

# Harmonization of Quality Data Model with HL7 FHIR to Support EHR-driven Phenotype Authoring and Execution: A Pilot Study

Guoqian Jiang<sup>1</sup>, Harold R. Solbrig<sup>1</sup>, Richard Kiefer<sup>1</sup>, Luke V. Rasmussen<sup>2</sup>, Huan Mo<sup>3</sup>, Jennifer A. Pacheco<sup>2</sup>, Enid Montague<sup>2</sup>, Jie Xu<sup>2</sup>, Peter Speltz<sup>3</sup>, William K. Thompson<sup>2</sup>, Joshua C. Denny<sup>3</sup>, Christopher G. Chute<sup>4</sup>, Jyotishman Pathak<sup>1</sup>

<sup>1</sup> Mayo Clinic College of Medicine, Rochester, MN; <sup>2</sup> Northwestern University, Chicago, IL; <sup>3</sup> Vanderbilt University, Nashville, TN; <sup>4</sup> Johns Hopkins University, Baltimore, MD

## Abstract

*The objective of this pilot study is to describe a crowdsourcing effort in harmonizing high-level data elements between Quality Data Model (QDM) and HL7 Fast Healthcare Interoperability Resources (FHIR) to support electronic health records (EHR)-driven phenotype authoring and execution. In total, 194 mapping pairs between the two models were identified, with a Fleiss's kappa statistics ( $k=0.24$ ) calculated for inter-rater agreement. We discuss challenging issues of the mappings.*

## Introduction

In previous studies, we developed a metadata repository and its services for data elements derived from Quality Data Model (QDM) and HL7 Fast Healthcare Interoperability Resources (FHIR) to support electronic health records (EHR)-driven phenotype authoring and execution (1-2). In the present study, we describe a crowdsourcing effort in harmonization of high-level data elements between the two models. The purpose of the harmonization is to understand the domain coverage of the two models and to establish the degree of interoperability between QDM-based phenotype algorithms and patient data populated with FHIR models.

## Methods

We extracted a collection of high-level data elements with their textual definitions from both QDM and HL7 FHIR using our metadata repository services (<http://projectphema.org>). We designed a mapping application using an Excel spreadsheet. We created five worksheets: instructions, mappings, QDM data elements with definitions, FHIR data elements with definitions, and SKOS mapping properties with definitions. One-to-many mappings are allowed by the spreadsheet application. We used five SKOS mapping properties (exactMatch, closeMatch, broadMatch, narrowMatch and relatedMatch) as the mapping types. We asked the project team members to complete the mappings individually and Fleiss' kappa statistics was calculated to assess inter-rater agreement.

## Results and Discussions

In total, 94 data elements from QDM (consisting of 18 QDM Categories and 76 QDM Datatypes) and 98 data elements from FHIR (all FHIR Resources) were extracted for mappings. We received the responses from 7 team members and 206 mapping pairs were created. All QDM data elements had at least one mapping suggested whereas 65 FHIR data elements did not have any mappings. Only fair agreement ( $kappa = 0.24$ ) was achieved. The QDM categories Communication, Condition/Diagnosis/Problem and Encounter were in relatively high agreement. The textual definitions of the data elements, along with multiple factors (e.g., attributes associated with each type, ambiguity in naming within two models), are important for creating correct mappings. In summary, the pilot study provides valuable insight into the challenging issues for a community-based harmonization of QDM and HL7 FHIR. In the future, we plan to extend the harmonization effort in collaboration with broader clinical research communities, aiming at producing a collection of complete and high-quality mappings.

**Acknowledgments:** This work has been supported in part by funding from PhEMA (R01 GM105688), eMERGE (U01 HG006379, U01 HG006378 and U01 HG006388), and caCDE-QA (1U01CA180940-01A1).

## References

1. Thompson WK, Rasmussen LV, Pacheco JA, Peissig PL, Denny JC, Kho AN, Miller A, Pathak J. An evaluation of the NQF Quality Data Model for representing Electronic Health Record driven phenotyping algorithms. AMIA Annu Symp Proc. 2012;2012:911-20.
2. Jiang G, Solbrig HR, Kiefer R, Rasmussen LV, Mo H, Speltz P, Thompson WK, Denny JC, Chute CG, Pathak J. A semantic framework for quality data model to support EHR-driven phenotype authoring and execution. AMIA Jt Summits Transl Sci Proc. 2015 (in press).