Modeling Problem-oriented Clinical Notes*

F. H. J. M. Cillessen; P. F. de Vries Robbé
Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands

1. Introduction

In a Dutch university medical centre we are developing a patient-centered Electronic Health Record (EHR). In this project, Health Level 7 (HL7) is a leading factor. The project started in 2007 with the elicitation of requirements; we are now at the end of the implementation phase of this application.

Most of the clinical notes are progress notes, which should reflect the events during a patient’s care. Progress notes hold the medical facts, the reasoning, opinions and plans about the patient’s condition, demands and needs. They are key pieces of the EHR.

In the late 1960s, Weed published his ideas about the Problem-Oriented Medical Record (POMR) [1]. He also suggested that computerization is a necessity for the full POMR. The POMR itemizes and monitors each medical problem. Weed also published his ideas about the structuring of the progress notes under the headings Subjective data, Objective data, Assessment and Plan (SOAP) [2].

In the POMR, progress notes are linked to problems, and the record should contain a complete list of all the patient’s problems.

This “problem list” should be a dynamic “table of contents”, which can be updated at any time [1]. Weed also proposed that the list be separated into active and inactive problems. The term “problem” is used by Weed for medical and non-medical ‘diagnosis’, for ‘intermediate’ diagnosis and for aspects of worry. For the intermediate diagnosis one can think of the evolving idea about the underlying disease. The aspects of worry are just a list of things requiring attention.

The POMR and SOAP concepts offer a generic way to record progress notes. The purpose of this research is to provide a validated model for progress notes as a generic component of an Electronic Health Record (EHR) based on the Health Level Seven (HL7) Version 3 standard as a starting-point for the software engineering process.

2. Background

We use EHR as a generic concept, as defined in the ISO/TR 20514:2005 Health informatics – Electronic health record – Definition, scope and context [3]. The EHR in its basic generic form is defined as "a repository of information regarding the health status of a subject of care, in computer-processable form". Here 'subject of care' is used synonymously with 'patient'. Ideally, an EHR includes all the information derived from the 'subject of care' such as medical history, imaging, allergies, and - as the key piece – clinical notes for documenting patient morbidity, treatment, and care over time. The main objectives of an EHR are to improve quality and efficiency and to reduce errors and costs. Making EHRs interoperable should amplify these goals [4], though some reports suggest

* This work was performed at the Department of Medical Informatics and the Electronic Health Record Group of the Radboud University Nijmegen Medical Centre
otherwise [5]. Interest in improving health care quality, efficiency, and costs with health information technology has increased, but how institutions can achieve those benefits, and at what price, are unclear [6]. Information technology changes the nature of the work and has implications for organizational design. Knaup et al. assert that many changes are necessary to move from paper-based records to EHRs [7]. Along with the EHR, specifications must be developed that include the data elements, logic and definitions required to interpret the data unambiguously.

To maximize the benefit of an EHR an underlying structure is required [8]. According to Häyrinen et al. only a few papers have described the structure of the EHR in combination with the SOAP concept [9]. However, none of them describes how to model the POMR and SOAP concept with HL7.

Bayegan et al. stated that problem orientation is the foundation for a well-structured EHR that supports the way clinicians work, but they suggested a knowledge-based approach with no mandatory binding to the SOAP format [10]. Salmon et al. listed some limitations of the POMR, but also suggested some areas in which it may be extended [11]. De Clercq suggested a flexible link between data and problems to extend the POMR architecture [12]. The Institute of Medicine report recommends that an EHR should contain a problem list from which clinicians can easily obtain a concise view of all of a patient’s medical problems [13]. The “Meaningful Use” Regulation for EHRs is designed to ensure that healthcare providers adopt EHRs to achieve significant improvements in care [14, 15]. The ‘meaningful-use’ criteria underscore the significance of the POMR concept by the objective “Maintain up-to-date problem list of current and active diagnoses”.

