# Family History Information Exchange Services Using HL7 Clinical Genomics Standard Specifications

Amnon Shabo (Shvo), IBM Research Lab, Israel
Kevin S. Hughes, Massachusetts General Hospital, Partners Health Care, USA

## ABSTRACT

*A number of family history applications are in use by health care professionals (e.g., CAGENE, Progeny, Partners Health care Family History Program) as well as by patients (e.g., the US Surgeon General's Family History Program). Each has its own proprietary data format for pedigree drawing and for the maintenance of family history health information. Interoperability between applications is essentially non-existent. To date, disparate family history applications cannot easily exchange patient information. The receiving application should be able to understand the semantics of the incoming family history and enable the user to view and/or to edit it using the receiving applications interface. We envision that any family history application will be able to send and receive an individual's family history information using the newly created HL7 Clinical Genomics Specifications through the semantic Web, using services that will transform one format to the other through the HL7 canonical representation.*

*Keywords: Please provide*

## INTRODUCTION

The need to represent a patient's family history information associated with clinical and genomic data was introduced as a storyboard to the Clinical Genomics Special Interest Group (SIG) in HL7 (CG, 2005). The SIG develops HL7 standards

(HL7, 2005) to enable the exchange of interrelated clinical and personalized genomic data between disparate organizations (e.g., health care providers, genetic labs, research facilities, pharmaceutical companies). Agreed-upon standards to allow this exchange are crucial, as it is envisioned that the use of genomic data in health care practice will become ubiquitous in the near future. A few emerging cases for this include tissue typing, genetic testing (e.g., cystic fibrosis, BRCA1, BRCA2), and pharmacogenomics clinical trials. These cases are represented in the SIG storyboards, which have led to the development of the Genotype model as the basic unit of genomic data representation, focusing on a specific chromosomal locus.

It was determined that there was a set of basic information required to record a family history and to create a pedigree for the purpose of breast cancer risk assessment (Thull & Vogel, 2004). For each family member, this set included the information about his or her relationships to other members of the family and the information regarding his or her health. Relationship information included the type of relative, a personal identifier, and the identifier of the person's mother and father. Health information data included disease type, age at diagnosis, current age or age of death, genetic syndrome suspected, genetic test done, genetic test result as raw data, and interpretation of genetic test.

The explosion in our knowledge of genetics has increased our understanding of the hereditary basis of many diseases. While we present here the example of exchanging family history and risk information relative to breast cancer, we believe this model can be used for the exchange of any hereditary risk information.

An outline of the patient's family history is presented in Appendix A. Populating this data set with patient data results is the example shown in Table 1.

## Storyboard Presentation

The following fictitious scenario demonstrates the potential use of the semantic Web (Berners-Lee, Hendler & Lassila, 2001) in offering services of exchanging family history information. Note that this is an abridged version of the full presentation contained in the HL7 specifications (CG, 2005).

1. Martha Francis is 39 years old. Her mother had ovarian cancer and was found to have a deleterious BRCA1 mutation. She has two sisters, a husband, and a daughter. She is not of Ashkenazi Jewish descent.
2. She makes an appointment at a risk clinic. The clinic instructs her to use the Surgeon General's Family History (Yoon, 2002) Web-based tool to prepare for the visit. She brings up the Surgeon General's Family History tool and enters her family history.
3. She then sends the data to the risk clinic prior to her appointment with that clinic, where a CAGENE application receives the data.
4. The counselor at the risk clinic (nurse geneticist, nurse practitioner, genetic counselor, doctor, etc.) uses the CAGENE application (a pedigree drawing program that runs risk mod-

els), where the patient's family history already has been received. The counselor edits the data after confirming and clarifying various issues with the patient and adds additional information that was not entered at home.

5. The patient is considered to be at high risk, and she is told that she is a candidate for genetic testing. This includes a thorough discussion of the pros and cons of testing. The patient decides not to have testing and leaves.
6. The counselor at the risk clinic sends back the updated family history to the Surgeon General History tool, so that the patient can use it in future encounters, if needed.

In this fictitious scenario, both the Surgeon General and CAGENE programs use publicly available Web Services that transform the data to the format needed by the receiving application, if it is not yet complying with the HL7 specs. The various aspects by which this kind of storyboard relates to the current processes are described in Table 2, where the left column shows the current practice, and the right column shows the proposed improvements. The scenario described in Table 2 is more complex than the previous, as it includes the use of three more family history programs with their own data formats as well as results from a genetic testing facility.

## THE FAMILY HISTORY EXCHANGE MODEL

Following the previous analysis of a patient's family history outline as well as

*Table 1. Example of a cancer patient's family history instance*

| ID | Father ID | Mother ID | Relationship | Vital Status | Current Age or Age of Death | Disease | Age of Diagnosis | Genetic Test | Result | Interpretation |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 3 | 2 | Client-F | ALIVE | 47 | NONE | 0 | N/A | N/A | N/A |
| 2 | 5 | 4 | Mother | DEAD | 72 | Ovarian | 40 | N/A | N/A | N/A |
| 3 | 7 | 6 | Father | ALIVE | 75 | NONE | 0 | N/A | N/A | N/A |
| 4 | 0 | 0 | Maternal Grandmother | ALIVE | 98 | NONE | 0 | N/A | N/A | N/A |
| 5 | 0 | 0 | Maternal Grandfather | ALIVE | 67 | NONE | 0 | N/A | N/A | N/A |
| 6 | 0 | 0 | Paternal Grandmother | ALIVE | 78 | NONE | 0 | N/A | N/A | N/A |
| 7 | 0 | 0 | Paternal Grandfather | ALIVE | 87 | NONE | 0 | N/A | N/A | N/A |
| 8 | 3 | 2 | Sister | DEAD | 67 | Ovarian | 60 | N/A | N/A | N/A |
| 9 | 3 | 2 | Sister | DEAD | 55 | Ovarian | 80 | N/A | N/A | N/A |
| 1 | 0 | 0 | Husband | ALIVE | 57 | NO | 0 | N/A | N/A | N/A |
| 11 | 1 | 1 | Daughter | DEAD | 33 | Breast | 30 | BRCA1 | 185delAG | DELETERIOUS MUTATION |

