



Published in final edited form as:

Sex Transm Dis. 2024 December 01; 51(12): 794–802. doi:10.1097/OLQ.0000000000002064.

Validating ICD-10-CM Diagnostic Codes with Laboratory Test Results for Use in Identifying Chlamydial and Gonococcal Infections Among American Indians and Alaska Natives: Indian Health Service, 2016–2021

Dana L. Haberling, MSPH^{1,2}, Kerry Mauk, MSPH³, Ethan Bornstein, MD⁴, J. Pekka Nuorti, MD, PhD², Andria Apostolou, PhD, MPH⁵

¹Division of High-Consequence Pathogens and Pathology, Centers for Disease Control and Prevention, Atlanta, GA

²Health Sciences Unit, Faculty of Social Sciences, Tampere University, Finland

³Division of STD Prevention, Centers for Disease Control and Prevention, Atlanta, GA

⁴Navajo Area, Indian Health Service, Chinle, AZ

⁵Division of Epidemiology and Disease Prevention, Indian Health Service, Rockville, MD

Abstract

Background—National case rates of chlamydia and gonorrhea (CT/GC) among American Indian and Alaska Native (AI/AN) persons are disproportionately high. The Indian Health Service (IHS), which provides healthcare to members of federally recognized tribes, does not currently have a dedicated CT/GC surveillance system. The purpose of this study was to validate the use of CT/GC diagnostic codes for estimating diagnosed CT/GC infections among AI/AN persons that use IHS services.

Methods—We conducted a retrospective study using IHS medical records from all persons aged 15 years and older from 2016 to 2021. We linked records with CT (A56, A74) and GC (A54, O98.2) ICD-10-CM diagnostic codes to laboratory results within 30 days for each person. We calculated the sensitivity, specificity, and positive and negative predictive values (PPV, NPV) of CT/GC diagnostic codes using laboratory test results as the reference standard.

Results—We identified over 1.6 million CT/GC laboratory tests, and 52,815 CT and 19,971 GC diagnostic codes. Diagnostic code sensitivity was slightly higher for CT (54%) than GC (50%). Specificity, PPV, and NPV were high for CT and GC (range: 83.3–99.8%). About one-third of CT/GC diagnostic codes could not be linked to a test result.

Conclusions.—The validation indicates that diagnostic codes align well with linked laboratory test results. However, due to the relatively large number of diagnostic codes and positive tests that

Corresponding author: Dana Haberling, 1600 Clifton Rd NE, MS H24-12, Atlanta, GA 30333, Fax: 404-235-0503, dhaberling@cdc.gov.

The authors declare no conflicts of interest. No specific funding was provided from any agency for the preparation of this manuscript.

could not be linked, combining the two would inform more reliable estimates of diagnosed CT/GC infections among AI/AN persons who use IHS for healthcare.

Summary

Diagnostic codes should be combined with laboratory results to estimate diagnosed chlamydial and gonococcal infections among American Indians and Alaska Natives served by the Indian Health Service.

Keywords

chlamydia; gonorrhea; American Indian and Alaska Native; sensitivity; ICD-10-CM

Introduction

American Indian and Alaska Native (AI/AN) persons have had the second highest rates of chlamydia and gonorrhea (CT/GC) among all racial and ethnic groups in the United States since 1996.^{1,2} As of 2021, 8.7 million persons self-reported their race(s) to the US Census as AI/AN³ and over 2.5 million AI/AN members of 574 federally recognized tribes are eligible for care from Indian Health Service (IHS).⁴ Previous reports have estimated CT/GC case rates among AI/AN persons served by the IHS using the laboratory-based US national surveillance data and US Census denominators for counties where IHS provides services.^{5,6} However, this is not all sourced directly from IHS and there is a need for more regular and timely reporting. The use of IHS electronic medical records data could be a convenient source for reporting the burden of CT/GT where there is currently no dedicated IHS CT/GC surveillance program. This information could be utilized for assessing the needs of IHS with regards to identifying outreach focus areas and for projecting needs for testing and treatment, staffing, and budgeting.

IHS is a federal agency that provides healthcare at IHS, Tribal, and Urban Indian healthcare centers (I/T/U) in 37 states and by funding contracted care from other providers.⁴ IHS maintains routinely collected medical records data submitted by these I/T/U sites,⁷ which contain International Classification of Diseases, 10th Revision, Clinical Modification (ICD-10-CM)⁸ diagnostic codes. However, there is a need to assess the validity of using ICD-10-CM diagnostic codes to estimate diagnosed infections from this source. One approach to determine if the appropriate code was recorded is by using other relevant information, such as confirmatory test results. Previous studies combining state surveillance laboratory results and diagnostic codes from private insurance data sources in the United States have found the sensitivity of CT/GC ICD-CM 9th and 10th revision diagnostic codes to be very low at 6–25%, indicating that diagnostic codes are often not captured in administrative data after a positive laboratory test result.^{9,10} The specificity and negative predictive values (NPV) in these studies were generally very high at over 96% each, and positive predictive value (PPV) ranged from 33–87%. They indicated that, while CT/GC ICD-9/10-CM diagnostic codes were found to be very accurate, they may grossly underestimate CT/GC infections when used alone.

The purpose of this study was to validate the use of ICD-10-CM CT/GC diagnostic codes for estimating CT/GC infections in IHS medical records. We estimated the sensitivity, specificity, PPV, and NPV of ICD-10-CM CT/GC diagnostic codes using CT/GC laboratory results as the reference standard. We describe select characteristics of persons with ICD-10-CM CT/GC diagnostic codes with and without linked laboratory results and for laboratory results with and without linked diagnostic codes.

Materials and Methods

The National Data Warehouse (NDW) is the IHS national data repository, which is comprised of medical records data submitted by I/T/U sites. These records contain ICD-10-CM diagnostic codes from care settings including hospitals, outpatient facilities, and emergency departments (ED), and testing results from laboratories.⁷ Demographic information including sex, age, and IHS region (location of residence) are included by date of discharge (hospital) or service (outpatient, ED, laboratory). The IHS regions are an aggregation of the 12 IHS administrative areas: Alaska (Alaska), East (Nashville), Northern Plains East (Bemidji), Northern Plains West (Great Plains, Billings), Southern Plains (Oklahoma City), Southwest (Albuquerque, Navajo, Phoenix, Tucson), and West (California, Portland).

