Due to collinearity between the product terms and the component covariates, all 4 interaction terms were removed from the model in the early assessment stages. This did not demonstrate the absence of interaction between year and the covariates, but rather interactions could not be properly assessed because of collinearity. However, when the model was run using a stratified approach (women aged 15 to 19 years, women aged 20 to 24 years, black women, white women, Hispanic women), findings were not meaningfully different (data not shown).

IPP primarily serves socio-economically disadvantaged young women; however, clinics participating in IPP have diverse populations and screen a substantial population of young women of various races and ages. Moreover, the IPP mean clinic-specific chlamydia positivity of 7.0% reported here is similar to an estimate of chlamydia prevalence among a nationally-representative (NHANES) sample of sexually active 14 to 19 year-old women (7.1%) (68). Although the ages represented are not identical, this suggests that the IPP family planning clinic population in this analysis may be somewhat similar to the general population of sexually active young women in the U.S. However, women tested in family planning clinics likely represent a heterogeneous risk group, seeking healthcare for a variety of reasons; other clinic types in IPP, such as prenatal

clinics where women are less impacted by healthcare seeking behaviors, may better approximate the general population.

In addition to some of the aforementioned strengths of the modeling approach, this analysis revealed consistent results using well-defined methodology, regardless of clinic participation level. Utilization of clinic-level summary data likely accounted for some unmeasured confounding and allowed for trend assessment in this population. Use of a correlated analysis approach with a random intercept included likely minimized the impact of varying clinic populations and policies. Moreover, the analytic approach applied in this study is easily reproducible; a correlated analysis with a random intercept addresses the study question using the best available data and moves beyond limitations of current IPP analyses. Other approaches, such as utilizing a multi-level model that included variables at the individual level as well as at the clinic level, may have also been appropriate if more extensive individual test data were available, such as sexual behavior data.

In summary, findings suggest that, in U.S. family planning clinics reporting chlamydia tests to IPP, chlamydia positivity did not change substantially among 15-24 year-old women from 2004 to 2008. These findings are consistent with other national prevalence analyses, suggesting that chlamydia prevalence in the U.S. is not increasing, despite increases in chlamydia case reports.

Chapter 3

Is the prenatal population an alternate population for monitoring chlamydia trends?

Chlamydia trachomatis Infections Among Women Attending Prenatal Clinics: United States, 2004-2009

Catherine Lindsey Satterwhite[1], MSPH, MPH; Alyson M. Gray[2], MPH; Stuart Berman[1], MD, SCM; Hillard Weinstock[1], MD, MPH; David Kleinbaum[3], PhD, MA; Penelope P. Howards[3], PhD, MS

[1]Centers for Disease Prevention and Control, Atlanta, Georgia
[2]CDC Foundation, Atlanta, Georgia
[3]Department of Epidemiology, Emory University, Atlanta, Georgia

ABSTRACT

Objective: Describe chlamydia screening practices, positivity, and trends from 2004-2009 in publicly-funded prenatal clinics.

Methods: A phone-based survey assessing chlamydia screening practices was conducted among a random sample of clinics providing prenatal services (prenatal, family planning, and integrated clinics: "prenatal clinics") that reported data to the Infertility Prevention Project (IPP) in 2008. Using existing data from IPP, chlamydia positivity and trends were assessed among women aged 15-24 years seeking care in any prenatal clinic reporting ≥3 years of data to IPP from 2004-2009. Linear trends of the effect of year (a continuous variable) on positivity were evaluated using a correlated modeling approach with a random intercept where the unit of analysis was the individual clinic performing chlamydia tests (clinic-based analysis). Covariates included race, age, test technology, and geography.

Results: Of 210 sampled clinics, 166 (79%) completed the survey. Of these, 163 (98.2%) had documented chlamydia screening criteria. Most clinics screened all women during their first trimester and reported 100% screening coverage. From 2004-2009, 267,416 tests among women aged 15-24 years were reported to IPP from eligible prenatal clinics. Overall chlamydia positivity was 8.3%. Controlling for all covariates, positivity decreased from 2004-2009 (odds ratio: 0.93 per year, 95% confidence interval: 0.92, 0.95, 35% decrease overall).

Conclusions: The substantial burden of chlamydia among young women tested in prenatal clinics reporting data to IPP suggests the continued need for routine screening.

Decreasing trends from 2004-2009 in the IPP prenatal population provides supporting evidence of overall decreasing chlamydia prevalence in the U.S.

BACKGROUND

Chlamydia is the most commonly reported notifiable disease in the U.S. Over 1.2 million cases were reported in 2009 (2). Among women, chlamydia, a sexually transmitted disease (STD), may cause PID, ectopic pregnancy, infertility, or chronic pelvic pain. Among pregnant women, chlamydia can lead to adverse birth outcomes (premature rupture of membranes, low birthweight) and infection of the infant (neonatal ophthalmia) (3, 53, 54, 55). Estimates of chlamydia prevalence among pregnant women are limited. In a U.S. population-based survey conducted from 1999 to 2002, 2.0% of pregnant women aged 14-39 years had chlamydia (46).

CDC currently recommends that all pregnant women, regardless of age, be screened for chlamydia at their first prenatal visit (69). The United States Preventive Services Task Force (USPSTF) recommends first trimester screening of pregnant women aged <25 years and older women at increased risk (e.g., new or multiple sexual partners or inconsistent condom use) (29). Both CDC and USPSTF recommend later additional screening during pregnancy for at-risk pregnant women.

Despite the risks of chlamydial infections in pregnancy and established screening recommendations, published reports on screening practices and screening coverage are not widely available. Screening coverage measures, such as those from commercial health plans and Medicaid, do not directly address pregnant women, possibly because chlamydia screening among pregnant women is assumed to be high (33). Screening for

syphilis, another STD, among pregnant women has been demonstrated to be widespread (70), suggesting that chlamydia screening may also be routine.

The Infertility Prevention Project (IPP) is a national screening program to detect and treat *Chlamydia trachomatis* and *Neisseria gonorrhoeae* infections among sexually-active young women to prevent sequelae leading to infertility. While the majority of clinics participating in IPP are designated as family planning clinics, prenatal clinics also participate in IPP.

IPP data represent a unique opportunity to assess chlamydia screening and positivity in a large group of women seeking prenatal care in publicly-funded clinics. In order to better describe the clinics serving these women, a survey assessing chlamydia screening policies and practices was conducted among prenatal clinics reporting data to IPP. Existing IPP data were then used to describe chlamydia tests conducted among women seeking prenatal services, including chlamydia positivity and trends from 2004 to 2009.

MATERIALS AND METHODS

Using data submitted through IPP, clinics providing care to women seeking prenatal services were identified. Each row of IPP data contains information on a chlamydia test result, the test technology used to conduct the test, basic demographics (age, race, sex), the state where the test was conducted, the type of clinic where the test was conducted, and the type of visit associated with the test conducted, defined as the reason a woman

sought care that day. Three types of clinics were eligible for inclusion, prenatal clinics, family planning providing prenatal services, and integrated clinics providing prenatal services. Prenatal clinics were defined as clinics whose primary mission is to provide health care and education to pregnant women. Family planning and integrated clinics offer multiple services; for these clinics, a visit type designated as "prenatal" was used to identify women eligible for inclusion. For simplicity, all three clinics will be referred to as "prenatal clinics" (clinics providing prenatal services) for this analysis.

Survey of Prenatal Clinics Reporting Data through IPP

From December 2009 to July 2010, a survey of prenatal clinics was conducted among prenatal clinics reporting data to IPP in 2008; due to funding limitations, a simple random sample of clinics, rather than all eligible clinics, was surveyed. Of the 559 prenatal clinics participating in IPP in 2008, 210 (38%) were invited to participate in the survey. The objective was to evaluate chlamydia screening policies and practices in publicly-funded prenatal clinics participating in IPP. Questions addressed age criteria used, frequency of testing, timing of test (trimester), and whether or not women seeking a pregnancy test only are routinely screened for chlamydia (survey instrument, Appendix A). Survey questions also included basic clinic census information aggregated at the clinic level (number of women seeking care at the clinic and the number of those women tested for chlamydia) to assess chlamydia screening coverage. The survey was administered to clinic administrators (or equivalent) via phone, with optional email follow-up. No incentives were offered for participating in the survey, which typically

took 10-15 minute to complete. The survey was determined to not constitute human subjects research (CDC and Emory University Institutional Review Boards).

Survey responses to clinic chlamydia screening policies and practices were summarized and reported in a descriptive manner. Open-ended responses were reviewed to determine common responses and create summary data groupings.

Analysis of Chlamydia Positivity and Trends

Using data from IPP, the study population consisted of all prenatal clinics reporting at least 3 years of data to IPP from 2004 to 2009. Only individual test results from women aged 15-24 years were included. In order for a clinic to contribute data for a given calendar year, the clinic must have reported at least 25 tests designated as prenatal (positive and negative) during that year; such a restriction reduces the influence of outlier clinics. National prenatal screening guidance is consistent in recommending that women aged <25 years be routinely screened for chlamydia in the first trimester; thus, an age-restricted analysis (women aged <25 years) was conducted. Of the subset of 166 prenatal clinics responding to the prenatal survey described above, 71 (43%) were eligible to be included in the trend analysis. The remaining clinics (95 of 166) either did not report enough years of data or reported data of insufficient volume to be included in the trend analysis.

Where available, pregnancy status was included in descriptive statistics, as well as trend analyses; however, pregnancy status is not part of the standard IPP dataset, and is

therefore not uniformly reported. It was thus not possible to determine if the visit was made because a woman was already aware of her pregnancy or if the visit was made primarily to seek a pregnancy test. Women seeking a pregnancy test only are a high-risk group, with substantially higher chlamydia rates than other groups, particularly women seeking standard prenatal care (56).

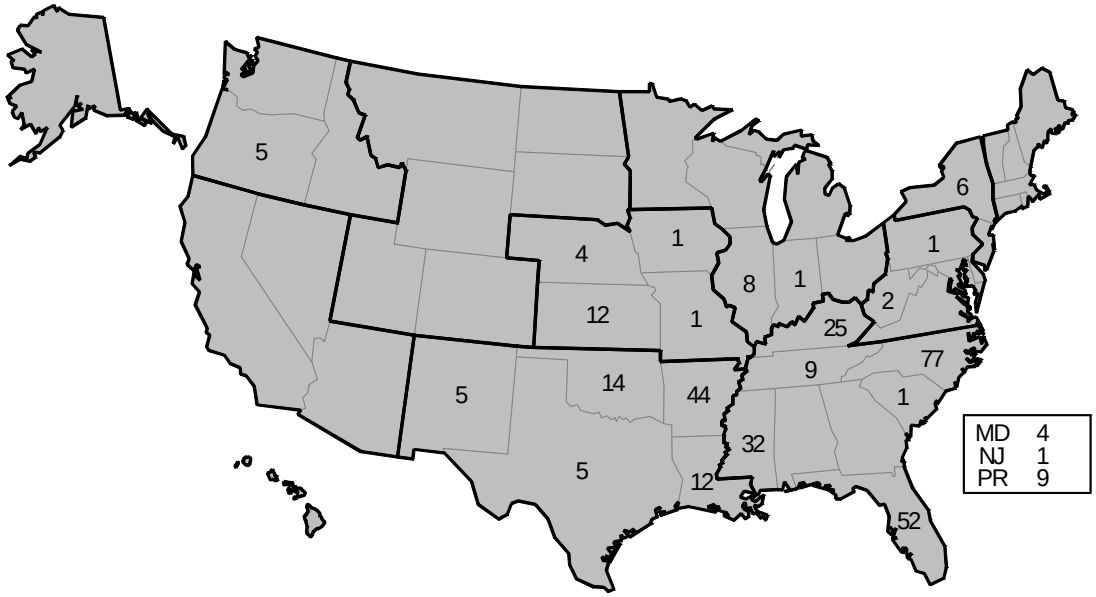
IPP data consist of chlamydia test results. No personal identifiers are available; therefore, women who had multiple test results over the course of a year contribute multiple observations to the dataset. Chlamydia positivity was calculated by dividing the total number of positive tests by the total number of positive and negative tests.