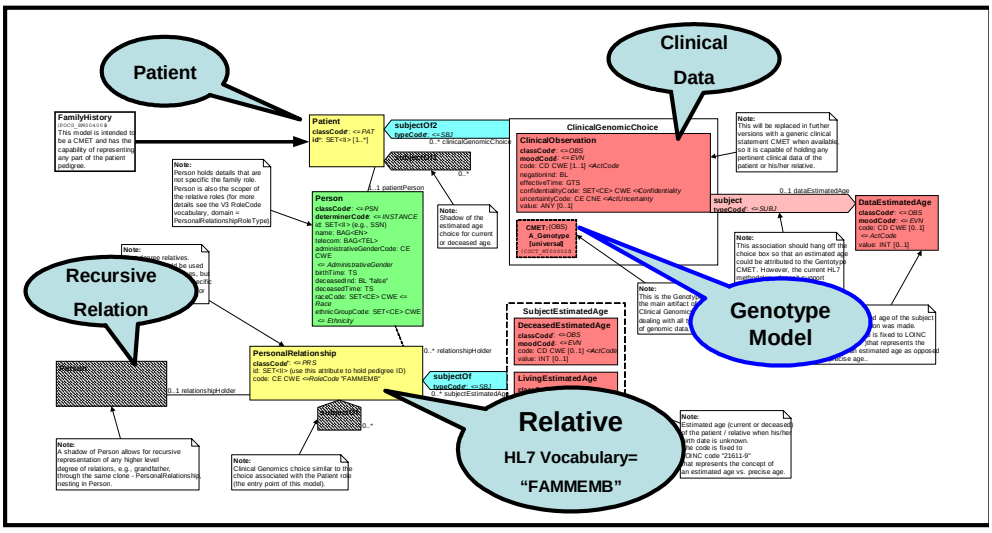
*Table 2. How semantic interoperability of family history information can improve current practices*

| CURRENT MEDICAL APPROACH | ENVISIONED APPROACH |
|---|---|
| Martha Francis is a 39-year-old woman with ovarian cancer. She has a family history of breast and ovarian cancer and believes she may be carrying a BRCA1 or BRCA2 mutation (which predisposes to breast and ovarian cancer). | Martha Francis is a 39-year-old woman with ovarian cancer. She has a family history of breast and ovarian cancer and believes she may be carrying a BRCA1 or BRCA2 mutation (which predisposes to breast and ovarian cancer). |
| She downloads the Surgeon General's Family History tool onto her computer at home and enters her family history. | She uses the Surgeon General's Family History Web-based tool from her home and enters her family history. |
| She then prints out her information onto paper and brings the paper to her clinician. | She then sends the data to her clinician (the Surgeon General's tool uses Web Services to export its data to the HL7 format and then to transform it to the clinician's system format). |
| Her clinician types the information from the paper into the homegrown electronic medical record (EHR). | Her clinician is able to see her family history as part of the homegrown electronic health record (EHR) system used in the clinician's office. |
| The clinician reviews the family history with the patient and makes corrections and additions in the EHR. | The clinician reviews the family history with the patient and makes corrections and additions in family history information of the patient's EHR. |
| The patient is considered to be at high risk of having a mutation, and this information is given to her. | The patient is considered to be at high risk of having a mutation, and this information is given to her. |
| She is referred to a risk clinic. | She is referred to a risk clinic. |
| Francis' family history details are printed on paper and sent to the risk clinic. | Francis' family history details are sent to the risk clinic (the clinician's system uses Web Services to export its data to the HL7 format and then to transform it to the risk clinic's required format). |
| The counselor at the risk clinic (nurse geneticist, nurse practitioner, genetic counselor, doctor, etc.) types the data into a number of programs: (1) Progeny (Progeny, 2005) to draw a pedigree; (2) CAGENE (CaGene, 2005) to run risk models; and (3) a homegrown Microsoft Access database to hold various and sundry other data. The counselor then reviews the family history information collected by the primary clinician, edits it, reviews results of the risk model algorithms, decides what genetic syndrome her family might have, and categorizes the patient as to degree of risk. | The counselor at the risk clinic (nurse geneticist, nurse practitioner, genetic counselor, doctor, etc.) imports the patient's family history information into a number of programs: (1) Progeny (Progeny, 2005) to draw a pedigree; (2) CAGENE (CaGene, 2005) to run risk models; and (3) a homegrown Microsoft Access database to hold various and sundry other data. The counselor then reviews the family history information collected by the primary clinician, edits it, reviews results of the risk model algorithms, decides what genetic syndrome her family might have, and categorizes the patient as to degree of risk. |
| The counselor speaks with the patient and adds additional information to the databases. | The counselor speaks with the patient and adds additional information to the databases. |
| If there have been any changes or additions to the family history, the counselor runs the computer models again. | If there have been any changes or additions to the family history, the counselor runs the computer models again. |

*Table 2. How semantic interoperability of family history information can improve current practices (cont.)*

| CURRENT MEDICAL APPROACH | ENVISIONED APPROACH |
|---|---|
| The patient is considered to be at high risk, and she is told she is a candidate for genetic testing. This includes a thorough discussion of the pros and cons of testing. | The patient is considered to be at high risk, and she is told she is a candidate for genetic testing. This includes a thorough discussion of the pros and cons of testing. |
| The order for testing is issued, and the family history information is included with the lab requisition (required by the testing laboratory). The family history information is transcribed by hand onto a paper lab requisition, which is sent to the testing facility along with a blood sample. | The order for testing is issued, and the family history information is included with the lab requisition (required by the testing laboratory). Data are sent to the testing facility through the aforementioned family history Web Services, along with a delivery of a blood sample. |
| At the central testing facility, the family history data are typed into the database (homegrown). | At the central testing facility, the HL7 message received from the family history Web Services is imported into the database (homegrown). |
| Testing of the BRCA1 and BRCA 2 genes for mutations is undertaken. | Testing of the BRCA1 and BRCA 2 genes for mutations is undertaken. |
| The results are entered into the database. | The results are entered into the database. |
| Identified mutations are assessed for functional significance by determining if they are truncating (deleterious), or if they are irrelevant (no change in amino acid coded by that codon). All other mutations are compared to known mutations to determine if information is available on their functional significance. In this case, a mutation is identified in BRCA1 and the mutation is deleterious. | Identified mutations are assessed for functional significance by determining if they are truncating (deleterious) or if they are irrelevant (no change in amino acid coded by that codon). All other mutations are compared to known mutations to determine if information is available on their functional significance. In this case, a mutation is identified in BRCA1 and the mutation is deleterious. |
| The actual mutations and the assessment of functional significance are printed on paper, which is sent to the counselor. | The actual mutations and possibly the entire gene sequences as well as the assessment of functional significance are exported using the HL7 Genotype model, which is part of the family history standard specification. The Genotype model is known to clinical genomics Web Services that annotate the genomic data by the most updated knowledge and to associate it with the patient clinical history. The annotated results are sent to the counselor. |
| The counselor types the results into his or her databases, makes comments, and then prints a final report, which is sent to the primary provider and to the patient. | The counselor receives the results through his or her family history program and further annotates it. The counselor then sends the information to the primary provider and to the patient so both can update their records. As in all information exchanges thus far, this is seamlessly accomplished through publicly available Web Services that can transform all known family history formats through HL7 standards specifications. |

*Figure 1. A bird's eye view of the HL7 family history model*




the contextual storyboard presentation, we have developed an HL7 model to allow the representation of a pedigree with an unlimited depth of generations. The model addresses the storyboard requirements while making use of the Genotype model for embedding genomic data at any level of granularity available and needed.

The modeling effort is based on the new HL7 Reference Information Model (RIM) from which all HL7 V3 specs are derived (e.g., labs, pharmacy, clinical documents, clinical trials, etc.). The HL7 RIM (RIM, 2005) has four core classes that basically allow the representation of an *entity* playing a *role* that has a *participation* in an *act*. For example, a person is playing a role of a relative that has a participation in an observation act of clinical and genomic data. By using the dedicated HL7 tools, we created UML-like models, where we refined the core RIM classes and associated them in a way that represents a pedigree. We then were able to generate automatically an XML schema from these models and to experiment with family history information exchange.

