Learning Health Systems

**TECHNICAL REPORT**

# Requirements and validation of a prototype learning health system for clinical diagnosis

Derek Corrigan[1] | Gary Munnelly[2] | Przemysław Kazienko[3] | Tomasz Kajdanowicz[3] | Jean-Karl Soler[4] | Samhar Mahmoud[5] | Talya Porat[5] | Olga Kostopoulou[6] | Vasa Curcin[5] | Brendan Delaney[6] (ID)

[1] Royal College of Surgeons in Ireland, Dublin, Ireland

[2] Trinity College Dublin, Dublin, Ireland

[3] Wrocław University of Science and Technology, Wrocław, Poland

[4] Mediterranean Institute for Primary Care, Malta

[5] Kings College London, London, UK

[6] Imperial College London, London, UK

**Correspondence**
Derek Corrigan, Royal College of Surgeons in Ireland, Dublin, Ireland.
Email: derekcorrigan@rcsi.ie

**Funding information**
European Commission–DG INFSO, Grant/Award Number: FP7 247787

## Abstract

**Introduction:** Diagnostic error is a major threat to patient safety in the context of family practice. The patient safety implications are severe for both patient and clinician. Traditional approaches to diagnostic decision support have lacked broad acceptance for a number of well-documented reasons: poor integration with electronic health records and clinician workflow, static evidence that lacks transparency and trust, and use of proprietary technical standards hindering wider interoperability. The learning health system (LHS) provides a suitable infrastructure for development of a new breed of learning decision support tools. These tools exploit the potential for appropriate use of the growing volumes of aggregated sources of electronic health records.

**Methods:** We describe the experiences of the TRANSFoRm project developing a diagnostic decision support infrastructure consistent with the wider goals of the LHS. We describe an architecture that is model driven, service oriented, constructed using open standards, and supports evidence derived from electronic sources of patient data. We describe the architecture and implementation of 2 critical aspects for a successful LHS: the model representation and translation of clinical evidence into effective practice and the generation of curated clinical evidence that can be used to populate those models, thus closing the LHS loop.

**Results/Conclusions:** Six core design requirements for implementing a diagnostic LHS are identified and successfully implemented as part of this research work. A number of significant technical and policy challenges are identified for the LHS community to consider, and these are discussed in the context of evaluating this work: medico-legal responsibility for generated diagnostic evidence, developing trust in the LHS (particularly important from the perspective of decision support), and constraints imposed by clinical terminologies on evidence generation.

**KEYWORDS**

diagnostic decision support systems, knowledge discovery, knowledge representation, learning health systems

## 1 | INTRODUCTION

A recent Academy of Medicine report has highlighted the importance of diagnostic error in medicine.[1] Diagnosis in family practice relates to a wide range of poorly differentiated problems. In specialist care, this means achieving sufficiently precise diagnoses in a world where therapies are increasingly targeted at specific phenotypes. Diagnostic error has been shown to be the single biggest source of malpractice

claims in family practice in both the United States and the United Kingdom[2,3] and to be the commonest reported safety incident in UK primary care.[4] A recent report also highlighted the financial impact of such errors with an average pay out for each individual malpractice claim in that study calculated to be $442 000, a huge financial burden at a time when health systems can least afford it.[2] The development of diagnostic decision support systems has attempted to address these problems with limited success and acceptance in actual clinical practice. Some of the challenges in delivering effective diagnostic decision support are technical in nature and well understood: poor integration with electronic health records (EHRs) and clinician workflow, static black-box rule–based evidence that lacks transparency and trust, and use of proprietary technical standards hindering wider interoperability.[5-9]

There is a strong link between general practitioners' initial diagnostic impressions and their subsequent diagnosis and management of common presentations.[10] Recent data show that this phenomenon is not simply one of hypothesis generation but also of subconscious information distortion.[11-13] The fact that family practice acts as a gatekeeper in many countries to direct patients to services within the wider health system means that these errors also have knock-on effects.[14,15] Diagnostic error therefore constitutes a recognised threat to patient safety in family practice with profound impacts for both patient prognosis and clinician professional reputation.

Traditional channels for generating and disseminating clinical evidence have not translated well into decision support tools. The current translational process is measured in years and highly dependent on outputs from traditional clinical trials.[16] To the great frustration of many clinical staff, the effective practice of evidence-based diagnosis relies on the management of a corpus of clinical knowledge in the form of static, document-based guidelines. This evidence base is growing without tools to refine that knowledge appropriately for use in any presenting patient case.[17] A more fundamental question is yet to be addressed relating to the means of production of clinical evidence: How do we facilitate an efficient cycle of derivation, curation, and dissemination of clinical evidence to support clinical decision-making? These are fundamental questions that belong at the heart of a learning health system (LHS).[18,19]

The TRANSFoRm project has developed an LHS for diagnosis as part of the development of a broader LHS technical infrastructure to drive electronic research in family practice.[20] This paper describes the design, development, and implementation of the diagnostic evidence service focussing on both evidence dissemination and evidence derivation from aggregated sources of electronic family practice clinical data. We describe the implementation of the following requirements for the TRANSFoRm LHS for diagnosis:

1. A generalizable model for the representation of diagnostic evidence: This model captures the necessary concepts required to fully support diagnostic reasoning in family practice. It is independent of any specific diagnostic condition and flexible enough to be populated with evidence supporting diagnosis of any given diagnostic condition.

2. Model support for evidence binding to coded clinical terminologies: The evidence model represents clinical evidence independently of clinical coding schemes, whilst supporting optional binding to recognised clinical classifications or terminologies. As will be demonstrated later, support for coded data allows for easier integration with existing EHRs along with collection of unambiguous and structured diagnostic data that can in itself be aggregated and used to generate data-mined clinical evidence.

3. A method of data mining aggregated coded EHR data sets: To support dynamic generation of clinical evidence, an LHS learning loop is necessary to provide for generation and update of quantified clinical evidence with diagnostic knowledge obtained on the basis of coded data captured through the EHR.

4. A clinical evidence service for disseminating diagnostic evidence independently of EHRs: This dissemination mechanism is based on open technical interoperability standards to provide for easier integration and use by a variety of clinical EHR systems.

5. An EHR integrated decision support tool that supports presentation of diagnostic clinical evidence: This diagnostic evidence should be fully integrated with the EHR and present evidence in a manner that will positively influence clinical decision-making.

6. Integrated provenance: Versioning of evidence and recording of decision support system (DSS) recommendations in the EHR are required along with ongoing monitoring of accuracy of rules to facilitate evidence learning and trust.

## 2 | METHODS

Existing approaches to implementing decision support for diagnosis such as ISABEL and DXplain rely on a knowledge engine to define a series of rules in the form of a proprietary database of knowledge tied to a single application.[21,22] Other approaches to decision support for interventions and disease management have used rules triggered or combined together in the form of guidelines based on statements using rule engines such as DROOLS and knowledge rule languages like Arden Syntax, GLIF, and GELLO.[23-28] A more recent approach provides for separation of the actual representation of clinical knowledge using an underlying ontology model and the dissemination of that knowledge using semantic web technologies.[29-31] These models define the named directional relationships between evidence facts in a unified model structure. Proponents argue that a number of desirable benefits result from such a model-based approach: easier maintenance through separation of application from knowledge, ability to reason from cue to diagnosis and from diagnosis to cues using the same knowledge structure, and explicit definition of modelling assumptions along with "open" access to underlying knowledge.[32] A core design requirement for TRANSFoRm was to apply such a model and service-based approach to implementing a learning diagnostic decision support tool.