HL7 International is the global authority on standards for interoperability of health information technology. One of the primary goals of HL7 Version 3 is to deliver standards that enable semantic interoperability. For people new to HL7 the most frequently asked questions can be found on their website [16]. HL7 Version 3 is based on an object-oriented data model. This model, the Reference Information Model (RIM), is the cornerstone of the HL7 Version 3 development process. HL7 also provides a methodology [17] which describes the process of message development through generation of models and graphs. The RIM is the top level model represented as a set of classes and is the ultimate source from which all HL7 Version 3 specification standards draw their information-related content. By restriction other models are derived finally to XML schemas for the exchange of messages. Thus HL7 as language has its own grammar specified by the RIM.

In this research we focus on the Act and ActRelationship as core classes of the RIM. Every happening is an Act and represents all of the actions documented by a health care professional. Acts are related through ActRelationships. The Act class is further represented by specialized classes such as Observation or Assessment. An ActRelationship is an association between a pair of Acts such as Component or Support.

The HL7 Clinical Statement Pattern is a refinement of the RIM, and is the common foundation for the Clinical Document Architecture (CDA) that provides an exchange model for clinical documents. CDA is an XML-based document markup standard that specifies the structure and semantics of a clinical document. The CDA supports the exchange of clinical documents and the re-use of clinical data. The CDA can contain any type of clinical content such as progress notes. There have been several attempts to use CDA as a predecessor of a more fine-grained structure [18]. The CDA Release 2 model is richly expressive, enabling clinical statements to be formally represented [19]. This research however, is focused on the relationship between constituting parts of a progress note and specially between progress notes by linking a progress note to previous recorded conditions that are part of an earlier progress note. Johnson et al. describe a structured-narrative EHR model based on the HL7 CDA standard [20]. Their proposed model uses CDA for clinical notes, allowing full freedom of expression and to reach some level of semantic interoperability. The CDA defines the structure and semantics of clinical documents however is not intended to represent relationships between the elements of different CDA messages, as is needed in our model.

2.1 Definitions

We avoid the term ‘problem’ because of its ambiguity. Since the introduction of the POMR there have been discussions about the term ‘problem’. In our view the term is used not only for diagnosis, but for all matters of worry that need attention (concern), as is the case for some medications, social status, or behavior.

Furthermore, the term ‘diagnosis’ is used ambiguously; it denotes the patient’s disease (condition) on the one hand, and a statement about this disease by the physician (assessment) on the other. For example, a physician can state that a disease is ruled out.

In ISO/TR 20514:2005 it is noted that differences in health paradigms or models are likely to lead to major differences in EHR content. In ISO/TR 20514:2005 the term ‘health problem’ is defined as a ‘health condition that results in some disability, pain and/or activity limitation’. Diagnosis is not defined in ISO/TR 20514:2005.

In the HITSP Glossary, version 2.0, July 8, 2009, the term ‘diagnosis’ is defined as ‘the condition for which the patient is seeking care’ [21]. Types of diagnoses range from early and indefinite to final and definitive. Early diagnoses may be vague, or even expressed in natural language rather than coded. The term ‘problem’ is not used in this glossary.

We use three key concepts with the following definitions:

Concern: any piece of information about the patient, captured at a point in time, that needs attention. The list of concerns can differ between one doctor and another. However, the underlying condition is patient-specific.

Condition (used synonymously with ‘health condition’ in ISO/TR 20514:2005); alterations or attributes of the health status of an individual that may lead to distress, interference with daily activities, or contact with health services; it may be a disease (acute or chronic), disorder, injury or trauma, or reflect other health-related states such as pregnancy, aging, stress, congenital
anomaly or genetic disposition. However, not all conditions are concerns.

Assessment: represents the clinician’s conclusions and working assumptions that will guide treatment of the patient. This definition is from the Implementation Guide for CDA Release 2.0 Consolidated CDA Templates [22], which contains a library of CDA templates, incorporating and harmonizing previous efforts from HL7, Integrating the Healthcare Enterprise (IHE), and Health Information Technology Standards Panel (HITSP).

3. Methods

3.1 Leading Design Criteria

In the transition from domain-specific paper-based to patient-centered computerized clinical notes we followed three leading design criteria:
1. The model must be generic and patient-centered.
2. The model must be based on POMR and SOAP.
3. The model must be derivable from the HL7 v3 RIM.