Figure 1 shows a bird's eye view of the Family History model utilizing the Genotype model to represent optional genomic data for the patient and each of his or her relatives.

## UNDERLYING STANDARDS AND TECHNOLOGIES

The following sections describe standards and technologies that underlie the family history model and enable its potential use in the semantic Web.

### HL7 Standards as a Foundation for Semantic Interoperability

The HL7 standard specifications focus on the seventh layer of the ISO Open Systems Interconnection model (i.e., the

semantic level that defines the contents of the messages exchanged between disparate systems within and across enterprises). An outstanding example of such use of HL7 is the principle design of the UK NHS NPfIT (National Health Service – National Program for IT) (Williams et al., 2004). It is a radical approach to change the entire strategy for information service provision in England and Wales. The plan is to have a foundation layer of nationwide applications running over a new broadband infrastructure and exchanging information using HL7 version 3 messages extended and localized by the requirements of this program. Three main applications are built on these new infrastructures: (1) a national e-booking system that enables the patient to participate in where and when an appointment is made; (2) an electronic transmission of prescriptions that will enable prescriptions to be sent electronically between GPs and retail pharmacies; and (3) a national integrated patient care service. The national patient record system, comprising a medical snapshot of every patient, will be fed into a national spine on top of the IT infrastructure. The spine will link the full range of the IT services specified locally. As a result, electronic patient records will be held centrally and will be available from all parts of the NHS with improved debugging, duplication, and management facilities, compared to today's dispersed and fragmented systems. The UK multi-billion-dollar NPfIT implementation (started in 2004) is a quantum leap compared to the current health care situation in the UK, and it is all built around HL7 standard specifications — a crucial enabler of semantic interoperability on a large national scale. A similar initiative for National Health Information Infrastructure now is evolving in the US, although with different architecture — less centralized than the UK (ONCHIT, 2005). As in the UK, it also is built around health care standards as main enablers of semantic interoperability. More national health IT initiatives are going on these days around the globe, some of which are based on HL7 standards; for example, in Finland (Porrasmaa, 2004), the Netherlands (NICTIZ, 2005), Canada (InfoWay, 2005), and Australia (HealthConnect, 2005).

## Development Methodologies and Technologies

The development of HL7 specifications follows the HL7 Development Framework (HDF, 2005), a methodology that dictates the use of pure UML models to represent the domain analysis and activity in the storyboards of interest to each working group (e.g., laboratory, pharmacy, medical records, clinical genomics). After the completion of these domain-specific UML models to the satisfaction of the domain experts, the working group represents these models through the HL7 RIM building blocks, resulting in an HL7 Domain Information Model (a UML-like model with HL7-specific constraints). The latter then serves as a basis of creating several HL7 Message Information Models that can be serialized to Hierarchical Message Descriptions and organized into interaction sets. The aforementioned artifacts are being balloted and sent

to ANSI. For implementation purposes, the HL7 specifications could be translated to some implementable technology like XML. The resulting XML schemas are not considered part of the balloted content but rather one option of implementation, taking into account that at a later time, there might be new implementation technologies for the same standard specifications. Note that only HL7 Message Information Models can be translated to XML schemas, as they are serializable models as opposed to Domain Information Models that can be more complex, encompassing all relevant data and associations in the domain.

The process of creating HL7 specifications is facilitated by a suite of tools developed specifically for HL7; a drawing tool allows the designer to draw an HL7 model from a pallet of RIM core classes. It also allows the validation of the model and its storage in a design repository. Another tool serializes a Message Information Model into a Hierarchical Message Description (HMD) exported into an XML format. Finally, a schema generator, which is part of the HL7 XML ITS (Implementable Technology Specification), generates an XML schema out of the HMD representation.

## The *Genotype* Model

As aforementioned, the Family History model utilizes the HL7 Genotype model to carry genomic data relevant to the patient's family history. The Genotype model is intended to be used as a shared component in any HL7 specification that conveys genomic data. It embeds various types of genomic data relating to a chromosomal locus, including sequence variations, gene expression, and proteomics. Within the Genotype model, we have utilized existing bioinformatics markups that commonly are used by the genomic community (e.g., MAGE-ML for gene expression data or BSML for DNA sequences). Those bioinformatics markups represent the raw genomic data and are encapsulated in HL7 objects. On the other hand, only portions of the raw and mass genomic data are relevant to clinical practice. To that end, we have constrained the full-blown bioinformatics markup schemas and excluded areas that describe pure research data. More importantly, the Genotype model also includes specialized HL7 objects (e.g., SequenceVariation of SNP type) that hold those portions of the raw genomic data that seem to be significant to clinical practice. Those specialized objects have attributes that represent the essential genomic data along with optional annotation. They are populated through a bubbling-up process that dedicated applications carry out. The bubbling-up process should take into account the goals of clinical care, the patient-specific history, and the most current knowledge about relevant clinical-genomic correlations. Once populated, those specialized objects can be associated with Clinical Phenotypes, represented either internally within the Genotype model or elsewhere; for example, as diagnoses and allergies residing in the patient medical records.

Figure 2 shows a bird's eye view of the Genotype model, distinguishing between the encapsulating objects vs. bubble-up objects.

Figure 3 illustrates a possible use of the encapsulate and bubble-up paradigm in the case of family history data. The genetic testing lab sends raw genomic data encapsulated in the Genotype encapsulating objects (e.g., full sequencing of the BRCA1 gene, expressed with BSML and encapsulated in the Sequence object of the Genotype model). When this portion of the HL7 message arrives at the EHR system, it gets appended to the patient's family history. We then envision that specialized decision support services will parse the family history and bubble up those SNPs in the raw data that are most clinically significant to the goal of assessing patient risk, resulting in annotation and enrichment of the data to be more useful to clinical practice. Thus, we envision that services will not only enable the exchange of information from one proprietary format to the other but also leverage it to be more effective to the receiving user.

Note that several bubbling-up processes could be performed at the same time (e.g., different algorithms, ontologies, etc.) and in different times (e.g., when new discoveries become available and the same raw data can be interpreted differently). Therefore, it is important not to abstract away the raw genomic data of a specific patient but rather encapsulate it and make it available to any processes that attempt to associate it with clinical data and facilitate a clinical decision at the point of care.

## The Clinical Statement Model

The clinical data in the family history model might be represented using a shared model of a clinical statement, which is under development in HL7 by various working groups (e.g., the Structured Documents Committee, the Orders and Observations Committee, the Patient Care Committee). The clinical statement model provides the grammar of how various discrete acts (observations, procedures, substance administrations, etc.) are associated to generate a meaningful clinical statement.

## The EHR Functional Model

The HL7 EHR Functional Model (recently approved as a DSTU - Draft Standard for Trial Use) has a family history function stating the information found in Table 3.

It is expected that this draft standard will be mandatory in the near future and that every EHR system will comply with certain profiles derived from this standard. The US government is developing an incentives program (through the CMS, for example) to encourage providers to adopt EHR systems that comply with this standard (conformance metrics are being developed) as well as to encourage EHR vendors to offer their systems in accordance with the standards functions (Dickinson, Fischetti & Heard, 2003).