On the basis of the 6 requirements previously identified, the architecture of the TRANSFoRm diagnostic decision support system was developed as a number of distinct components shown in

Figure 1. The design and implementation of these requirements is subsequently discussed in detail.

## 2.1 | A generalizable model for the representation of diagnostic evidence

The diagnostic process in family practice requires formulation of a working diagnosis based on the primary presenting patient complaint or reason for encounter (RfE).[33] Consideration is given to each candidate differential diagnosis with a view to ruling it in or out based on the confirmed patient diagnostic cues identified through consultation.[34,35] The clinical concepts to support a diagnostic process have been modelled as an ontology of clinical evidence shown in Figure 2. The ontology design methodology used is based on the design practices advocated by the work of Gruninger and Fox.[36]

The model provides for representation of the relationships between a presenting patient RfE and the associated candidate differential diagnoses to consider. The evidence relating to any particular diagnosis is captured as associated diagnostic cues, of which there are cue sub-concepts to represent clinician-observed signs, patient-reported symptoms, risk factors, and clinical tests. The quantified measure of association, in the form of a likelihood ratio, for example, of an RfE or diagnostic cue relating to a particular diagnosis is represented by the quantification concept. Each quantification exists within the context of a particular demographic context. More than one quantification for different demographic contexts may exist for the same RfE/cue/diagnosis relationship.

## 2.2 | Model support for evidence binding to coded clinical terminologies

Semantic interoperability is provided for by binding terminology to the clinical concepts themselves. This is used to associate potentially many different clinical terminology codes for any single ontology RfE, cue, or diagnosis. Specifically, we have supported interoperability with a single EHR vendor in the United Kingdom using National Health Service (NHS) read codes Clinical Terms Version 3.[37] Localisation support to allow easier searching by a third-party consumer for ontology terms using locally defined synonyms is provided for by the "synonym" concept. An example of concepts and associated instances (in red) of the diagnostic cue concept for a patient history of irritable bowel syndrome, with an associated NHS read code "14CF.00" and local synonym "HO IBS" is shown in Figure 3. The ontology can support other coding schemes including ICPC2, ICD10, SNOMED, and UMLS.[38-41]

## 2.3 | A method of data mining aggregated coded EHR data sets

The data mining process and underlying methodology for interpretation of relative "strength" of evidence are illustrated and discussed
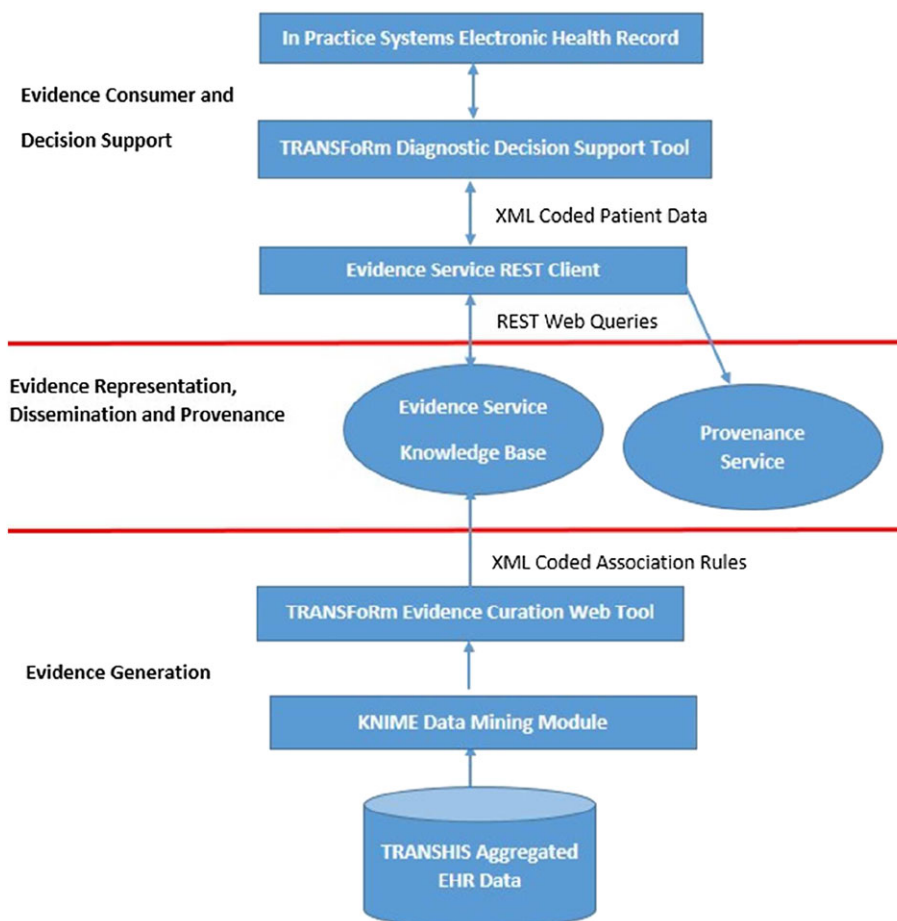


**FIGURE 1** Component architecture of the TRANSFoRm decision support system. EHR indicates electronic health record
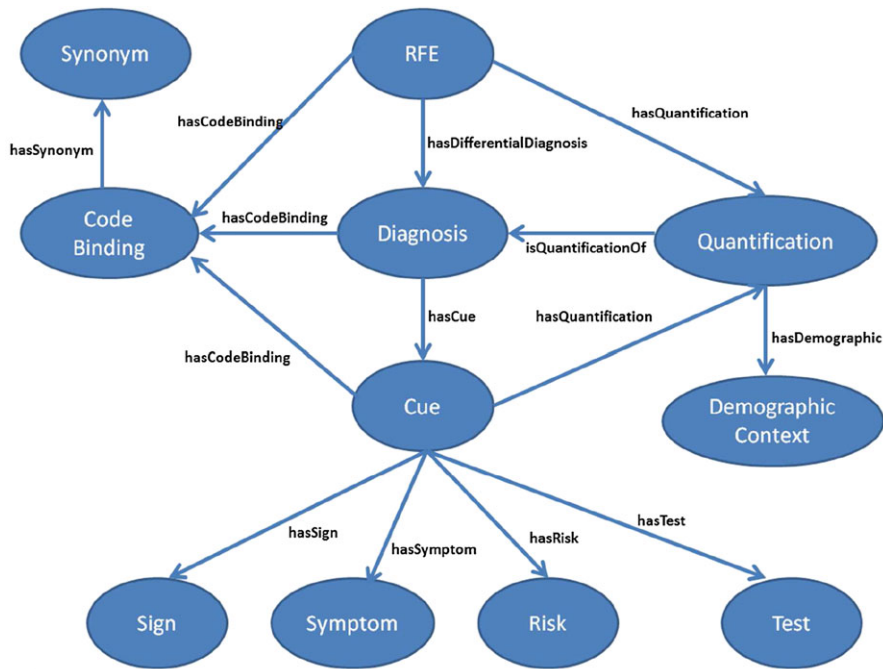
**FIGURE 2** Core ontology concepts and relationships for knowledge representation. RfE indicates reason for encounter
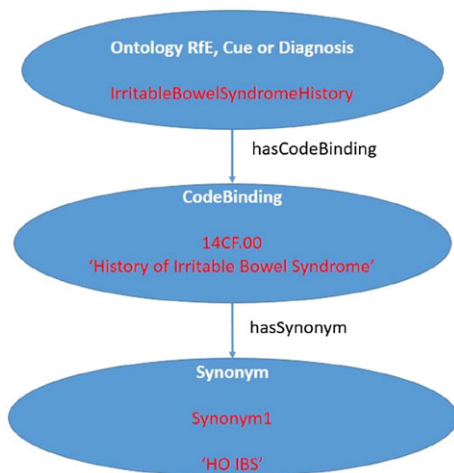


