Progress Report to the Funding Agencies

For the Period September 1, 2000 to April 30, 2001
Grant R01 LM06594-01
"A Shared Internet Server for Delivering Guidelines"
National Library of Medicine,
US Army Telemedicine and Advanced Technology Research Command (TATRC),
and Agency for Healthcare Research and Quality

The InterMed Collaboratory
Columbia, Harvard, and Stanford Universities
(including the Brigham and Women's Hospital)

Principal Investigators:
Robert A. Greenes, MD, PhD
Vimla L. Patel, PhD, DSc
Edward H. Shortliffe, MD, PhD
Samson W. Tu, MS

Co-Investigators:
Aziz A. Boxwala, MD, PhD
Mor Peleg, PhD
Qing Zeng, PhD
Omolola Ogunyemi, PhD
Elmer Bernstam, MD, MS
Dongwen Wang, MPhil
Nachman Ash, MD
William L Salomon, MD
James Q Yin, MD
Jeeyae Choi
Jose F. Arocha, PhD
Timothy Branch, MA
Melissa Diermeier, BA

Introduction (Work done during July 1999 – August 2000)

Since July 1999, members of the InterMed Collaboratory have been working on the development of the third version of the GuideLine Interchange Format, dubbed GLIF3. We have been analyzing the previously published version, GLIF2 [Ohno-Machado, Gennari 1998], looking at examples of GLIF-encoded guidelines, studying other guideline-modeling approaches, such as Arden Syntax [Clinical Decision Support & Arden Syntax Technical Committee of HL7 1999] [Peleg, Bernstam 2000], EON [Tu and Musen 1999], PRODIGY [Sugden, Purves 1999], PRO*forma* [Fox and Rahmanzadeh 1998], and Prestige [Gordon and Veloso 1999], and examining various guidelines from the National Guideline Clearinghouse (http://www.guidelines.org). Through this analysis, we identified a number of areas in which GLIF2 needed to be extended in order to

support fully the encoding of computer-interpretable and executable guidelines. We identified additional areas of work as a result of the invitational workshop on Modeling Clinical Guidelines that was hosted in Boston by InterMed in March 2000. During the workshop, we were exposed to views from about 80 people (representing 10 different countries) from academia, government agencies, professional organizations (including the ACP-ASIM, which continues as an active collaborator as outlined below), providers, and industry.

During the period November 1999 through March 2000, we compiled a list of **functional requirements** for a shared guideline representation [Boxwala, Tu 2001, Peleg, Boxwala 2001]. The requirements are necessary for a guideline representation to be expressive, comprehensive, and sharable among (1) several institutions that differ in their health care settings and clinical information systems, and (2) different applications (e.g., decision support, quality assurance, education). By August 2000, GLIF3 supported many of these functional requirements [Peleg, Boxwala 2000a] [Peleg, Boxwala 2000b]. In the rest of this section, we describe the work that we did on GLIF3 leading up to August 2000.

GLIF2 enabled modeling of a guideline as a flowchart of structured steps, representing clinical actions and decisions. However, the attributes of structured constructs were defined as text strings that could not be parsed, and such guidelines could not be used for computer-based execution involving automatic inference. Thus, this representation is a human-readable flowchart. GLIF3 extends this flowchart representation from GLIF2 to enable computability and to facilitate implementation; hence it is a Computer-Interpretable Guideline (CIG) model. GLIF3 supports modeling of guidelines at three levels of abstraction: a **conceptual** flowchart (as in GLIF2), a **computable** specification that can be verified for logical consistency and completeness, and an **implementable** specification that can be incorporated into particular institutional information systems. This last level is currently only partly supported and is an area of ongoing work during the upcoming year of our grant (see next section). The separation between these three layers is important for conceptual understanding of guidelines as well as for sharing of their encoded versions, since different institutions may share encodings of the conceptual and computable levels, whereas the implementable level specifications are likely to be different from one site to another.

CIGs base their recommendations on eligibility and decision criteria that relate medical-concepts to patient data. Therefore, a CIG model, like GLIF, should define a formal **expression language** for specifying eligibility and decision criteria, and a **domain ontology** – a data model of medical concepts and patient data to which the expressions can refer. In developing GLIF3, we used existing standards – The HL7 *Reference Information Model* (*RIM*), and Arden Syntax's logic grammar – as the patient data model and expression language respectively.

We investigated the HL7 RIM as a default patient data model. The **HL7 RIM** [Schadow, Russler 2000] is an object-oriented model that has three advantages: (1) since it is an object-oriented model, it encapsulates the attributes of a patient data item into a single complex data object, (2) it is general enough to represent the data structures for a wide

range of medical data and concepts in a uniform manner while using a small number of classes, and (3) HL7 is creating standards for messaging interfaces for EMRs based on the RIM. Nevertheless, we continue to examine the possibility of adopting other models for patient data, such as the *Virtual Medical Record* model [Johnson, Tu 2001] used by EON and PRODIGY.

We developed other layers of GLIF's *domain ontology* [Zeng, Tu 2000]. The **Core GLIF** layer has definitions of variable and literal data items that refer to codes taken from controlled vocabularies and to medical data model classes (such as RIM classes). Work included specification of another part of the domain ontology that will enable guideline authors to specify their own medical concepts, concept hierarchies, and data-model classes. This is needed for cases where standard vocabularies and standard data models for medical concepts fall short. Current work on this issue is covered in the next section.

We examined the possibility of using Arden Syntax's logic grammar as a formal expression language for expressing eligibility and decision criteria. By August 2000, it had become clear to us that we could not use *minor* extensions of Arden Syntax's logic grammar for formally representing criteria. We started developing a language called the Guideline Expression Language (GEL) that is based on Arden Syntax's logic grammar but can work with complex data types, needed in order to work with GLIF's medical ontology. GEL is covered in the next section.

To aid in the modeling of large complex guidelines, GLIF3 has three different mechanisms for the management of **complexity** [Peleg, Boxwala 2000]. They include **nesting** action and decision steps into sub-guidelines, using **macro steps** [Boxwala, Mehta 2000] to declaratively specify procedural patterns that occur in CIGs, and using **views**, which are filters that collapse segments of the guideline into a default view that is customized to a given kind of user or setting.