## The HL7 Health Care Services Standardization Effort

Further to the aforementioned EHR Functional Model, there is a new effort to define standard services for EHR systems, undertaken jointly by HL7 and OMG. This will lead to the realization of a semantic Web for health care, as functions defined

*Figure 2. A bird's eye view of the Genotype model (blue callouts point to encapsulating objects while yellow callouts point to bubbled-up objects)*

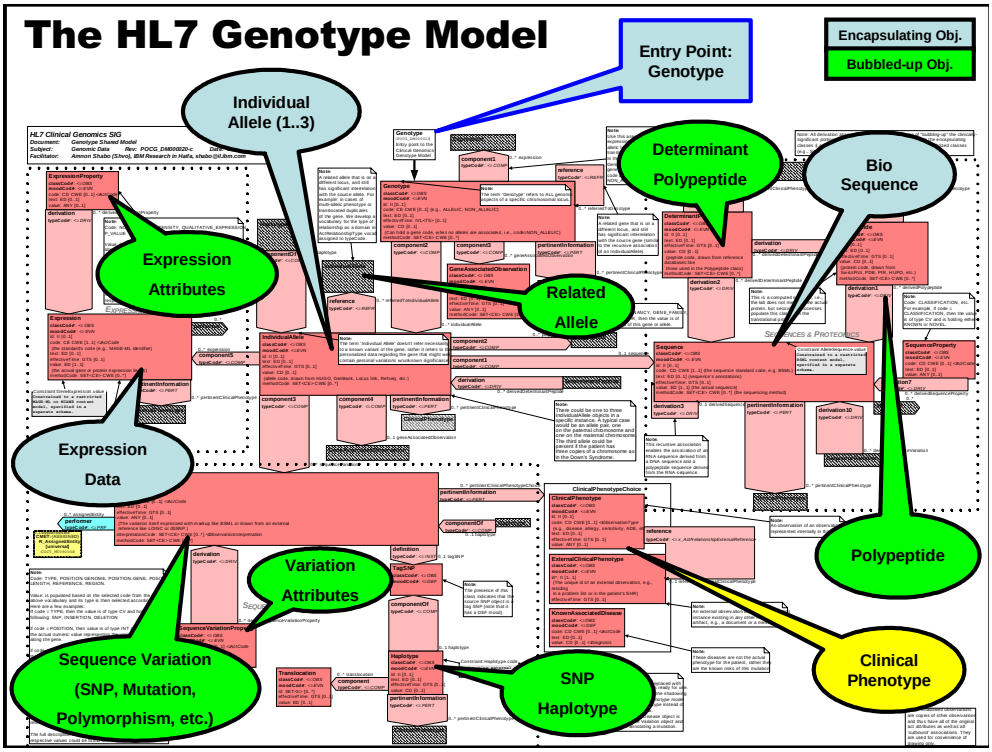



*Figure 3. Encapsulate and bubble up family history clinical genomics data*

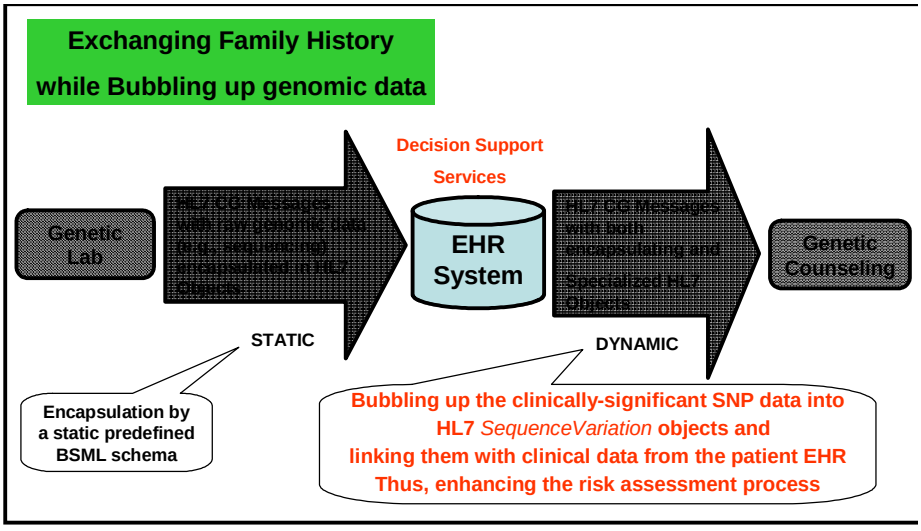


*Table 3. Family history function statements in the HL7 EHR functional model*

| **Function Name** | **Subject-to-Subject Relationship** |
|---|---|
| Function Statement | Capture relationships among patients and others to facilitate appropriate access to their health record on this basis (e.g., parent of a child), if appropriate. |
| Functional Description | A user may assign the relationship of parent to a person who is their offspring. This relationship may facilitate access to their health record as parent of a young child. |
| Rationale | Support delivery of effective healthcare; facilitate management of chronic conditions. |

in the functional model will be available as Web Services, either by the enterprise EHR system or by trusted third parties, who will provide Web Services for interoperability of EHR systems as well as for decision support and annotation.

In order to move forward in this direction, it is necessary for health care enterprises to implement a service-oriented architecture while establishing service offerings within such architecture. Realization of service-oriented health care interoperability is based on identifying the business functions and behavior being performed by a set of agreed-upon services as well as defining conformance metrics. This work is predicated on the availability of a robust semantic model describing precisely the information payload across organizations (Rubin, 2005).

## OWL and the HL7 Templates

The specifications developed thus far in HL7 are general-purpose specs and are not customized or constrained to specific requirements from the various clinical domains and their subspecialties. For example, there is a generic specification of a clinical document, but there is not a standard way of representing a discharge summary. The HL7 Templates SIG's mission is to address this issue of specialization and customization by offering mechanisms for constraining the generic specifications (Elkin & Kernberg, 2005). Template specifications eventually will constitute the majority of the HL7 standards, including specialized domains such as clinical genomics. One candidate formalism for expressing templates is OWL (Heflin, 2004), and once approved as a standard template mechanism (currently under ballot), the HL7 Clinical Genomics SIG will attempt to represent its specifications using OWL. Using OWL representations will allow designers to better constrain the generic specifications and to elaborate them in order to create a rich set of specialized data constructs in their clinical domain. The current HL7 methodology and tooling do not permit the structured representation of constraints over generic HL7 models. The current practice is to use a constraint box pointing to a class or attribute in the model and to specify the constraint in free text. This is a limited mechanism that will be improved once these balloted models can be constrained further using languages such as OWL. In addition, OWL representations will make the HL7 specs a better fit for the semantic Web.

## IMPLEMENTATION

We have implemented the Family History model by automatically generating the XML schema from the model and crafting XML samples with actual patients' family history data that validate that schema. We also added patient-specific BRCA sequences to illustrate the benefits of encapsulating raw genomic data in the context of BRCA risk assessment. We are now working with various stakeholders (owners of family history programs and diagnostic facilities, such as Myriad Genetics) to make their programs interoperable with the HL7 specification.

As an example, we show here fragments of a Family History sample XML that represents clinical and genomics data of a patient who has a mother and a father (each has two parents), two sisters, a husband, and a daughter. The full sample is presented in Appendix B, and it is recommended that you copy it into a separate file and open that file with a browser for best viewing of the structure as well as the content of the sample.