**FIGURE 3** Example of cue ontology concept instance for "history of irritable bowel syndrome." RfE indicates reason for encounter

with examples of quantified clinical evidence elsewhere but will be summarised here.[42,43] An open source data analysis tool, the Konstanz Information Miner (KNIME), was used to derive association rules based on ICPC2 coded data in an episode of care structure obtained from the openly available TransHIS EHR. The association rules identify all possible combinations of RfE, diagnostic cues, and demographic variables (antecedent variables) that are linked with a recorded diagnostic outcome (consequent variable).

The distinct steps implemented in the data mining process as shown in Figure 4 were as follows:

- Step 1—derivation of association rules linking RfE, diagnostic cues, and demographics to a recorded diagnosis made during the first encounter of a new episode of care.

- Step 2—calculation of association rule quality measures, such as likelihood ratios LR+ and LR−, to determine the relative strength
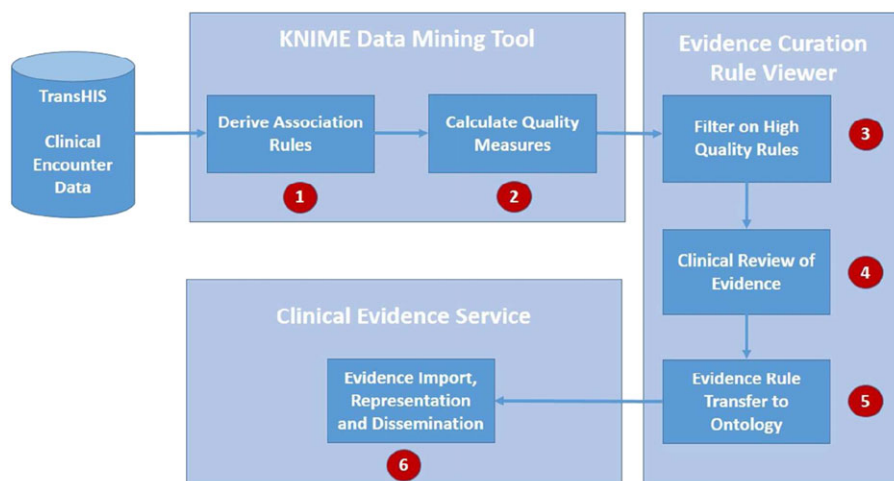


**FIGURE 4** The steps of the knowledge derivation process using data mining. KNIME indicates Konstanz Information Miner

of each rule association derived to support Bayesian reasoning in the form of a post-test probability calculation for each diagnosis.

- Step 3—curation and filtering of association rules to allow selection of "high-quality" association rules.
- Step 4— clinical review of selected rules to assess clinical validity of rules with respect to the wider clinical body of evidence and transfer to evidence service.
- Step 5—import of evidence rules to clinical evidence service.
- Step 6—evidence representation and dissemination through evidence ontology and service.

### 2.3.1 | Evidence curation web tool

A web-based association rule viewer allows curation and clinical review of all generated association rules from the KNIME tool.[44] The tool allows filtering of all rules generated from KNIME on any of the coded ICPC2 antecedent variables (RfEs or diagnostic cues/anams, and demographics), shown in Figure 5. The outcome diagnosis being examined can be filtered by selecting a consecutive variable. In addition, thresholds can be set on the defined quality measures to filter based on the relative strength of the rules required. These measures include the number of occurrences of a rule (support), the positive or negative likelihood ratios for the rule (including 90% confidence intervals), and the sensitivity or specificity of the rule. A scenario name can

be entered that identifies a particular snapshot of rules run on a certain date. This allows multiple copies of versioned rules at different points in time to be stored and retrieved on the basis of a scenario label.

The "filter" button selects the required rules into the main rule viewer screen in the centre of the screen shown in Figure 6. By highlighting a particular rule, the associated rule descriptions are shown along with 95% confidence intervals.

Rules of interest can be selected for deployment to the evidence service from the main screen as an Extensible Markup Language (XML)-exported format. An example of rules export generation is shown in Figure 7.

### 2.3.2 | Evidence service association rule import

The Ontology Updater tool is a command line tool developed as part of the TRANSFoRm project for the purpose of sequentially processing each generated XML rule and updating the clinical evidence service to reflect the updated evidence and quantity measures. A quantification is a unique combination of antecedent variables, consecutive, and demographic context. If a quantification with the precise combination of these elements is not found in the ontology, then it is created. This quantification is then linked with the relevant RfE objects and differential diagnoses objects in the ontology. If it exists already, then the quantification measure values are updated. The quantifications are then made available through the evidence service using a Representational state transfer (REST) query as shown in Figure 8.
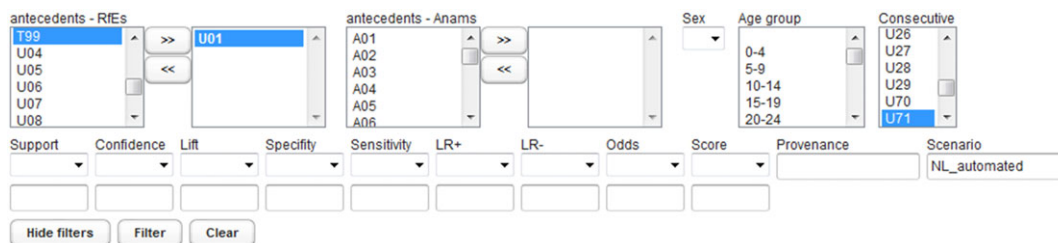


**FIGURE 5** Filters available in the web tool for clinical evidence review. RfE indicates reason for encounter
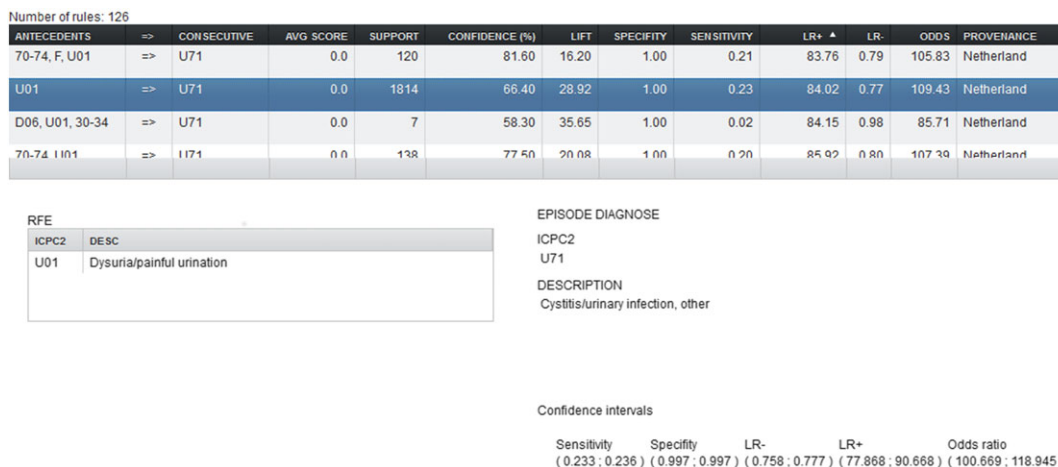