For this analysis, we examined de-identified medical records from the NDW from all persons aged 15 years and older who had a documented CT/GC laboratory test or ICD-10-CM diagnostic code during calendar years 2016 to 2021. Diagnoses were defined using the ICD-10-CM codes A56 and A74 for chlamydial infections and A54 and O98.2 for gonococcal infections.¹¹ Repeat ICD-10-CM diagnostic codes within 28 days were excluded as some tests within four weeks may provide false-positive results.¹² Laboratory tests, henceforth referred to as tests, were defined by Logical Observation Identifiers Names and Codes (LOINC) as previously described^{10,11} and by text strings containing appropriate labels indicating CT/GC tests when codes were not available. Tests were excluded from validation analyses when results were unknown, including indeterminate results, cancelled tests, results noted elsewhere in file and unavailable, or contained a value that could not be interpreted. Repeat CT/GC test results within 28 days were excluded, as test of cure is not usually recommended before four weeks;¹² the following hierarchy was used to select a result value from multiple results entered on the same date: positive, negative, unknown.

We attempted to “link” all ICD-10-CM diagnostic codes to a laboratory test result for the same person, where a laboratory test precedes an ICD-10-CM diagnostic code by 30 days or less. Diagnostic codes and tests were otherwise considered not linked. Tests 30 days before the beginning of 2016 were included, except where an ICD-10-CM diagnostic code from the same time period could be linked. The tests that could not be linked to a CT/GC diagnostic code were further examined by looking for any other hospital, outpatient facility, or ED visits within 30 days of the test for each person. This allowed us to see if any of the tests without linked CT/GC diagnoses could be linked to other with ICD-10-CM encounter codes⁸ used to indicate contact with health services that might be related to CT/GC testing.

We estimated the sensitivity, specificity, PPV, and NPV of the linked ICD-10-CM CT/GC diagnostic codes using positive or negative laboratory test results as the reference standard. Sensitivity of ICD-10-CM diagnostic codes was defined as the number of ICD-10-CM diagnostic codes linked to a positive test in the previous 30 days or less divided by the number of all positive tests and expressed as a percentage. Specificity is the number of negative tests without a linked diagnostic code divided by the number of all negative tests; PPV is the number of diagnostic codes that had a linked positive test divided by the number of diagnostic codes that had a linked positive or negative test; NPV is the number of negative tests without a linked diagnostic code divided by the number of positive and negative tests without a linked diagnostic code; and all expressed as percentages. Estimates were further stratified by age, sex, region, care setting, and calendar year, and 95% binomial confidence intervals (CI) were calculated. Differences in percentages by group were presented with 95% Wald asymptotic confidence intervals. Cochran-Armitage test for trend was used to assess association with linked status and age groups with $p < 0.05$ considered statistically significant. Small cell counts ($n < 5$) and surrounding cells were suppressed to avoid revealing exact small cell values. All analyses were completed using SAS[®] 9.4M6 (SAS Institute, Cary, NC). This activity was reviewed by CDC, deemed not research, and was conducted consistent with applicable federal law and CDC policy. Authors have no conflicts of interest to report. The findings and conclusions in this manuscript are those of the authors and do not necessarily represent the views of the funding agencies.

Results

During the study period of 2016 to 2021, 1.2 million AI/AN persons aged 15 years and older used healthcare services within the IHS system. An average of 673,000 persons had at least one laboratory test of any kind recorded in the NDW during 2.3 million hospital, outpatient, or ED visits each year.

We identified nearly 53,000 ICD-10-CM CT diagnoses, 20,000 ICD-10-CM GC diagnoses, and 820,000 tests each for CT/GC (Tables 1, 2). Nearly all CT ($n=784,747$; 95.7%) and GC ($n=776,561$; 94.8%) tests had a positive or negative result available. The majority of CT/GC diagnoses were among females and persons aged 15–39 years, and the largest number of CT/GC diagnostic codes were recorded in the Southwest IHS region (Tables 1, 2). Most ICD-10-CM diagnoses originated from an outpatient setting (CT: 95%, GC: 91%), and nearly all CT/GC tests (98%) were ordered in an outpatient setting.

The annual numbers of positive CT tests (minimum: 7,943 in 2020; maximum: 10,258 in 2019) were similar to, or slightly greater than, the numbers of ICD-10-CM CT diagnoses over time (minimum: 6,935 in 2020; maximum: 9,628 in 2019; Figure 1A). The annual percentage of positive CT tests among all CT tests remained steady (range: 6.6–7.1%). The numbers of positive GC tests (minimum: 2,597 in 2016; maximum: 4,505 in 2019) were similar to, or slightly greater than, the numbers of ICD-10-CM GC diagnoses over time (minimum: 2,703 in 2016; maximum: 3,906 in 2019; Figure 1B). The annual percentage of positive GC tests among all GC tests increased slightly over time (2.0% in 2016 and 3.6% in 2021).

The validation analyses used all the positive and negative CT/GC tests and 69% of the CT (Table 1) and 63% of the GC (Table 2) ICD-10-CM diagnoses, which were linked to a positive or negative test. Most linked ICD-10-CM diagnoses and tests occurred within 14 days (CT: 84%; GC 92%). Most tests linked to diagnostic codes were positive and most tests not linked to diagnostic codes were negative (Tables 3, 4).