3.2 Elicitation of Requirements

3.2.1 Literature and Other Sources

We performed a literature search for the POMR, SOAP, progress notes and HL7, and looked at the huge amount of available data about HL7 and several HL7-related sources [23, 24].

3.2.2 Clinicians

We formed a discussion group, consisting mainly of health care professionals from different disciplines, and organized regular meetings. We had six whole group meetings in which a total of fourteen clinicians participated, from eight different disciplines.

3.3 Analysis and Design

As a guideline we used the HL7 methodology specification as written in the Healthcare Development Framework (HDF) document [17]. We used HDF version 1.2 published on March 29, 2008. The HDF is a framework of modeling and administrative processes to produce specifications to develop interoperability standards among health care information systems. The HDF uses the Unified Modeling Language (UML) [25] as the preferred syntax for visualizing and documenting the artifacts. After the project initiation process we started with the domain analysis followed by specification design. With the HDF as guideline, a set of interrelated artifacts was produced such as business requirements, story boards, workflows, use cases, interaction diagrams, business rules and truth tables depicting the static and behavioral aspects of problem-oriented progress notes. We started with the RIM version 2.18 and looked at the universal domains that were of most interest such as Care Provision, Clinical Statement and Observations. The Domain Message Information Models (DMIMs) and Refined Message Information Models (RMIMs) of these domains were used as references. From the HL7 Project Homebase site we downloaded several tools for our modeling activities [26]. The most important were RoseTree and the RMIM designer, which works with Microsoft Visio.

3.4 Validation

The model was validated by the development of a clinical notes application and by explaining the three leading design criteria. We started in April 2007, with several stakeholders, to discuss the development of this clinical notes application. We ended our functional simulation at the end of 2007. We discussed our underlying model, supported by mock-up screens, with clinicians and other stakeholders such as allied health professionals, scientists, jurists, nurses, secretaries and managers, as appropriate. In discussion with our database engineers and HL7 experts, we have iteratively improved our model. A working prototype of the clinical notes application using a database structure according to our model was released in spring 2008 and was tested by several clinicians to evaluate applicability and usability. Also, several groups of medical students tested our prototype with real closed cases in classroom sessions. Informatics students tested it for consistency and coherence. The results were discussed by small groups of clinicians. All the findings were used repetitively to improve our artifacts. A second prototype was launched at the end of 2008, and was demonstrated to small groups. In June 2009 two departments started a trial. They worked with the clinical notes application in daily practice. Various modifications were processed during the trial in the second half of 2009. The application was finally released in February 2010.

4. Results

4.1 Leading Design Criteria

4.1.1 The Model Must Be Generic and Patient-centered

By using the SOAP structure we provided a generic classification of elements within progress notes. We used no specialization-specific components. Many objects are modeled as pertinent information related directly to the patient. The progress notes, the list of conditions and the list of concerns are shared by all disciplines.

4.1.2 The Model Must Be Based on POMR and SOAP

Clinical notes can be related to one or more conditions that are part of an earlier progress note. Combined with the SOAP structure and the relationship between constituting parts of a progress note we can reproduce the clinical reasoning. For summaries our model provides the possibility to link observations to assessments and to plans as part of different progress notes.

4.1.3 The Model Must Be Derivable from the HL7 v3 RIM

We derived our model from the RIM. We used the DMIMs and RMIMs from the Care Provision domain. Our model is depicted with MS Office Visio and the RMIM designer. We only used the available objects provided by the RMIM designer. We dis-
cussed the model regularly with a certified HL7 v3 RIM specialist and the EHR database administrator. We used the observation act for assessment, and the condition act for a medical condition as a refinement of an observation act.

4.2 Elicitation of Requirements

4.2.1 Literature and Other Sources

The information gathered from the literature was used as input for discussion. The most important issues where:

  - L1a – Complex to maintain
  - L1b – Not all headings or diagnoses are ‘problems’
  - L1c – Single ‘problem’ consultation is rare
  - L1d – Information linkage to more than one ‘problem’
  - L1e – Linking ‘problems’ in a causal way
  - L1f – ‘Problems’ tends to go in cycles
  - L1g – Information need is different among clinicians
  - L1h – Some ‘problems’ are complex and difficult to read
- L2 – Fragmentation and lack of overview [27]
- L3 – Implications of coding [28]

For L1a we generate the list of conditions, and by using the assessment act this list can be easily maintained. For L1b, L1c and L1f we distinguish concern and condition. For L1d a progress note can be linked to more than one condition. For L1e our condition list can be hierarchical by linking conditions to other conditions. For L1g, L1h and L2 we provide different views and filtering options. For L3 coding of conditions is possible with ICD-10.