Whereas GLIF2 supported only an if-then-else conditional step, GLIF3 has a **flexible decision model** through a hierarchy of decision step classes. This decision hierarchy distinguishes between decision steps that can be automated (**case steps**) and ones that have to be made by a physician or other health worker and cannot be automated **(choice steps**). A case step has a GEL criterion that is compared to mutually-exclusive values of the various decision options. Once a match is found, control flows to the guideline step that is specified in that decision option. Choice steps do not have a single criterion. Instead, for each decision options, rules in favor and against that decision option are specified to help the user to make a decision. The decision hierarchy can be extended in the future to support different decision models, such as models that consider uncertainty or patient preferences.

GLIF3 has an extended **action-specification hierarchy** that models two types of actions:(1) guideline-flow-relevant actions, such as calling of a sub-guideline, or retrieving data values from EMRs; and (2) clinically relevant actions, such as making recommendations. Clinically relevant actions reference the domain ontology for

representations of clinical concepts such as prescriptions, laboratory test orders, or referrals.

More additions to GLIF3 include constructs for expressing **iteration**, **patient states**, and **events** and **exceptions**.

The format in which GLIF-encoded guidelines will be stored and exchanged among different institutions is **RDF** (Resource Description Language), which is based on the eXtensible Markup Language (XML). *RDF* is a foundation for processing metadata; it provides interoperability among applications that exchange machine-understandable information on the Web. We created an RDF schema for GLIF3. This schema specifies the syntax of GLIF3. RDF files containing GLIF-encoded guidelines can be checked automatically for validity and correctness, when compared against this formal schema using available tools.

Another project undertaken by InterMed researchers has been the development of a guideline-classification scheme [Bernstam, Ash 2000]. The approach was based on the classification scheme used by the National Guideline Clearinghouse (http://www.guidelines.org). The axes of the proposed scheme have implications for designing formal methods and structures for representing, retrieving, and authoring clinical guidelines.

In order to test the robustness of GLIF3, we used it to encode 4 guidelines: (a) management of patients with chronic stable angina from ACP-ASIM [American College of Cardiology/American Heart Association/American College of Physicians-American Society of Internal Medicine 1999], (b) chronic cough [Irwin, Boulet 1998] from the American College of Chest Physicians, (c) a partial encoding of a low back pain guideline from AHCPR [AHCPR 1994], and (d) a partial encoding of a heart failure guideline from AHCPR [AHCPR 1994]. Guidelines (a) and (b) were encoded using Protégé as an authoring tool. The encoding of guidelines (c) and (d) concentrated on specifying decision and eligibility criteria, and on patient data, and was done without using automated authoring tools.

At Stanford, we have been using Protégé [Grosso, Eriksson 1999] as an authoring and viewing tool for GLIF3. We completed two "widgets" for Protégé for RDF support. One widget can translate guidelines, encoded in GLIF 3.0 and created using Protégé, into RDF format. The other widget can import RDF files containing GLIF-encoded guidelines into Protégé, and automatically lay out the flowcharts diagrammatically on the screen.

The research team at Harvard created designs for a guideline-authoring tool, a vocabulary-editing tool, an expression evaluator, and a guideline server. The design has a strong emphasis on reusability of the software for different applications and uses a component-based approach. We started the implementation of these tools in the Java programming language.

Vimla Patel's research team, in collaboration with Dr. Mottur-Pilson and other representatives from the ACP-ASIM, investigated the use of two clinical-practice guidelines (those for thyroid screening and diabetes management) at the point of patient care (in experiments at McGill University's endocrine clinic). They focused in particular on guideline use as a function of a) guideline format (written text versus algorithms versus flowcharts) and b) the levels of expertise of the physicians receiving the guidance (specialists, primary-care physicians, and house staff). The research team analyzed the data using a detailed propositional and semantic structural analysis of symbolic discourse (e.g., text, dialogue, and diagrams) in an effort to understand users' representations and interpretations [Patel, Arocha, et al, 2001]. The results showed that experts, house staff, and practitioners use guidelines for different purposes. They also interpret the guidelines differently, reflecting differences in their internal guideline representations. The development of representations appears to be non-monotonic (i.e., a U-shaped curve when plotting performance versus expertise from a novice to an expert). These studies have provided insights into how the nature and purpose of guidelines can be tuned to different users, and this knowledge can in turn be used in the design process.

The Columbia InterMed team also collaborated with another Columbia team (J. Cimino) and with Harvard (A. Boxwala) and Stanford (M. Peleg) to perform usability analyses of guideline encoding and application in clinical practice [Kushniruk, Patel, et al, 2001]. This work involved remote tracking of physicians as they accessed clinical guidelines via a Web-based EMR at Columbia. The results indicate that physicians use guidelines primarily for upgrading their general knowledge. In addition, we used innovative technology to collect remotely video recordings of users as they interacted with online guidelines. The Columbia InterMed group collected usability data over the Web, assessing user interactions with guidelines that are available via ACP-Online. We assessed the relationship between EMR technology and automatic prompts for guideline use during practice. This approach will also make critical information more accessible to physicians when they need it most. These laboratory-based and clinical-practice studies have important implications for the development of guidelines that are generic as well as specific for a particular case. They have informed us as to how guidelines should be written and encoded to minimize errors and misinterpretations by the eventual end users.

After the March 31, 2000 report, the McGill team concentrated on evaluating existing clinical-practice guidelines and their modes of use. In particular, working from a guideline for depression diagnosis and management from the ACP-ASIM, we did formal studies to investigate (a) physicians' and lay people's beliefs about depression and its diagnosis and treatment, and (b) the interpretation of depression guidelines by physicians and patients. Focal areas of examination were the following: (1) what beliefs about depression people hold and how their beliefs affect the decisions they make about treatment, and (2) how the prior beliefs held by both physicians and patients affect the way they interpret the guidelines intended for their use and, in turn, how that affects their subsequent interaction. Results from these studies continue to inform the InterMed team as we develop authoring environments and make knowledge-representation decisions for the underlying GLIF3 model.

Accomplishments during the period of September 1st, 2000 – April 30th, 2001

Our efforts this past year have concentrated on four areas: (1) further development of the GLIF model, (2) development of software tools, (3) collaborations with other groups who are working in the field of computer-interpretable guidelines, and (4) cognitive psychological studies related to guideline development. To test the GLIF model and authoring tools, we encoded 6 guidelines in GLIF3 in addition to the 4 guidelines that were previously encoded.