To assess positivity trends from 2004 to 2009, a correlated, longitudinal analysis was conducted. The unit of analysis was the individual clinic performing chlamydia tests (clinic-based analysis), as opposed to an analysis based only on individual test results because treating the clinic as a covariate in a multivariate model using individual test result data would produce a prohibitive number of parameters. Methodology using the clinic as the unit of analysis has been previously described and applied to IPP data; data summarized at the clinic level and treated as correlated likely minimizes the influence of some unmeasured confounders (chapter 2). The outcome of interest was chlamydia positivity. An events/trials approach was used to allow for the incorporation of clinic size (denominator) into the modeled outcome.

The primary independent exposure of interest was defined as calendar year, from 2004 to 2009. In order to determine if year should be treated as ordinal or categorical, a linearity assessment was performed. Estimated logit plots, as well as parameter estimates from a generalized estimating equation model including categorical indicator variables representing year, suggested a general linear pattern. Therefore, year was included as an ordinal variable in the final model. To account for the variation in the underlying prenatal chlamydia positivity between clinics at baseline, a random intercept was included. The data layout for the correlated, longitudinal analysis is provided in Appendix B.

Five possible covariates were considered. The proportion of tests conducted using NAAT technology performed by a given clinic was likely an important covariate because NAATs demonstrate substantially increased sensitivity over prior-generation tests (16). Two demographic covariates were also considered: the proportion of tests occurring among young women aged 15 to 19 years and the proportion of tests occurring among black women. The fourth covariate identified the region (Figure 3.1) where the clinic was geographically located. All two-way interaction terms (product of year with each of the 4 covariates) were assessed, in order to detect differences in trends over time in subgroups. The proportion of the tested population that was pregnant at the time of the test was also considered in the subset of clinics where this information was available.

Figure 3.1. Number of prenatal clinics reporting at least 3 years of data to the Infertility Prevention Project, by state and region*



*Region I: CT, ME, MA, NH, RI, VT; Region II: NJ, NY, PR, VI; Region III: DE, DC, MD, PA, VA, WV; Region IV: AL, FL, GA, KY, MS, NC, SC, TN; Region V: IL, IN, MI, MN, OH, WI; Region VI: AR, LA, NM, OK, TX; Region VII: IA, KS, MO, NE; Region VIII: CO, MT, ND, SD, UT, WY; Region IX: AZ, CA, GU, HI, NV; Region X: AK, ID, OR, WA

During trend ascertainment, data from four separate clinic scenarios were modeled: (1) data from all prenatal clinics, (2) data from prenatal clinics where all reported tests were conducted on women known to be pregnant, (3) data from all prenatal clinics where pregnancy test results were reported IPP, and (4) data from clinics where pregnancy status was not reported. The second scenario attempted to capture the specific population with the lowest risk (positivity). By selecting clinics who report chlamydia tests on pregnant women only, women not pregnant (and hypothetically only seeking a pregnancy test, a high-risk group) are excluded. The third scenario allowed for incorporation of pregnancy status into the model as a potential confounder. The fourth scenario was used to determine if trends among clinics not reporting pregnancy status were consistent with the other scenarios.

Prior to testing the four clinic scenarios, all covariates and interaction terms were entered into the initial model containing data from all prenatal clinics to assess for covariate collinearity using a SAS macro that produced diagnostics involving condition indices and variance decomposition proportions (collingenmodv9c.sas macro, Emory University, Atlanta, GA, modified). The macro to assess collinearity does not allow random effects; consequently, collinearity was evaluated for a logistic model with only fixed effects, the primary effects of interest, only. During collinearity assessments, all interaction terms were diagnosed to be sources of collinearity, and thus were removed from the model. To evaulate for confounding, odds ratio (OR) point estimates for year were compared across models including different covariate combinations. If models with subsets of all possible variables did not vary substantially from the full model, covariates were retained in the final model a priori due to demonstrated associations with the outcome (chlamydia positivity). The final model was also applied to population subgroups (age: 15-19, 20-24 years; race: black, white, Hispanic; tests conducted using NAAT technology, tests conducted using non-NAATs; pregnant women, non-pregnant women) to assess result consistency.

All statistical analyses were performed using SAS version 9.2 (SAS Institute Inc., Cary, NC). Modeling was conducted using the SAS GLIMMIX procedure, with a σ_0^2 (scalar) G-matrix and a $\sigma_1^2 I_5$ R-matrix correlation structure (exchangeable, with random intercept). All *P* values are two-sided.

RESULTS

Survey

Of the 210 clinics invited to participate in the survey, 166 (79%) completed questions about clinic chlamydia screening policies and practices. When compared to responders, non-responding clinics reported fewer tests to IPP (medians: 49 versus 21 tests reported in 2008) and had lower positivity among women aged <25 years (medians: 4.2% versus 0.9%). The majority of clinics not completing the survey (29 of 44, 66%) were not able to be contacted despite multiple attempts; failure to contact was due primarily to unreturned messages or lack of a useful phone number.

Overall, clinics from 22 states, Puerto Rico, and the Virgin Islands completed at least part of the survey. Of the 166 clinics completing the survey, 163 (98.2%) reported that their clinic had documented chlamydia screening criteria for prenatal women (Table 3.1). The majority of clinics did not have age-based screening criteria and screened women of all ages per CDC recommendations. Only four clinics (2.4% of 166 clinics) reported following USPSTF screening recommendations and only screened young pregnant women age <25 years routinely. Most clinics reported retesting women who were diagnosed with chlamydia earlier in pregnancy, per recommendations (74.1%). Over half of clinics rescreen "high-risk" women (57.8%, 96 of 166 clinics). When asked to define "high-risk", clinics included behavioral risk factors, such as having multiple (70.8%) or new (13.5%) partners, exchanging sex for money or drugs (4.2%), and having substance abuse problems (9.4%). Fourteen clinics (14.6% of 96 clinics) reported that risk was not

assessed, and conducted rescreening on all pregnant women. Only 12.7% of clinics (21 of 166) reported that they screen women aged <25 years who come into the clinic only seeking a pregnancy test.

Table 3.1. Description of reported chlamydia screening practices among sampled prenatal clinics participating in the Infertility Prevention Project, 2008

n=166	Number of Clinics (%)
Clinic has documented prenatal chlamydia screening criteria	163 (98.2)
Screen pregnant women aged <25 years only[1]	4 (2.5)
Clinic retests women who were diagnosed with chlamydia earlier in pregnancy	123 (74.1)
Retest after 1 month[2]	67 (54.5)
Retest after 2 months[2]	15 (12.2)
Retest after 3 months[2]	8 (6.5)
Retest in third trimester[2]	27 (22.0)
Clinics rescreen "high-risk" women	96 (57.8)
Multiple partners[3]	68 (70.8)
New partner[3]	13 (13.5)
Exchange sex for money or drugs[3]	4 (4.2)
Substance abuse[3]	9 (9.4)
Prior STD or PID[3]	35 (36.5)
Young age (adolescent/teen)[3]	5 (5.2)
Risk not assessed (rescreen all women)[3]	14 (14.6)
Clinic screens women aged <25 years who come into clinic seeking only a pregnancy test	21 (12.7)

[1]Percentage calculated using n=163 (number of with documented screening criteria).
[2]Percentage calculated using n=123 (number of clinics retesting). Only primary retesting time frame reported in table. Clinics may retest at multiple times.
[3]Percentage calculated using n=96 (number of clinics rescreening). Clinics may have identified multiple risk categories.

Of the 106 clinics reporting information on clinic census, only about half reported consistently analyzable data (e.g., did not respond to questions with "don't know"). Median clinic-reported screening coverage among women aged <25 years who were seeking prenatal care was 100% (mean: 86.2%, IQR: 77.6% to 100%, 55 clinics). Median chlamydia positivity in this group of women was 5.7% (mean: 10.0%, IQR: 4.0% to 10.6%, 40 clinics).

Chlamydia Positivity and Trends

From 2004 to 2009, 267,416 chlamydia tests conducted among young women aged 15-24 years seeking prenatal care were reported to IPP. These tests were administered in 335 prenatal clinics, each reporting data for at least 3 years during the analysis time frame: 213 clinics reported data 4 years or more; 154 reported data 5 years or more; and, 117 clinics reported data to IPP all 6 years. Eligible prenatal clinics were located in 23 states, Puerto Rico, and the District of Columbia, representing 7 of 10 Health and Human Services Public Health Regions (Figure 3.1). The mean number of tests reported per clinic per year was 180 (IQR: 61 to 227).

When examining mean clinic-specific proportions, the most tests were conducted among white women (37.0%), followed by Hispanic women (26.0%) and black women (24.8%) (Table 3.2). Most chlamydia tests were reported among women aged 20-24 years. While mean NAAT use increased over time, the median was 100.0% beginning in 2005, remaining at that level in subsequent years. All other covariates (race, age, region) were also stable over the 6-year analysis period (data not shown). Of the 335 prenatal clinics included in this analysis, 220 reported data on pregnancy status at the time of test administration (65.7%). The overwhelming majority of chlamydia tests conducted at these clinics were administered to pregnant women (mean: 94.4%).

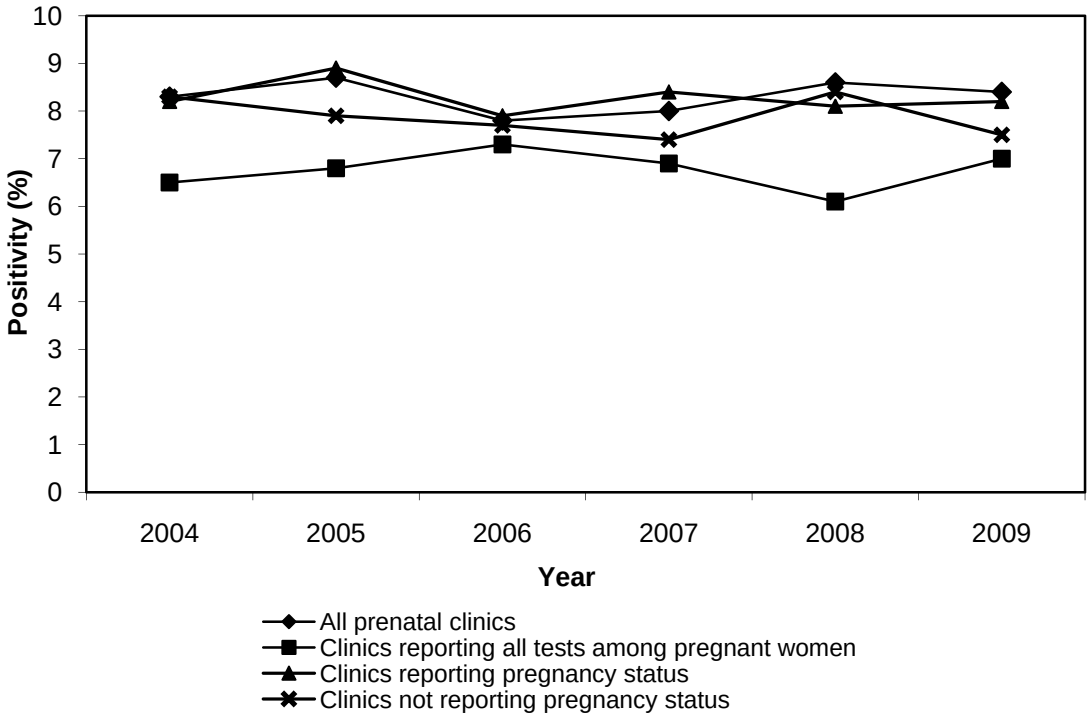
Table 3.2. Clinic-specific population characteristics of prenatal clinics reporting data to the Infertility Prevention Project on chlamydia tests conducted among women aged 15-24 years, 2004-2009

n=335	Mean Proportion* (Median)		IQR
Race/Ethnicity			
Hispanic	26.0	(10.2)	0.0-46.2
Non-Hispanic Black	24.8	(15.5)	2.4-42.0
Non-Hispanic White	37.0	(28.5)	10.6-59.8
Other	5.1	(0.6)	0.0-3.7
Unknown/Missing	7.1	(1.2)	0.0-5.0
Age Group			
15-19 Years	37.2	(37.2)	32.1-41.9
20-24 Years	62.1	(62.2)	57.5-66.7
NAAT Usage			
2004	47.9	(63.7)	0.0-74.9
2005	71.9	(100.0)	0.0-100.0
2006	73.9	(100.0)	0.4-100.0
2007	80.9	(100.0)	84.4-100.0
2008	92.9	(100.0)	100.0-100.0
2009	96.3	(100.0)	100.0-100.0
Pregnant[+]	94.4	(97.8)	93.9-99.6

NAAT=nucleic acid amplification test
*Mean and median of all clinic values
[+]Of 220 clinics with available data

Overall chlamydia positivity was 8.3% (22,097/267,416). When examining prenatal clinics by whether or not pregnancy status at the time of a chlamydia test was reported, trends in crude chlamydia positivity from 2004 to 2009 were fairly consistent, with no meaningful changes over time (Figure 3.2). Based on data from all prenatal clinics, positivity ranged from a minimum of 7.8% to a maximum of 8.7%. When assessing clinics where all chlamydia tests (100%) were reported as being conducted among pregnant women, the range was slightly lower, from 6.1% to 7.3%. Positivity in clinics reporting pregnancy status and clinics not reporting pregnancy status was similar.

