Project Development Phase

Code-Layout,Readability And Reusability

Team Id	NM2023TMID04415
Project Name	Block Chain Technology For Electronic Health Records

Introduction:

We describe the formulation, development, and initial expert review of 3x3 Data Quality Assessment (DQA), a dynamic, evidence-based guideline to enable electronic health record (EHR) data quality assessment and reporting for clinical research.
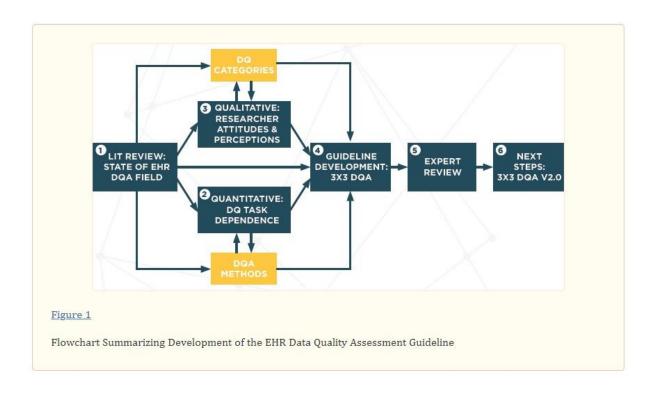
Methods:

3x3 DQA was developed through the triangulation results from three studies: a review of the literature on EHR data quality assessment, a quantitative study of EHR data completeness, and a set of interviews with clinical researchers. Following initial development, the guideline was reviewed by a panel of EHR data quality experts.

Results:

The guideline embraces the task-dependent nature of data quality and data quality assessment. The core framework includes three constructs of data quality: complete, correct, and current data. These constructs are operationalized according to the three primary dimensions of EHR data: patients, variables, and time. Each of the nine operationalized constructs maps to a methodological recommendation for EHR data quality assessment. The initial expert response to the framework was positive, but improvements are required.

Discussion:

The initial version of 3x3 DQA promises to enable explicit guideline-based best practices for EHR data quality assessment and reporting. Future work will focus on increasing clarity on how and when 3x3 DQA should be used during the research process, improving the feasibility and ease-of-use of recommendation execution, and clarifying the process for users to determine which operationalized constructs and recommendations are relevant for a given dataset and study.

Figure 1

Flowchart Summarizing Development of the EHR Data Quality Assessment Guideline

Literature Review: State of the EHR Data Quality Assessment Field:

A systematic review of the informatics literature was used to identify the most common constructs of data quality and methods of data quality assessment. The details of this study can be found in Weiskopf and Weng, 2013 [17]. Our inclusion criteria were that the papers must describe methods of data quality assessment, that they focus on EHR-derived data, and that they were peer-reviewed. A PubMed search was used to identify an initial pool of relevant papers, which were then used as the basis for an ancestor search of their references. From each paper we extracted: 1) the constructs of data quality assessed and 2) the

methods of assessment used. An iterative process was used to consolidate the extracted data into fundamental constructs. We quantified the strength of the relationships between data quality constructs and assessment methods by the frequency as indicated by the publication count, with which each method was used for each construct. These frequencies were used to select and prioritize methodological approaches to assessing each data quality construct during the guideline development.

Quantitative: Data Quality Task-Dependence:

A quantitative approach was used to explore the concept of task-dependence (fitness for use), specifically focusing on the data quality construct of completeness as an exemplar We hypothesized that EHR data completeness can be defined in multiple ways, depending upon intended use, and that, in turn, efforts to calculate rates of records completeness would vary based upon these different definitions and uses. The details of this study can be found in Weiskopf et al., 2013 [4]. We proposed four definitions of completeness.