1. The GLIF model

- *Medical domain ontology support*

We extended GLIF3's medical ontology. Specifically, we augmented Core GLIF by adding constructs for creating **user-defined medical concepts and data models**. This is needed for cases where standard vocabularies and standard data models for medical concepts fall short. We have encountered such examples in the course of encoding the chronic cough guideline [Irwin, Boulet 1998]. We also created a Medical Knowledge layer that defines medical concepts and medical knowledge in the form of relationships between concepts.

- *Expression Language for expressing eligibility and decision criteria*

We developed an expression language for eligibility and decision criteria, called *Guideline Expression Language* (*GEL*) based on Arden Syntax's logic grammar [Hripcsak, Ludemann 1994]. **GEL** [Peleg, Ogunyemi 2001] supports Arden data types and operators, and in addition supports complex objects and time intervals that are needed for it to interact with the HL7 Reference Information Model, that is used as GLIF's default medical data model. GEL is extendable, as it supports user-defined functions.

- *Initial draft of an Object-oriented expression language*

Lessons learned from the development of GEL led us to understand that an object-oriented expression language would work better with object-oriented domain ontologies. In addition, an object-oriented language utilizes the encapsulation of data and methods that are relevant for a medical concept, in a way that is not utilized by GEL. In addition, an object-oriented model is extendable, so users could define new classes as well as new methods for them. Methods are not currently supported by GLIF. We have started developing the object-oriented expression language for GLIF.

- *Action specification classes for data and knowledge retrieval*

The Get_Data_Action was developed to retrieve patient data from EMRs and transform it into Query_Results that GEL can work with. Get_Knowedge_Action was developed for knowledge retrieval from medical knowledge represented as Core GLIF classes. We experimented with using these constructs in two guidelines: chronic cough (Irwin, Boulet et al. 1998) from the American College of Chest Physicians, and Prevention, Detection,

Evaluation, and Treatment of High Blood Pressure from the National Institute of Health (National Institute of Health 1997). These guidelines were encoded using the Protégé authoring tool.

- *Augmenting GLIF with attributes taken from GEM*

We have augmented GLIF with attributes taken from the *Guideline Elements Model* (*GEM*). **GEM** [Shiffman, Karras 2000] is a document model of a clinical guideline that whose elements relate to the nature of the process used to develop the guideline, kinds of evidence, and explanations and references for recommendations. It is therefore complementary to GLIF, which is an execution model of a clinical guideline. Our preliminary evaluation, based on an Alzheimer's disease treatment guideline [California Workgroup on Guidelines for Alzheimer's Disease Management 2000], shows that the merged model has enhanced capacity to relate decision and actions to explanations and the evidence upon which they are based [Yin, Peleg 2001].

- *Guideline-classification scheme*

We conducted an initial evaluation of the guideline-classification scheme that was mentioned in the Introduction. A poster that summarizes the evaluation was submitted to AMIA 2001 [Bernstam, Ash 2001].

- *Corpus of decision criteria*

Recently, we prepared a corpus of decision criteria that were extracted from 24 randomly selected guidelines, which belong to different guideline categories (e.g., screening, prevention, diagnosis). The corpus can be used for different studies, such as a classification of decision criteria, test cases to test expression languages against, and characterizing data types that criteria usually reference for aiding in developing the implementation level of the GLIF model.

2. Guidelines encoded in GLIF3

In order to test the new constructs of GLIF3, we encoded 6 guidelines (in addition to the 4 guidelines that were encoded until August 2000) [Peleg, Boxwala 2000]. The guidelines include the following guidelines from ACP-ASIM: (a) Screening for Thyroid Disease [ACP-ASIM 1998], and (b) Acute Major Depression and Dysthymia [ACP-ASIM 2000]. We also encoded (c) an Alzheimer's disease treatment guideline [California Workgroup on Guidelines for Alzheimer's Disease Management 2000], developed by the California Workgroup on Guidelines for Alzheimer Disease Management, (d) the JNC VI Hypertension guideline [National Institute of Health 1997], (e) prevention and control of influenza [Advisory Committee on Immunization Practices 2000], and (f) Screening for Tuberculosis and Tuberculosis Infection in High-Risk Populations, developed by the Advisory Council for the Elimination of Tuberculosis [Advisory Council for the Elimination of Tuberculosis 1995]. Guidelines (a) through (e) were encoded using Protégé as an authoring tool. Guidelines a, b, and f were encoded using the Decision Systems Group (DSG) Guideline Authoring Tool.