Figure 3.2. Mean clinic-specific chlamydia positivity among women aged 15-24 years who attended prenatal clinics reporting data to the Infertility Prevention Project, by type of prenatal clinic, 2004-2009



*Includes prenatal clinics reporting at least 3 years of data from 2004-2009.

Contrary to crude positivity trends, trends assessed using a multivariate approach with year coded as a continuous variable showed a decrease in chlamydia positivity from 2004 to 2009 among women aged 15-24 years tested in all prenatal clinics reporting at least 3 years of data to IPP during this time frame (Table 3.3). The magnitude of effect for year did not change appreciably when different potential confounders were included or excluded from the model. Therefore, erring on the side of including all covariates, the full model controlled for age, race, test technology, and region, resulting in an OR of 0.93 (95% CI: 0.92, 0.95) per year. The odds of a positive chlamydia test declined by about 35% from 2004 to 2009 (average of 7% per year). OR values for the continuous covariates (age, race, NAAT) were near 1. The effect of region varied.

Table 3.3. Clinic-based model output assessing chlamydia positivity trends in prenatal clinics reporting data to the Infertility Prevention Project from chlamydia tests conducted among women aged 15-24 years, 2004-2009

Effect	Odds Ratio (95% CI)
Year	0.93 (0.92, 0.95)
Proportion 15-19	1.01 (1.01, 1.01)
Proportion Black	1.01 (1.01, 1.02)
Proportion NAAT	1.01 (1.00, 1.01)
Region II	1.93 (1.38, 2.70)
Region III	0.77 (0.58, 1.02)
Region IV	0.96 (0.78, 1.19)
Region V	0.98 (0.74, 1.30)
Region VI	1.42 (1.15, 1.75)
Region VII	1.03 (0.82, 1.31)
Region X	ref

CI=Confidence interval

Further sensitivity analyses varying the clinics included in analysis were consistent with these findings. Among the 33 prenatal clinics reporting that all tests were conducted on pregnant women, the OR associated with a single year change was 0.94 (95% CI: 0.86, 1.01). Likewise, clinics that reported pregnancy status, as well as clinics that did not report pregnancy status showed decreasing chlamydia positivity trends (data not shown). When the proportion of the population that was pregnant was entered as a covariate in the model, results did not change.

When data from each region were modeled separately, controlling for age, race, and test technology, some variations were seen in the effect of year on chlamydia positivity. Of the 7 regions included in the analysis, trends were relatively flat in 5 regions, as suggested by the OR associated with a single year change within a region and the corresponding CI: II (OR: 0.94, 95% CI: 0.88, 1.01), III (OR: 0.95, 95% CI: 0.84, 1.07), V (OR: 1.01, 95% CI: 0.95, 1.07), VI (OR: 1.00, 95% CI: 0.96, 1.07), and X (OR: 1.09,

95% CI: 0.89, 1.33). In Region VII, the OR was 1.04 (95% CI: 1.01, 1.08). In Region IV, where the majority of eligible prenatal clinics were located, a decrease was detected (OR: 0.92, 95% CI: 0.91, 0.94).

When population subgroups were independently modeled, adjusting for all other covariates, decreases of varying magnitude were consistently demonstrated by the OR values for a single year of change: women aged 15 to 19 years (OR: 0.92, 95% CI: 0.90, 0.94); women aged 20 to 24 years (OR: 0.95, 95% CI: 0.93, 0.96); black women (OR: 0.97, 95% CI: 0.95, 0.98); white women (OR: 0.96, 95% CI: 0.93, 0.99); Hispanic women (OR: 0.94, 95% CI: 0.91, 0.96); and, among women tested using NAAT technology (OR: 0.94, 95% CI: 0.92, 0.95). Among women tested using non-NAAT technology, the sample size was small, but a decrease was also suggested (OR: 0.96, 95% CI: 0.92, 1.01). Among only the subset of women known to be pregnant, the OR was 0.91 (95% CI: 0.89, 0.94). Although only 1,462 tests among non-pregnant women were reported from 6 clinics (reporting at least 25 of these tests for 3 or more years from 2004 to 2009), a decreasing trend in chlamydia positivity was also seen (OR: 0.92, 95% CI: 0.85, 0.99).

DISCUSSION

The chlamydia burden among women aged 15-24 years seeking prenatal services in clinics reporting data to IPP is substantial. The overall chlamydia positivity estimate of 8.3% is substantially higher than the population-based estimate of 2.0% (95% CI: 1.2%,

3.2%) (46). However, the latter estimate covers women tested at all stages of pregnancy, including many women who were likely already screened and treated for chlamydia at prior prenatal visits; thus, this may be an underestimate of the actual burden among pregnant women initially screened for chlamydia. In addition, the population-based estimate represents women aged 14 to 39, thus including older women at lower risk for chlamydia. Moreover, positivity in IPP prenatal clinics is likely an overestimate of the proportion of pregnant women with chlamydia because it includes women who may have sought a pregnancy test only, a higher risk group.

Although women seeking a pregnancy test only may perceive themselves to be at risk for pregnancy, in fact, they may also be at risk for an STD. In a recent study, Geisler et al. reported a chlamydia prevalence of 12% among women aged 16-45 years who were seeking a pregnancy test only (South Carolina family planning clinics) (56). Importantly, prevalence did not vary based upon whether or not the women tested positive for pregnancy. Similar risk has been reported among women seeking emergency contraception and induced abortions, groups where pregnancy was not desired and the risk behaviors (e.g., lack of barrier contraception) leading to that pregnancy may also lead to STD acquisition (57, 58).

While some women attending IPP prenatal clinics may have come only for a pregnancy test, the majority were pregnant at the time of chlamydia test administration. Among prenatal clinics where pregnancy status was available, almost all women tested were pregnant (mean 94.4%), suggesting that only a small proportion of women seeking

prenatal care were not aware of their pregnancy prior to the visit. In fact, chlamydia positivity in prenatal clinics reporting all chlamydia tests among pregnant women was only slightly lower than overall positivity, supporting this hypothesis.

Survey findings suggested that most clinics screen all pregnant women for chlamydia at least once prior to delivery regardless of age. Interestingly, only 21 (12.7% of 166 respondents) clinics reported routinely screening women aged <25 years who were only seeking a pregnancy test, despite this group being covered by general chlamydia screening recommendations. Pregnancy tests are performed using urine; testing the same sample for chlamydia using a nucleic acid amplification test (NAAT) requires no additional effort on the part of the patient, and little additional effort on the part of clinic staff. While prenatal clinics report high screening coverage among pregnant women, not actively screening women seeking a pregnancy test may be a missed opportunity for prevention, even if, relatively few women utilize prenatal clinics for this service.

The finding that chlamydia positivity decreased over the 6-year analysis timeframe supports a previous analysis of NHANES data assessing trends from 1999 to 2006 (47). Datta et al.'s analysis suggested an overall decrease among the general population of men and women in the U.S., aged 14-39 years. Likewise, an analysis of data from the National Job Training Program (NJTP), a program targeting socioeconomically disadvantaged men and women aged 16-24 years, showed significant declines in chlamydia prevalence among both sexes from 2003 to 2007 (8). These analyses run counter to reported trends in national chlamydia case report data (2). Increasing

screening coverage and improvements in test technology to use more sensitive tests make interpretation of such data challenging. Although chlamydia case rates have steadily increased over the past 20 years, these data likely do not represent true increases in disease incidence. Because of the limitations in case report data, examining chlamydia prevalence in specific populations, such as women in prenatal settings, is important in supplementing case report data.

When comparing regional trend findings in prenatal clinics (2004 to 2009) to trends in family planning clinics (2004 to 2008), results were similar (chapter 2). While small sample sizes limited the ability to detect changes in other regions, trends in Region IV in both types of clinics suggested decreasing positivity. Region IV is located in the Southeast U.S., where the burden of STDs is the highest (2).

While 23 states, Puerto Rico, and the District of Columbia were represented in this analysis, 3 regions (I, VIII, IX) of the U.S. were not included due to the absence of IPP data from any clinics designated as providing prenatal services. Findings presented in this analysis may not represent the prenatal populations in these states, located primarily in the upper Northeast, upper Midwest/Mountain region, and the West. Overall, the majority of prenatal clinics eligible for inclusion in this analysis were located in the eastern half of the U.S, specifically in Region IV.

Decreasing chlamydia prevalence suggests that public health intervention efforts may be having some impact, particularly in Region IV. Declining PID rates provide further

supporting evidence (2, 65), even given limitations in PID case detection. Chlamydia screening among the subset of sexually active young women who had healthcare visits has increased over time (33). However, mathematical modeling suggests that efforts beyond solely increasing screening may be necessary to substantially impact the burden of chlamydia in the U.S (71). Specifically, an increased emphasis on notifying and treating partners of patients with chlamydia has the potential to further reduce chlamydia prevalence.

This analysis has some limitations. The population of women represented in this analysis may be somewhat different than women in standard prenatal care who are already aware of and committed to their pregnancy. While it was not possible to determine if women knew their pregnancy status prior to their visit, nor to differentiate women seeking a pregnancy test only, chlamydia positivity trends were similar across different prenatal settings, suggesting that trends were likely not impacted by possible population differences. Only a limited number of covariates were available for analysis. Although use of data summarized at the clinic-level likely minimized the influence of some unmeasured confounders, additional data to allow for further classification of women by pregnancy status and healthcare seeking reason may have been useful. In addition, lack of personal identifiers necessitated the use of positivity calculations instead of prevalence. IPP data included in this analysis likely contained a small number of multiple test results for a single individual, especially since most surveyed prenatal clinics reported retesting pregnant women with an initial positive chlamydia test and

rescreening high-risk pregnant women. However, chlamydia positivity has been found to be a reasonable approximation of prevalence (48).

Findings of high chlamydia positivity in this analysis support the need for continued screening in the IPP population seeking prenatal care to prevent possible adverse outcomes of infection in both pregnant and non-pregnant women. The prenatal population may be a reasonable sentinel population to monitor trends in chlamydia prevalence. With the exception of population-based estimates, chlamydia prevalence surveillance is generally limited to higher risk populations, such as women attending family planning clinics (a healthcare-seeking population) and high-risk young men and women entering NJTP. Population-based estimates are not easily or economically obtained at state and local levels, nor are they reliable long-term national sources for chlamydia surveillance; if chlamydia prevalence continues to decline over time, standard errors will increase, limiting point estimate precision. The biases associated with service-based clinic prevalence are likely to be minimized when looking at the prenatal population compared to other populations captured in national data, because the prenatal population is likely to be a more stable population that is less impacted by general healthcare-seeking behaviors.