The XML instance starts with the *Patient* as the root element (see Table 4).

Note that the patient as well as each of her relatives has optional nested ID elements that identify their ID and role; for example, *NMTH* is an HL7 code that means *natural mother*.

The fragment in Table 5 describes the daughter of the patient who died of breast cancer. The genomic data appear first, identifying a specific allele of the BRCA2 gene.

The fragment in Table 6 shows an elaboration of that BRCA2 allele by encapsulating sequences from that allele that might consist of personal SNPs beyond those variations that identified this allele. Note that the DNA sequences below are presented for illustration purposes only and are not necessarily accurate.

Finally, the fragment in Table 7 shows a few of the SNPs from the BRCA2 allele, represented with BSML within the encapsulated object (using the *Isoform* element). In contrast, the element *derivedSequenceVariation* represents an object of the *Genotype* model that holds the results of the bubbling-up processing, picking on a specific SNP and representing it as a deleterious mutation. The mutation is then associated with clinical phenotypes (clinical observation from the patient medical records).

Note that only the observation id appears in this XML instance, because it is represented by the ISO OID (Object Identifier) standard, which ensures uniqueness of object identifiers across systems and organizations and thus enables services to resolve the location of an object such as this patient's diagnosis and to get it from where it is being stored.

The *value* element is the end of the encapsulation portion, because the raw genomic data are encapsulated in the value attribute of the HL7 *Sequence* object. The *reference* tag allows referencing back to the BSML *Isoform* element. This referencing enables the association between the bubbled-up object like the *185delAG* mutation to the encapsulated data as evidence. On the other hand, the mutation object is associated with clinical information to enable its usability within EHR systems.

*Table 4. Family history XML sample: The root element fragment*

```
<Patient xmlns="urn:hl7-org:v3" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="urn:hl7-org:v3POCG_MT004008.xsd">
  <id extension="555.001-SUBJ"/>
  <id extension="555.002-NMTH"/>
  <id extension="555.003-NFTH"/>
  <!-- PATIENT-->
  <patientPerson>
      <administrativeGenderCode code="F"/>
      <birthTime value="1957"/>
      …
```

*Table 5. Family history XML sample: The daughter fragment*

```
<!-- DAUGHTER-->
    <relationshipHolder>
        <id extension="555.011-SUBJ"/>
        <id extension="555.001-NMTH"/>
        <id extension="555.01-NFTH"/>
        <code code="DAU"/>
        <relationshipHolder>
            <deceasedInd value="true"/>
        </relationshipHolder>
        <!-- GENOMIC DATA-->
        <subjectOf>
            <Genotype>
                <component2>
                    <individualAllele>
                        <text>breast cancer 2, early onset</text>
                        <value code="U43746" displayName="BRCA2" codeSystemName
                                                                    ="HUGO"/>
        …
```

*Table 6. Family history XML sample: The sequencing fragment*

```
<sequence>
  <code code="BSMLcon3"/>
    <value>
      <Definitions>
        <Sequences>
          <Sequence id="seq1" molecule="dna"
              ic-acckey="U14680 REGION: 101..199"
            db-source="GenBank" title="BRCA1, exon 2" representation="raw"
            local-acckey="this could be used by the genetic lab">
              <Seq-data>
                GCTCCCA CTCCATGAGG TATTTCTTCA
                CATCCGTGTC CCGGCCCGGC CGCGGGGAGC CCCGCTTCAT
                CGCCGTGGGC TACGTGGACG ACACGCAGTT CGTGCGGTTC
                GACAGCGACG CCGCGAGCCA GAGGATGGAG CCGCGGGCGC
                CGTGGATAGA GCAGGAGGGG CCGGAGTATT GGGACCAGGA
                GACACGGAAT GTGAAGGCCC AGTCACAGAC TGACCGAGTG
                GACCTGGGGA CCCTGCGCGG CTACTACAAC CAGAGCGAGG
                CCG
              </Seq-data>
          </Sequence>
          …
```

*Table 7. Family history XML sample: The bubbling-up fragment*

```
            <Isoforms>
                <Isoform-set>
                    <Isoform id="SNP123" seqref="seq1" location="9" change="T"/>
                        <Isoform id="SNP456" seqref="seq1" location="32" change="C"/>
                        <Isoform id="SNP789" seqref="seq2" location="124" change="G"/>
                    </Isoform-set>
            </Isoforms>
        </Definitions>
    </value>
    <methodCode code="SBT"/>
    <derivation4>
        <derivedSequenceVariation>
            <code code="DNA"/>
            <text>
                <reference value="#SNP456"/>
            </text>
            <!--MUTATION-->
            <value xsi:type="CE" code="185delAG"/>
            <interpretationCode code="DELETERIOUS"/>
            <pertinentInformation>
                <pertinentClinicalPhenotype>
                    <reference typeCode="SUBJ">
                        <referredToExternalClinicalPhenotype>
                            <id root="2.16.840.1.113883" extension="diagnosis1"/>
                        </referredToExternalClinicalPhenotype>
                    </reference>
                </pertinentClinicalPhenotype>
            </pertinentInformation>
  …
```

## CONCLUSION

The vision of the semantic Web could play a major role in the current efforts to achieve interoperability of disparate health information systems. In this article, we have focused on the use case of exchanging family history data, which is crucial for breast cancer patients. In particular, elaborated family history with raw genomic data is becoming more important as clinical-genomics correlations are now a standard part of modern health care.

While we have focused on breast cancer family history, this model has the generalizability to be utilized to exchange family history information for any hereditary condition. Hereditary conditions (benign or malignant) tend to be defined by the number of relatives with a condition or conditions, the age at which those conditions occur, and the closeness of that relative to the patient (degree of relative). In addition, if genetic testing is undertaken, the genetic mutation is discovered. Our model collects this information in a uniform format.

For example, let us consider a family suspected of having hemochromatosis (Toland, 2000), a benign condition that causes the cells of the body to retain more iron than they need, which can lead to diabetes, heart disease, liver disease, and ar-

thritis. The important information would be the presence of liver disease, diabetes, and so forth in various relatives, the age of onset of these conditions in each relative, the bloodline and the ability to show the relative in a pedigree, and whether genetic testing was done, which test was done, what the actual result was, and what the interpretation is. All these data items have placeholders in the HL7 model described in this article.

As a more abstract conception, in thinking about other indicators of hereditary conditions, we can conceive of a condition where the genetic test is not available or was not performed, but a laboratory test might give useful information. In our model, laboratory test results can be transmitted for each individual and can be displayed on a pedigree or run through a model. Some conditions may require the interaction of multiple genes, and our model allows the representation of multiple genetic test results for each relative.

Whatever condition is suspected or whatever data is collected, our model allows the transmission of this data from clinician to clinician as a message and, in addition, allows this information to be used to draw pedigrees and to run computer models of risk.

The importance of drawing a pedigree should not be underestimated. A pedigree is a graphical display of the family history information that allows the clinician to visualize the diseases and the relationships, and thus to be able to interpret the data better. In addition, the data are in a format that can be imported easily into computer models of risk in order to provide quantitative analyses of the likelihood of the condition for various family members.

Our HL7 model allows the transmission of all pertinent information for any hereditary condition of which we currently can conceive and has the flexibility to be extended to future, more complex genetic conditions.

