Electronic Health Record Interoperability as Realized in Turkey's National Health Information System

A. Dogac[1,2], M. Yuksel[1,2], A. Avcı[3], B.Ceyhan[3], Ü. Hülür[3], Z. Eryılmaz[3],
S. Mollahaliloğlu[3], E. Atbakan[3] and R. Akdağ[3]
[1] Dept. of Computer Eng., Middle East Technical University, Ankara, Turkey
[2] SRDC Ltd., Ankara, Turkey
[3] Ministry of Health, Turkey

Summary

Objectives: The objective of this paper is to describe the techniques used in developing Turkey's National Health Information System (NHIS-T), a nation-wide infrastructure for sharing the Electronic Health Records (EHRs).

Methods: In designing the logical EHR structure, to be able to increase the re-use of common information blocks in EHRs, the UN/CEFACT Core Components Technical Specification (CCTS) methodology is applied.

Results: Turkey's NHIS (NHIS-T) became operational on January 15, 2009 and by June 2010, 99% of the public hospitals and 71% of the private and university hospitals were connected to NHIS-T with daily feeds of their patient EHRs. Out of the 72 million citizens of Turkey, electronic healthcare records of 43 million citizens have already been created in NHIS-T. Currently, only the general practitioners can access the EHRs of their patients. In the second phase of the implementation, once the legal framework is completed, the proper patient consent mechanisms will be available through the Personal Health Record system that is under development, so that these EHRs will also become accessible to the authorized healthcare professionals in the secondary and the tertiary healthcare systems.

Conclusions: There is a number of factors that affected the successful implementation of Turkey's NHIS. The first is the adoption of the specified standards by all the stakeholders. The second factor is the use of UN/CEFACT Core Components Technical Specification (CCTS) approach which facilitated the understanding of rather complex EHR schemas by the software developers. Finally, the comprehensive testing of vendor Hospital Information System (HIS) applications for their conformance to and interoperability with NHIS-T through an automated testing platform, namely, TestBATN, greatly enhanced the fast integration of the HISs to NHIS-T.

Keywords:
Electronic Healthcare Records, National Health Information System, UN/CEFACT CCTS, HL7 CDA, eHealth Conformance and Interoperability testing

Correspondence to:
Dr. Asuman Dogac

SRDC Software Research, Development and Consultation Ltd.
ODTU Silikon Bina, 1. Kat, No:14
Middle East Technical University Campus, 06531 Ankara Turkey
Phone: +90 (312) 210 1393
Fax: +90 (312) 210 1837
email: asuman@srdc.com.tr

1 Introduction and Objectives

The National Health Information System of Turkey (NHIS-T) [1] is a nation-wide infrastructure for sharing patient Electronic Health Records (EHRs). NHIS-T is a part of the national eHealth infrastructure, called Saglik-Net (Turkish for "Health-Net"). The motivation for Saglik-Net is the creation of standards-based Electronic Health Records as well as providing statistical data analysis support for decision making. The current implementation of NHIS-T supports the transfer of EHRs from the Hospital Information Systems (HIS) to NHIS-T servers at the Ministry of Health (MoH). General practitioners (GPs) can currently access the EHRs of their own patients that are managed through the Family Medicine Information System (FMIS), which is a different system that is linked with NHIS-T. Every morning a GP receives the list of his FMIS patients and their records from NHIS-T if they had a visit to a secondary or tertiary healthcare provider the previous day. In the second phase of NHIS-T implementation, once the legal framework is completed and the proper consent mechanisms are developed, the EHRs will also become accessible to the authorized healthcare professionals in the secondary and the tertiary healthcare systems. For this purpose, the implementation of a Personal Health Record system on top of NHIS-T is ongoing and will let patients define their consents and to access their own healthcare information.

The ISO/TR 20514:2005 standard defines a longitudinal EHR as "a repository of information regarding the health of a subject of care in computer processable form, stored and transmitted securely, and accessible by multiple authorised users. It has a commonly agreed logical information model which is independent of EHR systems." [2]

Any infrastructure implementation for exchanging EHRs needs to address the following challenges:

1. *Setting the EHR standard and developing the EHR content*: In order to provide interoperability of the EHRs, a standard interface such as the HL7 Clinical Document Architecture (CDA) [3] needs to be adopted. However, there can be many different ways of organizing the same clinical information even when the same EHR standard is used: the same content can be expressed through different components and the components can be aggregated differently. Therefore a common schema is necessary. A further complication is that the longitudinal EHR documents are created piecemeal such as "patient examination", "mouth and teeth examination" or "infant observation". Therefore a common schema for each type of document is needed to enforce interoperability. On the other hand, these document schemas share many common components and the common components need to be the same in different clinical document schemas for meaningful use. In order to make sure that the common EHR components are re-used, the UN/CEFACT Core Components Technical Specification (CCTS) [4] methodology is applied to the development of Electronic Health Records in Turkey.

2. *Determining the patient identification mechanisms:* Patient identifiers are used to locate the EHRs of a patient. Yet each healthcare organization may (and typically will) have a different patient identifier domain. To be able to share the EHRs, the different patient identifiers need to be mapped to one another. In Turkey, every

citizen has a unique identity number and these are used as patient identifiers and this helped to solve the problem easily. The citizen identity numbers are maintained in a system called MERNIS (The Central Civil Registration System) [5] and the patient identifiers in the EHRs are validated through this system.