In summary, survey findings suggested that the prenatal care population is being regularly screened for chlamydia. This analysis demonstrates a substantial burden of chlamydial infection in the population of women aged 15-24 years who received a chlamydia test during a prenatal care visit to a publicly-funded clinic reporting data to

IPP. While positivity was high, modeled trends showed a decrease from 2004 to 2009, providing further evidence to suggest overall decreasing chlamydia prevalence in the U.S.

Chapter 4

How can administrative data be better used to detect PID cases, an important adverse outcome of untreated chlamydial infections?

PID Case Detection: Development of an Automated Case-Finding Algorithm Using Administrative Data

Catherine Lindsey Satterwhite[1], MSPH, MPH; Onchee Yu[2], MS; Marsha Raebel[3], PharmD,; Stuart Berman[1], MD, ScM; Penelope Howards[4], PhD, MS; Hillard Weinstock[1], MD, MPH; David Kleinbaum[4], PhD, MA, Delia Scholes[2], PhD, MS

[1]Centers for Disease Control and Prevention, Atlanta, Georgia
[2]Group Health Research Institute, Group Health Cooperative, Seattle, WA
[3]Institute for Health Research, Kaiser Permanente, Denver, CO
[4]Department of Epidemiology, Emory University, Atlanta, Georgia

ABSTRACT

Background: Research and surveillance work addressing pelvic inflammatory disease (PID) often rely on use of ICD-9 diagnostic codes from automated data sources to identify potential cases. However, cases identified in this way may not be clinical PID. A PID case-finding algorithm incorporating additional administrative data in addition to ICD-9 codes may offer improvements.

Methods: Using ICD-9 codes, potential PID cases were identified among women aged 15-44 years enrolled in two large health maintenance organizations, Group Health Cooperative (GH) in the northwest U.S. (data from 2003-2007) and Kaiser Permanente Colorado (KPCO) (data from 2003-2008). Medical records were reviewed to verify clinical PID status: 393 potential cases for algorithm development (GH) and 500 for external validation (KPCO). Using information on demographics, diagnosis and procedure codes, and treatment, a classification and regression tree analysis was conducted to develop a PID case-finding algorithm. Algorithm performance was compared to PID case-finding based on ICD-9 codes alone.

Results: When using ICD-9 codes alone to identify PID cases, the positive predictive value (PPV) was 78.8% in GH and 79.1% in KPCO. The algorithm identified two main predictors of PID beyond ICD-9 codes: PID-appropriate treatment and age 15-25 years. Algorithm PPV was 86.9% in GH and 84.5% in KPCO.

Conclusions: Algorithm PPV was high in both sites and improvements were seen in both sites. While better approaches for detecting clinician-diagnosed PID cases from

administrative databases are desirable, approaches such as this may assist with surveillance efforts.

BACKGROUND

An estimated 770,000 cases of pelvic inflammatory disease (PID) are diagnosed annually in the United States (64). PID comprises infection and inflammation of the uterus, fallopian tubes, ovaries, and other adjacent tissue and has multiple infectious etiologies, many of which have been demonstrated to be sexually transmitted, including *Chlamydia trachomatis* and *Neisseria gonorrheae* (12). *C. trachomatis* has been isolated in approximately one-quarter of patients with a symptomatic PID diagnosis (59).

PID is the most immediate adverse outcome of chlamydial infection among females. Untreated chlamydia leads to PID in approximately 10% to 15% of cases (59, 60). PID of any etiology may lead to further adverse outcomes, including tubal-factor infertility, ectopic pregnancy, and chronic pelvic pain (12); about 10% to 20% of PID cases are associated with infertility and ectopic pregnancy (59). The specific contribution of chlamydia to each of these adverse outcomes is unknown (61). However, among infertile couples using assisted reproductive therapy, 10-20% are diagnosed with tubal infertility (72, 73). In an effort to prevent PID and subsequent infertility, chlamydia screening is recommended for all sexually active women aged <25 years (11, 29). Prior studies have suggested that screening can reduce the risk of PID development by up to 50% (26, 74).

While monitoring trends in PID is a critical component to evaluating the impact of chlamydia and gonorrhea prevention efforts, PID surveillance is challenging. In the absence of a laboratory-based case definition, PID is diagnosed on the basis of clinical signs and symptoms (69). The Centers for Disease Control and Prevention (CDC) recommends empiric treatment for PID when young women have lower abdominal pain

with no other clear cause, accompanied by either uterine or adnexal or cervical motion tenderness (11). Thus, the clinical diagnosis lacks specificity and is not easily standardized. The "gold standard" for diagnosing tubal infection is laparoscopy, an invasive procedure that is rarely performed in clinical practice (62). Once made, a clinical diagnosis of PID can be represented by several ICD-9 codes and coding practices and preferences vary. ICD-9 codes have been used to identify PID cases from administrative data for analysis and surveillance purposes. The most commonly referenced ICD-9 code, 614.9 (female pelvic inflammatory disease not otherwise specified) has a positive predictive value (PPV) of only 18.1% for the CDC PID surveillance case definition, a substantially stricter than the clinical definition used for empiric treatment[1] (63). When coupled with a positive chlamydia test, the PPV increases to 56%; however, laboratory test results are frequently unavailable in the administrative datasets used to examine PID rates and trends.

A PID case-finding algorithm that moves beyond exclusive reliance on ICD-9 codes may represent an improvement in the methodology used to identify PID cases and allow for more accurate burden and trend ascertainment. In this study, a PID case-finding algorithm was developed using administrative data and medical record reviews from a large health plan in Washington State. The performance of the algorithm was then evaluated by applying it to data from another large health plan in Colorado.

---

[1] The surveillance definition specifies that the patient must have lower abdominal tenderness,AND tenderness with motion of the cervix,AND adnexal tenderness AND one of the following: *C. trachomatis* infection or gonorrhea OR temperature >100.4 F (>38.0 C) OR leukocytosis >10,000 white blood cells/mm OR purulent material in the peritoneal cavity obtained by culdocentesis or laparoscopy OR pelvic abscess or inflammatory complex detected by bimanual examination or by sonography OR the patient is a sexual contact of a person known to have gonorrhea, chlamydia, or nongonococcal urethritis.

METHODS

Data from two mixed model healthcare organizations, Group Health Cooperative (GH, Seattle, WA) and Kaiser Permanente Colorado (KPCO, Denver, CO) were used in this analysis. In 2006, approximately 125,000 women between the ages of 15 and 44 years were enrolled in GH, and about 116,000 women of the same ages were enrolled in KPCO. Both organizations maintain extensive automated administrative and clinical data including enrollment information, demographics, health care utilization, diagnoses, procedures, laboratory tests ordered, and pharmacy records on each enrollee.

Data Collection

To develop the PID case-finding algorithm, potential PID cases were identified from GH data using a set of ICD-9 codes used in other evaluations of PID cases selected from administrative databases (Table 4.1) (64, 65). Only codes for acute PID cases were considered since acute cases may better represent PID cases associated with infectious causes, such as chlamydia, that could be prevented by screening efforts. PID diagnoses that occurred within 60 days of each other were considered the same PID episode. Using GH data from 2003 to 2007, 2,764 total potential cases were identified among women aged 15 to 44 years. From these, 393 potential cases were randomly selected for a medical record review to determine if the clinician diagnosis was actually PID. The original sample size of 400 was selected based on resource availability; 7 potential cases were subsequently omitted from the sample due to data discrepancies. If a woman had multiple PID diagnoses from 2003 to 2007, only the first PID episode was eligible for inclusion into the sample. The distribution of ICD-9 codes associated with the 393

potential cases is shown in Table 4.1; multiple ICD-9 codes could have been selected for the visit associated with each potential PID case.

Table 4.1. ICD-9 codes commonly utilized to identify acute pelvic inflammatory disease (PID) and code distribution among potential PID cases sampled from Group Health Cooperative

ICD-9 Code	**Description**	**Number of Potential Cases with Code* (%)**
098.10	Acute GC Upper GU tract, site unspecified	5 (1.3)
098.16	Acute GC Endometritis	
098.17	Acute GC Salpingitis	
098.19	Acute GC Upper GU tract, other site	
098.86	Acute GC Peritonitis	
099.56	Acute CT Peritonitis	0 (0.0)
614.0	Acute Salpingo-oophoritis	8 (2.0)
614.5	Acute or Unspecified Pelvic Peritonitis	
614.8	Other Specified Inflammatory Disease, Female Pelvic Organs	
614.2	Salpingitis/oophoritis, not acute or chronic	22 (5.6)
614.3	Acute Parametritis/PID	53 (13.5)
614.9	Unspecified Inflammatory Disease, Female Pelvic Organs	252 (64.1)
615.0	Inflammatory Disease of Uterus, except cervix	15 (3.8)
615.9	Unspecified Inflammatory Disease of Uterus	80 (20.4)

GC=gonorrhea
GU=genito-urinary
CT=chlamydia
* A single potential PID case may include multiple ICD-9 codes. 393 total potential PID cases were identified, and a total of 435 ICD-9 codes were used.

Determination of the actual PID case status (i.e., a valid clinical diagnosis) was made by reviewing the electronic medical record using a structured chart review instrument (Appendix C). Potential PID cases were confirmed as being clinician-diagnosed cases or not based on explicit clinician documentation of a PID diagnosis during the visit. The determination of clinical PID status was made regardless of what clinical signs or symptoms were indicated to support such a diagnosis and which case definition criteria may have been used (empiric treatment for suspected PID or diagnosis based on the stricter surveillance case definition). Cases where the clinical case status was uncertain

were further reviewed by another team member (DS). Once the medical record review was complete, personal identifiers were removed from the database.

In addition to the set of ICD-9 diagnosis codes for PID shown in Table 4.1, potential predictors that were evaluated in the PID case-finding algorithm included age at diagnosis, treatment for PID, inpatient admission, whether chlamydia testing was conducted, and other diagnoses occurring within 7 days of PID diagnosis date. These diagnoses were defined by ICD-9 codes and included appendicitis, ovarian cysts, ectopic pregnancy, pyelonephritis, pancreatitis, leiomyoma, and endometriosis. Treatment appropriate for PID was defined as levofloxacin (500mg orally once per day for 14 days) or ofloxacin (400mg orally twice per day for 14 days); other possible antimicrobial regimens for PID treatment were also included.

To evaluate the performance of the PID case-finding algorithm in another population, administrative data from KPCO was used as an external validation dataset. In the KPCO administrative data from 2003 to 2008, 2,685 potential PID cases were identified using ICD-9 codes alone among women aged 15 to 44 years. Of these, 500 were randomly selected for medical record review to determine the clinical PID case status. The same structured chart review instrument that was used in the GH development dataset was used during the medical record review at KPCO.

Statistical Analysis

A classification and regression tree (CART) analysis was performed to develop a PID case-finding algorithm using the GH development dataset, which consisted of data on each potential PID case, including the diagnosis based on medical record review and the potential predictors described above in addition to each of the 14 ICD-9 codes initially used to identify cases. The algorithm goal was to identify variables that would predict the clinical PID case status as defined by the medical record review. CART is a nonparametric, binary recursive partitioning method that builds a decision tree or a classification algorithm by splitting data into two groups at each branch (or "node") (75). Important predictors are hierarchically identified, and potential cases are classified as PID cases or not at each node. This process is repeated multiple times until the optimal tree is built. At each branch, data are optimally split to maximize the differentiation of observations based on the dependent variable; in this case, the dependent variable was a confirmed clinical PID diagnosis (yes/no) from medical record review.

After development, the PID-case finding algorithm was applied to the KPCO external validation dataset. Algorithm performance was assessed by comparing the PID case status predicted by the algorithm to the clinical PID case status determined by medical record review (clinical diagnosis) and calculating the positive predictive value (PPV, proportion of algorithm-classified PID cases that were confirmed to be PID by medical record review). Because the study population was selected based on ICD-9 codes, rather than a random sample of the covered women in GHC and KPCO, calculation of sensitivity, specificity, and negative predictive value was not possible.