We envision the use of services based on health standards over the Web that various family history specialized applications will be able to use to seamlessly exchange family history data. These services will be part of the entire set of health services being defined by major standardization bodies such as HL7 and OMG. In the development of the Family History model, we used the HL7 development methodology and the HL7 dedicated tooling. We thus defined the semantics of the payload of family history services, and as the aforementioned health services technical framework becomes available, our Family History model could be utilized by those services as the domain-specific ontology.

## ACKNOWLEDGMENT

The work described in this article has been carried out partly within the HL7 Clinical Genomics Special Interest Group, and the authors wish to thank its members for their contributions and reviews of the Family History specification.

## REFERENCES

Berners-Lee, T., Hendler, J., & Lassila, O. (2001, May). The Semantic Web. *Scientific American*.

Bray, T., Paoli, J., Sperberg-McQueen, C., Maler, E., & Yergeau, F. (2004, February). Extensible markup language (XML) 1.0 (3rd ed.) (W3C Recommendation).

CAGENE (CancerGene). (2005). Retrieved April 1, 2005, from The University of Texas Southwestern Medical Center at Dallas at http://www3.utsouthwestern.edu/cancergene/cancergene.htm

CG. (2005). Clinical genomics domain. V3 ballot package. Retrieved April 1, 2005, from http://www.hl7.org

Dickinson, G., Fischetti, L. & Heard, S. (2003). Functional model and standard, draft standard for trial use, Release 1.0. Retrieved April 1, 2005, from http://www.hl7.org

Elkin, P. L., & Kernberg, M. (2005). HL7 standard V3 template architecture 3.3. Retrieved April 1, 2005, from http://www.hl7.org

HDF. (2005). HL7 development framework. Retrieved April 1, 2005, from http://www.hl7.org

HealthConnect: A Health Information Network for all Australians. (n.d.). Retrieved July 2, 2005, from http://www.healthconnect.gov.au/index.htm

Heflin, J. (2004). OWL Web ontology language use cases and requirements (W3C Recommendation). Retrieved April 1, 2005, from http://www.w3.org/TR/webont-req/

HL7. (2005). *What is HL7?* Retrieved April 1, 2005, from http://www.hl7.org

InfoWay: Canada interoperable electronic health record (EHR) solutions and related telehealth development. (n.d.). Retrieved July 2, 2005, from http://www.infoway-inforoute.ca/home.php?lang=en

NICTIZ: National IT Institute for Health Care in the Netherlands. (n.d.). Retrieved July 2, 2005, from http://www.nictiz.nl

ONCHIT. (2004, July 21). *The decade of health information technology: Delivering consumer-centric and information-rich health care. Framework for strategic action*. Washington, DC: Department of Health & Human Services. Retrieved July 2, 2005, from http://www.hhs.gov/

Porrasmaa, J. (2004, October). *HL7 and common services in health information systems: Finnish experiences*. Paper presented at the 5th HL7 International Affiliates Meeting and 2nd International Conference on the Clinical Document Architecture, Acapulco, Mexico.

Progeny: A unique pedigree software (genogram) program. (2005). Retrieved April 1, 2005, from http://www.progeny2000.com/

Radhakrishnan, R., & Wookey, M. (2005). *Model driven architecture enabling service oriented architectures*. Retrieved April 1, 2005, from http://www.omg.org

RIM. (2005). Reference information model. Foundations V3 ballot package. Retrieved April 1, 2005, from http://www.hl7.org

Rubin, K. (2005). *Health care services specification (HSS) project: The impetus for collaboration between HL7 and OMG*. Retrieved April 1, 2005,

from http://www.hl7.org

Thull, D. L., & Vogel, V. G. (2004). Recognition and management of hereditary breast cancer syndromes. *The Oncologist, 9*(1), 13-24.

Toland E. A. (2000). *Hereditary hemochromatosis: Is genetic testing available?* Retrieved July 2, 2005, from http://www.genetichealth.com/HCROM_Genetic_Testing_for_Hemochromatosis.shtml

Williams, R., Bunduchi, R., Gerst, M., Graham, I., Pollock, N., Procter, R, Voß, A., et al. (2004, August). *Understanding the evolution of standards: Alignment and reconfiguration in standards development and implementation arenas*. Paper presented at the 4SEASST Conference, Paris, France.

Yoon, P. W., et al. (2002). Can family history be used as a tool for public health and preventive medicine? *Genetics in Medicine 2002, 4*(4), 304-310. Retrieved April 1, 2005, from http://www.cdc.gov/genomics/info/reports/research/famhist_yoon.htm

*Amnon Shabo (Shvo), PhD, works at IBM Research Lab in Haifa as a research staff member and is involved in various health care and life sciences projects. He is a co-chair and the modeling facilitator of the Clinical Genomics Special Interest Group in HL7 (an ANSI-accredited Standards Development Organization in health care), and a co-editor of the HL7 CDA Release 2 standard for clinical documents. In addition, Amnon specializes in longitudinal and cross-institutional Electronic Health Records (EHR). Amnon was the visionary and a co-author of the mEHR proposal made by a consortium of nineteen partners to the European Commission's Sixth Framework Program, based on his vision of Independent Health Records Banks for addressing the lifetime EHR sustainability challenge. In his academic studies he has built on his BSc degrees in biology, mathematics, and computer sciences to conduct research at the Hebrew University of Jerusalem in the areas of educational technology for biology and medicine, as well cognitive sciences and case-based reasoning through his postdoc research at the College of Computing of Georgia Institute of Technology in the USA. Amnon has extensive experience in education through 20 years of teaching basic programming to life sciences university students.*

*Dr. Hughes is a member of the Department of Surgical Oncology at the Massachusetts General Hospital where he is the surgical director of the Breast Screening Program, surgical director of the Breast and Ovarian Cancer Genetics and Risk Assessment Program, and co-director of the Avon Comprehensive Breast Evaluation Center. He is a graduate of Dartmouth College and Medical School, and trained at Mercy Hospital of Pittsburgh for general surgery. He did a fellowship in surgical oncology at the National Cancer Institute. Dr. Hughes is an assistant professor of surgery at Harvard Medical School and was formerly on the faculty*

*of Tufts University, the University of California, Davis and Brown University. Dr. Hughes has served on numerous national and regional committees and is actively involved in research regarding the genetics, screening, and diagnosis and treatment of breast cancer. He is the author of numerous papers and book chapters on the subjects of breast cancer, screening, diagnosis and treatment, and risk assessment.*

# APPENDIX A.

## Outlining Family History Data of a Breast Cancer Patient

| | | | | |
|---|---|---|---|---|
| Patient ID | | | | |
| | Relative type (Self) | | | |
| | Cancer | | | |
| | | Year diagnosed | | |
| | | Age diagnosed | | |
| | Genetic syndrome suspected | | | |
| | | Genetic test done | | |
| | | | Genetic test result specific | |
| | | | Genetic test result interpretation | |
| | Mother ID number | | | |
| | Father ID number | | | |
| | | | | |
| | Relative ID number | | | |
| | | Relative type (Brother, sister…) | | |
| | | Cancer | | |
| | | | Year diagnosed | |
| | | | Age diagnosed | |
| | | Genetic syndrome suspected | | |
| | | | Genetic test done | |
| | | | | Genetic test result specific |
| | | | | Genetic test result interpretation |
| | | Mother ID number | | |
| | | Father ID number | | |
| | Relative ID number | | | |
| | | Relative type (Brother, sister…) | | |
| | | Cancer | | |
| | | | Year diagnosed | |
| | | | Age diagnosed | |
| | | Genetic syndrome suspected | | |
| | | | Genetic test done | |
| | | | | Genetic test result specific |
| | | | | Genetic test result interpretation |
| | | Mother ID number | | |
| | | Father ID number | | |