The overall sensitivity of ICD-10-CM diagnoses was slightly higher for CT (54%) than GC (50%; Tables 3, 4). Sensitivity varied by characteristic for both CT and GC: higher among females than males; highest among the youngest age group and decreased with age. CT sensitivity was highest in the Southwest and lowest in the West regions, and GC sensitivity was highest in the Northern Plains East and lowest in the West regions (Tables 3, 4). Sensitivity decreased over time for CT (60.3%, 95% CI: 59.3–61.3 in 2016 and 49.0%, 95% CI: 48.0–50.1 in 2021) and GC (55.7%, 95% CI: 53.8–57.6 in 2016 and 49.2%, 95% CI: 47.7–50.7 in 2021; Figure 1). Specificity was > 99% for both CT and GC, and there was little variability by sex, age, or region. PPV was high, but slightly lower for CT (83%) than GC (91%; Tables 3, 4). For CT, PPV was similar between sexes; highest among the youngest age group; and highest in the Northern Plains West and lowest in the Southern Plains (Table 3). For GC, PPV was slightly higher among females; and highest among the youngest age group (Table 4). NPV was very high, but slightly lower for CT (97%) than GC (99%) and was slightly higher for females than males for both (Tables 3, 4).

Nearly one-third of CT/GC ICD-10-CM diagnoses were not linked to a test and could not be used in the validation analyses (Tables 1, 2). Males had a slightly larger proportion of ICD-10-CM diagnoses not linked to a test when compared to females for both CT (difference=6.8%, 95% CI: 5.8–7.8) and GC (difference=3.0%, 95% CI: 1.5–4.4) and the proportion of ICD-10-CM diagnoses not linked to a test increased with increasing age group for both CT/GC ($p < 0.001$; Tables 1,2). The Alaska region accounted for 42% of both CT/GC ICD-10-CM diagnoses not linked to a test (Tables 1, 2). On average, the Alaska region reported 65.0% (95% CI: 64.1–65.9) more CT and 64.7% (95% CI: 63.4–66.0) more GC ICD-10-CM diagnoses not linked to tests than the other regions (Tables 1,2).

Among all tests, 3% of CT tests and 1% of GC tests were positive and not linked to ICD-10-CM CT/GC diagnoses (Tables 1–4). Nearly all of these positive CT tests (91.3%; $n=23,088$) and positive GC tests (93.9%; $n=10,703$) not linked to CT/GC diagnosis codes could be linked to hospital, outpatient, or ED visits with other ICD-10-CM diagnostic codes. Among these, encounter codes were linked to 83.5% of CT and 80.1% of GC tests; the most common codes included encounter for issue of a repeat prescription (Z76.0), encounter for other administrative examinations (Z02.89), encounter for administrative examinations (Z02.9), and encounter for screening for infections with a predominantly sexual mode of transmission (Z11.3).

Discussion

The results of this study indicate that ICD-10-CM diagnostic codes should be combined with positive laboratory test counts to inform more reliable estimates of diagnosed CT/GC infections among AI/AN persons using IHS healthcare services. The high PPV and very

high specificity results indicate that ICD-10-CM codes align well with laboratory test results: most diagnostic codes that could be linked have positive tests and very rarely do negative tests have linked diagnostic codes. The high NPV shows that where no CT/GC diagnostic codes were linked to positive or negative tests, the tests were mostly negative, as expected. However, observed sensitivities show that CT/GC diagnostic codes were only linked to about half of the positive test results. We did find that the overall absolute number of CT/GC ICD-10-CM diagnostic codes and positive tests was similar, but after linking the two for the validation analysis, about a third of CT/GC diagnostic codes were not linked to any test results. Therefore, these validation results alone are also not enough to understand the true burden of CT/GC.

High specificity, PPV, and NPV values for CT/GC ICD-10-CM diagnoses in the NDW data were similar to previous reports.^{9,10} The sensitivity level for CT/GC ICD-10-CM diagnoses was much higher than previous studies that combined state and private insurance data sources,^{9,10} which may be a result of IHS clinical and laboratory records coming from a singular administrative data system with the capability to easily link person-based records. The highest sensitivity values for both CT/GC diagnostic codes were among those recommended for routine screening including females and younger age groups,¹³ and who may interface more with healthcare visits that encourage routine screening for CT/GC infections. While the sensitivity of ICD-10-CM CT/GC diagnostic codes did decrease over time, the number of tests reported was generally increasing until 2020 and could continue to increase in parallel with rising cases in the US general population, any increases in infections within IHS, increases in screening rates, or due to increased availability of local laboratory testing,¹⁴ therefore, validation of CT/GC ICD-10-CM diagnostic codes should be intermittently assessed to measure stability of these estimates. There were larger differences in sensitivity between regions for both CT/GC, which is likely related to varying medical record keeping practices or other data reporting practices from different facilities, so we recommend that this be considered when estimating CT/GC infections by region in future analyses.

A major strength of the IHS NDW data, that may not be available in other administrative data sources, is that person-records can easily be linked across multiple data sources and over time. This was necessary for linking test results to ICD-10-CM diagnoses and for removing duplicate diagnostic codes or tests in a defined period for a specific person. Unlike previous studies,^{9,10} the IHS NDW could be used to enumerate and describe all CT/GC ICD-10-CM diagnoses that were unable to be validated because they could not be linked to tests. We were also able to link most positive tests without a corresponding CT/GC diagnostic code to other hospital, outpatient, or ED records; almost all were linked to other visits, and many contained encounter ICD-10-CM codes. Some of these encounter codes may be related to screening for sexually transmitted infections and could be considered for a similar validation study in the future.

Consistent with US general population surveillance and guidance,^{1,2} females and younger age groups had the greatest absolute number of CT/GC diagnoses and tests, regardless of linked status. Females may have more frequent healthcare contact in the younger age groups seeking services for contraception or prenatal care. Males had slightly higher percentage

of diagnostic codes without linked tests, which could be because they have symptomatic infections more often and may be receiving presumptive treatment at higher levels.¹³ The Southwest region had the highest levels of CT/GC testing and diagnosing, which is at least partially attributed to population size (approximately 32% of those aged ≥ 15 years reside in the Southwest). The distribution of IHS CT/GC diagnostic codes, regardless of linked test status, aligns with previous reporting of US laboratory-based case surveillance in counties where IHS provides healthcare services, including cases primarily among females and younger age groups, and higher CT/GC case counts in the Alaska, Northern Plains West and Southwest regions,^{5,6} however, counts of diagnostic codes or positive tests cannot be directly compared until there is a more current surveillance report available.