**FIGURE 6** The rule detail selection window showing textual descriptions of dysuria as a presenting complaint indicating urinary tract infection for the Netherlands population

```xml
<?xml version="1.0" encoding="UTF-8"?>
- <RuleSet RuleSetSource="RulesAssesment" RuleSetDate="Thu Sep 11 13:46:32 CEST 2014" RuleSetComment="Rule exported fr
    xsi:noNamespaceSchemaLocation="RuleSetV1.xsd" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  - <Rule RuleScenario="NL_automated" RuleProvenance="Netherland">
      <RuleAntecedent Description="Dysuria/painful urination" AntecedentType="RFE">U01</RuleAntecedent>
      <RuleConsecutive Description="Cystitis/urinary infection, other" ConsecutiveEncoding="ICPC2">U71</RuleConsecutive>
      <RuleRemark/>
      <RuleScore ScoreType="Support">1814.000</RuleScore>
      <RuleScore ScoreType="Confidence">66.400</RuleScore>
      <RuleScore ScoreType="Lift">28.918</RuleScore>
      <RuleScore ScoreType="Specifity">0.997</RuleScore>
      <RuleScore ScoreType="Sensitivity">0.234</RuleScore>
      <RuleScore ScoreType="LR+">84.025</RuleScore>
      <RuleScore ScoreType="LR-">0.768</RuleScore>
      <RuleScore ScoreType="Odds">109.426</RuleScore>
    </Rule>
  - <Rule RuleScenario="NL_automated" RuleProvenance="Netherland">
      <RuleAntecedent Description="Flank/axilla symptom/complaint" AntecedentType="RFE">L05</RuleAntecedent>
      <RuleConsecutive Description="Cystitis/urinary infection, other" ConsecutiveEncoding="ICPC2">U71</RuleConsecutive>
      <RuleRemark/>
      <RuleScore ScoreType="Support">52.000</RuleScore>
      <RuleScore ScoreType="Confidence">4.700</RuleScore>
      <RuleScore ScoreType="Lift">2.030</RuleScore>
      <RuleScore ScoreType="Specifity">0.997</RuleScore>
      <RuleScore ScoreType="Sensitivity">0.007</RuleScore>
      <RuleScore ScoreType="LR+">2.080</RuleScore>
      <RuleScore ScoreType="LR-">0.997</RuleScore>
      <RuleScore ScoreType="Odds">2.088</RuleScore>
    </Rule>
  - <Rule RuleScenario="NL_automated" RuleProvenance="Netherland">
      <RuleAntecedent Description="Urination problems, other" AntecedentType="RFE">U05</RuleAntecedent>
      <RuleConsecutive Description="Cystitis/urinary infection, other" ConsecutiveEncoding="ICPC2">U71</RuleConsecutive>
      <RuleRemark/>
      <RuleScore ScoreType="Support">57.000</RuleScore>
      <RuleScore ScoreType="Confidence">13.000</RuleScore>
      <RuleScore ScoreType="Lift">5.657</RuleScore>
      <RuleScore ScoreType="Specifity">0.999</RuleScore>
      <RuleScore ScoreType="Sensitivity">0.007</RuleScore>
      <RuleScore ScoreType="LR+">6.352</RuleScore>
      <RuleScore ScoreType="LR-">0.994</RuleScore>
      <RuleScore ScoreType="Odds">6.391</RuleScore>
    </Rule>
```

**FIGURE 7** Generation of urinary tract infection association rules export Extensible Markup Language from the web rule sender

## 2.4 | A clinical evidence service for disseminating diagnostic evidence independently of EHRs

The clinical evidence service consists of 3 implementation layers. The ontology is implemented as a Web Ontology Language model using Protégé version 4.3 and hosted on a Sesame triple store.[45-47] The implementation technologies for the 3 layers are summarised in Figure 9. The persistence layer provides a data store for ontology hosting upon which the evidence service is constructed. This provides a platform for multi-user access and dynamic update of ontology clinical content through a programmable interface. The service layer provides a fully functional Jersey REST-based web interface with defined end points to allow parameterised querying of diagnostic questions based on patient data supplied from a third-party consumer.[48]

Using structured evidence service end points, we can access any required ontology concept with results returned as XML (default), JavaScript Object Notation, or Resource Description Framework formats.[49-51] The REST query to access the differentials to consider for a patient presenting with abdominal pain for example is

http://host/ClinicalEvidencerRESTService/interfaces/query/rfes/differentials/AbdominalPainRFE

To access the cues supporting diagnosis of urinary tract infection, shown in Figure 10, the query is

http://host/ClinicalEvidenceRESTService/interfaces/query/differentials/cues/UrinaryTractInfection

Sesame also provides flexibility beyond the defined end points by providing functionality to process custom ad hoc Sesame triple store queries executed directly against its own accessible web service interface.

The REST client handles exchange of patient data between the decision support tool with appropriate calls sent to the back-end evidence service. The client accepts patient data in the form of an XML patient evidence set describing the patient RfE, demographics (extracted from the EHR), and the underlying cues confirmed through consultation with the patient (extracted from the integrated decision support tool itself) shown in Figure 11.

## 2.5 | An EHR integrated decision support tool that supports presentation of diagnostic clinical evidence

The evidence service was made available for use with the TRANSFoRm decision support tool that was developed separately. The design requirements and clinical motivation for implementing the decision support tool are described elsewhere and summarised here.[52,53] The tool supported appropriate diagnostic cues for 78 diagnostic conditions relating to 3 presenting patient complaints: abdominal pain, chest pain, and dyspnoea.

The evidence service provides the decision support tool with ontology-driven coded prompting and recording of coded patient diagnostic cues as shown in Figure 12. The diagnostic decision support tool is embedded and interoperable with the In Practice Systems Vision 3 EHR using the EHR application programmable interface (API). The tool allows for bottom-up input of observed patient cues independent of associated diagnosis (left window) or top-down drilling into and selection of evidence cues supporting specific diagnoses (right window). The differential diagnoses list is based on the presenting RfE and ordered in descending cue count based on the number of patient cues confirmed present for each differential along with the supporting

```
http://localhost:8080/ClinicalEvidenceRESTService/interfaces/query/rfes/quantifications/DysuriaRFE

<?xml version="1.0" encoding="UTF-8" standalone="true"?>
- <quantificationCollection>
  - <quantification>
    - <quantifyDiagnosis>
      - <codeBindings>
          <code>U71</code>
          <codingScheme>ICPC2</codingScheme>
          <description>Urinary tract infection</description>
          <isPrimary>true</isPrimary>
          <synonym>UTI - urinary tract infection</synonym>
        </codeBindings>
        <differentialId>UrinaryTractInfection</differentialId>
        <prevalence>Common</prevalence>
        <severity>Normal</severity>
      </quantifyDiagnosis>
    - <context>
        <contextType>Country</contextType>
        <contextValue>Netherlands</contextValue>
      </context>
    - <value>
        <quantificationType>Confidence</quantificationType>
        <value>66.40</value>
      </value>
    - <value>
        <quantificationType>EvidenceSource</quantificationType>
        <value>NL_automated</value>
      </value>
    + <value>
    - <value>
        <quantificationType>NegativeLR</quantificationType>
        <value>0.77</value>
      </value>
    - <value>
        <quantificationType>PositiveLR</quantificationType>
        <value>84.02</value>
      </value>
    - <rfes>
      - <codeBindings>
          <code>U01</code>
          <codingScheme>ICPC2</codingScheme>
          <description>Dysuria/painful urination</description>
          <isPrimary>true</isPrimary>
        </codeBindings>
        <rfeId>DysuriaRFE</rfeId>
      </rfes>
      <quantifyID>QuantifyRFE1</quantifyID>
    </quantification>
    + <quantification>
    + <quantification>
  </quantificationCollection>
```