3. *Developing the healthcare professional identification mechanisms:* The healthcare professional identity is also an indispensable part of the EHRs because of the medico-legal requirement that a healthcare professional has to assume responsibility for a clinical document. Therefore, it is necessary to validate this information in an EHR instance against a master registry of health professionals. In Turkey, the physician related information is kept in a database, namely, *Doctor Data Bank (DDB)* [6] and the identifiers of the healthcare professionals that appear in the EHRs are validated against this DDB.
4. *Developing mechanisms to verify the coded elements that appear in the EHRs*: Coded elements are used in EHRs to precisely express the medical terms and concepts independent of the words used to describe them. Medicine has a long tradition in structuring its domain knowledge through terminologies and coding schemes for diseases, medical procedures and anatomical terms such as SNOMED [7], LOINC [8], READ Codes [9], MeSH [10] and ICD-10 [11]. In order to validate the content of an EHR instance, the coded elements need to be verified against the related terminology systems. In NHIS-T implementation, a publicly accessible server, namely, the Health Coding Reference Server (HCRS) [12] is used that contains all the international and national coding systems used in Turkey. The codes in the EHRs are validated through this server.
5. *Developing mechanisms to provide the security and privacy of the EHR*: EHRs contain highly sensitive data and therefore the security and privacy of the patients must be preserved. To protect the data collected at the MoH NHIS-T servers from unauthorized disclosures while ensuring its availability for authorized uses, a comprehensive set of security and privacy measures have been implemented. There are two types of administrators in the system: Security Administrator and the Database Administrator. Security Administrator is in charge of granting rights to the Database Administrators, but they themselves have no right to access the database. Various "view" mechanisms are developed to hide the patient demographic data from unauthorized users. Additionally, access to NHIS-T data is audited by logging all the user events. An Identity Management System provides group, role and rule based management and authorization options.
6. *Deciding on the transport protocols*: An EHR is a document and constitutes a message payload. In other words, document instances need to be carried over the network using transport protocols and HL7 Version 3 provides three transport specifications [13] - ebXML, Web Services and Minimal Lower Layer Protocol (MLLP) - for the exchange of HL7 based content, messages and documents. In NHIS-T implementation, HL7 Web Services Profile [14] is used as the communication infrastructure. For the security of the Web services, WS-Security Username Token Profile [15] over Secure Sockets Layer (SSL) [16] is used.

7. *Enforcing the Business rules*: A two-phase validation technique is applied for the validation of the business rules on the incoming messages to NHIS-T. In the first phase, which is called syntax validation phase, an incoming document instance is validated against the related XML Schema Definition (XSD). If successful, the message is conveyed to the second phase where the values in the data elements and the relationships between them are checked against the previously defined rules.

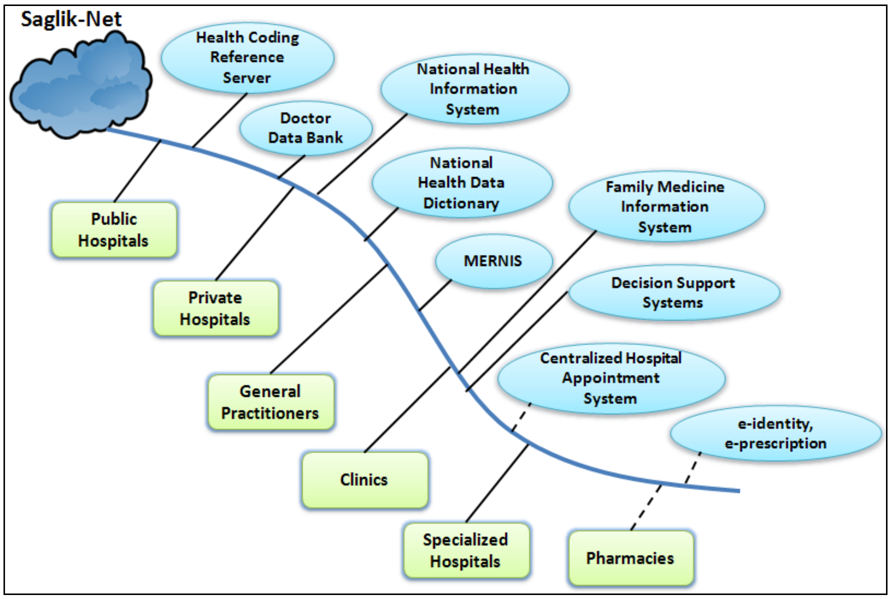

Figure 1 Saglik-Net 2010

Figure 1 shows the overall architecture of the Saglik-Net that is comprised of the National Health Data Dictionary [17], the Health Coding Reference Server [12], the Central Civil Registration System (MERNIS) [5], the National Health Information System (NHIS-T) [1], the Family Medicine Information System (FMIS) [18], the Decision Support Systems (DSS), the Centralized Hospital Appointment System (CHAS) [19] and the e-identity and e-prescription mechanisms. The CHAS, the e-identity and the e-prescription projects are currently in their pilot phases in some provinces of Turkey, but the coverage is expanding rapidly. Among the users of the network, currently the Family Medicine Information System (FMIS) clients and HISs are fully connected, whereas the rest of users (payer institutions, pharmacies etc.) are not yet connected.

As a summary, in realizing the NHIS-T, in order to increase the re-use of common information blocks in EHRs, the UN/CEFACT Core Components Technical Specification (CCTS) methodology is applied. The data types and the data elements used in EHRs are identified and then a set of reusable building blocks of the EHRs, termed the "Minimum Health Data Sets (MHDS)", are created. The data elements are coded with coding systems that are available from the National Health Coding Reference Server (HCRS). Both the data elements and the MHDSs are stored in a National Data Dictionary. These reusable building blocks are then assembled into aggregate document components, called "Transmission Data Sets" (piecemeal EHRs). This conceptual EHR architecture is then serialized into XML based on the HL7 CDA structure. The healthcare professional identities in the EHRs are validated against the Doctor Data Bank (DDB). The patient identification is realized by using the nation-wide citizen identity numbers. HL7 Web service Profile is used for the transportation of "Transmission Schema" instances.

The paper is organized as follows: Section 2 briefly summarizes the enabling technologies and the standards. In Section 3, we describe how EHR content is designed by using the UN/CEFACT Core Components Technical Specification and how it is serialized to HL7 CDA. Section 4 presents the communication infrastructure of NHIS-T and the validation of the EHRs. Section 5 explains the functional and semantic testing of NHIS-T and its current deployment status. Finally, Section 6 concludes the article.

2 Enabling Standards and Technologies

This section briefly introduces the standards and technologies used in this work.

2.1 UN/CEFACT Core Components Technical Specification

In order to increase the interoperability among electronic business documents, UN/CEFACT Core Components Technical Specification (CCTS) [4] has specified a methodology to define the structure and the properties of abstract document components. CCTS methodology identifies a set of reusable building blocks, called *Core Components*, which represent the common data elements of business documents such as "Address", "Amount", or "Line Item". These atomic document components are called Basic Core Components (BCC) and get their data types from a fixed list of Core Data Types (CDTs) defined by UN/CEFACT CCTS. Currently, there are 14 Core Data Types such as "Amount", "Identifier" or "Measure" and they are built from primitive data types such as binary, decimal, integer or string.