Analyses were conducted using SAS version 9.1.2 (SAS Institute Inc., Cary, NC), R (R Foundation for Statistical Computing, Vienna, Austria), and OpenEpi (76, 77). The CART analysis was performed using 'rpart' in the R package. All study procedures received human subjects review and approval at each institution.

RESULTS

Of the 393 potential PID cases identified using ICD-9 codes alone from GH where chart abstractions were performed, 275 (70.0%) potential cases were confirmed to be clinical PID based on medical record review; 74 (18.8%) were determined to not be PID; 6 (1.5%) were determined to be of uncertain case status, and 38 (9.7%) potential cases had no information available regarding the visit where the PID ICD-9 code was recorded (Table 4.2). Of the 500 potential KPCO PID cases, 349 (69.8%) were confirmed to be PID, 92 (18.4%) were not PID, 5 (1.0%) were classified as uncertain, and 54 (10.8%) had no information available on the visit.

Table 4.2. Results from medical record reviews to assess PID cases status at Group Health Cooperative (GH) and Kaiser Permanente Colorado (KPCO)

	PID Diagnosis Based on Medical Record Review (%)				Total
	PID	Not PID	Uncertain	No Information	
GH Development Dataset	275 (70.0)	74 (18.8)	6 (1.5)	38 (9.7)	393
KPCO Validation Dataset	349 (69.8)	92 (18.4)	5 (1.0)	54 (10.8)	500

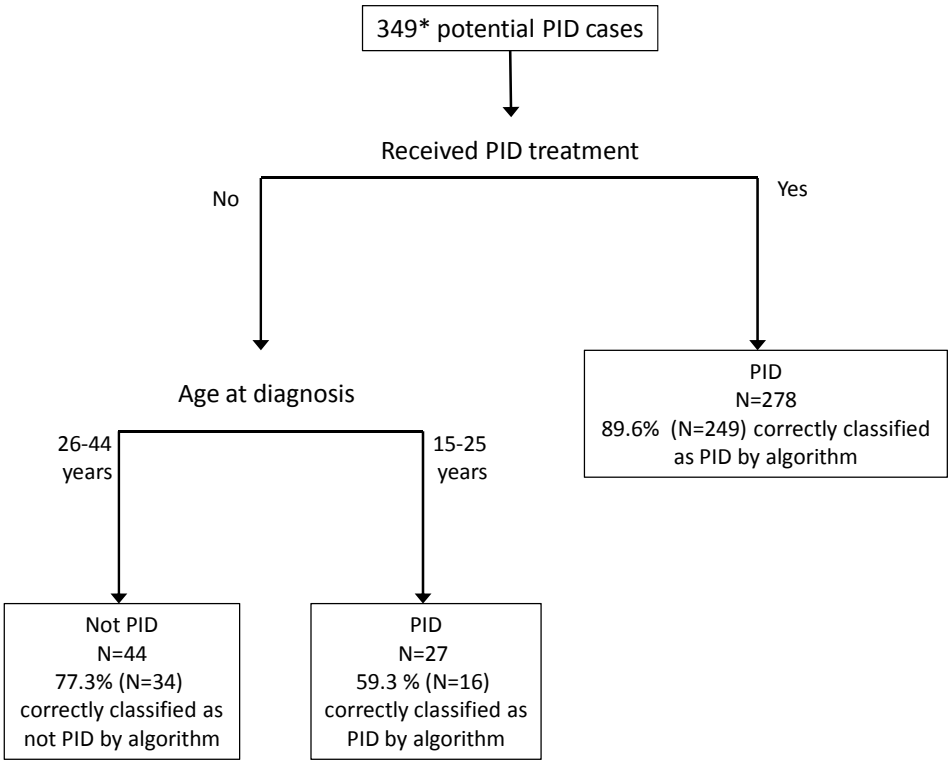
PID=pelvic inflammatory disease

Fourteen ICD-9 codes were used to identify potential PID cases from GH; 614.9 was the most common, associated with 64.1% of the 393 potential cases (Table 4.1). The majority of visits where a potential PID case was identified had only one ICD-9 selected (92.4%); 5.9% had two codes, and 1.8% had three or more. Three ICD-9 codes (614.3, 614.9, 615.9) identified 96.4% of all confirmed PID cases (265/275) in GH. However, 55 of the 74 potential cases confirmed to not be PID (74.3%) also were coded with one of these three codes. Of the 500 potential PID cases in KPCO, 50.4% were coded with the 614.9 ICD-9 code. Similarly, of the 441 confirmed PID cases, 51.7% used the 614.9 code.

The PID case-finding algorithm is shown in Figure 4.1. Of the 393 potential PID cases in the GH development dataset, 44 were excluded because medical record review failed to confirm or reject a clinical diagnosis of PID. Thus, 349 potential PID cases were used to develop the algorithm. Two predictors of clinical PID were identified by the algorithm. The strongest predictor identified was the presence of treatment appropriate for PID. Of the 278 potential cases where treatment was documented in administrative data, 249 (89.6%) were confirmed as clinically diagnosed PID by the medical record review. However, 29 (10.4%) potential cases with PID treatment were not confirmed as PID. Among those women with no PID treatment recorded, younger age was found to be the most important predictor. Specifically, young women between 15-25 years of age who had not received PID treatment were classified by the algorithm as PID cases. Among 27 such women, 14 (59.3%) were confirmed PID cases. Among 44 women who had no PID treatment and were aged 26-44 years, 34 (77.3%) were correctly classified by the

algorithm as not having PID. No specific ICD-9 code was a stronger predictor than PID treatment and age. In summary, if any antimicrobial treatment appropriate for PID was included in the administrative data, the potential case was classified by the algorithm as being a PID case; if no treatment was recorded, age was considered next; potential cases occurring in women aged 15-25 years were classified as PID cases, and potential cases in women aged 25-44 years were classified as not PID.

Figure 4.1. PID case-finding algorithm developed using automated administrative data from Group Health Cooperative



PID=pelvic inflammatory disease
*Of 393 potential PID cases, 44 were not included due to uncertainty of PID case status after medical record review.

Of the 305 potential PID cases classified by case-finding algorithm as PID, 265 were confirmed by medical record review, which resulted in a PPV of 86.9% (265/305, 95% CI: 82.9-90.5%) (Table 4.3). When using ICD-9 codes alone to identify PID cases, the PPV was 78.8% (275/349, 95% CI: 74.1-83.0%). Overall, 21.2% of potential PID cases identified using ICD-9 codes alone were false positives (74/349), classified as PID cases but found to not be PID upon medical record review. When the PID case-finding algorithm was applied, the proportion of false positives improved, decreasing to 13.1% (95% CI: 9.7-17.3%).

Table 4.3. Performance statistics comparing PID case identification from administrative data using ICD-9 codes* alone versus algorithm developed by CART analysis

Accuracy of PID case-finding algorithm: GH Development Dataset

		New Algorithm Classification		Total
		Not PID	**PID**	
Chart-confirmed Diagnosis	**Not PID**	34	40	74
	PID	10	265	275
	Total	44	305	349

Accuracy of PID case-finding algorithm: KPCO Validation Dataset

		New Algorithm Classification		Total
		Not PID	**PID**	
Chart-confirmed Diagnosis	**Not PID**	34	58	92
	PID	34	315	349
	Total	68	373	441

Performance Statistics (95% CI)	**GH Development Dataset**	**KPCO Validation Dataset**
New algorithm PPV	86.9% (82.6-90.5%)	84.5% (80.4-88.0%)
ICD-9 codes* alone PPV	78.8% (74.1-83.0%)	79.1% (75.0-82.8%)

* ICD-9 codes shown in Table 4.1. Only potential cases with complete chart-review information are included.

CART=classification and regression tree
GH=Group Health
KPCO=Kaiser Permanente Colorado
PID=pelvic inflammatory disease
CI=Confidence interval
PPV=positive predictive value

The distribution of the two predictors included in the algorithm was similar for PID treatment between GH and KPCO potential PID cases, but different for age at PID diagnosis. In GH, 90.6% (249/275) of confirmed PID cases had documented PID treatment, compared to 39.2% (29/74) cases found not to be PID. Likewise, in KPCO, 84.0% (293/349) of confirmed PID cases had received antimicrobial treatment, compared to 38.0% (35/92) of non-PID cases. When examining age at diagnosis in GH, 49.1% (135/275) of confirmed PID cases had a diagnosis age of <26 years, compared to 28.4% (21/74) of cases confirmed by medical record review to not be PID. However, in KPCO, a different pattern was seen; women aged <26 years accounted for 38.1% (133/349) of confirmed PID cases, and 41.3% (38/92) of cases that were not PID.

When the algorithm was applied to the external validation dataset from KPCO, the PPV of the PID case-finding algorithm was 84.5% (95% CI: 80.4-88.0%); of the 373 potential cases classified as PID by the algorithm, 315 were confirmed to be PID by medical record review (Table 4.3). When using ICD-9 codes alone to identify PID cases, the PPV was 79.1% (349/441, 95% CI: 75.0-82.8%). The proportion of potential PID cases misclassified as false positives using ICD-9 codes alone was 20.9% (95% CI: 17.2-25.0%); when applying the PID case finding algorithm, the proportion of false positives improved to 15.5% (12.1-19.5%).

Discussion

One of the primary goals in STD prevention is to reduce the burden of STD-associated infertility. Monitoring trends in PID, an intermediate adverse outcome between STD

acquisition and the development of infertility, may help identify progress in STD prevention. However, surveillance of PID has been historically difficult.

The clinical diagnosis of PID is imprecise and not standardized. Symptoms of PID can be very mild, and subclinical tubal infection and inflammation are known to occur. Only two-thirds of women with a clinical diagnosis of PID actually have salpingitis documented by laparoscopy (visual confirmation of fallopian-tube infection) (12). In addition, the case definition for PID used for surveillance purposes is more restrictive than criteria for empiric treatment of PID. A PID case is defined by CDC and the Council of State and Territorial Epidemiologists as the presence of lower abdominal pain of no known cause plus lower abdominal tenderness, cervical motion tenderness, and adnexal tenderness plus at least one other indication[1] (78). Conversely, empiric PID treatment is recommended for patients presenting with lower abdominal pain and only one additional symptom. Because of the possibility for the development of severe adverse outcomes, clinicians are encouraged to maintain a low threshold for the diagnosis and treatment of PID (11).

To identify clinical diagnoses of PID from medical records data, administrators and researchers have traditionally relied solely on ICD-9 codes. However, no single ICD-9 code is universally used for indicating a PID diagnosis. While subsets of ICD-9 codes used have been similar across studies, use of ICD-9 codes has substantial limitations, including a lack of specificity (63). ICD-9 codes are not applied in a standard fashion

and are subject to varying usage by individuals and healthcare sites in selecting which ICD-9 codes to indicate for diagnostic purposes.

The algorithm developed in this analysis, incorporating additional automated data elements, was an attempt to improve the identification of clinically diagnosed PID cases and move beyond exclusive reliance on ICD-9 codes to enhance surveillance efforts that rely on administrative data. Neither of the two predictors identified by the algorithm, antimicrobial treatment for PID and age at diagnosis, were individual ICD-9 codes, even though each ICD-9 code used to initially identify potential cases was entered as a possible predictor. While this supports the idea that there is no single ICD-9 code that predicts a clinical PID diagnosis well and that a set of ICD-9 codes should continue to be used to identify potential PID cases, three ICD-9 codes (614.3, 614.9, 615.9) identified 96.4% of all confirmed PID cases in GH. Overall, using ICD-9 codes to detect potential PID cases is simple and fairly accurate, with a PPV of about 79% in both sites.

Overall, the algorithm offered some improvement in case identification in a both settings Using GH data, 21.2% of potential PID cases identified using ICD-9 codes were not confirmed by medical record review (i.e., were false positives), compared with 13.1% of those classified by the algorithm (an absolute improvement of 8.1%). In KPCO, 20.9% of potential cases identified using ICD-9 codes alone were false positives, and 15.5% were false positives using the PID case-finding algorithm (an absolute improvement of 5.4%). Because the study population consisted of potential PID cases identified using ICD-9 codes, no information on the population with PID but not identified by ICD-9

codes (false negatives) was available; thus, the sensitivity, specificity, and negative predictive value of the case-finding algorithm could not be determined.