**FIGURE 8**    Association rule available evidence made available as a quantified value through the clinical evidence service

underlying evidence cues for each diagnosis. Upon exiting the tool, a working diagnosis can be selected. The evidence cues supporting the current working diagnosis collected during the consultation are coded, saved, and accessible from the patient's EHR for future reference. The recording of this diagnostic consultation data supports a feedback loop to enable further data mining of EHR data to enable further evidence generation and update.

## 2.6 | Integrated provenance

Close attention to automatically captured computable provenance is critical. This operates at 2 levels. Firstly, at patient level, versioning of the evidence and recording the recommendations of the DSS in the EHR are needed. Secondly, at system level, monitoring of ongoing

accuracy of the rules against long-term outcomes allows the system to improve (a critical part of learning) and facilitates trust. Our system used a novel provenance architecture described elsewhere. The architecture is based on W3C PROV standard and the concept of provenance templates to automatically capture the audit trail of the recommendations made, actions performed, and rules and data sets used.[54]

## 3 | RESULTS AND ARCHITECTURAL VALIDATION

A clinical evaluation of the decision support tool itself has been published separately showing an 8% to 9% absolute improvement in diagnostic accuracy, and general practitioners coded significantly more
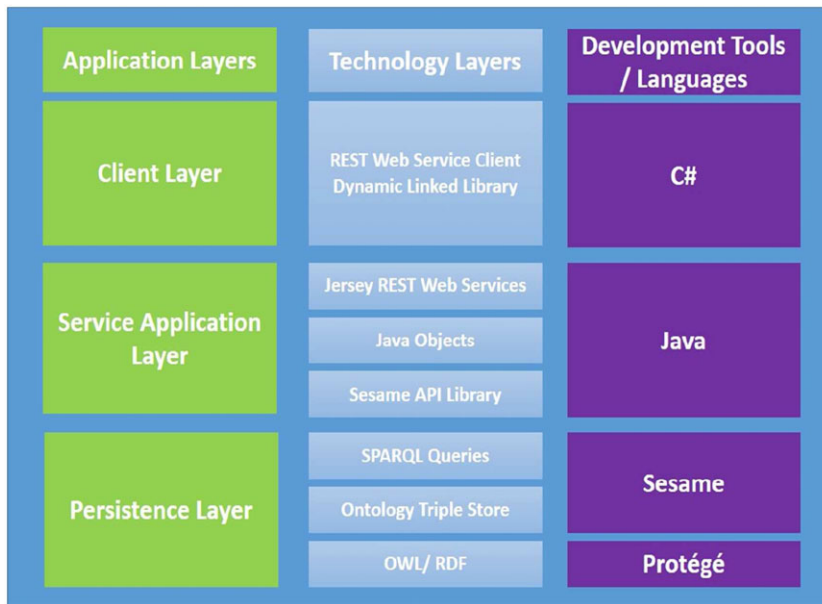
**FIGURE 9** Evidence service implementation technologies. API indicates Application Programmable Interface; OWL, Web Ontology Language; RDF, Resource Description Framework; SPARQL, Sesame triple store query language

data when using the DSS (a prerequisite for producing more primary care clinical evidence).[55] In addition, clinicians found the tool easy to use and learned it quickly. The average length of consultations did not significantly increase (averaged around 10 min per consultation with and without use of the tool).[55]

A validation of this work focussed on the architecture itself and was done on the basis of a comparison to technical best practice in diagnostic decision support development as described in reviews of desirable architectural features for ensuring successful decision support.[5,56] These reviews of features for implementing successful decision support identify desirable functions for comparison of our architecture with other leaders in this area. Each desirable feature is discussed in the context of this work to assess if it has been satisfied.

**Avoids vocabulary issues**—(satisfied)—the decision support system should avoid conflicts with terminology and vocabulary. The clinical evidence service specifically addresses this through the representation of clinical evidence independently as a model with explicit support for binding of vocabulary terms to that evidence. Any desired vocabulary or terminology may bound to that model.

**Shareable**—(satisfied)—the knowledge base should be shareable across many systems. This is explicitly satisfied through provision of a separate clinical evidence service and support for delivery of clinical evidence to local sites using openly available technical standards. Some local integration using a local EHR API is required to enable saving of consultation data to the local EHR.

**Can view decision support content**—(satisfied)—decision support knowledge can be viewed in a number of ways using this architecture. Queries can be run directly using the Sesame triple store query language. Queries can also be run using the web service itself and obtained in XML format as demonstrated. To generate human interpretable versions of the XML, it will be relatively straightforward (but not done here) to generate an application that applies XML formatting using stylesheets to present the content as a document. Data-mined rule content is available through the web-based data mining tool. Finally, manually curated knowledge could be viewed using the source ontology file through the Protégé modelling tool.

**Content separate from code**—(satisfied)—the explicit model-based approach using an ontology ensures that clinical evidence content is separate from any coding done and implemented in the client layer.

**Automatic central updates**—(satisfied)—the discussion of the data mining approach with support for data-mined rule export and import into the centralised clinical evidence demonstrates support for centralised updates of evidence. Once updates are made, all client decision support tools that consume the service can receive the updates.

**Content integrated into workflows**—(satisfied)—the ability to programmatically combine evidence service calls as part of a client layer that implements a diagnostic workflow was implemented as part of the evidence service REST client. A separate evaluation study showed that clinicians found the EHR integrated tool useful and easy to use from a workflow perspective, without imposing an overhead on their consultation time.[55]

**Supports event-driven clinical decision support**—(satisfied)—the TRANSFoRm decision support tool demonstrated an event-driven form of decision support based on entering a presenting RfE that triggers calls to the back-end evidence service.

**Supports non–event-driven clinical decision support**—(satisfied)—the ability to trigger a decision support process based on a batch scheduled process has not been explicitly demonstrated as part of this work. The architecture however can support such a process through the Jersey REST-based interface and the Sesame API that allow programmatic update or querying of the knowledge base.

**Supports decision support over populations**—(satisfied)—as per the discussion on non–event-driven clinical decision support, the architecture can support batch processing over a population of patients (rather than a single patient triggering a single event).