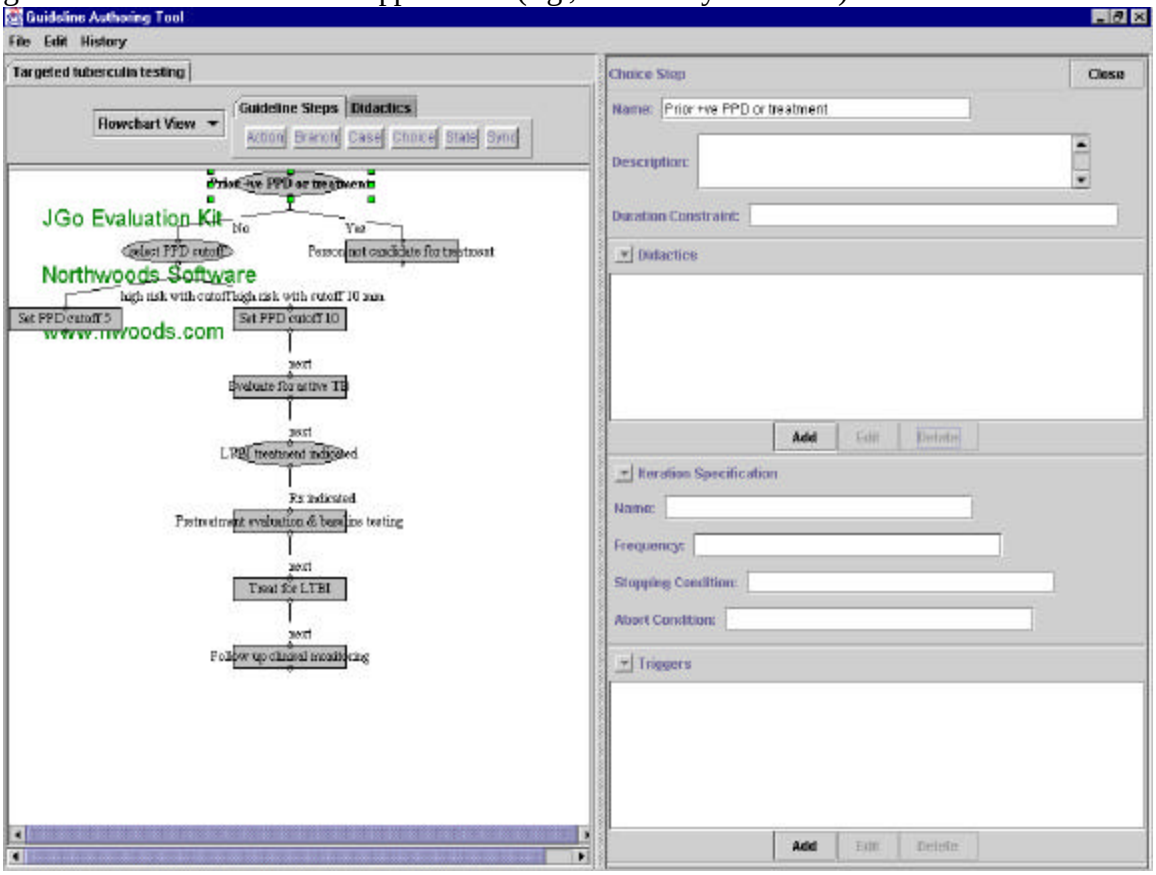
3. Software tools

We have developing a set of software tools aimed at facilitating creation, distribution, and management of GLIF guidelines and implementation of guideline-based clinical applications. We have chosen an object-oriented, component-based strategy for the design of the software. The design for the software is being specified in the Unified Modeling Language. The implementation has largely been in the Java programming language.

3.1 Tools

- Guideline *authoring* tool

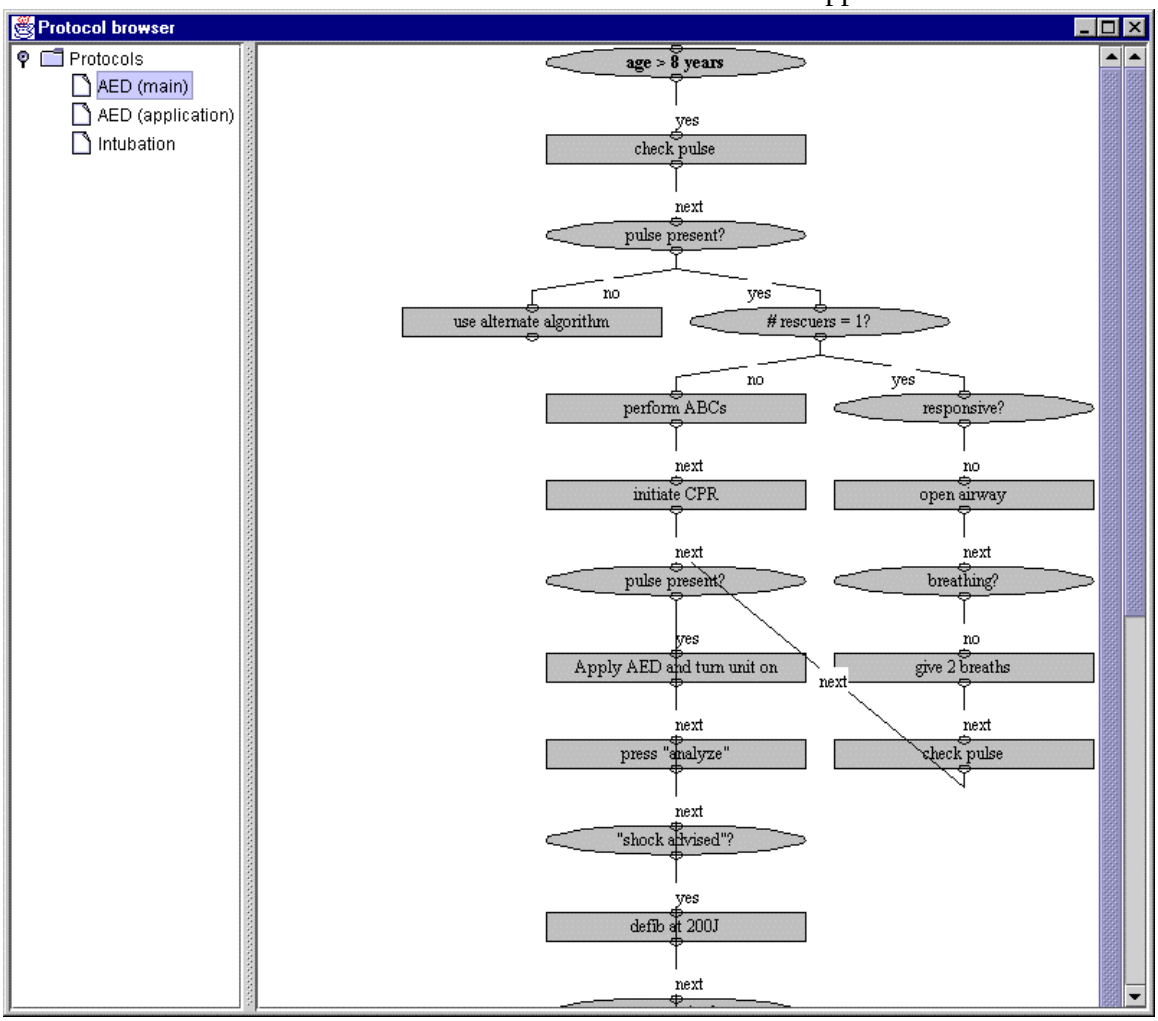
A guideline authoring-tool was implemented that provides a graphical environment for creating and editing guidelines in GLIF. The tool provides flowchart and tree views of guidelines. Steps can be encoded using forms. The encoded guidelines can be saved to disk as an RDF format file or can be submitted to a guideline server for storage.
The authoring tool has been designed so that it can be easily customized for different types of guidelines and can be extended with more features such as specialized views of guidelines or links to external applications (e.g., vocabulary databases).



Screenshot of guideline authoring tool displaying the tuberculosis screening guideline from the Centers for Disease Control and Prevention.