Our model invites the writer to structure the information and stimulates clinical reasoning. Because each condition is recorded separately and by the possibility of linking a progress note to one or more conditions, a condition can be used as an index so we can assume each discipline has recorded about one or more conditions in an organized way. Furthermore because of the relationships between constituting parts of a progress note it is easy to read all the observations that supports a condition and to read all the plans as a reason because of a condition.

4.2.2 Clinicians

We spoke iteratively and flexibly to many people from different domains, such as anesthesiology, surgery, cardiology, internal medicine, pediatrics, hematology, psychiatry, neurology, pulmonary diseases and urology.

The main functional needs of the clinicians were:

- F1 – Representation of different views
- F2 – Filtering of relevant data
- F3 – Critical data visibility
- F4 – Generation of different kind of letters
- F5 – Representation of diagnoses and differential diagnoses
- F6 – Revision and refinement of diagnoses
- F7 – The integration of different disciplines

For F1, F2 and F4 we provide different views and filtering options in our application based on the underlying model. For F3 we distinguish concern and condition. For F5, and F6 our condition list can be hierarchical by linking conditions to other conditions, and by using the assessment act this list can be easily maintained. For F7 many objects are modeled as pertinent information. Also the progress notes, the list of conditions and the list of concerns are shared by all disciplines.

4.3 Analysis and Design

For RIM we used version 2.18 from September 2007, 2.19 from March 2008 and 2.20 from April 2008. Besides the RIM, the most important models and documents from HL7 we studied were: Clinical Statement [29], Care Provision [30], Care Structures Topic [31], Observations [32], and Care Provision Domain Models: Explanation & Guidance [33]. The classes, attributes, state-machines and relationships in the HL7 v3 RIM were used to derive domain-specific information models. These models were further transformed through a series of constraining refinement processes, eventually to yield a static model of the information content. The related artifacts produced with the HDF as guideline and the references from HL7, especially the Clinical Statement Domain, ultimately resulted in our HL7 v3-based model of problem-oriented clinical progress notes.

The complete ProgressNote model is shown in Figure 1 as an HL7-based model. The ProgressNote as Organizer groups the SOAP items. The reason (RSON) for a progress note can be one or more previously-initiated Conditions. This according with the POMR concept. Subjective Observation and Objective Observation are the two forms of observations linked with a Component (COMP) ActRelationship to the ProgressNote. The observations, if stated together with an assessment, are linked to the Assessment act with a Support (SPRT) ActRelationship. The Assessment act is also a Component of the ProgressNote organizer. If, in the same ProgressNote, a Plan is stated, then the Plan act is linked through a Reason (RSON) ActRelationship with the Assessment act. An Assessment act is always linked with a Reference (REFR) ActRelationship to a Condition act. All the Condition acts can be related to each other because of the recursive Episode link (ELNK) ActRelationship. These relationships between conditions are used to refine conditions. This Episode link ActRelationship between conditions also enables us to build a hierarchical condition tree. All the Condition acts are linked with the Component ActRelationship to the ConditionList act. To specify a condition as a concern as part of a ConcernList, we use the Concern act with a Subject (SUBJ) ActRelationship to the Condition act. The concerns are linked by the Component ActRelationship to the ConcernList act. Every item in an episode of care can be linked with a Concern act to specify that it needs extra attention. Coding according to the model is done in the background based on the selected data entry fields. The ConditionList and ConcernList are WorkingList acts. A working list is a dynamic list of individual instances of an Act which reflect the needs of somebody to view groups of Acts for clinical or administrative reasons.
Thus, we designed a flexible and applicable paradigm by using acts for each assessment that refer to a condition, and by separating conditions from concerns, and finally, by using working lists to group acts for specific purposes.