**Enables separation of responsibilities**—(satisfied)—this functionality specifically addresses the need to separate programmer responsibilities from clinical content curation responsibilities. These requirements have been explicitly provided for through the separation of clinical content in the form of models and the provision of curation tool for reviewing evidence before import.

```
- <symptomCollection>
    + <symptomItem>
    + <symptomItem>
    + <symptomItem>
    - <symptomItem>
        - <codeBindings>
            <code>1A1..12</code>
            <codingScheme>ReadCodeBinding</codingScheme>
            <description>Polyuria</description>
            <isPrimary>false</isPrimary>
        </codeBindings>
        - <codeBindings>
            <code>1A1..13</code>
            <codingScheme>ReadCodeBinding</codingScheme>
            <description>Urinary frequency - high</description>
            <isPrimary>true</isPrimary>
        </codeBindings>
        <symptomId>UrinaryFrequency</symptomId>
    </symptomItem>
    - <symptomItem>
        - <codeBindings>
            <code>1A45.00</code>
            <codingScheme>ReadCodeBinding</codingScheme>
            <description>Blood in urine - haematuria</description>
            <isPrimary>true</isPrimary>
        </codeBindings>
        - <codeBindings>
            <code>1A45.11</code>
            <codingScheme>ReadCodeBinding</codingScheme>
            <description>Blood in urine - symptom</description>
            <isPrimary>false</isPrimary>
        </codeBindings>
        - <codeBindings>
            <code>1A45.12</code>
            <codingScheme>ReadCodeBinding</codingScheme>
            <description>Haematuria - symptom</description>
            <isPrimary>false</isPrimary>
        </codeBindings>
        <symptomId>Haematuria</symptomId>
    </symptomItem>
</symptomCollection>
```

**FIGURE 10**   An evidence service reply describing the symptoms of urinary tract infection including frequency and haematuria with associated code bindings

**Free choice of knowledge representation syntax—(satisfied)—**this functionality addresses whether users of the decision support functionality need to have an explicit knowledge of the underlying proprietary knowledge representation format to use it. In the case of this architecture, the underlying knowledge is represented as an ontology but is made available through the web service, which can return results in XML, JavaScript Object Notation, and Resource Description Framework formats. These formats are all based on open and widely used technical standards that are not proprietary in nature.

## 4 | DISCUSSION

A number of important development issues arose during this work. A crucial choice involved giving due consideration to the granularity of coding schemes used at both ends of the LHS infrastructure. The data-mined evidence was based on coarsely granular ICPC2 EHR data.

The target EHR for decision support was based on a much finer granularity in the form of NHS read codes. We demonstrated success in deriving quantified ICPC2-coded diagnostic knowledge that was consistent with the clinical literature for some common clinical conditions along with model-based representation and service-based access to it. The ICPC2 was not a suitable terminology in itself for representing epidemiological knowledge at the required level of low granular detail to support diagnostic decision-making. The tension lies in the requirement of capturing evidence in the form of non-ambiguous, non-overlapping classifiers, able to drive clinical prediction rules and at sufficient granularity of diagnostic classification that the available data are able to fill the categories adequately. As it stands the TransHIS data are too small in volume to support UK practice. Manually curated evidence based on clinical review of best practice guidelines was necessary to fill that gap to construct clinical scenarios with sufficient levels of clinical detail to demonstrate the utility of the decision support tool itself. Careful consideration

**FIGURE 11**   An Extensible Markup Language patient evidence case submitted to the evidence service for a female patient presenting with chest pain and symptoms including fatigue
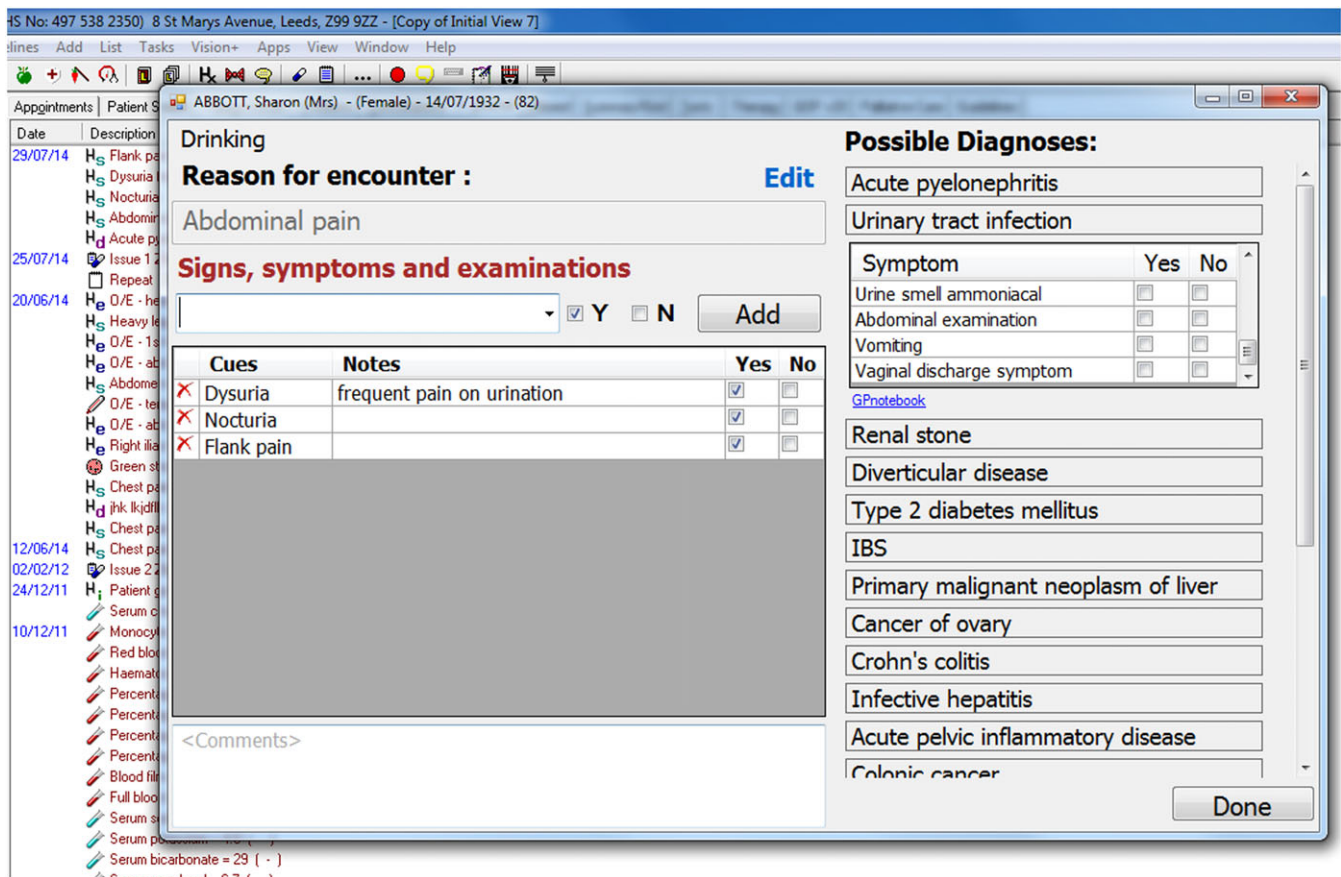


**FIGURE 12**   The integrated diagnostic decision support window accessible from the patient electronic health record shown in the background (the data presented are of a simulated patient)

therefore needs to be given to the underlying EHR structure and coding schemes used for knowledge generation along with the ability to map between coding schemes used by EHRs where the knowledge will be deployed.