- Guideline *viewing* tools

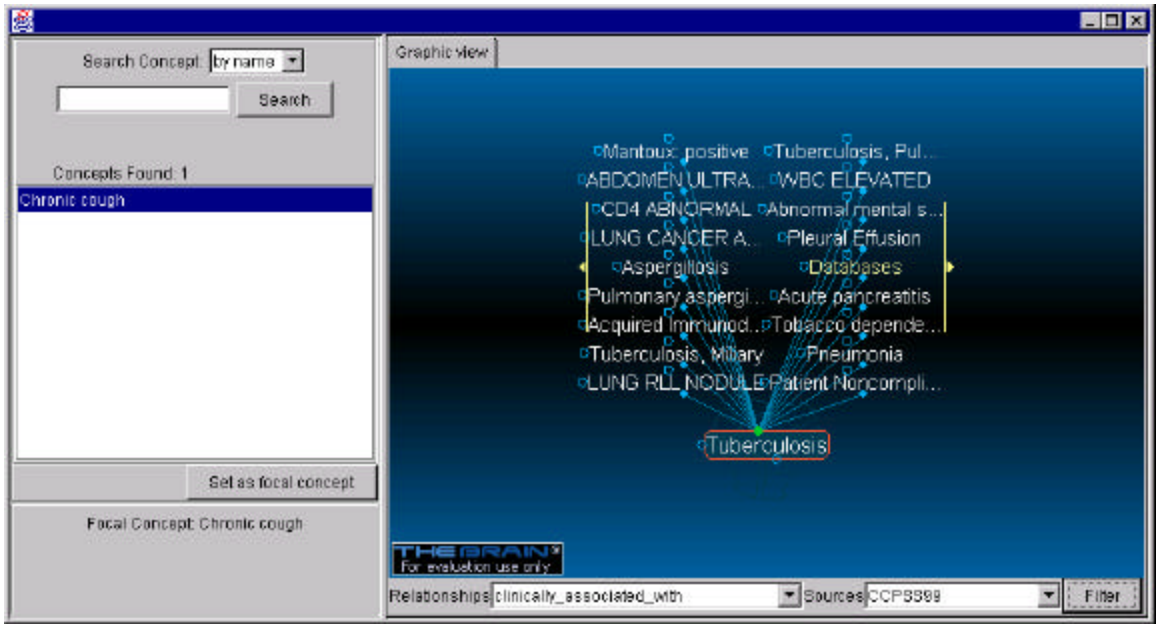
A prototype guideline viewing application was developed. The application provides a list of guidelines available in a server. The user can select a guideline that is then displayed as a flowchart. This viewer can be modified to use as a web-based applet.



Screenshot of guideline viewing application developed for the Digital EMS project at the University of Texas Houston. The left side of the screen displays guidelines and protocols in the server. The right side displays guidelines in a flowchart form.

- Vocabulary *editing* tool

A tool for editing vocabularies was developed. The tool provides graphical views of the vocabularies (as networks and hierarchies). Users can add, delete, and modify concepts, terms, and relationships. The tool also provides a read-only view of the UMLS meta-thesaurus. We will link in this tool into the guideline authoring-tool.



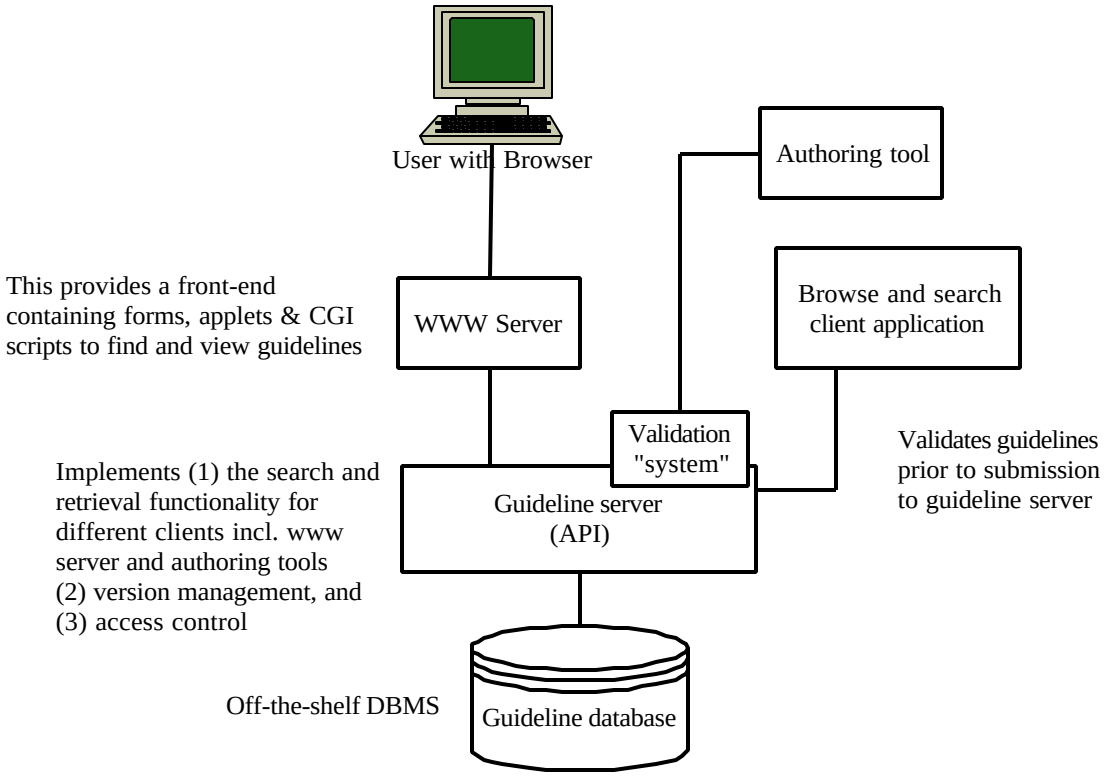
Screenshot of vocabulary browsing tool.

- Guideline *execution* engine

We developed a framework of a Guideline Execution Engine (GEE) that attempts to facilitate its integration with a host institutional information system, as well as its maintenance. Client-server communication part of this framework is already implemented. We are in the process to refine the GEE framework.