The changes between differential diagnosis, working diagnosis and other types of diagnosis are gradual. Therefore, we chose not to model these types separately. The appearance of a condition in the hierarchical condition tree together with the assessments is used for the interpretation of a type. If a condition X has two or more subsequently hypothesized conditions, we interpret those subsequent conditions as a differential diagnosis. If there is only one subsequent hypothesized condition, we interpret that condition as a working diagnosis. A ruled-in condition is interpreted as a definite diagnosis. Because of the assessment act one can easily modify the interpretation by performing re-assessments.

It is often difficult to derive the intended functioning from a static model, especially when recursive relationships are in place. Figure 2 shows the coherence of the assessment, condition, concern and list acts. On the basis of Figure 2 we can explain several types of assessments, which are numbered from Assessment-1 through Assessment-6. As part of the progress note an assessment is recorded. In our model this is always an assertion about a condition. It can be the initiation of a hypothesized condition, but also an assertion about the condition, e.g. ruled in, ruled out or resolved. The arrows are ActRelationships linked to each assessment. If a condition, or another piece of information, needs attention, it is linked with a concern. A list can contain acts that form the content of the list. We specified the following lists:

- condition list
- concern list
- other working lists, such as allergy or history

Assessment-1 is the assertion of the presence of a Condition-1. Assessment-2 is the assertion that Condition-1 needs special attention. Assessment-3 is an assertion of a new Condition-2 as a refinement of Condition-1. Assessment-4 records a new Condition-3. Assessment-5 is the assertion that Condition-3 is subsumed under a refinement of Condition-1. Assessment-6 is the assertion that Condition-3 is placed on a WorkingList.

4.4 Validation

We have developed a fully functional prototype application according to our model that fulfills the explicit leading design criteria and the main functional needs of the clinicians. Furthermore, the clinical
notes application is now used throughout our hospital in daily clinical practice by more than 1000 physicians and 2500 other health care professionals.

We use a database that is RIM-based. Components and mutual relationships must be modeled properly to record and retrieve the right subsets of information. We have discussed the model with our database engineers and HL7 experts against the interface design and the program code that stores and retrieves the data. In this case we assume that we have also achieved sufficient technical validation.

In Figure 3 a screenshot of the "View cond" tab of the clinical notes application is shown. The conditions are displayed as a hierarchical tree from which one or more conditions may be selected. The radio buttons can be used to select which progress notes are to be displayed. Notes are displayed in relation to (1) the selected conditions, (2) the selected and previous, (3) the selected and following conditions, or (4) both directions. In Figure 3 the condition "shortness of breath" is selected. The selected radio button means that all the notes linked to this particular condition will be displayed.

5. Discussion

The RIM, as the ultimate source from which restricted models such as Clinical Statement and Care Provision are derived, leaves many degrees of freedom about how to model clinical notes. This can be seen as one of its strengths, but also as one of its weaknesses because it makes interoperability between systems complicated. An example pertinent to our interests was the Discussion on Condition Tracking [34]. It seems that, at this point, there is no consensus about these concepts. We also communicated our ideas to the HL7 in November 2008 [35]. In our approach we use condition, concern and assessment and avoid the concepts 'problem' and 'diagnosis'. In fact we implemented condition tracking and no concern tracking.

The level of detail that may be expressed in a model is also a major point. According to HL7 the Clinical Statement model is de-

![Fig. 2](image-url) This figure shows the dynamics of the assessment, condition, concern and list acts.
liberately broad and encompassing. Within the constraints of the Clinical Statement Pattern, similar clinical information can be represented in different ways. At this point it seems to be inconsistent or incomplete, and needs further research as performed in this case to contribute and reach consensus.

Within our medical informatics group we had many discussions about the static and abstract view of HL7 models to identify all the intentions. We found that the HL7 models were not useful for discussion with clinicians because they are too abstract and cannot be fitted into daily practice. These models are crucial, but must be targeted to people with a medical informatics background. Our experience is not to denigrate the work of HL7; on the contrary, we admire their work in striving for true semantic interoperability.