Despite these challenges in case identification, small improvements, such as those demonstrated by the algorithm's added value in reducing false positives, offer opportunities to move beyond the current practice of identifying PID cases based only on ICD-9 codes for surveillance purposes. If resources permit and PID surveillance or research is a priority, it may be reasonable for an individual health plan, or a group of health plans, to develop a unique algorithm to identify potential PID cases best suited to their data. The algorithm developed would depend on geographic, plan-based, or provider-based variations in ICD-9 code usage, depending upon the best predictors of a clinically diagnosed PID case that are identified. In this analysis, treatment appropriate for PID was the strongest predictor of clinical PID in both study settings; using the set of ICD-9 codes plus only treatment would likely offer improvements in case identification at other sites as well. In GH, this approach resulted in a PPV of 89.6%; in KPCO, PPV was 84.0% (data not shown).

Currently, due to widespread data limitations, public health professionals must rely primarily on ecologic comparisons of STD incidence trends, PID diagnosis trends, and concurrent sexually transmitted disease (STD) prevention activities in order to evaluate programmatic impact. However, as data systems improve, ascertainment of STD-specific PID diagnoses may be possible with better automated linkages between laboratory data, clinical data, and other administrative data. The expanded use of electronic medical

records will likely further enhance surveillance of STD-associated PID. The identification of possible methods to improve PID case-finding will be a contributing factor.

This analysis has several limitations. Due to budget and time restrictions, only a limited number of chart abstractions were possible. However, a random sample of potential PID cases was selected, so the data used to develop and validate the algorithm should be generally representative of the entire population of potential PID cases in GH and KPCO during the years that were included. In this analysis, potential PID cases were identified using ICD-9 codes. It is possible that the ICD-9 codes used missed some clinically diagnosed PID cases that were not coded with one of these selected codes (false negatives). These diagnoses were not included in the datasets and may have limited our ability to fully assess the performance of case-finding by ICD-9 codes alone and the algorithm. Prior literature on PID case identification has faced similar challenges to those discussed here in ascertaining the sensitivity, specificity, and negative predictive value of ICD-9 based approaches.

Strengths of this analysis include the utilization of CART methodology, which allowed for a comprehensive evaluation of all available predictors of clinically diagnosed PID cases and all possible value splits of those predictors without the necessity of making assumptions about underlying variable distributions (non-parametric approach). Interpretation of the CART findings was straight-forward and easily applied to another external setting (KPCO) after initial algorithm development. In addition, the algorithm

was tested at a second geographic site to allow for an assessment of robustness. Finally, this analysis provided new evidence on the performance of using ICD-9 codes alone to identify PID.

Monitoring PID is important in assessing STD prevention and control efforts, particularly prevention of chlamydia and gonorrhea. Although the approach utilized in this analysis can help improve efforts to monitor PID, findings also highlight the difficulties in identifying PID cases. While efforts to explore novel approaches to identifying PID should continue, using ICD-9 codes alone currently appears to be an acceptable approach to identify potential PID cases and monitor ongoing trends.

Chapter 5

Discussion and Summary

Chlamydia prevention programs, based primarily on screening, have been in place for 15 years, yet program evaluation remains challenging. National case report rates continue to steadily increase, reflective primarily of expanded screening efforts. The asymptomatic nature of chlamydia necessitates screening for detection of infection; case report data will continue to be subject to screening programs: who gets tested? Therefore, data sources that can be used to calculate chlamydia positivity and prevalence will play ongoing important roles in assessing the burden of chlamydial infection and allowing ecologic analyses of the impact of prevention efforts, when compared to screening coverage data and adverse reproductive outcome data. The studies presented in this dissertation contribute to the existing body of literature on chlamydia trends by reporting new findings on chlamydia trends and demonstrating a novel analytic approach to trend ascertainment, in addition to evaluating a new algorithm for identifying the most immediate adverse outcome of chlamydial infection, PID.

Population-based estimates of chlamydia prevalence from the National Health and Nutrition Examination Survey (NHANES), a nationally representative sample of the non-institutionalized U.S. civilian population, have significantly declined from 1999 to 2006 among men and women aged 14 to 39 years (47). New data, recently submitted for publication, suggest that this trend continues through 2008 (79). An analysis of a high-risk population, men and women aged 16 to 24 years who were uniformly screened when entering the National Job Training Program (NJTP), also reported significant declines in chlamydia prevalence (8). Likewise, the analysis of chlamydia positivity trends in prenatal clinics reported in this dissertation (chapter 3) showed a decreasing trend. When

considered together, these three unique sources provide strong evidence that the overall burden of chlamydia in the U.S. is decreasing.

While the analysis of chlamydia positivity trends in family planning clinics did not show a decrease, findings still confirm that increasing trends in national case report rates are likely more representative of programmatic activities rather than actual disease prevalence, particularly when taking the other analyses of prevalence into account. There are several possible reasons that trends in family planning clinics did not appear to be decreasing. For one, although the analytic approach treated observations in each clinic as correlated over time to minimize the impact of unmeasured confounders, changes may have occurred in clinic policies or practices over time. For instance, in the 2009 Comprehensive STD Prevention System (CSPS) grant, the CDC instructed programs to focus chlamydia screening efforts among sexually active young women (aged <26 years) to clinics (e.g., family planning clinics, STD clinics, etc.) where chlamydia positivity was three percent or higher (80). If clinic positivity was below three percent, the program needed to either shift resources away from that clinic, to a clinic with higher prevalence, or develop plans to target screening to those women at highest risk (e.g., adolescents). To accommodate these requests, it is possible that relatively low risk clinics began to selectively test riskier women, thus increasing their clinic-based positivity. Such changes may have had an impact on the family planning trend assessment and masked a possible decrease in general clinic positivity. Another possible explanation for the lack of decreasing chlamydia positivity trends in the family planning analysis centers on the population seeking care at family planning clinics. Generally, women seeking health care

at family planning clinics do so because of some perceived risk, either STD related or related to pregnancy prevention. In other populations where chlamydia prevalence was assessed, including prenatal clinics, health care seeking behaviors are likely to have much less potential effect on observed prevalence.

While valuable, data from NHANES are not easily or economically reproduced at state and local levels, nor will NHANES be a reliable long-term national source for chlamydia surveillance if chlamydia prevalence continues to decline over time. With decreases and low prevalence, standard errors increase, limiting point estimate reliability. Issues such as those surrounding family planning clinics (and other service-based settings, such as STD clinics) and concerns about the future utility of NHANES to provide reliable population-based estimates suggest the need for novel approaches to surveillance. One approach might be to create a sentinel clinic system for surveillance, where all clients are screened, regardless of risk; clinics participating in this system would not be subject to programmatic influences, but measures will still be affected by population health care seeking behaviors. Another option, supported by findings in this dissertation, is to further develop similar surveillance in the prenatal population.

Findings from the prenatal analysis are particularly important. Although somewhat limited geographically, findings were generally consistent with decreases seen in NHANES. Positivity decreases in the prenatal population suggest that this population may be a stable population less impacted by general health care seeking behaviors and therefore a population where chlamydia prevalence trends may mirror the general

population. In the 65.7% of clinics where pregnancy status was available (chapter 3), over 94% of women tested were concurrently pregnant; in addition, survey results suggested that all of these women are routinely tested for chlamydia. Assuming a steady sample of women who get pregnant over time, the population seeking prenatal care, whether in public or private healthcare settings, is consistent. Likewise, screening recommendations are unlikely to change, and the prenatal population should continue to be well-screened. In a setting such as a family planning clinic, the population attending the clinic (and subsequently tested) is more unknown; reasons for seeking care vary and may change over time. Likewise, as described earlier, screening policies and practices within a family planning setting are dependent upon resource availability. Even given high resources, screening coverage among non-pregnant sexually active young women is low. Future chlamydia surveillance efforts should be explored using data from the prenatal population, including identifying mechanisms to utilize data from the privately-insured prenatal population.

In addition to the contribution of the findings presented in this analysis, the analytic approach employed offers opportunities to improve ongoing surveillance in publicly-funded family planning and prenatal clinics participating in the Infertility Prevention Project. The analytic approach applied is easily reproducible; a correlated analysis with a random intercept addresses the study question assessing clinic-based chlamydia positivity trends using the available best data and moves beyond limitations of current IPP analyses. Other approaches, such as utilizing a multi-level model that included variables at the individual level as well as at the clinic level, may have also been appropriate if more

extensive individual test data were available, such as sexual behavior data. Even if modeling is not performed, evaluating crude chlamydia positivity among continuously participating clinics likely presents a more accurate picture of burden than calculating positivity broadly (either by state or at a national level). Trends seen in median state-specific positivity among women aged 15-24 years who were tested in family planning participating in IPP show upward trends (2); conversely, median clinic-specific positivity in the same group of women is fairly flat, as demonstrated in chapter 2. Treating clinics as correlated through a longitudinal modeling approach further strengthens the analytic approach.

Another critical consideration in chlamydia surveillance is the active monitoring of trends in adverse reproductive sequelae associated with chlamydia: PID, ectopic pregnancy, and infertility. As the most proximal adverse outcome, PID trends may offer some insight into the impact of chlamydia and gonorrhea prevention efforts. Monitoring adverse outcomes of chlamydia may be more important than monitoring chlamydia prevalence; the goal of chlamydia screening is to prevent adverse reproductive outcomes. Although data have limitations, the National Survey of Family Growth suggests that infertility decreased from 2002 to 2008 (81). Ectopic pregnancy trends assessed using data from a large database of administrative claims were flat from 2002 to 2007 (82). Decreases in PID have also been reported (2, 65). While analyses such as these are important and somewhat suggestive of a possible ecologic impact of chlamydia screening efforts, surveillance of adverse outcomes remains challenging.

The analysis assessing a new PID case-finding algorithm presented in this dissertation identified a mechanism to improve the detection of clinically diagnosed PID cases from an existing insurance database. While using ICD-9 codes alone to identify clinically-diagnosed PID cases was adequate, applying the PID case-finding algorithm further improved the PPV. While a statistical test assessing changes in the PPV would have been useful, such a test was not readily available, so conclusions reached in the analysis were limited only to clinical significance. The changes in the PPV were likely clinically meaningful, with absolute improvements of 8.1% in Group Health Cooperative (78.8% to 86.9%) and 5.4% in Kaiser Permanente Colorado (79.1% to 84.5%). Efforts such as these are important steps towards advancing surveillance of reproductive sequelae of chlamydial infection. A critical next step is to identify mechanisms to conduct surveillance on chlamydia-associated sequelae.

As electronic medical records improve, opportunities to create linkages between laboratory test results and concurrent diagnoses will increase, and chlamydia surveillance can move beyond a reliance on ecologic analyses to assess programmatic impacts. Maximizing the utility of administrative data, both in assessing adverse outcomes and in examining the privately insured prenatal population, offers substantial opportunities to improve surveillance around this important public health issue.