- Guideline server

We are in the process of specifying the design of a guideline server for storage of guidelines. A schematic of the server design is shown below. In the next section, we describe a component also known as *Guideline Server* that performs the low level tasks of guideline storage and management that will be used by this application.



3.2 Components

The software tools, described above, have been developed using the components listed in this section.

GLObject library

The GLObject library contains classes (as Java beans) that represent the different objects in a GLIF guideline. This library forms the foundation for most of the tools and many of the other components.

GDF parser and writer

The GDF library provides a parser for reading guidelines in RDF files into GLObject objects and a writer for serializing guidelines in memory as GLObjects into an RDF file. The design of this library is flexible to accommodate with minimal work, any revisions to the GLIF specification.

Guideline server

The guideline server library manages storage of and access to guideline files. Currently, two instantiations of the abstract server API exist: one instance stores guidelines in the native file system of the server and the other instance stores guidelines in a relational database management system. These servers are implemented in Java. In future versions, the servers may be implemented using CORBA. The search functions of the database-linked server are relatively unsophisticated in this version. We expect to improve search and retrieval capabilities in future versions.

Flowchart viewer

This component displays guidelines in a flowchart form. The viewer dynamically updates the display when a guideline is modified. The viewer can be extended to accommodate new subclasses of guideline steps. It can also be customized to add different flowchart rendering algorithms. It also supports drag and drop capability. This component has been used in guideline browsing and authoring tools.

Tree viewer

The tree viewer displays guidelines in a hierarchical form. This form may be suited for particular types of guidelines where the compositional hierarchy of the guideline is of importance such as protocols. Like the flowchart viewer, the tree viewer component dynamically updates itself, is completely extensible and supports drag and drop.

Expression parser and interpreter

We developed a Java-based parser and interpreter that evaluates expressions written in GEL, creating an Abstract Syntax Tree for each expression and executing the actions corresponding to the expression. To ensure extensibility and flexibility in the language, the parser and interpreter are able to process and execute user-defined functions, thereby allowing operations that are not provided by GEL's set of built-in operators.

Vocabulary library

This library builds an interface for a semantic network for vocabularies. The package is independent of any particular vocabularies and databases.

Vocabulary semantic net viewer

This component renders a graphical view of a vocabulary (previous section) as a hyperbolic tree. It has been developed using a commercial SDK called theBrain. The viewer is used in vocabulary editing and browsing applications.

Vocabulary tree viewer

The tree viewer is suitable for displaying vocabularies that are hierarchically organized or for displaying hierarchical views of vocabularies.

Vocabulary database library

This library provides access to vocabularies stored in relational database management systems using a Java API. It also allows updating of these vocabularies. Current instantiation includes access to the UMLS 2000 metathesaurus stored in MS SQL Server version 6.5.

4. Collaborations with other groups

4.2 HL7

- Following the March 2000 workshop in Boston, we contacted HL7 regarding the possibility of establishing a guideline standardization effort within it, that would relate to its work on supporting the Arden Syntax

standard, and would enable us to build on GLIF with a consensus-based process.

- In September, 2000, the HL7 Board agreed to establish a GLIF Special Interest Group, under a newly formed Clinical Decision Support Technical Committee, which would also be the parent for an Arden Syntax SIG.
- In January, 2001, a first meeting of the GLIF SIG occurred in Orlando, FL, as part of the HL7 meeting, attended by Greenes, Tu, Boxwala, and Peleg. At that time, for political reasons, it was decided to change the name of the SIG to the Clinical Guidelines SIG, in order to be more "ecumenical". It was decided that initial foci should be on evolution of the expression language, along the lines discussed herein, and on the adoption of a "virtual patient record" construct as an intermediary between expressions and the HL7 RIM.
- The next meeting of the SIG will be in Minneapolis, MN, May 8-10, 2001, to be attended again by Greenes, Tu, Boxwala, and Peleg.

4.3 Comparing GLIF 3.0 to other guideline modeling approaches

The comparison process is under way. We are collaborating with researcher from 5 other teams: EON [Musen, Tu 1996], PRODIGY [Johnson, Tu 2000], PRO*forma* [Fox and Rahmanzadeh 1998], Asbru [Shahar, Miksch 1996], and the Pavia group [Quaglini, Dazzi 1998].

4.4 Collaboration with Digital EMS project at University of Texas Houston

We worked with the Digital EMS Project team at the University of Texas Medical Center in Houston on use of GLIF-based protocols in the EMS units. A demonstration application for displaying protocols in GLIF was developed. Currently, we are discussing the implementation of a guideline execution engine for online decision support for EMS personnel based on data entered in the "run record".

5. Cognitive Psychological Studies

The McGill research team, led by Dr. Vimla Patel (who moved to Columbia to join the faculty there late in 2000) has continued investigating the process of guideline encoding of text-based guidelines into GLIF. To this end, a study was conducted which evaluates the translation of an updated version of the Guideline Interchange Format(GLIF), GLIF3, into an electronically encoded form [Patel, Branch, et al, 2001]. The goal of GLIF3 is to support the effective sharing of such guidelines among various clinical institutions and settings by investigating the cognitive processes underlying such translation. The encoding of two clinical guidelines both into GLIF3 and into GLIF2 by two medical informaticians was videotaped and analyzed. GLIF3 was found to be more robust than GLIF2 for representing the content and logical structure of the clinical guidelines studied. This formative evaluation demonstrated that GLIF3's intended improvements in expressiveness were achieved. This research, together with the tele-evaluation of Web-based information can provide set of methodological tools for the study of information system' use [Kushniruk, Patel, et al, 2001].

The team has also continued the development and refinement of a theoretical and methodological framework [Patel, Arocha, et al, 2001] for the use of cognitive analysis to support the representation of biomedical knowledge and the design of clinical systems, using clinical-practice guidelines. A hypothesis underlying such development is that propositional and semantic analyses, when used as part of the system-development process, can improve the validity, usability and comprehension of the resulting biomedical applications. The framework is based on a large body of research [Patel, Arocha, and Kaufman, 2001] on the study of how people mentally represent information and subsequently use it for solving problems and making decisions in the health domains. This research encompasses many areas of psychology, but the more important ones are the study of memory and the study of comprehension. Of particular relevance is research devoted to investigating the comprehension and memory of language, expressed verbally or in text. In addition, research on how contextual variables affect performance is informative because these psychological processes are influenced by situational variables (e.g., setting, culture).