The use of ICPC2 was driven by TransHIS as an openly available aggregated EHR data source. The use of richer controlled data sources of health care data for secondary use is more problematic with concerns expressed relating to patient privacy and access by commercial interests.[57] If the LHS is to be successful at the scale that is envisioned, the issue of ownership of patient data needs to be definitively addressed in the public health interest to ensure access to data at a large scale is a reality; otherwise, data silos will continue to exist. In addition, considerable care and clinical input are required in mapping fine-grained terminologies to wider diagnostic concepts in the ontology.

A crucial issue to consider is quality assurance in the form of clinical reliability of the knowledge that is generated and disseminated using these tools. Deployment of an LHS for diagnosis requires trust in the evidence being built. The use of the provenance service is therefore a fundamental requirement for developing trust through evidence monitoring and improvement.[54] In Europe and the United States, this fact has been recognised in legislation that treats software that has been developed for diagnostic purposes as a "medical device" in its own right.[58-60] This places consideration of quality assurance issues and appropriate regulation to the forefront for such categories of software.

## 5 | CONCLUSION

The 6 core requirements for implementing a diagnostic LHS as identified at the outset were successfully implemented as part of this research work. A number of unique aspects distinguish this work from other DSS efforts:

- a focus on gathering evidence to support early diagnosis in primary care (rather than secondary care),
- demonstrable integration with a real commercial primary care EHR system (In Practice Systems in the United Kingdom),
- identification and implementation of best practices in decision support design and architecture, and
- evaluation of the impact of the system on primary care clinical decision-making.

This work satisfies the 2 main requirements of the LHS supporting both evidence generation and dissemination as part of an iterative cycle. In addition, a validation of the architecture demonstrates that the implementation supports features defined as current best practices in decision support design.

The use of EHR data as provided for research purposes is not without limitations. An underlying assumption that the EHR data provided is provided "as-is" and was of sufficiently "good quality" without explicitly exploring what that means or doing a detailed investigation to establish it.[61] This was done on the basis that the data

set had already demonstrably been used to do data mining using manual methods rather than information and communication technology data mining methods.[62,63] A data mining approach at the scale of volume envisaged by the LHS would ideally ensure that quality procedures are followed to ensure consistency of coding of patient data across distributed locations. This would allow for meaningful comparisons and aggregation of that underlying data without risk that coding schemes are used differently in local contexts introducing data bias. The curation of data-mined evidence supported "versioning" of evidence, and a limited comparison was done against clinical literature to assess its clinical accuracy.[42] More research is needed to improve upon the manual nature of that curation process with a view to supporting an iterative cycle of ongoing evidence improvement that is sustainable and accurate.

The design and implementation of our evidence service is consistent with recent initiatives to define service-based provision and access to EHRs and decision support systems such as the Fast Healthcare Interoperability Resources initiative. Our work on primary care data mining and evidence curation tools can also be viewed as having applications to the wider problem of identifying clinically significant features necessary for the authoring of openly portable electronic phenotype definitions. These can be used for identification of research cohorts across clinical networks for the purpose of conducting genotype-phenotype association studies, as demonstrated by initiatives such as the eMerge network in the United States, for example.[64] Recent calls have also highlighted the need for more research on diagnostic error in practice. The Committee on Diagnostic Error in Health at the Institute of Medicine made a number of important recommendations in this area in 2015.[1] Specifically, they recommended the development of "a coordinated research agenda on the diagnostic process and diagnostic errors by the end of 2016." They make reference to the inclusion of "a broad range of stakeholders including health IT industries" to develop decision support tools that

> "ensure that health IT used in the diagnostic process demonstrates usability, incorporates human factors knowledge, integrates measurement capability, fits well within clinical workflow, provides clinical decision support, and facilitates the timely flow of information among patients and health care professionals involved in the diagnostic process."[1]

The TRANSFoRm approach to diagnostic decision support can be seen as a practical contribution to addressing those recommendations.

TRANSFoRm has demonstrated that the goals of the LHS are closely aligned with the development of next generation decision support tools. Further policy work is needed to ensure that these tools can make use of the growing bodies of EHR data at scale, to support derivation and dissemination of diagnostic evidence in practice, supporting clinicians and driving real improvements in patient safety.

## REFERENCES

1. Balogh EP, Miller BT, Ball JR, eds. *Improving Diagnosis in Health Care.* Washington, DC: The National Academies Press; 2015:450.

2. 2014 CBS report: malpractice risks in the diagnostic process: CRICO; 2015. Available from: https://www.rmf.harvard.edu/Clinician-Resources/Article/2014/CBS-Intro

3. Wallace E, Lowry J, Smith SM, Fahey T. The epidemiology of malpractice claims in primary care: a systematic review. *BMJ Open*. 2013;3(7):e002929

4. Carson-Stevens A, Hibbert P, Williams H, et al. Characterising the nature of primary care patient safety incident reports in the England and Wales National Reporting and Learning System: a mixed-methods agenda-setting study for general practice. *Health Serv Deliv Res*. 2016;4(27):1-76.

5. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *BMJ*. 2005;330 (7494):765

6. El-Kareh R, Hasan O, Schiff GD. Use of health information technology to reduce diagnostic errors. *BMJ Qual Saf*. 2013;22(2):1-12.

7. Shibl R, Lawley M, Debuse J. Factors influencing decision support system acceptance. *Decis Support Syst*. 2013;54(2):953-961.

8. Patel VL, Kaufman DR, Kannampallil TG. Diagnostic reasoning and decision making in the context of health information technology. *Rev Hum Factors Ergon*. 2013;8(1):149-190.

9. Sittig DF, Wright A, Osheroff JA, et al. Grand challenges in clinical decision support. *J Biomed Inform*. 2008;41(2):387-392.

10. Kostopoulou O, Rosen A, Round T, Wright E, Douiri A, Delaney B. Early diagnostic suggestions improve accuracy of GPs: a randomised controlled trial using computer-simulated patients. *Br J Gen Pract*. 2015;65(630):e49-e54.

11. Elder NC, Dovey SM. Classification of medical errors and preventable adverse events in primary care: a synthesis of the literature. *J FAM PRACTICE*. 2002;51(11):927-932.

12. Kostopoulou O, Delaney BC, Munro CW. Diagnostic difficulty and error in primary care—a systematic review. *Fam Pract*. 2008;25 (6):400-413.

13. Nurek M, Kostopoulou O. What you find depends on how you measure it: reactivity of response scales measuring predecisional information distortion in medical diagnosis. *PLoS One*. 2016;11(9):e0162562.

14. Starfield B. *Primary Care: Balancing Health Needs, Services, and Technology*. USA: Oxford University Press; 1998.

15. Donaldson MS, Yordy KD, Lohr KN, Vanselow NA. *Primary Care: America's Health in a New Era*. Washington DC: The National Academies Press; 1996.

16. Lang ES, Wyer PC, Haynes RB. Knowledge translation: closing the evidence-to-practice gap. *Ann Emerg Med*. 2007;49(3):355-363.

17. Greenhalgh T, Howick J, Maskrey N. Evidence based medicine: a movement in crisis? *BMJ*. 2014;348:

18. Delaney BC, Ethier J-F, Curcin V, Corrigan D, Friedman C (Eds). *International perspectives on the digital infrastructure for The Learning Healthcare System*. Washington DC, USA: AMIA annual symposium 2013.

19. Friedman C, Rigby M. Conceptualising and creating a global learning health system. *Int J Med Inform*. 2013;82(4):e63-e71.