From the IHE wiki [36] we looked at 'problem entry', 'problem concern', 'concern entry', 'problem status observation', 'severity' and the 'Progress Note'. From the openEHR foundation [37] we looked at the archetype approach. Because our use of HL7 both sources were found to bring little or no value so were not included. We performed also a small-scale market research but that was only used as an indicator that oriented us generally.

By explicitly representing relationships in our model it is possible to derive different hierarchical structures. For example, we are able to create a problem-oriented view out of a series of progress notes. Also, by linking conditions, we are able to create a hierarchical condition list. Furthermore, our condition list is generated automatically from the conclusions in the progress notes. The user interface design is important for supporting the clinician's work and minimizing the cognitive load.

To abstract and classify facts from the health care domain, to represent them in objects and describe the mutual relationships, is often difficult because clear definitions of those objects and relationships are lacking. Therefore, during HL7 modeling activities, we faced a matching problem because it is not easy to choose the corresponding act classes, specializations and attributes. By separating assessments from conditions we provide an option to re-assess previously recorded conditions. We think this is necessary because a clinician can rethink previously recorded data, or can change his mind because of new data.

As we used ideas from the POMR and SOAP as starting points, other research concluded that this could lead to more time spent in documenting clinical work, fragmentation of problems, and a lack of overview [27]. Bossen also suggested that the POMR should be supplemented with chronological, source- and task-oriented views. The SOAP concept has also been criticized [38]. Tange et al. expect most benefit if progress notes are divided into problem

![Screenshot of the clinical notes application. The hierarchical condition tree with a differential diagnosis of "shortness of breath" is shown. Progress notes related to "shortness of breath" are listed.](Fig. 3)

Clinical notes are mostly expressed in natural language, which can be seen as the most comprehensive and practical way considering today’s situation. Surrounding words are used to describe the precise context and the patient’s history. We analyzed approximately 30 paper-based medical records from different disciplines in our hospital and found an enormous variation of content with little structure. The shortcomings of the paper-based records are visible in the main functional needs of the clinicians. The change of using terminology in software for patient-centered systems is often underestimated [28].

In our experience, the step is far too big for clinicians to go from domain-specific written narratives to a digital patient-centered fine granularity of recording discrete data, as described in several HL7 documents, in order to strive for full semantic interoperability.

Started as a communication standard, HL7 Version 3 is maturing towards an architectural approach to semantic interoperability [40]. Each code set is represented as a vocabulary domain, which is a set of all concepts that can be taken as valid in an instance of a coded field or attribute. We need standards to achieve interoperability, but also empirical projects to conclude how far this interoperability reaches [41]. Additionally, good solutions for the preservation of clinical meaning across heterogeneous systems remain to be explored [42]. Furthermore, due to a missing HL7 ontology, essential concept representations have not been adequately integrated [43]. At this point we need consensus about diagnosis, problems, conditions and concerns.

Kay and Purves introduced an abstract Narratological Model for Medical Informatics [44].

Our model can be seen as a way to document the ‘clinician’s story’. By using a combined hierarchical condition list for all disciplines, with the possibility of referring to progress notes, the ‘sickness story’ can be derived at any moment. The ‘clinician’s story’ is further enhanced by modeling assertions on conditions, and by modeling concerns separated from conditions.

We see POMR and SOAP as invariant concepts, proven by experience over 40 years. The POMR gives us a logical ordering principle for the excessive amount of data in the EHR for the benefit of all.

6. Conclusion

We have modeled problem-oriented clinical progress notes according to the HL7 standard and to the needs of medical practitioners. First and foremost we achieved this by modeling separate acts for assessments, conditions and concerns. Furthermore, we are able to interpret different types of diagnosis by the appearance of conditions in the hierarchical condition tree. To evaluate its usefulness we developed a clinical notes application using a database structured according to our model.

References

2. Weed LL. Medical Records, Medical Education and Patient Care; The Problem-Oriented Record as a Basic Tool. Cleveland, USA: Case Western Reserve University Press; 1969.
29. HL7 V3 CS, R1, HL7 Version 3 Standard: Clinical Statement Pattern, Release 1, Last Ballot: November
F. H. J. M. Cillessen, P. F. de Vries Robbé: Modeling Problem-oriented Clinical Notes