The theoretical and methodological framework has been applied to the study of guideline utilization [Patel, Arocha, Diermeier, and Mottur-Pilson, 2001], in continuing collaboration with the ACP. This collaboration has continued to investigate the process of guideline development from creation to utilization. Furthermore, research has been undertaken in an attempt to understand the nature of impact of guideline use on physician performance. To this end, we investigated the impact of (a) algorithmic-based and (b) text-based practice guidelines on clinical decision making by physicians at varying levels of expertise. Data were collected using clinical scenarios and a think-aloud paradigm, both with (primed) and without (spontaneous) the use of the guidelines. The two guidelines used in the study were management of diabetes and screening for thyroid disease. The results showed that guidelines were used as reminders for both experts and non-experts and that guidelines acted as educational tools for non-experts by assisting in knowledge reorganization. Furthermore, text and algorithmic guideline formats were both useful to physician performance depending on the purpose of use—solving clinical problems or learning. This type of research serves to provide insights into how guidelines can be fine-tuned for different users and for different purposes and to demonstrate that empirical research, coupled with design principles from the cognitive sciences can form an essential component of guideline design and development, as well as its use [Patel, Arocha, and Kushniruk, 2001].

The framework has also been applied to patient-guideline interpretation by lay people [Patel, Diermeier, and Chaudhari, 2001]. This study aims to explore the extent to which lay people perceive a depression guideline as helpful in reasoning through a patient problem. A second question that this research attempts to address is the extent to which lay people are equipped with the knowledge needed to comprehend and use the patient guidelines. Finally, a third issue is whether lay person's beliefs match those that are presupposed in the patient guideline.

6. Plans for the next project period

6.1 Development of the GLIF model

1) Develop the object-oriented expression language for GLIF, mentioned in the previous section
2) Modify GLIF's default RIM into a virtual Medical Record, by working with HL7 Decision Support Technical Committee
3) Define the implementation level of GLIF
4) Further develop GLIF's macro step

6.2 Encoding Guidelines

Encode 1-2 guidelines in the next version of GLIF that will include the object-oriented expression language and the virtual EMR

6.3 Software development

Authoring tool

- Continue refinement of the software for synchronization with evolving GLIF specification, and bug removal
- Integration with guideline server
- Integration with vocabulary browsers
- Enhancement of authoring tool API so that custom authoring tool applications based on domain-specific macros can be developed. For example, we plan to develop authoring tools for screening guidelines and referral criteria.

Guideline server

- Implement a guideline server based on the design scheme described in an earlier section of this document
- Build a web-based front-end for this guideline server that allows users to access text versions and GLIF-encodings of guidelines using the browser
- Integrate with authoring tool and guideline browsing tools

Guideline execution engine

- Complete development of the execution engine. This includes:
  (a) A core part that will handle basic guideline execution tasks such as logic flow control and decision criteria evaluation.
  (b) A standard interface to the guideline representation model that will facilitate maintenance of the GEE.
  (c) A standard interface to the host institution application at the front-end that will provide the user interface of the GEE.

(d) A standard interface to the host institution information system at the back-end that will provide services to the GEE, such as patient data retrieval and automatic action performing.

- Revise parser for the new object-oriented expression language
- Integrate execution engine into guideline server web-site so that users can run the guidelines interactively by entering data

Guideline validation tools

We plan to develop a set of constraints that well-formed guidelines should satisfy (e.g., a branch with only one branch destination or a decision with only one next step) and develop validation tools that will check if an encoded guideline satisfies these constraints.

6.4 Collaboration with other teams

We will complete the study that compares GLIF 3.0 to other guideline modeling approaches.

We will continue to work within HL-7 on the expression language and virtual medical record.

6.5 Cognitive Psychological Studies

We are now ready to follow a guideline's development from inception at the ACP-ASIM, having identified a guideline (hemochromotosis management) and expert participants who will be studied as they review the literature and formulate evidence-based recommendations.

We will also study the process at the ACP-ASIM whereby College staff physicians formulate an algorithmic description of a guideline based on the evidence-based review that is prepared by the outside expert.

To assist in the design of authoring tools and to assure robustness in the GLIF3 representation scheme, we will introduce computer-science modeling (design) considerations into the guideline-development studies that are conducted in collaboration with the ACP-ASIM. Dr. Peleg from Stanford will work closely with Dr. Patel and the ACP-ASIM collaborators in this effort.

7. InterMed Publications

Publications and technical documentation (**GLIF3 User's Guide** and example encoded guidelines (http://www.smi.stanford.edu/projects/intermed-web/guidelines/GLIF_TECH_SPEC.pdf) created by InterMed team members (see references), and information about the guideline workshop that was held in March and about task forces that were created at that meeting, can now be found on our project web site, http://www.GLIF.org.

Peer-reviewed journal papers

Boxwala AA, Tu S, Zeng Q, Peleg M, Ogunyemi O, Greenes RA, Shortliffe, EH, and Patel, VL. Towards a representation format for sharable clinical guidelines. In preparation for submission to the *Journal of Biomedical Informatics, April* 2001.

Kushniruk AW, Patel, C, Patel VL, Cimino JJ. 'Televaluation' of clinical information systems: an integrative approach to assessing Web-based systems. *International Journal of Medical Informatics*. 2001;61:45-70.

Patel VL, Arocha JF, Diermeier M, How J, Mottur-Pilson C. Cognitive psychological studies of representation and use of clinical practice guidelines. *International Journal of Medical Informatics, in press* 2001.

Patel VL, Arocha JF, Diermeier M, Shortliffe EH, Greenes RA. Methods of cognitive analysis to support the design and evaluation of biomedical systems: The case of clinical practice guidelines. *Journal of Biomedical Informatics* 2001; 34:xxx-yyy.

Patel VL, T Branch,Wang D, Peleg M, Boxwala AA. Analysis of the Process of Encoding Guidelines: An Evaluation of GLIF3. Submitted to *Methods of Information in Medicine, March 2001.*

Patel VL, Arocha JF, Kaufman DR. A primer on aspects of cognition for medical informatics. *Journal of the American Medical Informatics Association*. in press:187-252.