These reusable building blocks are then assembled into aggregate document components that are composed of either basic document components, or they refer to other aggregate document components through association document components called the Association Core Components (ASCC). For example, a "Person" class describing an aggregate document component may contain a "name" basic document component and an "OfficialAddress" association document component, which refers to another aggregate document component called "Address". In other words, CCTS establishes the relationship between document components only through association components rather than nesting them inside each other. Considerable number of Core Components is available from the UN/CEFACT Core Component Library (CCL) for discovery and reuse and more will be available as their work progresses.

A Core Component is designed to be context-independent so that it can later be adapted to different contexts and reused. When a Core Component is restricted to be used in a specific business context, it becomes a Business Information Entity (BIE) and is given its own unique name.

2.2 HL7 CDA

HL7 Clinical Document Architecture (CDA) is a document markup standard that specifies the structure and the semantics of a clinical document (such as a discharge summary or progress note) for the purpose of exchange [3]. By definition, a valid CDA document is encoded in Extensible Markup Language (XML) and is validated against the CDA Schema once all extensions have been removed from the instance.

A CDA document has two main parts, the header and the body. A CDA header provides information on the authentication, the encounter, the patient, and the involved providers whereas the CDA body includes the clinical report. The body part can be either an unstructured blob or a structured hierarchy which involves one or more sections. Within a section, narrative blocks and CDA entries are defined. Machine processable clinical statements are represented by these CDA entries whereas the narrative blocks are human readable forms.

So far, HL7 has released two versions of CDA. In CDA Release 1 (R1) [20], only the header part is derived from the HL7 Reference Information Model (RIM) [21]. In CDA Release 2 (R2) [3], in addition to the header part, the clinical content in the document body is also derived from the RIM. Therefore, CDA R2 model enables the formal representation of clinical statements through CDA Entry classes, namely, "Act", "Observation", "ObservationMedia", "SubstanceAdministration", "Supply", "Procedure", "Encounter", "RegionOfInterest" and "Organizer".

3 Developing the EHR Content

Electronic Healthcare Records are "documents" describing the patient related healthcare data and, as such, they are organized into document components such as sections, entries and data elements. Indeed, the prominent EHR standards like HL7 CDA and CEN 13606-1 [22] structure the clinical documents from basic document components like "Sections" and "Entries" in HL7 CDA or "Compositions", "Sections" and "Entries" in 13606-1 and aggregating these basic components as needed.

Clearly, there can be many different ways of organizing the same clinical information even when the same EHR standard is used: the same information can be expressed through different components and the components can be aggregated differently. Hence, adhering to an EHR standard interface can make the EHR content machine processable but not interoperable unless a common schema and the necessary restrictions are used. Although it is possible to resolve such structural differences by providing the detailed mappings among the XML schemas, this is a very tedious and expensive process requiring human expertise and labor.

In order to make sure that the EHR components common in the clinical documents are re-used, NHIS-T employs the UN/CEFACT Core Components Technical Specification (CCTS) [4] methodology in developing Electronic Health Records. For this purpose, first the data types and the data elements used in the EHRs are identified, such as the "Address" or "Main Diagnosis" and then a set of meaningful and reusable grouping of data elements that represent the common data blocks of EHRs, called "Minimum Health Data Sets (MHDS)", are created (e.g., the "Medical Examination" or the "Prescription"). The data elements are coded using coding systems that are available from the Health Coding Reference Server (HCRS) [12]. Both the data elements and the MHDSs are stored in a National Data Dictionary [17].

These MHDSs are then assembled into aggregate document components called "Transmission Data Sets" (piecemeal EHRs), such as the "Examination Transmission

Data Set”. This conceptual architecture is then serialized into XML based on the HL7 CDA structure.

A point to note here is that the core components developed in Turkey are not context independent; in fact they are developed directly for the eHealth considering Turkey’s particular context, and, thus, would better be called Business Information Entities (BIEs). Yet, in order not to complicate the terminology any further, we stick to the term “core components” in this article.

3.1 National Health Data Dictionary

As already mentioned, the first step in developing electronic documents with the CCTS methodology is to define the core data types. Turkey first built its National Health Data Dictionary (NHDD) [17] by defining the commonly used healthcare data elements such as “Address”, “Name”, “Main Diagnosis”, “Vaccination”, and “Treatment Method”. The format of these data elements is defined according to the rules and guidelines given in ISO/IEC 11179-4 Standard [23]. Currently, there are 261 data elements in the data dictionary.

Following this, the “Aggregate Core Components”, which are called Minimum Health Data Sets (MHDSs), are formed using these data elements. MHDSs define the data sets that emerge at the time of presenting a specific healthcare service, for example, Infant Monitoring Data Set or Pregnant Monitoring Data Set. Currently, there are 46 Minimum Health Data Sets in the NHDD and Figure 2 gives an overall view of them.

