

Use of Administrative Health Care Data for Sexually Transmitted Disease Surveillance

Elaine W. Flagg* and Hillard Weinstock

Centers for Disease Control and Prevention, Atlanta, GA, USA

Objective

To evaluate the potential usefulness of 3 sources of administrative health care data for sexually transmitted disease (STD) surveillance.

Introduction

Surveillance for STD in the United States (US) relies primarily on case reports from clinicians and laboratories and sentinel surveillance; however, nationwide reporting is not required for viral STD and clinical sequelae of STD.

Methods

The MarketScan Commercial Claims and Encounters (CCAЕ) data from Truven Health Analytics contains annual beneficiary enrollment information and inpatient admissions, outpatient encounters, and outpatient pharmacy dispensing claims from approximately 100 employee-sponsored private insurance plans. The 2012 CCAЕ contains 1.1 billion inpatient and outpatient claims records for 53 million employees and their dependents.

The Healthcare Cost and Utilization Project Kids' Inpatient Database (KID) from the Agency for Healthcare Research and Quality is the largest publicly-available all-payer pediatric inpatient database in the US. KID data have been available every 3 years since 1997, and are weighted to provide national estimates. The 2012 KID contains 3.2 million birth and hospitalization discharge records for children and adolescents.

National Medicaid Analytic Extract (MAX) data from the Centers for Medicare and Medicaid Services contain claims records for enrollees in Medicaid, a joint state and federal health insurance program for low-income individuals in the US. MAX data are structured similarly to the CCAЕ. The 2009 MAX contains 2.3 billion inpatient and outpatient claims records for almost 68 million child and adult beneficiaries.

Results

CCAЕ data were used to examine trends in genital wart (GW) prevalence in enrollees aged 10-39 years. For GW, a combination of diagnosis, procedure, and prescription drug codes were utilized to create a case definition based on either (1) diagnosis codes specific for GW; or (2) less-specific diagnosis codes for viral warts or prescription drug codes for ≥ 1 GW medications, combined with either a procedure code for destruction or excision of a genital lesion, or a diagnosis code for a benign genital neoplasm, within 30 days of the viral warts diagnosis or prescription.

KID data were used to estimate national and regional incidence of neonatal herpes simplex virus (nHSV) infection. Laboratory test results are not included in inpatient discharge data and procedure codes specific to nHSV treatment do not exist. The case definition initially used hospital admissions within the first 60 days of life with ≥ 1 diagnosis codes for herpes simplex infection. However, because the recommended course of treatment for nHSV is >7 days, we further restricted cases to patients hospitalized at least that long or who died within 7 days, to help eliminate ruled-out diagnoses.

MAX data were used to estimate incidence of congenital syphilis (CS) in a cohort of infants continuously enrolled during the first year

of life. Procedure codes for syphilis tests exist, but codes specific to treatment of this infection do not. Several other diagnostic procedures may be used, and follow-up tests of cure should be conducted. All continuously enrolled infants with ≥ 1 syphilis diagnosis codes were first identified; possible CS cases were then defined as those with a syphilis test or diagnosis >2 months after the first diagnosis. Probable cases were identified as the subset of possible cases which also had a syphilis test or other diagnostic procedure within 30 days of the first diagnosis.

Strengths of all these sources include availability of diagnosis and procedure codes for large numbers of records. KID data are nationally representative, while MAX data represent the entire population of Medicaid enrollees. All data sources offer standardized data values. However, none of the sources include information on laboratory test results or inpatient medications, although CCAЕ and MAX do contain outpatient prescription claims. Codes may not exist for tests, procedures, or treatments sufficient to construct a plausible case definition for some STD. Access to these data is not timely; currently, KID and CCAЕ are available through 2012 and MAX through 2009. Race/ethnicity information is available in only KID and MAX, and these data are incomplete; 8% of race/ethnicity values are missing for KID 2012 and for MAX 2009.

Conclusions

Administrative health care data provide new opportunities to monitor STD among large numbers of health care consumers. These data may be particularly useful for assessing non-reportable STD and STD clinical sequelae, but their delayed availability may limit their utility for public health response.

Keywords

Sexually transmitted disease (STD); Public health surveillance; Claims data

*Elaine W. Flagg

E-mail: ewf2@cdc.gov