# APPENDIX B.

## The Family History XML Sample

```
<?xml version="1.0" encoding="UTF-8"?>
<!--Sample of Family History model showing a flat version of a patient's pedigree as well as the ability to represent clinical and
genomic data of the patient and any of his or her relatives. The pedigree represented in this sample file is as follows:
Patient has a mother and a father (each has two parents), two sisters, a husband, and a daughter.

This file is valid against the schema that was generated using the HL7 Schema Generator with the input of the HMD resulting from
the Visio model with the Genotype model plugged in as a CMET, which, in turn, includes the BSML and MAGE-ML constrained
schemas for the raw genomic data.

For comments, please e-mail Amnon Shabo (Shvo) at shabo@il.ibm.com (IBM Research Lab in Haifa).
-->
<Patient xmlns="urn:hl7-org:v3" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="urn:hl7-org:v3
POCG_MT004008.xsd">
    <id extension="555.001-SUBJ"/>
    <id extension="555.002-NMTH"/>
    <id extension="555.003-NFTH"/>
    <!-- PATIENT-->
    <patientPerson>
        <administrativeGenderCode code="F"/>
        <birthTime value="1957"/>
        <!-- MOTHER-->
        <relationshipHolder>
            <id extension="555.002-SUBJ"/>
            <id extension="555.004-NMTH"/>
            <id extension="555.005-NFTH"/>
            <code code="NMTH"/>
            <relationshipHolder>
                <!-- The value 'true' means that this person is dead. Default value is 'false'-->
                <deceasedInd value="true"/>
            </relationshipHolder>
            <subjectOf>
                <clinicalGenomicChoiceClinicalObservation>
                    <!-- Ovarian Cancer observation of the patient's mother-->
                    <code code="V1043" codeSystemName="ICD" displayName="HX OF OVARIAN MALIGNANCY"/>
                    <!-- The following construct represents the estimated age at which the above diagnosis was made
                        (40)-->
                    <subject>
                        <estimatedAge>
                            <value value="40"/>
                        </estimatedAge>
                    </subject>
                </clinicalGenomicChoiceClinicalObservation>
            </subjectOf>
            <!-- The following construct represents the estimated deceased age (72)-->
            <subjectOf>
                <clinicalGenomicChoiceEstimatedDeceasedAge>
                    <value value="72"/>
                </clinicalGenomicChoiceEstimatedDeceasedAge>
            </subjectOf>
        </relationshipHolder>
        <!-- end of MOTHER data-->
        <!-- FATHER-->
        <relationshipHolder>
            <id extension="555.003-SUBJ"/>
            <id extension="555.006-NMTH"/>
            <id extension="555.007-NFTH"/>
            <code code="NFTH"/>
            <!-- The following construct represents the estimated age (75)
                Note that the code element will be fixed in the schema to the LOINC code below,
                so there is no need to send it in each instance, and it appears here for illustration purposes.-->
            <subjectOf>
                <clinicalGenomicChoiceEstimatedAge>
                    <code code="21611-9" displayName="ESTIMATED AGE" codeSystemName="LOINC"/>
                    <value value="75"/>
                </clinicalGenomicChoiceEstimatedAge>
            </subjectOf>
        </relationshipHolder>
        <!-- end of FATHER data-->
```

```
<relationshipHolder>
    <!-- MATERNAL GRANDFATHER -->
    <id extension="555.004-SUBJ"/>
    <code code="GRFTH"/>
    <subjectOf>
        <clinicalGenomicChoiceEstimatedAge>
            <value value="98"/>
        </clinicalGenomicChoiceEstimatedAge>
    </subjectOf>
</relationshipHolder>
<!-- end of maternal grandfather data-->
<relationshipHolder>
    <!-- MATERNAL GRANDMOTHER -->
    <id extension="555.005-SUBJ"/>
    <code code="GRMTH"/>
    <subjectOf>
        <clinicalGenomicChoiceEstimatedAge>
            <value value="67"/>
        </clinicalGenomicChoiceEstimatedAge>
    </subjectOf>
</relationshipHolder>
<!-- end of maternal grandmother data-->
<relationshipHolder>
    <!-- PATERNAL GRANDFATHER -->
    <id extension="555.006-SUBJ"/>
    <code code="GRFTH"/>
    <subjectOf>
        <clinicalGenomicChoiceEstimatedAge>
            <value value="78"/>
        </clinicalGenomicChoiceEstimatedAge>
    </subjectOf>
</relationshipHolder>
<!-- end of paternal grandfather data-->
<relationshipHolder>
    <!-- PATERNAL GRANDMOTHER -->
    <id extension="555.007-SUBJ"/>
    <code code="GRMTH"/>
    <subjectOf>
        <clinicalGenomicChoiceEstimatedAge>
            <value value="87"/>
        </clinicalGenomicChoiceEstimatedAge>
    </subjectOf>
</relationshipHolder>
<!-- end of paternal grandmother data-->
<!-- SISTER-->
<relationshipHolder>
    <id extension="555.008-SUBJ"/>
    <id extension="555.002-NMTH"/>
    <id extension="555.003-NFTH"/>
    <code code="SIS"/>
    <relationshipHolder>
        <deceasedInd value="true"/>
    </relationshipHolder>
    <subjectOf>
        <clinicalGenomicChoiceClinicalObservation>
            <!-- Ovarian Cancer observation of the patient's sister-->
            <code code="V1043" codeSystemName="ICD" displayName="HX OF OVARIAN MALIGNANCY"/>
            <subject>
                <estimatedAge>
                    <value value="60"/>
                </estimatedAge>
            </subject>
        </clinicalGenomicChoiceClinicalObservation>
    </subjectOf>
    <subjectOf>
        <clinicalGenomicChoiceEstimatedDeceasedAge>
            <value value="67"/>
        </clinicalGenomicChoiceEstimatedDeceasedAge>
    </subjectOf>
</relationshipHolder>
<!-- end of first SISTER data-->
```

```
<!-- SISTER-->
<relationshipHolder>
    <id extension="555.009-SUBJ"/>
    <id extension="555.002-NMTH"/>
    <id extension="555.003-NFTH"/>
    <code code="SIS"/>
    <relationshipHolder>
        <deceasedInd value="true"/>
    </relationshipHolder>
    <subjectOf>
        <clinicalGenomicChoiceClinicalObservation>
            <!-- Ovarian Cancer observation of the patient's sister-->
            <code code="V1043" codeSystemName="ICD" displayName="HX OF OVARIAN MALIGNANCY"/>
            <subject>
                <estimatedAge>
                    <value value="50"/>
                </estimatedAge>
            </subject>
        </clinicalGenomicChoiceClinicalObservation>
    </subjectOf>
    <subjectOf>
        <clinicalGenomicChoiceEstimatedDeceasedAge>
            <value value="55"/>
        </clinicalGenomicChoiceEstimatedDeceasedAge>
    </subjectOf>
</relationshipHolder>
<!-- end of second SISTER data-->
<!-- HUSBAND-->
<relationshipHolder>
    <id extension="555.01-SUBJ"/>
    <code code="HUSB"/>
    <subjectOf>
        <clinicalGenomicChoiceEstimatedAge>
            <value value="57"/>
        </clinicalGenomicChoiceEstimatedAge>
    </subjectOf>
</relationshipHolder>
<!-- end of HUSBAND data-->
<!-- DAUGHTER-->
<relationshipHolder>
    <id extension="555.011-SUBJ"/>
    <id extension="555.001-NMTH"/>
    <id extension="555.01-NFTH"/>
    <code code="DAU"/>
    <relationshipHolder>
        <deceasedInd value="true"/>
    </relationshipHolder>
    <!-- GENOMIC DATA-->
    <subjectOf>
        <Genotype>