20. TRANSFoRm Project Homepage. Available from: http://www.transformproject.eu

21. Isabel Healthcare Homepage: Isabel Healthcare. Available from: http://www.isabelhealthcare.com/home/default

22. DXPlain Open Clinical. Available from: http://www.openclinical.org/aisp_dxplain.html

23. Hripcsak G, Ludemann P, Pryor TA, Wigertz OB, Clayton PD. Rationale for the Arden syntax. *Comput Biomed Res*. 1994;27(4):291-324.

24. Peleg M, Boxwala AA, Bernstam E, Tu S, Greenes RA, Shortliffe EH. Sharable representation of clinical guidelines in GLIF: relationship to the Arden syntax. *J Biomed Inform*. 2001;34(3):170-181.

25. Wang D, Peleg M, Tu SW, et al. Design and implementation of the GLIF3 guideline execution engine. *J Biomed Inform*. 2004;37(5):305-318.

26. Samwald M, Fehre K, De Bruin J, Adlassnig K-P. The Arden syntax standard for clinical decision support: experiences and directions. *J Biomed Inform*. 2012;45(4):711-718.

27. Ogunyemi O, Zeng Q, Boxwala A, Greenes R. Object-oriented guideline expression language (GELLO) specification: Brigham and Women's Hospital, Harvard Medical School, 2002. Decision Systems Group Technical Report DSG-TR-2002-001.

28. Li D, Endle CM, Murthy S, et al., eds. Modeling and executing electronic health records driven phenotyping algorithms using the NQF Quality Data Model and JBoss® Drools engine. AMIA annual symposium proceedings; 2012: American Medical Informatics Association.

29. Ethier J-F, Curcin V, Barton A, et al. Clinical data integration model. *Methods Inf Med*. 2015;54(1):16-23.

30. Ethier J-F, Dameron O, Curcin V, et al. A unified structural/terminological interoperability framework based on LexEVS: application toTRANSFoRm. *J Am Med Inform Assoc*. 2013;20(5):986-994.

31. Corrigan D, McDonnell R, Zarabzadeh A, Fahey T. A multi-step maturity model for the implementation of electronic and computable diagnostic clinical prediction rules (eCPRs). *eGEMs (Generating Evidence & Methods to improve patient outcomes)*. 2015;3(2):8.

32. Alper J, Geller A, eds. *How Modeling Can Inform Strategies to Improve Population Health: Workshop Summary*. Washington, DC: The National Academies Press; 2015:120.

33. Weed LL. Special article: medical records that guide and teach. *N Engl J Med*. 1968;278(12):593-600.

34. Kostopoulou O. Diagnostic Errors: Psychological Theories and Research Implications. In: Hurwitz B, Sheikh A, eds. *Health Care Errors and Patient Safety*. Oxford: Wiley-Blackwell; 2009:95-111.

35. Heneghan C, Glasziou P, Thompson M, et al. Diagnosis in general practice: diagnostic strategies used in primary care. *BMJ*. 2009; 1003-1006.

36. Gruninger M, Fox MS, editors. Methodology for the design and evaluation of ontologies. Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing, IJCAI; 1995.

37. Health and Social Care Information Centre—NHS Read Codes. 2014. Available from: http://systems.hscic.gov.uk/data/uktc/readcodes

38. Organisation WH. International classification of primary care, second edition (ICPC-2). Available from: http://www.who.int/classifications/icd/adaptations/icpc2/en/

39. Organisation WH. International classification of diseases version 10. Available from: http://apps.who.int/classifications/icd10/browse/2016/en

40. Organisation TIHTSD. SNOMED CT. Available from: http://www.ihtsdo.org/snomed-ct.

41. UMLS Terminology Services: National Institutes of Health. Available from: https://uts.nlm.nih.gov/home.html

42. Soler JK, Corrigan D, Kazienko P, et al. Evidence-based rules from family practice to inform family practice; the learning healthcare system case study on urinary tract infections. *BMC Fam Pract*. 2015;16(1):63

43. Danger R, Corrigan D, Soler JK, et al. A methodology for mining clinical data: experiences from TRANSFoRm project. *Stud Health Technol Inform*. 2014;210:85-89.

44. KNIME Homepage. Available from: http://www.knime.org/

45. Web Ontology Language Standard Homepage. Available from: http://www.w3.org/2004/OWL/

46. Protégé Homepage. Available from: http://protege.stanford.edu/

47. Sesame Homepage. Available from: http://www.openrdf.org/

48. Jersey REST Services. Available from: https://jersey.java.net/

49. XML Technology Standard: W3C. Available from: http://www.w3.org/standards/xml

50. The JSON Interchange Standard. Available from: http://json.org/

51. Resource Description Framework. Available from: http://www.w3.org/RDF/

52. Porat T, Kostopoulou O, Woolley A, Delaney BC. Eliciting user decision requirements for designing computerized diagnostic support for family physicians. *J Cogn Eng Decis Mak*. 2016;10(1):57-73.

53. Kostopoulou O, Rosen A, Round T, Wright E, Douiri A, Delaney B. Early diagnostic suggestions improve accuracy of GPs: a randomised controlled trial using computer-simulated patients. *Br J Gen Pract*. 2015;65(630):e49-e54.

54. Curcin V, Fairweather E, Danger R, Corrigan D. Templates as a method for implementing data provenance in decision support systems. *J Biomed Inform*. 2017;65:1-21.

55. Kostopoulou O, Porat T, Corrigan D, Mahmoud S, Delaney BC. Diagnostic accuracy of GPs when using an early-intervention decision support system: a high-fidelity simulation. *Br J Gen Pract*. 2017;67 (656):e201-e208.

56. Wright A, Sittig DF. A framework and model for evaluating clinical decision support architectures. *J Biomed Inform*. 2008;41(6):982-990.

57. Boiten E. Patients will resist medical record sharing if NHS bosses ignore their privacy fears. 2015.

58. Directive 2007/47/EC of the European parliament and of the council of 5 September 2007 amending Council Directive 90/385/EEC on the approximation of the laws of the Member States relating to active implantable medical devices, Council Directive 93/42/EEC concerning medical devices and Directive 98/8/EC concerning the placing of biocidal products on the market: European Union; 2007. Available from: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:247:0021:0055:en:PDF

59. Code of Federal Regulations Title 21 Part 820: US Food and Drug Administration. Available from: http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfCFR/CFRSearch.cfm?CFRPart=820

60. Code of Federal Regulations Title 21 Part 860: U.S. Food and Drug Administration. Available from: http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfCFR/CFRSearch.cfm?CFRPart=860

61. Kahn MG, Brown JS, Chun AT, et al. Transparent reporting of data quality in distributed data networks. *eGEMs*. 2015;3(1):

62. Soler JK, Okkes I, Oskam S, et al. An international comparative family medicine study of the Transition Project data from the Netherlands, Malta and Serbia. Is family medicine an international discipline? Comparing diagnostic odds ratios across populations. *Fam Pract*. 2012;29(3):299-314.

63. Soler JK, Okkes I, Oskam S, et al. An international comparative family medicine study of the Transition Project data from the Netherlands, Malta and Serbia. Is family medicine an international discipline? Comparing incidence and prevalence rates of reasons for encounter and diagnostic titles of episodes of care across populations. *Fam Pract*. 2012;29(3):283-298.

64. Gottesman O, Kuivaniemi H, Tromp G, et al. The electronic medical records and genomics (eMERGE) network: past, present, and future. *Genet Med*. 2013;15(10):761-771.