The greatest challenge in this study was capturing tests and test results from the NDW and would be difficult to replicate frequently without more resources. Laboratory test names were obtained using a combination of LOINC codes and systematically searching editable text fields, which is time and labor intensive and requires specialized skills. Further, lack of uniform labeling introduces potential for misclassifying results, which is most likely to lead to underreporting of test results. We recommend an increased use of LOINC codes for laboratory documentation or lists of standardized laboratory test names with limited editing capabilities to simplify analyses for regular use. The relatively small percent of ICD-10-CM diagnoses in the Alaska region that could be linked to a test result indicates there may be systematic differences in reporting laboratory data to IHS NDW by region; this could be related to data reporting practices, such as where LOINC or other laboratory data is not required to be reported by IHS tribal facilities using non-HL7 file formats.⁴ Generally, missing linked tests may be a result of: empiric treatment based on clinical characteristics and known risk factors; patient follow up after expedited partner therapy for sexual partners; or if treatment is given within IHS after outside testing. The laboratory data was otherwise found to be very complete for CT/GC based on the relatively small number of tests with unknown results but, given the potential value and ease of using ICD-10-CM diagnostic codes for CT/GC surveillance, we encourage more robust case management of positive lab results that refer patients to follow-up appointments for further evaluation or treatment response assessment that would also improve the ICD-10-CM sensitivity.

We defined several assumptions for this validation analysis. We preferentially included positive test results if multiple test results were found for the same person, which risks including a small number of false positive test results in our analysis. We chose a 30-day cutoff between test results and diagnoses to represent a reasonable amount of time between test results and recording diagnoses in a real-world clinical setting. While it is possible that some CT/GC reporting of diagnostic codes occurred after the 30-day cutoff, which would likely artificially decrease our sensitivity, we found that most diagnoses were documented within 14 days of testing. Additionally, we used a 28-day window for exclusion of duplicate test results that precludes description of reinfection and test-of-cure diagnoses during the study period. However, true positives occurring in this 28-day window are difficult to identify without further information and could lead to misclassification if included. It is possible that we may not have identified all CT/GC-related visits and tests for our study population if someone obtained CT/GC-related testing or follow-up services outside the IHS system which may not be captured in the NDW. Our recommendation to combine diagnostic

codes with laboratory results assumes that codes without links to tests are not fundamentally different than those with links and does not account for type of laboratory test, type of specimen tested, or imperfect testing results in the validation results.

Future work includes describing and estimating the burden of CT/GC infections among AI/AN persons who use IHS healthcare services to provide a more accurate assessment of needs for testing, treatment, outreach, staffing and funding that support of AI/AN persons who use IHS healthcare services. The results of this study should be considered when estimating diagnosed CT/GC infections, for which we suggest a more sophisticated approach, such as a capture re-capture analysis for the same time period. This type of analysis would combine the two data sources to estimate the CT/GC burden, which in turn will also provide a sense of the difference in magnitude between estimated burden and use of IHS ICD-10-CM diagnoses alone or the laboratory-based national surveillance estimates for counties where IHS provides service. Next, ICD-10-CM diagnostic codes can be used to create a cohort to describe CT/GC risk factors, co-diagnoses and sequelae, trends over time, and treatment.

This is the first study to validate the use of ICD-10-CM CT/GC diagnostic codes for estimating diagnosed CT/GC infections among AI/AN persons utilizing the IHS healthcare network. While we found diagnostic codes aligned well with test results, we also discovered that the use of diagnostic codes alone would underestimate the true burden of CT/GC. However, IHS epidemiologists could consider using diagnostic codes alone to describe diagnosed CT/GC infections, particularly in resource-limited circumstances where laboratory data are unavailable or difficult to analyze.

References

- Centers for Disease Control and Prevention. Sexually Transmitted Disease Surveillance, 2021. Published April 11, 2023. Available from: <https://www.cdc.gov/std/statistics/2021/default.htm> Accessed August 11, 2023.
- Centers for Disease Control and Prevention (CDC) NCHHSTP. AtlasPlus. Available from: <https://gis.cdc.gov/grasp/nchhstpatlas/tables.html> Accessed August 17, 2023.
- U.S. Census Bureau. 2017–2021 American Community Survey 5-Year Estimates. Published 2021. Available from: <https://data.census.gov/table?t=Two+or+More+Races&y=2021> Accessed August 17, 2023.
- Indian Health Service. IHS Profile | Fact Sheets. Newsroom. Available from: <https://www.ihs.gov/newsroom/factsheets/ihsprofile/> Accessed February 3, 2022.
- US Department of Health and Human Services. 2015 Indian Health Surveillance Report - Sexually Transmitted Diseases.; 2018. Available from: https://www.cdc.gov/std/stats/ihs/18ihs-dedp102_report_std_m_508.pdf Accessed January 8, 2024.
- Walker FJ, Llata E, Doshani M, et al. HIV, Chlamydia, Gonorrhea, and Primary and Secondary Syphilis among American Indians and Alaska Natives Within Indian Health Service Areas in the United States, 2007–2010. *J Community Health*. 2015;40(3):484–492. [PubMed: 25371109]
- National Patient Information Reporting System | Indian Health Service (IHS). National Patient Information Reporting System. Available from: <https://www.ihs.gov/npirs/> Accessed April 7, 2022.
- ICD-10-CM Guidelines FY2024. Available from: <https://www.cms.gov/files/document/fy-2024-icd-10-cm-coding-guidelines.pdf> Accessed October 20, 2023.
- Ho YA, Rahurkar S, Tao G, et al. Validation of International Classification of Diseases, Tenth Revision, Clinical Modification Codes for Identifying Cases of Chlamydia and Gonorrhea. *Sex Transm Dis*. 2021;48(5):335–340. [PubMed: 32740450]

10. Mauk KC, Torrone EA, Flagg EW. Can Diagnostic Codes in Health Care Claims Data Identify Confirmed Chlamydial and Gonococcal Infections? A Retrospective Cohort Study, 2003 to 2017. *Sex Transm Dis.* 2021;48(8S):S26–S31. [PubMed: 33967233]
11. Indian Health Service. IHS Clinical Reporting System.; 2023. Available from: https://www.ihs.gov/sites/crs/themes/responsive2017/display_objects/documents/crsv23/SelectedMeasuresV230.pdf Accessed August 18, 2023.
12. Centers for Disease Control and Prevention. Chlamydial Infections - STI Treatment Guidelines. Published August 15, 2022. Available from: <https://www.cdc.gov/std/treatment-guidelines/chlamydia.htm> Accessed January 13, 2023.
13. Workowski KA, Bachmann LH, Chan PA, et al. Sexually Transmitted Infections Treatment Guidelines, 2021. *MMWR Recomm Rep.* 2021;70.
14. United States Government Accountability Office. Indian Health Service Relief Funding and Agency Response to COVID-19 Pandemic.; 2022:67. Available from: <https://www.gao.gov/assets/gao-22-104360.pdf> Accessed November 7, 2023.