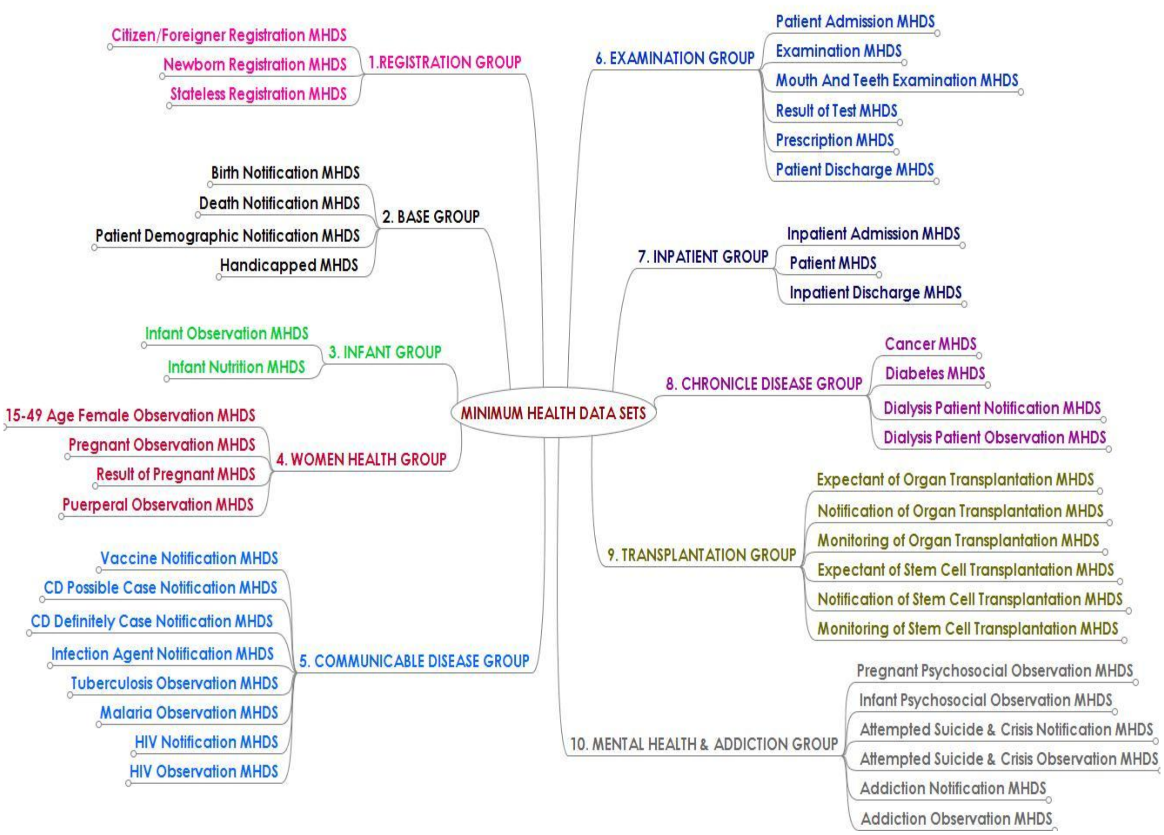

Figure 2 Minimum Health Data Sets in the NHDD

The data elements within the Minimum Health Data Sets are mostly coded using coding systems that are available from the Health Coding Reference Server (HCRS). If a data element is defined in the National Health Data Dictionary as coded or classified, then the related coding/classification system is given both within the definition of the data element and in the "HCRS System Code" field. There are two possibilities for a coded element: either the value is gathered from a coding system such as ICD-10 or the value is categorical such as gender, or marital status.

3.2 Transmission Data Sets

While aggregating the Minimum Health Data Sets into the Transmission Data Sets (piecemeal Electronic Healthcare Records), again the CCTS approach is followed. The MHDSs being the Aggregate Core Components themselves, the association between MHDSs and the Transmission Data Sets are established via Association Core Components (ASCCs). Figure 3 demonstrates how the "Examination Transmission Data Set" is aggregated by re-using the related Minimum Health Data Sets and by defining the associations: First, for representing the patient demographics information, a "Patient Registration MHDS" is linked to the "Examination Transmission Data Set" through "subject" ASCC (i.e. relationship). Then the other MHDSs where the actual healthcare information is represented are linked. "Examination MHDS" is linked through "examination details" ASCC, "Patient Admission MHDS" is linked through "admission details" ASCC and the "Patient Discharge MHDS" is linked through "discharge details" ASCC. "Test Result MHDS" and "Prescription MHDS" are linked similarly; the only difference is that these MHDSs are optional within the "Examination Transmission Data Set".

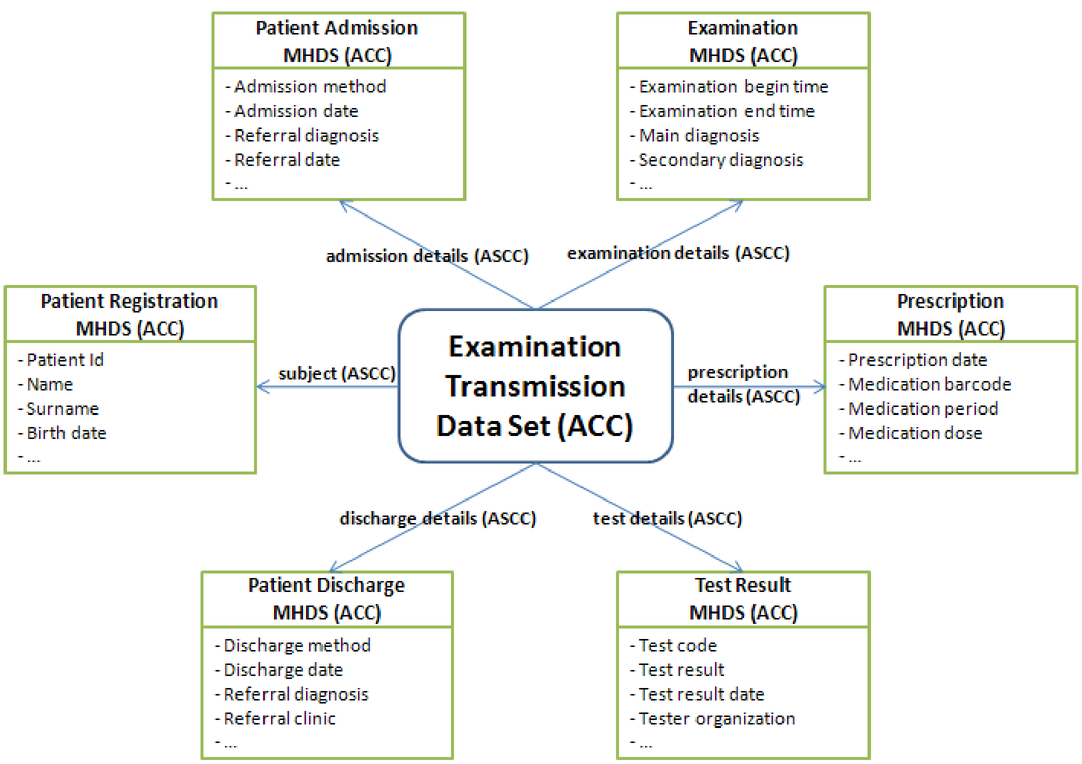

Figure 3 The Examination Transmission Data Set of NHIS-T

In the current version of NHIS-T implementation there are 41 Transmission Data Sets. The most important 25 are (details provided in Section 4) : “15-49 Age Female Observation”, “Mouth and Teeth Examination”, “Vaccine Notification”, “Infant Nutrition”, “Infant Observation”, “Infant Psychosocial Observation”, “Communicable Disease Definite Case Notification”, “Communicable Disease Probable Case Notification”, “Diabetes”, “Dialysis Notification”, “Dialysis Observation”, “Birth Notification”, “Pregnant Observation”, “Pregnancy Termination”, “Pregnant Psychosocial Observation”, “Patient Demographics Notification”, “Cancer”, “Puerperal Observation”, “Examination”, “Death Notification”, “Test Result”, “Citizen/Foreigner Registration”, “Stateless Person Registration”, “Newborn Registration” and “Inpatient”.

