

Evaluation of Clinical and Administrative Data to Augment Public Health Surveillance

Joseph R. Egger*¹, Brooks Lowe¹, Erica Johnson¹, Randy Durbin², Erica Griffin², Sanjaya Dhakal³, Achintya Dey³ and Sam Groseclose³

¹SciMetrika, LLC, Research Triangle Park, NC, USA; ²The Altarum Institute, Atlanta, GA, USA; ³Centers for Disease Control and Prevention, Atlanta, GA, USA

Objective

To assess the utility of inpatient and ambulatory clinical data compiled by public and commercial sources to enhance the Centers for Disease Control and Prevention's surveillance activities.

Introduction

Medical claims and EHR data sources offer the potential to ascertain disease and health risk behavior prevalence and incidence, evaluate the use of clinical services, and monitor changes related to public health interventions. Passage of the HITECH Act of 2009 supports the availability of standardized EHR data for use by public health officials to obtain actionable information. While full adoption of EHRs is still years away, there are presently publicly- and commercially-available EHR and medical claims data sets that could enhance public health surveillance at a national, regional and state level. The purposes of this evaluation were to i.) demonstrate the feasibility of gaining access to such data, ii.) evaluate their ability to augment current surveillance activities by developing measures for twenty separate healthcare indicators (e.g., HIV screening), iii.) evaluate each data source across a set of criteria needed for an effective surveillance system, and iv.) assess the ability of the data sources to evaluate changes in healthcare utilization and preventive services that may be a result of the 2009 Health Reform legislation.

Methods

Ten separate data sources were selected for inclusion in the study based on a number of criteria, including availability, representativeness, population, data structure and content, cost, and longitudinality. In collaboration with staff from seven Divisions across the CDC, detailed specifications were developed for twenty separate indicators of healthcare utilization or preventive services using best practices in healthcare quality measurement. Specifications were developed separately for EHR and medical claims data due to their differing structure, content and use of medical code sets and terminologies. Specifications for EHR data sources relied on the National Quality Forum (NQF) Meaningful Use (MU) clinical quality measure specifications. The use of NQF MU specification guidelines allowed us to gauge the current ability of each data source to measure healthcare utilization and preventive services as recommended by NQF, the national leader in healthcare measurement. Each of the data sources was also evaluated across established public health surveillance criteria, including data quality, representativeness, and flexibility, among others. Data analysis was performed using SAS 9.3 (SAS Institute, Cary, NC).

Results

All twenty of the healthcare indicators were developed for at least one data source; however, many of the indicator specifications had to be modified due to the low frequency of certain code sets (e.g., CPT-4 II, LOINC). The observed strengths of medical claims data were the relatively low cost, ability to track patients longitudinally, and the standardized representation of procedures and diagnoses through use of medical codes, such as ICD-9-CM, CPT-4 and HCPCS. The observed strengths of EHR data sources were the availability of information related to health behavior (e.g., current smoker), health assessment (e.g., BMI), prognostic indicators (e.g., vital signs, laboratory result), diagnostic testing, and functional status. While EHR data also capture diagnoses using ICD-9-CM, procedures such as medical and laboratory procedures remain documented through use of free text or semi-structured text fields, making it difficult to process.

Conclusions

Currently available healthcare data can improve the timeliness of health outcome monitoring and add complementary information on healthcare utilization to improve our interpretation of traditional public health surveillance data. Medical claims data support measurement of health outcomes and healthcare services provided to patient populations; however, without clinical encounter information, they cannot develop measures estimating the impact of services received on quality of care. EHR data have richer clinical information; however, the continued use of non-standards-based medical codes and free and semi-structured text fields make it difficult to analyze data at scale. Meaningful Use and other HITECH initiatives are changing this by incentivizing the standardization and aggregation of electronic healthcare data. In time, these data may yield timely, accurate and actionable information for public health surveillance.

Keywords

Surveillance; Evaluation; Healthcare; Electronic Health Record

*Joseph R. Egger

E-mail: joseph.egger@gmail.com