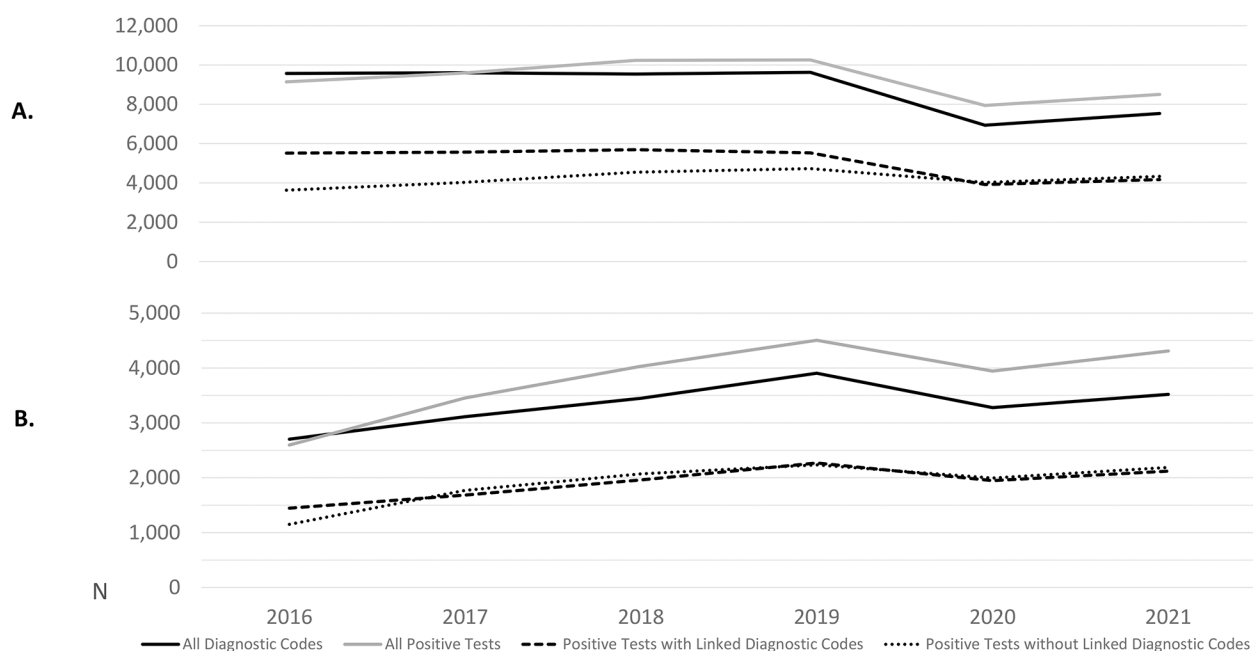


Figure 1.
Annual Chlamydia (A) and Gonorrhea (B) ICD-10-CM Diagnostic Codes and Testing
Counts: Indian Health Service, 2016–2021

Table 1. Chlamydia ICD-10-CM Diagnoses Overall and by Linked Test Status and Tests Overall with Results, IHS 2016–2021