3.3 Mapping the Transmission Data Sets to HL7 CDA Schema

The Transmission Data Sets are mapped to HL7 CDA to create the “Transmission Schemas”. For this purpose, first, three new code systems are created to give a unique code for each of the artifacts:

1. NHIS-T Document Type Code System (DocumentType-CS) gives the codes for the “Transmission Schemas”.
2. NHIS-T Data Set Code System (DataSet-CS) contains the codes for data sets in MHDS (Minimum Health Data Sets).
3. NHIS-T Data Section Code System (DataSection-CS) specifies the codes for data elements in NHDD (National Health Data Dictionary).

```
<examination classCode="DOCCLIN" moodCode="EVN">
   <id root="2.16.840.1.113883.3.129.2.1.3" extension="11333439-
08ab-42c5-ec2d-17064c153456" />
   <code code="MUAYENE" codeSystem="2.16.840.1.113883.3.129.2.2.1"
codeSystemName="Döküman Tipi" codeSystemVersion="1.0"
displayName="Muayene MSVS (Vatandaş/Yabancı)" />
```

Figure 4 The Beginning of the Examination Transmission Schema

During the mapping of the Transmission Data Sets to HL7 CDA, some modifications have been made to the original HL7 CDA XML Schema Definition (XSD), which broke the full conformity of Transmission Schemas to CDA. However, these modifications are mostly restricted to the renaming of XML tags, which can be easily converted back to the original schema [24].

Each “Transmission Schema” is wrapped with a root element named after the main data set in the transmission. For example, as shown in Figure 4, the root tag of the “Examination Transmission Schema” is <examination>, which maps to the root <ClinicalDocument> tag in original CDA Schema. In this example, the document type (“Doküman Tipi” in Turkish) is “Examination” (“MUAYENE” in Turkish) and this code is obtained from the DocumentType-CS whose object identifier (OID) is “2.16.840.1.113883.3.129.2.2.1”.

The “Minimum Health Data Sets (MHDSs)” in the “Transmission Data Sets” correspond to the first-level “sections” in the CDA structured body. The name of the opening tag of the MHDS is obtained by concatenating the name of the dataset with the “Dataset” keyword. The “code” of a MHDS is retrieved from the DataSet-CS. For instance, in the “Examination Transmission Schema” excerpt given in Figure 5, the first MHDS is given in <testResultDataset> block and the code is specified as “TestResult” (“TETKIKSONUCU”). This code is obtained from the DataSet-CS (“Veriseti”) whose OID is “2.16.840.1.113883.3.129.2.2.2”.

While mapping the Transmission Data Sets and the Minimum Health Data Sets to HL7 CDA, the various Association Core Components such as “subject”, “examination details” or “admission details” are mapped to the relationships that are allowed within the HL7 CDA Schema. The mapping of “subject” ASCC is easy; it directly maps to “recordTarget” participation of CDA Header. However, the established way of expressing other associations in the CDA Body is to use the “component” relationships among the sections. The meaning of the association together with the MHDS’ meaning is then represented through coded values as explained in the previous paragraph. An excerpt from the “Examination Transmission Schema” is given in Figure 5.

```
<examination classCode="DOCCLIN" moodCode="EVN">
  [CDA Header]
  <component typeCode="COMP" contextConductionInd="true">
      <structuredBody classCode="DOCBODY" moodCode="EVN">
         <component1 typeCode="COMP" contextConductionInd="true">
            <testResultDataset classCode="DOCSECT" moodCode="EVN">
               <code code="TETKIKSONUCU"
codeSystem="2.16.840.1.113883.3.129.2.2.2"
codeSystemName="Veriseti" codeSystemVersion="1.0"
displayName="Tetkik Sonucu Veriseti"/>
                ...
            </testResultDataset>
         </component1>
         <component2 typeCode="COMP" contextConductionInd="true">
            <dischargeDataset classCode="DOCSECT" moodCode="EVN">
               <code code="CIKIS"
codeSystem="2.16.840.1.113883.3.129.2.2.2"
codeSystemName="Veriseti" codeSystemVersion="1.0"
displayName="Çıkış Veriseti"/>
                ...
            </dischargeDataset>
         </component2>
         <component3 typeCode="COMP" contextConductionInd="true">
            <examinationDataset classCode="DOCSECT" moodCode="EVN">
               <code code="MUAYENE"
codeSystem="2.16.840.1.113883.3.129.2.2.2"
codeSystemName="Veriseti" codeSystemVersion="1.0"
displayName="Muayene Veriseti"/>
                ...
            </examinationDataset>
         </component3>
         <component4 typeCode="COMP" contextConductionInd="true">
            <receptionDataset classCode="DOCSECT" moodCode="EVN">
               <code code="KABUL"
codeSystem="2.16.840.1.113883.3.129.2.2.2"
codeSystemName="Veriseti" codeSystemVersion="1.0"
displayName="Kabul Veriseti"/>
                ...
            </receptionDataset>
         </component4>
```

```
            <component5 typeCode="COMP" contextConductionInd="true">
               <prescriptionDataset classCode="DOCSECT"
moodCode="EVN">
                  <code code="RECETE"
codeSystem="2.16.840.1.113883.3.129.2.2.2"
codeSystemName="Veriseti" codeSystemVersion="1.0"
displayName="Reçete Veriseti"/>
                    ...
               </prescriptionDataset>
            </component5>
         </structuredBody>
      </component>
</examination>
```

Figure 5 An Excerpt from the Examination Transmission Schema

Then, within the "sections" representing MHDSs, the NHDD data elements are represented by nesting the new "section" elements together with CDA Entry classes such as "observation", "procedure", "substanceAdministration" or "encounter", in which the actual values of data elements are placed. The details of this process are given in [25].

