FP7-TRANSFoRm: Open source tools to support the integration of Randomized Controlled Trials with clinical systems. [EFMI PCIWG Workshop]

Brendan DELANEY, Vasa CURCIN, Theodoros ARVANITIS, and Simon DE LUSIGNAN.

Abstract The Learning Health System (LHS) proposes a full integration of clinical care with knowledge generation and knowledge translation activities via the electronic health record. The FP7 TRANSFoRm project has over five years developed and is in the process of evaluating, an infrastructure to support both research and decision support in primary care. Primary Care research networks in Europe play an essential role in supporting academic departments and in running large-scale multi-centre randomised controlled trials. The requirement in practice settings of support for the processes of the RCT, recruitment, data entry and follow up, and the widespread use of electronic health record (EHR) systems provides the opportunity to develop and test the LHS for RCTs. We will describe the TRANSFoRm approach to using and extending existing standards (CDISC-ODM and SDM) and a model-based approach to interoperability. Tools for capturing data from within the EHR, and for obtaining patient related outcome measures via smartphone, and for capturing provenance for the system will be demonstrated.

Keywords. Electronic Health Records, Randomised Controlled Trials

Introduction

The intended audience for this workshop is clinical researchers and those with an interest in clinical research informatics and the conduct of research using electronic health record systems, not just use secondary use of data. The ‘Learning Healthcare System’ (LHS) describes the construction of a healthcare infrastructure that explicitly places support of research and knowledge translation at its heart. The Institute of Medicine in the USA and the EU have both produced reports over the past few years supporting this concept. The FP7-TRANSFoRm project has over the past 5 years developed a digital infrastructure to support a LHS in European Primary Care. We have developed a model-based system to 1. Enable simultaneous queries and data extraction from multiple heterogeneous clinical and genotype datasets. 2. Integrate the requirements of a randomised controlled trial (recruitment, consent, data collection, follow up) into routine health record systems. And 3. Develop a decision support system for diagnosis in primary care.
This workshop will explore both the informatics advances developed by TRANSFoRm, building on ontologies, information models and existing standards, as well as provide some hands-on experience with the TRANSFoRm tools running embedded within the In Practice Systems Vision3 health record system and patient related outcome measures (PROMS) being collected by web and smartphone.

1. **Aim of the discussion**

   The aim of the discussion will be to:
   1. Consider the concept of the learning Healthcare System as a core component of promoting research activity in community-based clinics and general practices.
   2. Explore the informatics issues of integrating not just data but functionality across research and clinical care and across heterogeneous systems, with an emphasis on standards and on models.
   3. Obtain hands on experience of the TRANSFoRm RCT tools integrated with an EHR system. (A Technology Assessment Model questionnaire will be used during the workshop)

2. **Contribution from each speaker (15 mins each)**

   **Brendan Delaney – Introduction to the TRANSFoRm Project:**
   TRANSFoRm has developed a set of methods, models and tools that facilitate the learning healthcare system. TRANSFoRm has brought together a highly multidisciplinary consortium where three carefully chosen clinical ‘use cases’ were used to develop, evaluate and validate the approach to the ICT challenges. For RCTs we have designed a full practice-based study of continuous on-demand use of Proton Pump Inhibitors in Gastro-oesophageal Reflux Disease. The study tests recruitment via the EHR data, completion and pre-population of an electronic case report form (eCRF), with data stored in the EHR as e-Source, and PROMs collected via mobile devices. The project builds on existing work at international level in clinical trial information models service-based approaches to semantic interoperability and data standards data discovery, machine learning and electronic health records based on open standards.

   **Simon de Lusignan – Identifying Primary Care Research Networks in Europe**
   There is no standard way clinical data are recorded across Europe; though there are many sentinel and other networks looking to make use of clinical data, and things have changed little in the last decade. The TRANSFoRm International Research Readiness (TIRRE) survey aimed to identify data use through two linked data studies and by identifying clinical data repositories and genetic databases or disease registries prepared to participate in linked research. The survey had 56 valid responses, or which 29 were databases of primary care data, 12 were genetic databases, and 15 were cancer
registries. Projects such as TRANSFORM are needed because of the lack of readily interoperable data across Europe.

Theodoros Arvanitis - ODM and Archetypes

In TRANSFoRm, the reuse of routinely collected clinical data is achieved by pre-populating eCRFs with available data directly from EHR systems (e.g. pre-population of eCRFs in the GORD RCT study). To achieve this, clinical data needs to be semantically interoperable between the eCRF and the EHR system. We have adopted a two-level modelling approach to make this possible. On the first level, the Clinical Research Information Model (CRIM) provides an information model to depict the processes of clinical research (e.g. RCT workflows). CRIM is then combined with the Clinical Data Integration Model (CDIM), an ontology of clinical primary care domain, to capture the structural and semantic variability of data representations across data sources. At the second level, archetypes are then used to constrain the domain concepts provided by CRIM/CDIM and specify implementation details of data elements within EHR systems. To maintain the semantic meaning when using eCRF pre-population, the CDISC ODM standard has been extended to allow for archetypes to further define eCRF elements and the way associated pre-population queries will extract specific clinical data items from EHR systems.

Vasa Curcin – Provenance and EHR data.
Validation challenge is a growing problem in the scientific community, which is struggling to ensure correctness of published research. Provenance technologies establish automatic record of what happened in a certain task, e.g. connecting data with the process that led to the data, and in accordance with some specified domain information model. The challenge is to introduce provenance support in all heterogeneous software tools that participate in the LHS. To that end, TRANSFoRm developed graph templates that minimise the effort required by the tools to submit provenance information to the Provenance Server. Using a well-defined web-service interface, tools, such as the TRANSFoRm Study system and the Diagnostic Support plug-in, instantiate these templates and store resulting information in an RDF database, where it can be queried and analysed. Based on our experiences, we shall discuss the benefits and adoption issues of this technology and how it contributes to the vision of reproducible research.

Expected results

Firstly we will explore the potential for carrying out large-scale research in European Primary Care Research Networks using the TRANSFoRm tools. The workshop will be an important component of TRANSFoRm’s end of project dissemination and exploitation activities. The TRANSFoRm tools will be available as open-source via the European Clinical Research Infrastructures Network and the proposed European Institute for Health Technology Innovation. Secondly we will be conducting a TAM assessment with participants, who will have opportunity to explore the TRANSFoRm tools. This will be written up for submission to a journal.