            <component2>
                <individualAllele>
                    <text>breast cancer 2, early onset</text>
                    <value code="U43746" displayName="BRCA2" codeSystemName="HUGO"/>
                    <component1>
                        <sequence>
                            <!-- full sequence of the daughter's BRCA2 gene goes here so that applications
                            could look for more information such as SNPs that are not recognized as mutations.
                            (note that the actual sequences below are not accurate and are presented for
                             illustration purposes only) -->
                            <code code="BSMLcon3"/>
                            <value>
                                <Definitions>
                                    <Sequences>
                                        <Sequence id="seq1" molecule="dna" ic-acckey="U14680 REGION:
                                            101..199" db-source="GenBank" title="BRCA1, exon 2"
                                            representation="raw" local-acckey="this could be used by the
                                             genetic lab">
                                            <Seq-data>
```

```
                    GCTCCCA CTCCATGAGG TATTTCTTCA
                    CATCCGTGTC CCGGCCCGGC CGCGGGGAGC
                    CCCGCTTCAT CGCCGTGGGC
                    TACGTGGACG ACACGCAGTT CGTGCGGTTC
                    GACAGCGACG CCGCGAGCCA
                    GAGGATGGAG CCGCGGGCGC CGTGGATAGA
                    GCAGGAGGGG CCGGAGTATT
                    GGGACCAGGA GACACGGAAT GTGAAGGCCC
                    AGTCACAGAC TGACCGAGTG
                    GACCTGGGGA CCCTGCGCGG CTACTACAAC
                    CAGAGCGAGG CCG
            </Seq-data>
            </Sequence>
            <Sequence id="seq2" molecule="dna" ic-acckey="U14680 REGION:
                200..253" db-source="GenBank" title="BRCA1, exon 3"
                representation="raw" local-acckey="this could be used by the
                genetic lab">
                <Seq-data>
                    GTTCTCA
                    CACCATCCAG ATAATGTATG GCTGCGACGT
                    GGGGTCGGAC GGGCGCTTCC
                    TCCGCGGGTA CCGGCAGGAC GCCTACGACG
                    GCAAGGATTA CATCGCCCTG
                    AACGAGGACC TGCGCTCTTG GACCGCGGCG
                    GACATGGCGG CTCAGATCAC
                    CAAGCGCAAG TGGGAGGCGG CCCATGTGGC
                    GGAGCAGCAG AGAGCCTACC
                    TGGATGGCAC GTGCGTGGAG TGGCTCCGCA
                    GATACCTGGA GAACGGGAAG
                    GAGACGCTGC AGCGCACGG
            </Seq-data>
            </Sequence>
        </Sequences>
        <Isoforms>
            <Isoform-set>
                <!--The isoform tag in BSML can be used to represent an SNP.
                The 'seqref' attribute is used to refer to the sequence where the
                SNP occurs.
             (Note that the SNPs are not based on real data but rather were made
                up for illustration purposes only)-->
                <Isoform id="SNP123" seqref="seq1" location="9" change="T"/>
                <Isoform id="SNP456" seqref="seq1" location="32" change="C"/>
                <Isoform id="SNP789" seqref="seq2" location="124"
                                                        change="G"/>
            </Isoform-set>
        </Isoforms>
    </Definitions>
</value>
<!-- The following attribute belongs to the HL7 Sequence class and represents the
    sequencing method.
    Its vocabulary has not been nailed down yet, and several options are suggested
    in the Genotype documentation.-->
<methodCode code="SBT"/>
<derivation4>
    <derivedSequenceVariation>
        <code code="DNA"/>
        <text>
            <!-- The HL7 'text' attribute is of ED data type and this data type has a
                reference tag that allows the
                pointing to the BSML Isoform element.
                This referencing enables the linking between the bubbled-up
                object like this sequence variation one,
                to the encapsulated data in the Sequence class.-->
            <reference value="#SNP456"/>
        </text>
        <value xsi:type="CE" code="185delAG"/>
        <!-- The interpretationCode value should be drawn from the
            ObservationInterpretation vocabulary that doesn't have the
        DELETERIOUS value (abnormal is the closest)
        but has been proposed to RIM Harmonization in November 2004 and was
        accepted in principle.-->
        <interpretationCode code="DELETERIOUS"/>
```

```
                                        <pertinentInformation>
                                            <pertinentClinicalPhenotype>
                                                <!-- The use of the ID attribute populated with an OID value could
                                                 facilitate the access to the location where the actual instance of
                                                 the referred diagnosis resides (e.g., in the patient medical
                                                records)-->
                                                <reference typeCode="SUBJ">
                                                    <referredToExternalClinicalPhenotype>
                                                        <id root="2.16.840.1.113883" extension="diagnosis1"/>
                                                    </referredToExternalClinicalPhenotype>
                                                </reference>
                                            </pertinentClinicalPhenotype>
                                        </pertinentInformation>
                                        <derivation>
                                            <derivedSequenceVariationProperty>
                                                <code code="TYPE"/>
                                                <value xsi:type="CV" code="MUTATION"/>
                                            </derivedSequenceVariationProperty>
                                        </derivation>
                                    </derivedSequenceVariation>
                                </derivation4>
                            </sequence>
                        </component1>
                    </individualAllele>
                </component2>
            </Genotype>
        </subjectOf>
        <!-- CLINICAL DATA-->
        <subjectOf>
            <clinicalGenomicChoiceClinicalObservation>
                <!-- Ovarian Cancer observation of the patient's daughter-->
                <code code="V1043" codeSystemName="ICD" displayName="HX OF OVARIAN MALIGNANCY"/>
                <subject>
                    <estimatedAge>
                        <value value="30"/>
                    </estimatedAge>
                </subject>
            </clinicalGenomicChoiceClinicalObservation>
        </subjectOf>
        <subjectOf>
            <clinicalGenomicChoiceEstimatedDeceasedAge>
                <value value="33"/>
            </clinicalGenomicChoiceEstimatedDeceasedAge>
        </subjectOf>
    </relationshipHolder>
    <!-- end of DAUGHTER data-->
</patientPerson>
<!-- end of PATIENT data-->
</Patient>
```