NHIS-T is open to extensions: When there is a need for a new EHR document, the existing Minimum Health Data Sets are re-used if possible; the new Minimum Health Data Sets are constructed by using the existent Data Elements and whenever the need arises, the NHDD is expanded by defining new Data Elements.

3.4 The Health Coding Reference Server (HCRS)

In order to make the common coding/classification systems available to all healthcare players, the MoH developed the Health Coding Reference Server (HCRS), which encapsulates all the international and national coding systems used in Turkey within a publicly accessible server.

Some of the coding systems available from HCRS are international such as ICD-10 or ATC (Anatomical Therapeutic Chemical Classification System) [26] and some are locally defined for certain sets of information such as Clinics, Patient Discharge Type, Pregnancy Result, or Baby Monitoring Calendar. The local coding systems are developed specifically for the data elements in the NHDD; apart from only a few narrative data elements or data elements like timestamps, there exists a coding system for each data element defined in the NHDD. There are 178 coding systems in total.

The current version of HCRS is 2.0 and is shared online through Web Services (through SOAP requests and responses) that are available from http://212.175.169.165/SKRSServis2/service.asmx. A tabular form that allows querying through Web browsers is also available from http://sbu.saglik.gov.tr/SKRS2_Listesi/.

All the software companies doing business in the Turkish health market are obliged to use the codes from the HCRS in developing their products.

3.5 The Healthcare Professional Registry

Ministry of Health is authorized to provide the work licenses to the physicians in Turkey. The diploma/specialty information of the medical professionals are recorded together with their Turkish citizenship numbers in the Doctor Data Bank (DDB). As of

June 2010, there are 187,181 registered physicians in the DDB. This includes all the physicians who obtained a license since 1923. The DDB is accessed to verify the identity of the physicians in the Transmission Schemas.

4 NHIS-T Communication Infrastructure

A “Transmission Schema” instance constitutes a message’s payload. In other words, “Transmission Schema” instances should be encapsulated in the messages to be transported. For this purpose, “HL7 Transmission and Control Act Wrapper” is used. The transmission wrapper provides information on the ID, creation time, sender and receiver of the message.

HL7 Version 3 provides three transport specifications [13] - ebXML, Web Services and Minimal Lower Layer Protocol (MLLP) - for the exchange of HL7 based content, messages and documents. Among them, Web Services Profile [14] is the most promising, as it is based on widely used Web Services Technology. Therefore, in NHIS-T implementation, Web Services Profile is used for the communication infrastructure. The Basic Profile and the Security Profile of Web Service Profile have been implemented. For security, WS-Security Username Token Profile [15] over Secure Sockets Layer (SSL) [16] is used.

There is an HL7 Web Service for each Transmission Schema, and almost for each HL7 Web Service there are four operations; namely “Insertion”, “Update”, “Deletion” and “Query”. All of these individual operations are synchronous, but overall, NHIS-T behaves asynchronously. The clients, which are the vendor applications deployed at healthcare organizations’ premises, send their “Transmission Schema” instances via the insertion operation that performs only syntax validation against the related schema and responds with an acknowledgement about the result of the validation. If the invocation of the first operation is successful, then the message is stored at NHIS-T servers for semantic validation of the content of the message as explained later. Then, at any time, the client invokes the query operation to query the result of the semantic validation using the Universally Unique Identifier (UUID) of the original document submitted. If the semantic validation of the document is successful, a positive acknowledgement is received from the query operation; otherwise, the errors encountered in the semantic validation phase are received. The error messages are well-documented.

As a summary, during syntactic validation phase, NHIS-T Web Services report the errors to the Hospital Information Systems (HISs) that sent the message. During semantic validation phase, on the other hand, the errors are queried from NHIS-T by the HISs that sent the message. The error handling practices can vary among hospitals, but usually, the IT staff of the hospitals examine the errors daily and resolve the technical ones (e.g. syntactic errors) themselves. For errors related with the medical content, the responsible healthcare professionals are informed. After corrections, the corrected EHR documents are resent to NHIS-T Web Services by the IT staff. The clients are also able to update and delete the previously inserted documents with the help of update and delete operations.

In addition to the HL7 Web Services, there are 16 more proprietary Web Services. These are native synchronous Web Services that are usually developed for some risky communicable diseases. “Malaria Notification” and “Tuberculosis Notification” are two such examples. They have Web-based forms on the MoH central servers as well, and it is expected that healthcare professionals will send these comparably rare observations through these forms.

4.1 Validation of the Transmission Schema Instances

A two-phase validation technique is applied to validate the incoming messages to NHIS-T. In the first phase, which is called syntax validation phase, an incoming document instance is validated against the related XML Schema Definition (XSD) of the "Transmission Schema" by the insertion or update operations. If successful, the message is conveyed to the second phase, which is called the semantic validation phase.

The semantic validation phase checks the values of the data elements and the relationships between them. The semantic constraints are checked as follows:

1. *MERNIS Central Civil Registration System:* The patient identifiers in the messages are validated against MERNIS.
2. *Doctor Data Bank (DDB):* The identifiers of the healthcare professionals that appear in the messages are validated against DDB.
3. *Value formats:* Values of some data elements should conform to some specific formats. As an example, HL7 date for representing a day must be of the form YYYYMMDD.
4. *Coded elements:* The coded elements should have values from the Health Coding Reference Server (HCRS).
5. *Business rules:* There is a set of rules among the message elements, for example "the examination end date should be later than the examination start date". There are more complex clinical business rules as well. These rules are defined and documented in collaboration with the healthcare professionals and the administrative staff of the MoH.

5 The Testing and Deployment of NHIS-T

NHIS-T has been thoroughly tested in three different categories before deployment:

1. Functionality and semantic testing of NHIS-T Web services,
2. Conformance and interoperability testing of the HISs to be connected to NHIS-T, and
3. Performance testing of NHIS-T under heavy load.

5.1 Functionality and Semantic Testing of NHIS-T Web services

Functionality testing checks whether NHIS-T Web services' functionality conforms to the developed specifications. For this purpose, NHIS-T Web services are tested both at the transport and at the content layer. For the content layer, as already mentioned in Section 4.1, the Transmission Schema instances are validated by NHIS-T Web services. Therefore, this testing involves checking whether NHIS-T system itself correctly achieves two-phase validation. This is fundamentally a manual process where the Web services are invoked with thousands of EHR documents to check their response. Some of the tests included checking the conformity of the Web Services to the defined EHR XML schemas; control of mandatory, optional and conditional elements; control of coded elements against the HCRS; HL7 data type checks and the conformity of the EHR documents against the defined business rules. Functional testing was conducted between February 2008 and August 2008.

5.2 Conformance and Interoperability testing of the HISs

In addition to testing the server side of NHIS-T, the client side applications, that is, the Hospital Information Systems (HISs) from different vendors, are thoroughly tested as well. A generic testing framework, namely, TestBATN [27,28] is customized to test the HISs by developing the test scenarios specific to the requirements of NHIS-T. For this purpose, 200 test scenarios categorized under 25 test suites are developed, details of which are given in [29].

TestBATN is used both for conformance and the interoperability testing. In the conformance testing, TestBATN simulates NHIS-T Web Services; the clients send their EHR instances to TestBATN by simply changing their endpoints, and TestBATN generates a detailed test report listing all the problematic issues related with the communication and content. For interoperability testing, TestBATN eavesdrops the messages between the HISs and NHIS-T Web Services and reports any communication and content related problem to both of the parties.

MoH organized two connectivity marathons as in IHE Connectathons with participation of around 60 HIS vendors; one in July 2008 and the other one in December 2008. During these workshops more than 5000 test scenarios were executed through TestBATN by an average of 130 participants from 55 vendors over a 5-day period. Each vendor was supplied with a very detailed report on the test scenarios and the steps performed.

TestBATN has been online since June 2008 (http://www.srdc.com.tr/testbatn) and it is still being used by the vendors and the MoH staff.

5.3 Performance testing

Following functional testing, the performance of NHIS-T is tested under heavy load. Performance was conducted between September 2008 and December 2008. During this period, millions of EHR documents were exchanged with NHIS-T Web Services. These tests revealed that the system had some performance issues, such as memory, logging and response time problems. As a result, several enhancements of the services were made, most of which were software related.

5.4 Deployment

NHIS-T became operational on January 15, 2009. As of June 2010, 99% of public secondary and tertiary healthcare providers (1018 out of 1033) and 71% of private and university hospitals (411 out of 579) are integrated in NHIS-T. The average number of EHR instances that are sent by the Hospital Information Systems (HIS) and successfully recorded in NHIS-T has reached 1.5 million a day. There are already more than 200 million EHR instances in NHIS-T. The total number of connected nodes is 1429, and this number continues to increase as the remaining healthcare providers complete their integration with NHIS-T. Out of the 72 million citizens of Turkey, electronic healthcare records of 43 million citizens have already been created in the system.

For each NHIS-T Web service, there is only a single public endpoint address and all clients (Hospital Information Systems) send their EHRs to these addresses. The MoH maintains a server farm for handling these EHR transactions. There are three HL7 servers and three SOA (Service Oriented Architecture) servers that are responsible for accepting the messages from the clients, and for syntactical validation and load balancing. The number of CPUs of these servers can reach up to 64. In addition, there are four Domain Web Service Application Servers that perform semantic validation and persistence of the messages. There are two Oracle database nodes to manage persistent data.

Regarding actual usage, as mentioned previously, the number of successfully recorded EHR instances a day has already reached 1.5 million. After an insertion operation, the semantic validation results usually become available in about 10 minutes.

It should be noted that the data required by NHIS-T is automatically generated by the HISs and are sent by invoking NHIS-T Web services; hence the process is invisible to the physicians who are creating the EHR documents. The physicians continue using the legacy HISs as before; the HIS vendors have developed wrappers to collect the data from their own systems using the interface that NHIS-T has defined and send the data by invoking NHIS-T Web services.

The collected data is used for decision support for the high level management in the MoH. NHIS-T Decision Support System is based on Oracle Business Intelligence (BI) solutions and provides business intelligence reports at different levels such as healthcare professional, healthcare provider, county, province, region or country levels. It provides pre-defined queries but it is also possible to create customized queries. Example queries include reporting “diarrhea incidences in a single province in the previous month”, “the number of stillbirths in a province in the previous year” or “all cancer occurrences in the whole country grouped per county”.

Regarding incentives/penalties, the MoH published a regulation requiring all public/private secondary and tertiary care providers to send EHRs to NHIS-T. The healthcare organizations that are successfully connected to NHIS-T are publicly listed in the MoH Web portal to create an incentive.

6 Conclusions

NHIS-T has a centralized architecture rather than a distributed one. The main reason for this is that the system is being used not only for sharing the EHRs of patients, but also for running decision support systems on the collected data to provide input for the high level management in the MoH. Some countries prefer a distributed and federated architecture. A detailed discussion of European good practices is available in [30].

There is a number of factors that affected the successful implementation of NHIS-T. The MoH, being the national authority to decide on the eHealth standards in Turkey, was able to enforce the standards that have led to their fast adoption. The second factor is the use of UN/CEFACT Core Components Technical Specification (CCTS) approach that facilitated the understanding of rather complex EHR schemas by the software developers.

Building a national common data dictionary consisting of eHealth data elements and minimum health data sets has helped to clearly identify the meaning of data elements; giving their explanation in the local language and, more importantly, making it possible to share and re-use these components has turned out to be an effective enabler for nation-wide interoperability.

Finally, the comprehensive testing, especially the conformance and interoperability testing of vendor HIS applications to NHIS-T through an automated testing platform, namely, TestBATN, greatly contributed to the fast integration of the different HISs of the majority of the 65 vendors in Turkey.

While addressing the national level EHR interoperability may be relatively easy, it is much more difficult to do this in an international environment. MoH has recently joined the epSOS Project [31] consortium to be able to contribute to the ongoing efforts for the exchange medical summaries within the EU as well as the associated countries.

Acknowledgments

The work described in this article has highly benefited from the results of the "FP6 IST 027065 RIDE: A Roadmap for Interoperability of eHealth Systems in Support of COM 356 with Special Emphasis on Semantic Interoperability" Project [32] supported by the European Commission's eHealth Unit.