Patel VL, Arocha JF, Kushniruk AW. EMR: Re-engineering the organization of health information. *Methods of Information in Medicine*. in press

Peleg M, Bernstam E, Boxwala AA, Tu S, Greenes RA, Shortliffe EH. Defining the Relationship between GLIF and the Arden Syntax; 2000. In preparation for submission to the *Journal of Biomedical Informatics, May 2001.*

Non-peer-reviewed journal papers

Elkin P, Peleg M, Lacson R, Bernstam E, Tu S, Boxwala A, et al. Toward Standardization of Electronic Guideline Representation. *MD Computing* Vol. 17, No. 6, 2000, pp. 39-44.(MD Comput is peer-reviewed, I think?)

Conference papers

Bernstam E, Ash N, Peleg M, Tu S, Boxwala AA, Mork P, et al. Guideline classification to assist modeling, authoring, implementation and retrieval. Proceedings of the American Medical Informatics Association (AMIA) Annual Symposium 2000. p. 66-70.

Boxwala AA, Clemenson SG, Salomon WL, Greenes RA. Implementing a Hypertension Guideline in a Health Care Information System: Insights, Challenges and Implications. Submitted to AMIA Ann Symp; 2001.

Greenes RA, Peleg M, Boxwala AA, Tu SW, Patel VL, Shortliffe EH. Sharable Computer-Based Clinical Practice Guidelines: Rationale, Obstacles, Approaches, and Prospects. Accepted for publication in MedInfo; 2001.

Peleg M, Boxwala A, Ogunyemi O, Zeng Q, Tu S, Lacson R, et al. GLIF3: The Evolution of a Guideline Representation Format. Proc. AMIA Annual Symposium; 2000; p. 645-649.

Peleg M, Boxwala AA, Tu S, Greenes RA, Shortliffe EH, Patel VL. Handling Expressiveness and Comprehensibility Requirements in GLIF3. To be published in Proc. MedInfo 2001.

Peleg M, Ogunyemi O, Tu S, Boxwala AA, Zeng Q, Greenes RA, et al. Using Features of Arden Syntax with Object-Oriented Medical Data Models for Guideline Modeling. Submitted to AMIA Annual Symposium, 2001.

Phillips RC, Bernstam EV, Mork P, Karras BT. Modification of GEM to Accommodate Changes in a Classification Scheme. Submitted to AMIA 2001 Fall Symposium.

Wang D, Peleg M, Tu SW, Shortliffe EH, Greenes RA. Representation of Clinical Practice Guidelines For Computer-Based Implementations. To appear in MedInfo; 2001.

Posters

Bernstam E, Ash N, Peleg M, Tu S, Shortliffe EH, Greenes RA. Preliminary Evaluation of a Guideline Classification System. submitted to AMIA 2001.

Boxwala AA, Mehta P, Peleg M, Lacson R, Ash N, Bury J, et al. Modeling guidelines using domain-level knowledge representation components. AMIA Annual Symposium, poster presentation; 2000. p. 645.

Boxwala AA, Peleg M, Zeng Q, Tu S, Lacson R, Bernstam E, et al. GuideLine Interchange Format (GLIF): Extensions and Practical Applications Workshop. In: Annual Meeting of the American Telemedicine Association; Phoenix, Arizona; 2000.

Choi J, Boxwala AA. Representing domain-level knowledge components using primitives. Submitted to AMIA Ann Symp; 2001

Yin JQ, Peleg M, Boxwala AA, Greenes RA. Combining a Document Model and an Execution Model for Clinical Guidelines. Submitted to AMIA Ann Symp.; 2001.

Zeng Q, Tu S, Boxwala AA, Peleg M, Greenes RA, Shortliffe EH. A Three-Layer Domain Ontology for Guideline Representation and Sharing. AMIA Annual Symposium, poster presentation; 2000.

AMIA Tutorials

Peleg M, Boxwala A. An Introduction to Modeling and Representation of Clinical Guidelines. In: AMIA; 2000; Los Angeles, CA; 2000.

Technical Reports

Peleg M, Boxwala A, Ogunyemi O, Zeng Q, Tu S, Lacson R, et al. GLIF3 Technical Documentation. In: http://smi-web.stanford.edu/projects/intermed-web/guidelines/GLIF1.htm Peleg M, Boxwala A, Ogunyemi O, Zeng Q, Tu S, Lacson R, et al., editors. 2000.

Patel VL, Diermeier ME, Chaudhari S. Educating the public: The effectiveness of a depression guideline for patients. Technical report,. Department of Medical Informatics, Columbia University; 2001.

Book Chapters

A book chapter on GLIF in the book on Cancer Informatics that will appear in Marion Ball's Springer-Verlag series on Health Informatics was accepted for publication, subject to editorial revision.

References

ACP-ASIM. Screening for Thyroid Disease. *Ann Int Med* 1998;129:141-143.

ACP-ASIM. Acute Major Depression and Dysthymia. *Ann Intern Med* 2000;132:738-742.

Advisory Committee on Immunization Practices A. Prevention and Control of Influenza. *Morbidity and Mortality Weekly Report* 2000;49(RR03):1-38.

Advisory Council for the Elimination of Tuberculosis. Screening for Tuberculosis and Tuberculosis Infection in High-Risk Populations. Morbidity and Mortality Weekly Report 1995:18-34.

AHCPR. Acute Low Back Problems in Adults, Clinical Practice Guideline Number 14: AHCPR Publication No. 95-0642; 1994. Report No.: 95-0642.

AHCPR. Heart Failure: Evaluation and Care of Patients With Left-Ventricular Systolic Dysfunction, Clinical Practice Guideline Number 11: AHCPR Publication No. 94-0612; 1994. Report No.: 94-0612.

American College of Cardiology/American Heart Association/American College of Physicians-American Society of Internal Medicine. Guidelines for the Management of Patients with chronic Stable Angina. *J Am Col Cardiol* 1999;33:2092-2197.

Boxwala AA, Greenes RA, Deibel SR. Architecture for a multipurpose guideline execution engine. In: Proc AMIA Symp; 1999; 1999. p. 701-705.

California Workgroup on Guidelines for Alzheimer's Disease Management. Guideline for Alzheimer's disease Management; 2000.

Clinical Decision Support & Arden Syntax Technical Committee of HL7, inventor Arden Syntax for Medical Logic Systems, version 2.0. Draft revision. USA. 1999 July 7, 1999.

Fox J, