Optimizing Health Information Technology for Clinical Research: Developing an Infectious Disease Research Data Repository (a Multi-Site Electronic Data Infectious Diseases Consortium [MEDIC] study)

Lois Lamerato PhD; Laura A. Puzniak PhD, MPH; Steven Lawrence MD, MSc; Anthony D. Harris MD, MPH; Andrew Masica MD, MSCI; Marcus Zerwas MD

Infectious Diseases, Henry Ford Hospital, Detroit, MI; Pfizer Inc, Collegeville, PA; Washington University, St. Louis, MO; Epidemiology and Public Health, University of Maryland, Baltimore; Baltimore School of Medicine, Baltimore, MD; Baylor Health Care System, Dallas, TX

Abstract

Background: Infectious diseases remain a significant contributor to health care costs and are a leading cause of hospital readmission. Advances in health information technology provide an opportunity to leverage real world clinical data to answer clinically relevant research questions. MEDIC is an industry collaborative between Pfizer and four geographical and patient population diverse academic medical centers. Established in 2012, MEDIC includes researchers from Pfizer, Baylor Health Care System, Henry Ford Health System, and Washington University School of Medicine.

Methods: The participating institutions have leveraged their native information systems to identify a cohort of hospitals with the following disease and treatment characteristics: pneumonia, skin infections, intra-abdominal infections (IA), transplant, positive culture for selected organisms, patients treated with linezolid, tigecycline or pneumococcal vaccine, and all admissions of patients with pneumococcal disease from January 1, 2000 through December 31, 2012. Data include patient demographics, administrative data, ICD-9 CM diagnoses and procedures, CPT4 procedures, Diagnostic Related Group (DRG), selected clinical laboratory results, pharmacy, culture results, selected vital data and vital status changes. Readmissions between 2008 and 2012 were included. Data were de-identified and used in a controlled environment.

Results: Over 550,000 hospitalizations in the targeted health disease states were identified. These include 77,974 pneumonias, 65,121 diabetes, 39,753 UA, 39,753 skin infections, 83,497 specific pathogen-related and 38,065 treatment-related. Due to the comprehensive nature of the database, multiple research questions of interest are being evaluated from this real-world database including comparative effectiveness, disease burden, epidemiology and cost-effectiveness.

Conclusion: The MEDIC platform is an innovative collaborative approach to conducting infectious diseases research based on an aggregated data repository from multiple and diverse academic medical centers. The infrastructure and partnerships developed will contribute significantly to clinical knowledge.

Background and Objectives

• Infectious diseases remain a significant contributor to health care costs and are a leading cause of hospital readmission.
• Advances in health information technology provide an opportunity to leverage real world clinical data to answer clinically relevant research questions.
• Multiple efforts are underway, both in the U.S. and abroad, to build the infrastructure to enable data-sharing among EHR systems and to facilitate the use of EHR data for a variety of research goals.
• Most EHR systems are transactional in nature (and understandably) developed to promote clinical care. Since research is typically a secondary goal of these systems, utilizing the data for clinical research requires considerable investment, both in infrastructure and informatics expertise.
• MEDIC is an example of an industry-healthcare delivery organization collaboration leveraging EHR data for research. Sponsored by Pfizer, the consortium involves four health systems (Baylor, Henry Ford, University of Maryland, Washington University) that represent geographic, practice environment and patient population diversity. A centralized, de-identified database has been constructed utilizing data directly from the four institutions to answer clinical research questions.

Methods

MEDIC, a database repository of real-world clinical data was built to answer multiple research questions across a variety of clinical areas. The database repository includes administrative, pharmaceutical, laboratory, vital, microbiology and cost data for all hospital admissions from 4 academic medical centers between January 1, 2008 through December 31, 2012 who met any of the following inclusion criteria:

- Diagnosis driven: Patient had an ICD-9 code for diagnosis (240 codes) or procedures (143 codes) related to pneumonia, soft tissue infection, IAT, SSI, transplant, history of transplant or diabetes during the study period
- Pharmacology driven: Patient received linezolid, tigecycline or pneumococcal vaccine during the study period
- Pathogen driven: Patient had a positive culture for Staphylococcus aureus, enterococcus spp., staph aureus spp., enterococcus spp., klebsiella spp., R. Agostini, Clodstrum spp during the study period
- Longitudinal Pneumococcal disease: Patient had a pneumococcal vaccine or diagnosis of pneumococcal disease from January 1, 2000 through December 31, 2012
- Once a patient is identified based on the above criteria, any admissions during the study period is also added to the repository in order to evaluate previous exposures and readmissions.

This study was approved by all sites Institutional Review Boards.

Challenges with Developing MEDIC

- Varying data models to implement consistent accessibility and extraction

Site Challenges with Developing MEDIC

- Transforming existing data into standards

Acknowledgements

• This Collaborative Research study was sponsored by Pfizer