

Building a Diabetes Screening Population Data Repository Using Electronic Medical Records

Wen-Jan Tuan, M.S., M.P.H.,¹ Ann M. Sheehy, M.D., M.S.,²
and Maureen A. Smith, M.D., M.P.H., Ph.D.¹

Abstract

There has been a rapid advancement of information technology in the area of clinical and population health data management since 2000. However, with the fast growth of electronic medical records (EMRs) and the increasing complexity of information systems, it has become challenging for researchers to effectively access, locate, extract, and analyze information critical to their research. This article introduces an outpatient encounter data framework designed to construct an EMR-based population data repository for diabetes screening research. The outpatient encounter data framework is developed on a hybrid data structure of entity–attribute–value models, dimensional models, and relational models. This design preserves a small number of subject-specific tables essential to key clinical constructs in the data repository. It enables atomic information to be maintained in a transparent and meaningful way to researchers and health care practitioners who need to access data and still achieve the same performance level as conventional data warehouse models. A six-layer information processing strategy is developed to extract and transform EMRs to the research data repository. The data structure also complies with both Health Insurance Portability and Accountability Act regulations and the institutional review board's requirements. Although developed for diabetes screening research, the design of the outpatient encounter data framework is suitable for other types of health service research. It may also provide organizations a tool to improve health care quality and efficiency, consistent with the “meaningful use” objectives of the Health Information Technology for Economic and Clinical Health Act.

J Diabetes Sci Technol 2011;5(3):514-522

Author Affiliations: ¹Health Innovation Program, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin; and ²Department of Medicine, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin

Abbreviations: (EMR) electronic medical record, (ETL) extract, transform, and load, (HIPAA) Health Insurance Portability and Accountability Act, (HITECH) Health Information Technology for Economic and Clinical Health, (IRB) institutional review board, (WCHQ) Wisconsin Collaborative for Healthcare Quality

Keywords: diabetes screening, information processing, outpatient research data repository, patient encounter

Corresponding Author: Wen-Jan Tuan, M.S., M.P.H., Health Innovation Program, Room 210-18, 800 University Bay Dr., Madison, WI 53705; email address tuan@wisc.edu