# The Integration Challenges in Bridging Patient Care and Clinical Research in a Learning Healthcare System

**Conference Paper** · August 2014

CITATIONS
0

READS
58

**5 authors**, including:

Sarah N. Lim Choi Keung
The University of Warwick
**28** PUBLICATIONS **44** CITATIONS
SEE PROFILE

Jean-François Ethier
Université de Sherbrooke et INSERM
**23** PUBLICATIONS **63** CITATIONS
SEE PROFILE

Lei Zhao
The University of Warwick
**23** PUBLICATIONS **39** CITATIONS
SEE PROFILE

Theodoros Arvanitis
The University of Warwick
**228** PUBLICATIONS **2,200** CITATIONS
SEE PROFILE

# The Integration Challenges in Bridging Patient Care and Clinical Research in a Learning Healthcare System

S.N. LIM CHOI KEUNG[a,1], J.-F. ETHIER[b], L. ZHAO[a], V. CURCIN[c] and T.N. ARVANITIS[a]
[a] *Institute of Digital Healthcare, WMG, University of Warwick, UK;*
[b] *INSERM UMR_S 872, France;*
[c] *Department of Primary Care and Public Health, Imperial College London, UK*

**Abstract** Routinely collected clinical data can be reused in a number of research tasks, including cohort identification and pre-population of electronic case report forms. TRANSFoRm adopts a model-based, mediation approach to address the data integration challenges of such reuse and bridge the semantic gap between clinical and research domains.

**Keywords.** electronic health records, secondary use, integration

## Introduction

The TRANSFoRm project aims to improve patient safety in Europe and increase the volume of clinical research. A dual-layer modelling approach separates the more stable domain information from the technical implementations of heterogeneous systems [1]. We describe the integration challenges in reusing primary care data for pre-populating electronic case report forms (eCRFs) from electronic health records (eHRs).

## 1. Methods

The dual-level modelling organises health information into information models that allow deployment in different technical formats. On Level 1, the Clinical Research Information Model (CRIM) depicts the workflow and data requirements for the clinical research task, while the Clinical Data Integration Model (CDIM) ontology presents a coherent and unified representation of the primary care domain. Level 2 uses archetypes, defined in openEHR Archetype Definition Language (ADL), for the individual specifications to represent the data elements. Those can be newly entered or extracted from eHR systems [1]. CDIM references allow usage of the TRANSFoRm unified interoperability framework [2]. As part of this framework, Data Source Models (DSM) are developed to define the data schema of each data source, and DSM mappings define how the CDIM ontology concepts map to the data source schema. The

[1] Corresponding Author: Dr Sarah N. Lim Choi Keung; Email: s.n.lim-choi-keung@warwick.ac.uk

DSM and its maps are used to translate the archetypes into a query executable at the data source, e.g. to retrieve eCRF elements available from the eHR. Archetypes are embedded into CDISC Operational Data Model (ODM), a standard for clinical research data collection, to ensure reusability of the approach.

## 2. Results and Discussion

Figure 1 shows an example of the transformation process that allows eHR data to be reused in a clinical study. While the DSM definition and maps need to be developed only once per data source, the correct semantic meaning needs to be precisely modelled in close collaboration with the data source team. Supplementary study information can be entered during research visits, without being added to the eHR.

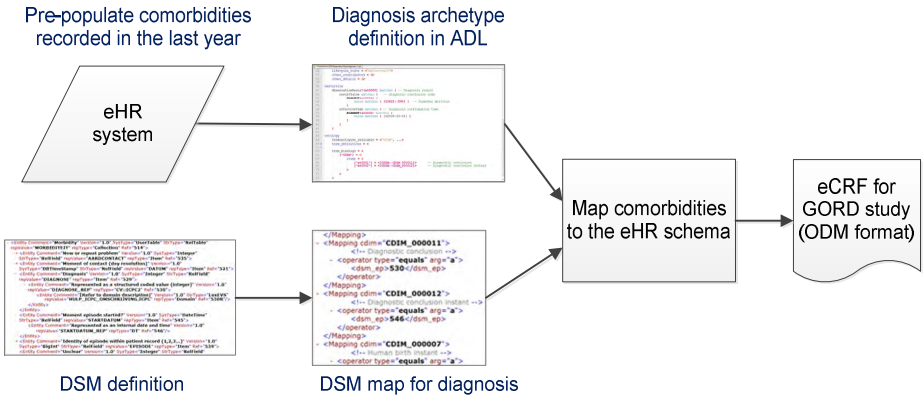


**Figure 1.** Example of pre-population of comorbidities for a gastro-oesophageal reflux disease (GORD) study.

## 3. Conclusions

We have presented a dual-level modelling approach to enable semantic interoperability between clinical research tasks and primary care data sources, that is flexible enough to handle new archetypes and data sources, while keeping the information model stable.


## Acknowledgements

The authors thank all TRANSFoRm project staff for their contribution. The TRANSFoRm project is partially funded by the European Commission under the 7th Framework Programme (Grant Agreement 247787).


## References


[1] S.N. Lim Choi Keung et al. "Detailed Clinical Modelling Approach to Data Extraction from Heterogeneous Data Sources for Clinical Research". 2014 AMIA Clinical Research Informatics Summit, San Francisco, April 2014 (Accepted).

[2] J.-F. Ethier et al. "A unified structural/terminological interoperability framework based on LexEVS: application to TRANSFoRm". J Am Med Inform Assoc. 2013;20(5):986–94.