Characteristic	ICD-10-CM Diagnoses									
	Linked Test Result					Tests				
	Total N (%)	Positive or Negative N (%)	Unknown* N (%)	Not Linked N (%)	Total N (%)	Positive N (%)	Negative N (%)	Unknown* N (%)	Positive N (%)	Negative N (%)
Total	52,815 (100.0)	36,467 (69.0)	1,031 (2.0)	15,317 (29.0)	819,582 (100.0)	55,974 (100.0)	728,773 (88.9)	34,835 (4.3)	55,974 (100.0)	728,773 (88.9)
Sex										
Female	41,499 (78.6)	29,240 (70.5)	826 (2.0)	11,433 (27.6)	646,436 (78.9)	42,144 (75.3)	576,015 (89.1)	28,277 (4.4)	42,144 (75.3)	576,015 (89.1)
Male	11,316 (21.4)	7,227 (63.9)	205 (1.8)	3,884 (34.3)	173,142 (21.1)	13,830 (24.7)	152,758 (88.2)	6,554 (3.8)	13,830 (24.7)	152,758 (88.2)
Age group (y)										
15–19	12,019 (22.8)	8,501 (70.7)	310 (2.6)	3,208 (26.7)	107,656 (13.1)	12,688 (22.7)	90,630 (84.2)	4,338 (4.0)	12,688 (22.7)	90,630 (84.2)
20–24	16,299 (30.9)	11,375 (69.8)	311 (1.9)	4,613 (28.3)	157,980 (19.3)	17,344 (31.0)	134,144 (84.9)	6,492 (4.1)	17,344 (31.0)	134,144 (84.9)
25–29	11,064 (21.0)	7,673 (69.4)	185 (1.7)	3,206 (29.0)	156,682 (19.1)	11,803 (21.1)	138,674 (88.5)	6,205 (4.0)	11,803 (21.1)	138,674 (88.5)
30–34	6,686 (12.7)	4,535 (67.8)	113 (1.7)	2,038 (30.5)	121,740 (14.9)	6,964 (12.4)	109,810 (90.2)	4,966 (4.1)	6,964 (12.4)	109,810 (90.2)
35–39	3,292 (6.2)	2,199 (66.8)	49 (1.5)	1,044 (31.7)	86,881 (10.6)	3,647 (6.5)	79,719 (91.8)	3,515 (4.0)	3,647 (6.5)	79,719 (91.8)
40–44	1,649 (3.1)	1,080 (65.5)	31 (1.9)	538 (32.6)	59,074 (7.2)	1,750 (3.1)	54,898 (92.9)	2,426 (4.1)	1,750 (3.1)	54,898 (92.9)
45–54	1,342 (2.5)	833 (62.1)	23 (1.7)	486 (36.2)	77,488 (9.5)	1,346 (2.4)	72,502 (93.6)	3,640 (4.7)	1,346 (2.4)	72,502 (93.6)
55–64	398 (0.8)	237 (59.5)	--**	--**	40,750 (5.0)	374 (0.7)	38,010 (93.3)	2,366 (5.8)	374 (0.7)	38,010 (93.3)
65+	66 (0.1)	34 (51.5)	--**	--**	11,331 (1.4)	58 (0.1)	10,386 (91.7)	887 (7.8)	58 (0.1)	10,386 (91.7)
Regions										
Alaska	7,548 (14.3)	1,113 (14.7)	39 (0.5)	6,396 (84.7)	26,207 (3.2)	2,047 (3.7)	22,458 (85.7)	1,702 (6.5)	2,047 (3.7)	22,458 (85.7)
East	1,226 (2.3)	962 (78.5)	9 (0.7)	255 (20.8)	29,035 (3.5)	1,567 (2.8)	26,881 (92.6)	587 (2.0)	1,567 (2.8)	26,881 (92.6)
Northern Plains East	1,382 (2.6)	716 (51.8)	334 (24.2)	332 (24.0)	23,963 (2.9)	1,193 (2.1)	13,866 (57.9)	8,904 (37.2)	1,193 (2.1)	13,866 (57.9)
Northern Plains West	9,946 (18.8)	7,609 (76.5)	395 (4.0)	1,942 (19.5)	119,393 (14.6)	12,758 (22.8)	101,002 (84.6)	5,633 (4.7)	12,758 (22.8)	101,002 (84.6)
Southern Plains	7,144 (13.5)	5,256 (73.6)	35 (0.5)	1,853 (25.9)	153,989 (18.8)	8,162 (14.6)	135,957 (88.3)	9,870 (6.4)	8,162 (14.6)	135,957 (88.3)
Southwest	22,204 (42.0)	18,219 (82.1)	159 (0.7)	3,826 (17.2)	367,134 (44.8)	24,811 (44.3)	337,939 (92.0)	4,384 (1.2)	24,811 (44.3)	337,939 (92.0)
West	3,365 (6.4)	2,592 (77.0)	60 (1.8)	713 (21.2)	99,861 (12.2)	5,436 (9.7)	90,670 (90.8)	3,755 (3.8)	5,436 (9.7)	90,670 (90.8)

Abbreviations: International Classification of Diseases, 10th Revision, Clinical Modification (ICD-10-CM)

* Results were unknown, including indeterminate results, cancelled tests, results noted elsewhere in file and unavailable, or contained a value that could not be interpreted.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Numbers suppressed due to small counts (n<5) for „unknown“, linked test results among those aged 65+ years.
**

Table 2.
Gonorrhea ICD-10-CM Diagnoses Overall and by Linked Test Status and Tests Overall with Results, IHS 2016–2021

Characteristic	ICD-10-CM Diagnoses									
	Linked Test Result					Tests				
	Total N (%)	Positive or Negative N (%)	Unknown* N (%)	Not Linked N (%)	Total N (%)	Positive N (%)	Negative N (%)	Unknown* N (%)		
Total	19,971 (100.0)	12,629 (63.2)	635 (3.2)	6,707 (33.6)	819,542 (100.0)	22,936 (100.0)	753,625 (92.0)	42,981 (5.2)		
Sex										
Female	13,962 (69.9)	8,959 (64.2)	438 (3.1)	4,565 (32.7)	647,244 (79.0)	13,975 (60.9)	598,302 (92.4)	34,967 (5.4)		
Male	6,009 (30.1)	3,670 (61.1)	197 (3.3)	2,142 (35.6)	172,293 (21.0)	8,961 (39.1)	155,323 (90.2)	8,009 (4.6)		
Age group (y)										
15–19	2,571 (12.9)	1,701 (66.2)	100 (3.9)	770 (29.9)	107,527 (13.1)	3,042 (13.3)	98,514 (91.6)	5,971 (5.6)		
20–24	4,610 (23.1)	2,930 (63.6)	165 (3.6)	1,515 (32.9)	157,330 (19.2)	5,420 (23.6)	143,533 (91.2)	8,377 (5.3)		
25–29	4,618 (23.1)	2,940 (63.7)	152 (3.3)	1,526 (33.0)	156,466 (19.1)	5,345 (23.3)	143,315 (91.6)	7,806 (5.0)		
30–34	3,513 (17.6)	2,162 (61.5)	112 (3.2)	1,239 (35.3)	121,853 (14.9)	3,925 (17.1)	111,892 (91.8)	6,036 (5.0)		
35–39	2,225 (11.1)	1,417 (63.7)	56 (2.5)	752 (33.8)	87,018 (10.6)	2,498 (10.9)	80,375 (92.4)	4,145 (4.8)		
40–44	1,124 (5.6)	722 (64.2)	26 (2.3)	376 (33.5)	59,155 (7.2)	1,292 (5.6)	55,051 (93.1)	2,812 (4.8)		
45–54	961 (4.8)	570 (59.3)	18 (1.9)	373 (38.8)	77,545 (9.5)	1,041 (4.5)	72,558 (93.6)	3,946 (5.1)		
55–64	286 (1.4)	167 (58.4)	--**	--**	40,890 (5.0)	319 (1.4)	38,007 (92.9)	2,564 (6.3)		
65+	63 (0.3)	20 (31.7)	--**	--**	11,758 (1.4)	54 (0.2)	10,380 (88.3)	1,324 (11.3)		
Regions										
Alaska	3,174 (15.9)	378 (11.9)	--**	--**	30,032 (3.7)	792 (3.5)	27,983 (93.2)	1,257 (4.2)		
East	440 (2.2)	330 (75.0)	--**	--**	28,878 (3.5)	595 (2.6)	27,258 (94.4)	1,025 (3.5)		
Northern Plains East	543 (2.7)	289 (53.2)	40 (7.4)	214 (39.4)	27,702 (3.4)	486 (2.1)	15,847 (57.2)	11,369 (41.0)		
Northern Plains West	4,923 (24.7)	3,494 (71.0)	392 (8.0)	1,037 (21.1)	118,928 (14.5)	6,690 (29.2)	100,593 (84.6)	11,645 (9.8)		
Southern Plains	2,346 (11.8)	1,646 (70.2)	12 (0.5)	688 (29.3)	150,346 (18.4)	3,270 (14.3)	139,742 (92.9)	7,334 (4.9)		
Southwest	7,305 (36.6)	5,753 (78.8)	41 (0.6)	1,511 (20.7)	364,584 (44.5)	9,416 (41.1)	350,507 (96.1)	4,661 (1.3)		
West	1,240 (6.2)	739 (59.6)	147 (11.9)	354 (28.5)	99,072 (12.1)	1,687 (7.4)	91,695 (92.6)	5,690 (5.7)		