References

1. Centers for Disease Control and Prevention. Summary of notifiable diseases -- United States, 2008. MMWR 2010;57:1-94.
2. Centers for Disease Control and Prevention. Sexually transmitted disease surveillance, 2009. Atlanta, GA: U.S. Department of Health and Human Services, November 2010.
3. Schachter J, Stephens RS. Biology of *Chlamydia trachomatis*. In: Holmes KK, Sparling PF, Stamm WE, et al., eds. Sexually Transmitted Diseases, Fourth Edition: McGraw Hill, 2008:555-74.
4. Farley T, Cohen D, Elkins W. Asymptomatic sexually transmitted diseases: The case for screening. Prev Med 2003;36:502-9.
5. Miller W, Ford C, Morris M, et al. Prevalence of chlamydial and gonococcal infections among young adults in the United States. JAMA 2004;291:2229-36.
6. Sutton T, Martinko T, Hale S, et al. Prevalence and high rate of asymptomatic infection of *Chlamydia trachomatis* in male college reserve officer training corps cadets. Sex Transm Dis 2003;30:901-4.
7. Cecil JA, Howell MR, Tawes JJ, et al. Features of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* infection in male army recruits. J Infect Dis 2001;184:1216-9.
8. Satterwhite CL, Tian LH, Braxton J, et al. Chlamydia prevalence among women and men entering the National Job Training Program: United States, 2003-2007. Sex Transm Dis 2010;37:63-7.
9. Stamm WE. *Chlamydia trachomatis* infections of the adult. In: Holmes KK, Sparling PF, Stamm WE, et al., eds. Sexually Transmitted Diseases, Fourth Edition: McGraw-Hill, 2008:575-93.
10. Lau C-Y, Qureshi A. Azithromycin versus doxycycline for genital chlamydial infections: A meta-analysis of randomized clinical trials. Sex Transm Dis 2002;29:497-502.
11. Centers for Disease Control and Prevention. Sexually transmitted diseases treatment guidelines, 2010. MMWR 2010;59:1-110.
12. Paavonen J, Westrom L, Eschenbach D. Pelvic inflammatory disease. In: Holmes KK, Sparling PF, Stamm WE, et al., eds. Sexually Transmitted Diseases, Fourth Edition: McGraw-Hill, 2008.
13. Hadgu A, Dendukuri N, Hilden J. Evaluation of nucleic acid amplification tests in the absence of a perfect gold-standard test: A review of the statistical and epidemiologic issues. Epidemiology 2005;16:604-12.
14. Moncada J, Schachter J, Liska S, et al. Evaluation of self-collected glans and rectal swabs from men who have sex with men for detection of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* by use of nucleic acid amplification tests. J Clin Microbiol 2009;47:1657-62.
15. Masek B, Arora N, Quinn N, et al. Performance of three nucleic acid amplification tests for detection of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* by use of self-collected vaginal swabs obtained via an internet-based screening program. J Clin Microbiol 2009;47:1663-7.

16. Black C, Marrazzo J, Johnson R, et al. Head-to-head multicenter comparison of DNA probe and nucleic acid amplification tests for *Chlamydia trachomatis* infection in women performed with an improved reference standard. J Clin Microbiol 2002;40:3757-63.
17. Schachter J. NAATs to diagnose *Chlamydia trachomatis* genital infection: A promise still unfulfilled. Expert Rev Mol Diagn 2001;1:137-44.
18. Dicker L, Mosure D, Steece R, et al. Laboratory tests used in US public health laboratories for sexually transmitted diseases, 2000. Sex Transm Dis 2004;31:259-64.
19. Dicker L, Mosure D, Steece R, et al. Testing for sexually transmitted diseases in U.S. public health laboratories in 2004. Sex Transm Dis 2007;34:41-6.
20. Yee E, Satterwhite C, Braxton J, et al. Current STD laboratory testing and volume in the United States among public health laboratories, 2007. Presented at 18th ISSTDR (International Society for STD Research), London, England, 06/28-07/01/2009.
21. Lee S-E, Nauschuetz W, Jordan N, et al. Survey of sexually transmitted disease laboratory methods in US Army laboratories. Sex Transm Dis;37:44-8.
22. Stamm WE. *Chlamydia trachomatis*--the persistent pathogen: Thomas Parran award lecture. Sex Transm Dis 2001;28:684-9.
23. Centers for Disease Control and Prevention. Policy guidelines for prevention and control: *Chlamydia trachomatis* infections. MMWR 1985;34:53S-74S.
24. Centers for Disease Control and Prevention. Recommendations for the prevention and management of *Chlamydia trachomatis* infections, 1993. MMWR 1993;42:1-38.
25. Dicker LW, Mosure DJ, Levine WC, et al. Impact of switching laboratory tests on reported trends in *Chlamydia trachomatis* infections. Am J Epidemiol 2000;151:430-5.
26. Scholes D, Stergachis A, Heidrich FE, et al. Prevention of pelvic inflammatory disease by screening for cervical chlamydial infection. NEJM 1996;334:1362-6.
27. U.S. Preventive Services Task Force. Screening for chlamydial infection: Recommendations and rationale. Am J Prev Med 2001;20:90-4.
28. Nelson HD, Helfand M. Screening for chlamydial infection. Am J Prev Med 2001;20:95-107.
29. U.S. Preventive Services Task Force. Screening for chlamydial infection: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2007;147:128-34.
30. Meyers D, Halvorson H, Luckhaupt S. Screening for chlamydial infection: An evidence update for the U.S. Preventive Services Task Force. Ann Intern Med 2007;147:135-42.
31. Maloney SK, Johnson C. Why screen for chlamydia? An implementation guide for healthcare providers. Partnership for Prevention, Washington, DC, 2008.
32. Maciosek M, Coffield A, Edwards N, et al. Priorities among effective clinical preventive services: Results of a systematic review and analysis. Am J Prev Med 2006;31:52-61.

33. Centers for Disease Control and Prevention. Chlamydia screening among sexually active young female enrollees of health plans-United States, 2000-2007. MMWR 2009;58:362-5.
34. National Committee for Quality Assurance. (http://www.ncqa.org/tabid/59/Default.aspx). Accessed March 11, 2011.
35. National Committee for Quality Assurance. The state of healthcare quality. Washington, DC: National Committee for Quality Assurance, 2009.
36. Centers for Disease Control and Prevention. Male chlamydia screening consulation: Meeting report (Atlanta, GA), 2007. (http://www.cdc.gov/std/chlamydia/ChlamydiaScreening-males.pdf). Accessed March 11, 2011.
37. Dunne E, Gift T, Stamm W. What about the men? Sex Transm Dis 2008;35:S1-S2.
38. Rietmeijer C, Hopkins E, Geisler W, et al. *Chlamydia trachomatis* positivity rates among men tested in selected venues in the United States: A review of the recent literature. Sex Transm Dis 2008;35:S8-S18.
39. Joesoef MR, Mosure D. 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;33:636-9.
40. Satterwhite C, Joesoef MR, Datta SD, et al. Estimates of chlamydia trachomatis infections among men: United States. Sex Transm Dis 2008;35:S3-S7.
41. Gift T, Gaydos C, Kent C, et al. The program cost and cost-effectiveness of screening men for chlamydia to prevent pelvic inflammatory disease in women. Sex Transm Dis 2008;35:S66-S75.
42. Peterman T, Newman D, Goldberg M, et al. Screening male prisoners for *Chlamydia trachomatis*: Impact on test positivity among women from their neighborhoods who were tested in family planning clinics. Sex Transm Dis 2009;36:425-9.
43. Weinstock H, Berman S, Cates W. Sexually transmitted diseases among American youth: Incidence and prevalence estimates, 2000. Perspect Sex Reprod Health 2004;36:6-10.
44. Miller WC. Epidemiology of chlamydial infection: Are we losing ground? Sex Transm Infect 2008;84:82-6.
45. Centers for Disease Control and Prevention. Sexually transmitted disease surveillance, 2008. Atlanta, GA: U.S. Department of Health and Human Services, November 2009.
46. Datta SD, Sternberg M, Johnson RE, et al. Gonorrhea and chlamydia in the United States among person 14 to 39 years of age, 1999 to 2002. Ann Intern Med 2007;147:89-96.
47. Datta SD, Sternberg M, Satterwhite C, et al. Trends in *Chlamydia trachomatis* prevalence in the U.S., 1999-2006: Results from the National Health and Nutrition Examination Survey (NHANES). Presented at 48th Annual ICAAC/IDSA 46th Annual Meeting, Washington, DC, 10/25-28/2008.
48. Dicker LW, Mosure DJ, Levine WC. Chlamydia positivity versus prevalence. What's the difference? Sex Transm Dis 1998;25:251-3.

49. Mertz KJ, Ransom RL, St Louis ME, et al. Prevalence of genital chlamydial infection in young women entering a national job training program, 1990-1997. Am J Prev Med 2001;91:1287-90.
50. Joesoef MR, Mosure D. Prevalence trends in chlamydial infections among young women entering the National Job Training Program, 1998-2004. Sex Transm Dis 2006;33:571-5.
51. Hosenfeld C, Workowski K, Berman S, et al. Repeat infection with chlamydia and gonorrhea among females: A systematic review of the literature. Sex Transm Dis 2009;36:478-89.
52. Dunne E, Chapin J, Rietmeijer C, et al. Rate and predictors of repeat chlamydia trachomatis infection among men. Sexually Transmitted Diseases 2008;35:S40-S4.
53. Hitti J, Watts DH. Bacterial sexually transmitted infection in pregnancy. In: Holmes KK, Sparling PF, Stamm WE, et al., eds. Sexually Transmitted Diseases, Fourth Edition: McGraw-Hill, 2008:1529-61.
54. Chow J, Kang M, Samuel M, et al. Assessment of the association of *Chlamydia trachomatis* infection and adverse perinatal outcomes with the use of population-based chlamydia case report registries and birth records. Public Health Rep 2009;124:24-30.
55. Blas M, Canchihuaman F, Alva I, et al. Pregnancy outcomes in women infected with *Chlamydia trachomatis*: A population-based cohort study in Washington state. Sex Transm Infect 2007;83:314-8.
56. Geisler W, James A. Chlamydial and gonococcal infections in women seeking pregnancy testing at family-planning clinics. Am J Obstet Gynecol 2008;198:502.e1-.e4.
57. Kettle H, Cay S, Brown A, et al. Screening for *Chlamydia trachomatis* infection is indicated for women under 30 using emergency contraception. Contraception 2002;66:251-3.
58. Renton A, Thomas BM, Gill S, et al. *Chlamydia trachomatis* in cervical and vaginal swabs and urine specimens from women undergoing termination of pregnancy. Int J STD AIDS 2006;17:443-7.
59. Haggerty C, Gottlieb S, Taylor B, et al. Risk of sequelae after *Chlamydia trachomatis* genital infection in women. JID 2010;201 Suppl 2:S134-S55.
60. Oakeshott P, Kerry S, Aghaizu A, et al. Randomised controlled trial of screening for *Chlamydia trachomatis* to prevent pelvic inflammatory disease: The POPI (prevention of pelvic infection) trial. BMJ 2010;340:c1642.
61. Wallace LA, Scoular A, Hart G, et al. What is the excess risk of infertility in women after genital chlamydia infection? A systematic review of the evidence. Sex Transm Infect 2008;84:171-5.
62. Simms I, Warburton F, Westrm L. Diagnosis of pelvic inflammatory disease: Time for a rethink. Sex Transm Infect 2003;79:491-4.
63. Ratelle S, Yokoe D, Blejan C, et al. Predictive value of clinical diagnostic codes for the CDC case definition of pelvic inflammatory disease (PID): Implications for surveillance. Sex Transm Dis 2003;30:866-70.