References

1. Kose I., Akpinar N., Gurel M., Arslan Y., Ozer H., Yurt N., Kabak Y., Yuksel M., Dogac A., "Turkey's National Health Information System (NHIS)", in the Proc. of the eChallenges Conference, Stockholm, October 2008, pp. 170-177.
2. ISO/TR 20514:2005 (Health informatics -- Electronic health record -- Definition, scope and context)
3. Clinical Document Architecture, Release [cited June 18, 2010]. Available from: http://www.hl7.org/v3ballot/html/infrastructure/cda/cda.htm
4. UN/CEFACT Core Components Technical Specification [cited June 18, 2010]. Available from: http://www.unece.org/cefact/ebxml/CCTS_V2-01_Final.pdf
5. The Central Civil Registration System (MERNIS), Turkish Ministry of the Interior, General Directorate of Civil Registration and Nationality [cited June 18, 2010]. Available from: http://www.nvi.gov.tr/English/Mernis_EN,Mernis_En.html
6. Doctor Data Bank, the Ministry of Health, Turkey [cited June 18, 2010]. Available from: http://sbu.saglik.gov.tr/drbank/drbilgis.aspx
7. The Systematized Nomenclature of Medicine (SNOMED) [cited June 18, 2010]. Available from: http://www.ihtsdo.org/snomed-ct/
8. Logical Observation Identifiers Names and Codes (LOINC) [cited June 18, 2010]. Available from: http://www.regenstrief.org/loinc/
9. READ Codes [cited June 18, 2010]. Available from: http://www.connectingforhealth.nhs.uk/systemsandservices/data/readcodes/
10. Medical Subject Headings (MeSH) [cited June 18, 2010]. Available from: http://www.nlm.nih.gov/mesh/2010/mesh_browser/MBrowser.html
11. International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10) [cited June 18, 2010]. Available from: http://apps.who.int/classifications/apps/icd/icd10online/
12. The Health Coding Reference Server (HCRS), the Ministry of Health [cited June 18, 2010]. Available from: http://sbu.saglik.gov.tr/SKRS2_Listesi/
13. HL7 Version 3 Standard: Transport Specifications Overview [cited June 18, 2010]. Available from: http://www.hl7.org/v3ballot/html/infrastructure/transport/transport-intro.htm
14. HL7 version 3 standard: Transport specification - web services profile, release 2 [cited June 18, 2010]. Available from: http://www.hl7.org/v3ballot/html/infrastructure/transport/transport-wsprofiles.htm
15. Web Services Security UsernameToken Profile 1.1, OASIS Standard Specification, 1 February 2006 [cited June 18, 2010]. Available from: http://www.oasis-open.org/committees/download.php/16782/wss-v1.1-spec-os-UsernameTokenProfile.pdf
16. Transport Layer Security (TLS) and Secure Socket Layer (SSL) [cited June 18, 2010]. Available from: http://en.wikipedia.org/wiki/Transport_Layer_Security
17. The National Health Data Dictionary (NHDD) of Turkey [cited June 18, 2010]. Available from: http://www.sagliknet.saglik.gov.tr/portal_pages/notlogin/bilisimciler/docs/usvs_sozluk_1.1.rar

18. The Family Medicine Model of Turkey [cited June 18, 2010]. Available from: http://www.ailehekimligi.gov.tr/
19. The Centralized Hospital Appointment System (CHAS) [cited June 18, 2010]. Available from (in Turkish): http://www.e-saglik.gov.tr/merkezi-hastane-randevu-sistemi-mhrs-id25-18.html
20. HL7 Clinical Document Architecture Release 1 [cited June 18, 2010]. Available from: http://www.hl7.org/implement/standards/cda.cfm
21. HL7 Reference Information Model [cited June 18, 2010]. Available from: http://www.hl7.org/v3ballot/html/infrastructure/rim/rim.htm
22. CEN EN 13606-1, "Health informatics – Electronic health record communication – Part 1: Reference model," European Committee for Standardization, Brussels, Belgium, EN 13606-1, 2007.
23. ISO/IEC 11179-4:2004 (Information technology -- Metadata registries (MDR) -- Part 4: Formulation of data definitions) [cited June 18, 2010]. Available from: http://www.iso.org/iso/catalogue_detail.htm?csnumber=35346
24. Yuksel M., Dogac A., Akdag R., Atbakan E., Hulur U. Automatic Transformation of Local EHRs to the International Standards: Application in Turkey's NHIS. eChallenges Conference, October 2009, Istanbul, Turkey.
25. Kabak Y., Dogac A., Kose I., Akpinar N., Gurel M., Arslan Y., Ozer H., Yurt N., Ozcam A., Kirici S., Yuksel M., Sabur E., "The Use of HL7 CDA in the National Health Information System (NHIS) of Turkey," in *9th International HL7 Interoperability Conference (IHIC)*, October 2008.
26. The Anatomical Therapeutic Chemical (ATC) Classification System, WHO Collaborating Centre for Drug Statistics Methodology [cited June 18, 2010]. Available from: http://www.whocc.no/atc/structure_and_principles/
27. Namli T., Aluc G., Dogac A., "An Interoperability Test Framework for HL7 based Systems", IEEE Transactions on Information Technology in Biomedicine Vol.13, No.3, May 2009, pp. 389-399.
28. Namli, T., Dogac, A., "Testing Conformance and Interoperability of eHealth Applications", Methods of Information in Medicine, Vol. 49, No.3, May 2010, pp. 281–289.
29. Namli T., Aluc G., Sinaci A., Kose I., Akpinar N., Gurel M., Arslan Y., Ozer H., Yurt N., Kirici S., Sabur E., Ozcam A., Dogac A., "Testing the Conformance and Interoperability of NHIS to Turkey's HL7 Profile", 9th International HL7 Interoperability Conference (IHIC) 2008, Crete, Greece, October, 2008, pp. 63-68.
30. RIDE - A Roadmap for Interoperability of e-Health Systems in Support of COM 356 with Special Emphasis on Semantic Interoperability, Deliverable D2.1.4 - European Good Practices [cited August 9, 2010]. Available from: http://www.srdc.metu.edu.tr/webpage/projects/ride/deliverables/RIDE_D2.1.4_EuropanGoodPractices_v1.2.doc
31. Smart Open Services for European Patients (epSOS), [cited August 9, 2010]. Available from: http://www.epsos.eu/
32. FP6 IST 027065 RIDE: A Roadmap for Interoperability of eHealth Systems in Support of COM 356 with Special Emphasis on Semantic Interoperability [cited September 30, 2010]. Available from: http://www.srdc.metu.edu.tr/webpage/projects/ride/