Abbreviations: International Classification of Diseases, 10th Revision, Clinical Modification (ICD-10-CM)

* Results were unknown, including indeterminate results, cancelled tests, results noted elsewhere in file and unavailable, or contained a value that could not be interpreted.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Numbers suppressed due to small counts (n<5) for “unknown,” linked test results among those aged 55–64 and 65+ years and in the Alaska and East regions.

**

Table 3:

Chlamydia ICD-10-CM Diagnoses Sensitivity, Specificity, Positive Predictive Value, Negative Predictive Value Using Linked Tests with Positive or Negative Results, IHS 2016–2021

Characteristic	ICD-10-CM Diagnoses Linked to Test Results		Test Results				Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
	Not Linked to ICD-10-CM Diagnosis		Positive	Negative	N (%)					
	Positive	Negative				N (%)				
Total	30,385 (83.3)	6,082 (16.7)	25,589 (3.4)	722,691 (96.6)	54.3 (53.9–54.7)	99.2 (99.1–99.2)	83.3 (82.9–83.7)	96.6 (96.5–96.6)		
Sex										
Female	24,402 (83.5)	4,838 (16.5)	17,742 (3.0)	571,177 (97.0)	57.9 (57.4–58.4)	99.2 (99.1–99.2)	83.5 (83.0–83.9)	97.0 (96.9–97.0)		
Male	5,983 (82.8)	1,244 (17.2)	7,847 (4.9)	151,514 (95.1)	43.3 (42.4–44.1)	99.2 (99.1–99.2)	82.8 (81.9–83.7)	95.1 (95.0–95.2)		
Age group (y)										
15–19	7,308 (86.0)	1,193 (14.0)	5,380 (5.7)	89,437 (94.3)	57.6 (56.7–58.5)	98.7 (98.6–98.8)	86.0 (85.2–86.7)	94.3 (94.2–94.5)		
20–24	9,393 (82.6)	1,982 (17.4)	7,951 (5.7)	132,162 (94.3)	54.2 (53.4–54.9)	98.5 (98.5–98.6)	82.6 (81.9–83.3)	94.3 (94.2–94.4)		
25–29	6,359 (82.9)	1,314 (17.1)	5,444 (3.8)	137,360 (96.2)	53.9 (53.0–54.8)	99.1 (99.0–99.1)	82.9 (82.0–83.7)	96.2 (96.1–96.3)		
30–34	3,727 (82.2)	808 (17.8)	3,237 (2.9)	109,002 (97.1)	53.5 (52.3–54.7)	99.3 (99.2–99.3)	82.2 (81.1–83.3)	97.1 (97.0–97.2)		
35–39	1,847 (84.0)	352 (16.0)	1,800 (2.2)	79,367 (97.8)	50.6 (49.0–52.3)	99.6 (99.5–99.6)	84.0 (82.5–85.5)	97.8 (97.7–97.9)		
40–44	878 (81.3)	202 (18.7)	872 (1.6)	54,696 (98.4)	50.2 (47.8–52.5)	99.6 (99.6–99.7)	81.3 (79.0–83.6)	98.4 (98.3–98.5)		
45–54	658 (79.0)	175 (21.0)	688 (0.9)	72,327 (99.1)	48.9 (46.2–51.6)	99.8 (99.7–99.8)	79.0 (76.2–81.8)	99.1 (99.0–99.1)		
55–64	189 (79.7)	48 (20.3)	185 (0.5)	37,962 (99.5)	50.5 (45.5–55.6)	99.9 (99.8–99.9)	79.7 (74.6–84.9)	99.5 (99.4–99.6)		
65+	26 (76.5)	8 (23.5)	32 (0.3)	10,378 (99.7)	44.8 (32.0–57.6)	99.9 (99.9–100.0)	76.5 (62.2–90.7)	99.7 (99.6–99.8)		
Regions										
Alaska	911 (81.9)	202 (18.1)	1,136 (4.9)	22,256 (95.1)	44.5 (42.4–46.7)	99.1 (99.0–99.2)	81.9 (79.6–84.1)	95.1 (94.9–95.4)		
East	810 (84.2)	152 (15.8)	757 (2.8)	26,729 (97.2)	51.7 (49.2–54.2)	99.4 (99.3–99.5)	84.2 (81.9–86.5)	97.2 (97.1–97.4)		
Northern Plains East	623 (87.0)	93 (13.0)	570 (4.0)	13,773 (96.0)	52.2 (49.4–55.1)	99.3 (99.2–99.5)	87.0 (84.5–89.5)	96.0 (95.7–96.3)		
Northern Plains West	6,766 (88.9)	843 (11.1)	5,992 (5.6)	100,159 (94.4)	53.0 (52.2–53.9)	99.2 (99.1–99.2)	88.9 (88.2–89.6)	94.4 (94.2–94.5)		
Southern Plains	4,020 (76.5)	1,236 (23.5)	4,142 (3.0)	134,721 (97.0)	49.3 (48.2–50.3)	99.1 (99.0–99.1)	76.5 (75.3–77.6)	97.0 (96.9–97.1)		
Southwest	15,113 (83.0)	3,106 (17.0)	9,698 (2.8)	334,833 (97.2)	60.9 (60.3–61.5)	99.1 (99.0–99.1)	83.0 (82.4–83.5)	97.2 (97.1–97.2)		
West	2,142 (82.6)	450 (17.4)	3,294 (3.5)	90,220 (96.5)	39.4 (38.1–40.7)	99.5 (99.5–99.5)	82.6 (81.2–84.1)	96.5 (96.4–96.6)		