64. Sutton M, Sternberg M, Zaidi A, et al. Trends in pelvic inflammatory disease hospital discharges and ambulatory visits, United States, 1985-2001. Sex Transm Dis 2005;32:778-84.
65. Bohm MK, Newman L, Satterwhite CL, et al. Pelvic inflammatory disease among privately insured women, United States, 2001-2005. Sex Transm Diseases 2010;37:131-6.
66. Fine D, Dicker L, Mosure D, et al. Increasing chlamydia positivity in women screened in family planning clinics: Do we know why? Sex Transm Dis 2008;35:47-52.
67. Centers for Disease Control and Prevention. Sexually transmitted disease surveillance, 2004. Atlanta, GA: U.S. Department of Health and Human Services, September 2005.
68. Forhan S, Gottlieb S, Sternberg M, et al. Prevalence of sexually transmitted infections among female adolescents aged 14 to 19 in the United States. Pediatrics 2009;124:1505-12.
69. Centers for Disease Control and Prevention. Sexually transmitted diseases treatment guidelines, 2006. MMWR 2006;55:7,38-40.
70. Rosenman M, Tao G, Szucs K, et al. Prenatal syphilis screening rates measured using Medicaid claims and electronic medical records. Sex Transm Dis 2008;35:387-92.
71. Kretzschmar M, Satterwhite C, Leichliter J, et al. Effects of screening and partner notification on chlamydia prevalence in the United States: A modeling study. Sex Transm Dis Submitted 2011.
72. Macaluso M, Wright-Schnapp TJ, Chandra A, et al. A public health focus on infertility prevention, detection, and management. Fertil Steril 2010;93:16 e1-0.
73. Centers for Disease Control and Prevention, American Society for Reproductive Medicine. Assisted reproductive therapy success rates: National summary and fertility clinic reports. Atlanta, GA: Department of Health and Human Services, December 2009.
74. Ostergaard L, Andersen B, Mller JK, et al. Home sampling versus conventional swab sampling for screening of *Chlamydia trachomatis* in women: A cluster-randomized 1-year follow-up study. Clin Infect Dis 2000;31:951-7.
75. Breiman L, Friedman J, Stone CJ, et al. Classification and regression trees. Belmont, CA: Taylor and Francis, 1984.
76. Dean A, Sullivan K, Soe M. OpenEpi: Open source epidemiologic statistics for public health, version 2.3.1. (www.OpenEpi.com, updated 2010/19/09) Accessed March 11, 2011.
77. RDC Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2008. (www.R-project.org). Accessed March 11, 2011.
78. Centers for Disease Control and Prevention. Case definitions for infectious conditions under public health surveillance, 1997. MMWR 1997;46:52.
79. Datta S, Torrone E, Kruszon-Moran D, et al. *Chlamydia trachomatis* trends in the United States among persons 14 to 39 years of age, 1999-2008. Sex Transm Dis Submitted 2011.

80. Centers for Disease Control and Prevention. CDC-RFA-PS09-902. Department of Health and Human Services, 2009 (http://www.cdc.gov/od/pgo/funding/PS09-902.htm). Accessed March 11, 2011.
81. Chandra A, Mosher W, Copen C, et al. Sexual behavior, sexual attraction, and sexual identity in the United States: Data from the 2006-2008 National Survey of Family Growth. Natl Health Stat Report 2011;36:1-49.
82. Hoover K, Tao G, Kent C. Trends in the diagnosis and treatment of ectopic pregnancy in the United States. Obstet Gynecol 2010;115:495-502.

Appendix A: Survey Instrument for Survey of Prenatal Clinics Participating in the Infertility Prevention Project in 2008

--CDC USE ONLY--	
Unique Facility ID	*<from Facility Reference File: facility_link_id>*
Facility Name	*<from Facility Reference File: facility_name>*
Facility State	*<from Facility Reference File: facility_state_name>*
Facility County	*<from Facility Reference File: facility_county_name>*
Facility City	*<from Facility Reference File: facility_city_name>*
Facility Phone Number	*<from various sources>*
Facility Type	*<from Facility Reference File: facility_type_code>*
Date Survey Administered	*<collected with survey data: variable name – date>*

--START SURVEY HERE--

Hello, my name is _____, and I'm calling on behalf of the Centers for Disease Control and Prevention. The Division of STD Prevention is conducting a brief survey to evaluate data collected through the Infertility Prevention Project, sometimes called IPP. We are also going to assess chlamydia screening practices in a sample of publicly-funded prenatal clinics. Your clinic was randomly selected to participate in this survey. This survey will only take about 10 minutes. I will ask questions about your clinic's chlamydia screening policies, as well as several questions about the number of women who come to your clinic.

First, are you the correct person to speak to about participating in this survey?

<*If no*> Could you please let me know who I should speak to? If possible, could you transfer me? *<name/contact information, if provided>*

<Once correct party identified>
Would your clinic be interested in participating in this important survey?
□ Yes □ No

QUESTION	RESPONSE
1. Before we begin, how may I address you?	
2. What is your title? *<clinic administrator, data manager??>*	

Now, I have a few questions to ask to make sure that I have your basic clinic information correct.

QUESTION	RESPONSE
3. For confirmation, could you please tell me the official name of your clinic?	
4. In what state is your clinic located?	
5. In what county is your clinic located?	
6. In what city is your clinic located?	

<Confirm that information on questions 3-6 matches the CDC-provided information in the gray box, above. Note that the name may be slightly different, and this is acceptable. If information does not match, please state:

Our information about your clinic appears to be incorrect, and we may have contacted you in error. We need to review our information before proceeding. Thank you for your time, and we apologize.>

Thanks. The next few questions that I will ask will help with our assessment of chlamydia screening practices.

7. Do you have documented, facility-specific chlamydia screening criteria?
   - ☐ Yes (*go to question 7a.*)
   - ☐ No (*go to question 8*)
   - ☐ Not sure (*go to question 8*)

7a. Could you please describe your documented, facility-specific chlamydia screening criteria?

7b. Could you send your documentation to us via email?

☐ Yes ☐ No

<*If yes*> The email address that you should use is ITT1@CDC.GOV.
<*If no, go to question 8*>

8. In your clinic, do you follow age-based screening criteria when screening pregnant women for chlamydia?
   - ☐ Yes, we screen all pregnant women under the age of 25
   - ☐ Yes, other
     Could you please describe?
   - ☐ No, we screen all pregnant women, regardless of age
   - ☐ Other
     Could you please describe?
   - ☐ Not sure

9. In your clinic, do you screen women under the age of 25 for chlamydia when they come in for a pregnancy test only?
   - ☐ Yes
   - ☐ No
   - ☐ Other
     Could you please describe?
   - ☐ Not sure

10. In your clinic, do you retest women for chlamydia if they tested positive for chlamydia earlier in their pregnancy?
    - ☐ Yes (*go to question 10a.*)
    - ☐ No (*go to question 11*)
    - ☐ Not sure (*go to question 11*)

    10a. When do you retest? <*describe below*>

11. Do you rescreen "high-risk" women for chlamydia, even if they tested negative for chlamydia earlier in their pregnancy?
    - ☐ Yes (*go to question 11a. & 11b.*)
    - ☐ No (*go to question 12*)
    - ☐ Not sure (*go to question 12*)

    11a. How does your clinic define "high-risk"? <*describe below*>

    11b. When do you rescreen "high-risk" women? <*describe below*>

Finally, these last few questions will help us assess IPP data quality. I'm going to ask some questions about your clinic population in 2008. If you don't have access to data from 2008, I can follow-up with you at another time.

I'm going to start asking the questions, but we can pause or stop at any time, if you need to find some information.

If you aren't sure of or do not have data from 2008 at this time, I can call you back at another time that's convenient for you. If you would prefer, I can email you the data questions.

<If the participant would like to be called back, set a date and time with the participant before ending the call.
Enter date and time:

If the participant would like to be emailed, state: I will send the questions to you via email after we finish the call. The email I send will include instructions for completing the questions, as well as when and how you should send your responses back to me. Could you please provide your email address >
Enter email address:

<If Facility Type is Prenatal Clinic, then ask:>

QUESTION	RESPONSE
12. How many total women under the age of 25 did you see at your clinic in 2008?	
13. How many of these women were tested at least once for chlamydia in 2008?	
14. How many of the women you saw under the age of 25 in 2008 were seeking just a pregnancy test? *<pregnancy test only>*	
15. How many total chlamydia tests did your clinic conduct in 2008 among women under the age of 25?	
16. How many positive chlamydia tests did your clinic have in 2008 among women under the age of 25?	
17. Out of the positive tests among women under the age of 25, how many of them were associated with pregnancy? *<how many of the women with a positive chlamydia test were also pregnant>*	

18. Approximately what percentage of women who come to your clinic are coming for prenatal care?
    - ☐ Less than 25%
    - ☐ 25-50%
    - ☐ 51-75%
    - ☐ More than 76%
    - ☐ Don't know

<If Facility Type is Integrated Clinic, then ask:>

QUESTION	RESPONSE
12. How many total women under the age of 25 who were seeking prenatal care did you see at your clinic in 2008?	
13. How many of these women were tested at least once for chlamydia in 2008?	
14. How many of the women seeking prenatal care that you saw under the age of 25 in 2008 were seeking just a pregnancy test? *<pregnancy test only>*	
15. How many total chlamydia tests did your clinic conduct in 2008 among women under the age of 25 who were seeking prenatal care?	
16. How many positive chlamydia tests did your clinic have in 2008 among women under the age of 25 who were seeking prenatal care?	
17. Out of the positive tests among women under the age of 25 who were seeking prenatal care, how many of them were associated with pregnancy? *<how many of the women with a positive chlamydia test were also pregnant?>*	

THANK YOU FOR YOUR PARTICIPATION IN THIS SURVEY!

Appendix B. Sample data layout for prenatal clinics reporting three or more years of data to the Infertility Prevention Project, 2004-2009

FACILITY_ID	YEAR	CT_TOTAL_POSITIVE	CT_TOTAL_TESTS	PROP_15TO19	PROP_BLACK	PROP_NAAT	REGION
0101101	2004	19	414	58.5942	42.3564	0.000	04
0101101	2005	32	485	56.2887	44.5798	0.000	04
0101101	2006	18	598	57.8662	47.9933	4.849	04
0101101	2007	40	347	52.8242	46.1931	100.000	04
0101101	2008	8	43	51.8605	47.2093	100.000	04
0101101	2009	20	201	55.2564	45.6893	100.000	04
0101104	2004	.	.	.	.	.	.
0101104	2005	.	.	.	.	.	.
0101104	2006	1	44	52.2727	59.0909	0.000	04
0101104	2007	7	29	58.6207	48.2759	100.000	04
0101104	2008	41	343	54.5190	41.9825	100.000	04
0101104	2009	.	.	.	.	.	.
0102501	2004	.	.	.	.	.	.
0102501	2005	1	63	52.3810	6.5219	0.000	04
0102501	2006	.	.	.	.	.	.
0102501	2007	3	161	59.0062	4.3478	100.000	04
0102501	2008	11	86	56.9767	8.7907	100.000	04
0102501	2009	28	181	58.9184	6.1791	100.000	04

FACILITY_ID: unique facility ID
YEAR: 2004-2009 (calendar year)
CT_TOTAL_POSITIVE: total number of positive chlamydia tests
CT_TOTAL_TESTS: total number of positive or negative chlamydia tests
PROP_15TO19: proportion of tests conducted among women aged 15-19 years
PROP_BLACK: proportion of tests conducted among black women
PROP_NAAT: proportion of tests conducted using NAAT technology
REGION: HHS Region (see Figure 2.1)

Appendix C. Abstraction form used for medical record review of potential pelvic inflammatory disease (PID) cases

ID: _____ IndexDate: ____/____/____

Was this a PID episode? *(Select one of A.-D.)*

A. ___**YES**, PID diagnosis
If Yes:

1. Was dxdate (date of PID diagnosis) correct?
   __ Yes, exact date
   __ Yes, but not exact; ≤7 days either side (date: _____)
   __ No, correct dxdate >7 days either side (actual date: _____)

2. What symptoms are noted?
   __ Low abdominal pain __ Uterine/cervical motion tenderness
   __ Adnexal tenderness __ Fever
   __ Painful intercourse
   __ Other (specify)_____ __ No symptoms noted

3. What was the treatment setting/type of treatment?

Treatment setting	Type(s) of Treatment					
	Surgical	Medical (antibiotics)	Surg+Med	Other*	None	Unk.
Inpatient						
Outpatient: Day surgery						
Outpatient: ER						
Outpatient: Urgent care						
Outpatient: Office visit						

* Specify:_____

B. ___ **NO**, not PID
If No, what health event occurred on/near dxdate (what made this look like PID in GH database)?

1. __ Rule out PID
   Diagnosis: _____ Date of diagnosis: _____
2. __ Follow-up of earlier PID episode Date of episode: _____
3. __ Other condition/health event
   Diagnosis: _____ Date of diagnosis: _____

C. ___ **UNCERTAIN**
If Uncertain, supply any information on why this date selected to be PID @ indexdate in GH database
Specify: _____

D. ___ **NO VISIT INFORMATION RECORDED IN CHART ON/NEAR DXDATE**
Specify possible reason (if any) _____