Abbreviations: International Classification of Diseases, 10th Revision, Clinical Modification (ICD-10-CM); Positive Predictive Value (PPV); Negative Predictive Value (NPV); Confidence Interval (CI)

Table 4:

Gonorrhea ICD-10-CM Diagnoses Sensitivity, Specificity, Positive Predictive Value, Negative Predictive Value Using Linked Tests with Positive or Negative Results, IHS 2016–2021

Characteristic	ICD-10-CM Diagnoses Linked to Test Results		Test Results Not Linked to ICD-10-CM Diagnosis		Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
	Positive	Negative	Positive	Negative				
	N (%)	N (%)	N (%)	N (%)				
Total	11,433 (90.5)	1,196 (9.5)	11,503 (1.5)	752,429 (98.5)	49.8 (49.2–50.5)	99.8 (99.8–99.9)	90.5 (90.0–91.0)	98.5 (98.5–98.5)
Sex								
Female	8,166 (91.1)	793 (8.9)	5,809 (1.0)	597,509 (99.0)	58.4 (57.6–59.3)	99.9 (99.9–99.9)	91.1 (90.6–91.7)	99.0 (99.0–99.1)
Male	3,267 (89.0)	403 (11.0)	5,694 (3.5)	154,920 (96.5)	36.5 (35.5–37.5)	99.7 (99.7–99.8)	89.0 (88.0–90.0)	96.5 (96.4–96.5)
Age group (y)								
15–19	1,574 (92.5)	127 (7.5)	1,468 (1.5)	98,387 (98.5)	51.7 (50.0–53.5)	99.9 (99.8–99.9)	92.5 (91.3–93.8)	98.5 (98.5–98.6)
20–24	2,637 (90.0)	293 (10.0)	2,783 (1.9)	143,240 (98.1)	48.7 (47.3–50.0)	99.8 (99.8–99.8)	90.0 (88.9–91.1)	98.1 (98.0–98.2)
25–29	2,669 (90.8)	271 (9.2)	2,676 (1.8)	143,044 (98.2)	49.9 (48.6–51.3)	99.8 (99.8–99.8)	90.8 (89.7–91.8)	98.2 (98.1–98.2)
30–34	1,954 (90.4)	208 (9.6)	1,971 (1.7)	111,684 (98.3)	49.8 (48.2–51.3)	99.8 (99.8–99.8)	90.4 (89.1–91.6)	98.3 (98.2–98.3)
35–39	1,269 (89.6)	148 (10.4)	1,229 (1.5)	80,227 (98.5)	50.8 (48.8–52.8)	99.8 (99.8–99.8)	89.6 (88.0–91.1)	98.5 (98.4–98.6)
40–44	662 (91.7)	60 (8.3)	630 (1.1)	54,991 (98.9)	51.2 (48.5–54.0)	99.9 (99.9–99.9)	91.7 (89.7–93.7)	98.9 (98.8–99.0)
45–54	502 (88.1)	68 (11.9)	539 (0.7)	72,490 (99.3)	48.2 (45.2–51.3)	99.9 (99.9–99.9)	88.1 (85.4–90.7)	99.3 (99.2–99.3)
55–64	--*	--*	--*	37,988 (99.6)	46.4 (40.9–51.9)	100.0 (99.9–100.0)	88.6 (83.8–93.4)	99.6 (99.5–99.6)
65+	--*	--*	--*	10,378 (99.7)	33.3 (20.8–45.9)	100.0 (100.0–100.0)	90.0 (76.9–100.0)	99.7 (99.5–99.8)
Regions								
Alaska	342 (90.5)	36 (9.5)	450 (1.6)	27,947 (98.4)	43.2 (39.7–46.6)	99.9 (99.8–99.9)	90.5 (87.5–93.4)	98.4 (98.3–98.6)
East	307 (93.0)	23 (7.0)	288 (1.0)	27,235 (99.0)	51.6 (47.6–55.6)	99.9 (99.9–100.0)	93.0 (90.3–95.8)	99.0 (98.8–99.1)
Northern Plains East	270 (93.4)	19 (6.6)	216 (1.3)	15,828 (98.7)	55.6 (51.1–60.0)	99.9 (99.8–99.9)	93.4 (90.6–96.3)	98.7 (98.5–98.8)
Northern Plains West	3,260 (93.3)	234 (6.7)	3,430 (3.3)	100,359 (96.7)	48.7 (47.5–49.9)	99.8 (99.7–99.8)	93.3 (92.5–94.1)	96.7 (96.6–96.8)
Southern Plains	1,475 (89.6)	171 (10.4)	1,795 (1.3)	139,571 (98.7)	45.1 (43.4–46.8)	99.9 (99.9–99.9)	89.6 (88.1–91.1)	98.7 (98.7–98.8)
Southwest	5,132 (89.2)	621 (10.8)	4,284 (1.2)	349,886 (98.8)	54.5 (53.5–55.5)	99.8 (99.8–99.8)	89.2 (88.4–90.0)	98.8 (98.8–98.8)
West	647 (87.6)	92 (12.4)	1,040 (1.1)	91,603 (98.9)	38.4 (36.0–40.7)	99.9 (99.9–99.9)	87.6 (85.2–89.9)	98.9 (98.8–98.9)

Abbreviations: International Classification of Diseases, 10th Revision, Clinical Modification (ICD-10-CM); Positive Predictive Value (PPV); Negative Predictive Value (NPV); Confidence Interval (CI)

* Numbers suppressed due to small counts (n<5) for “negative” linked test results among those aged 65+ years.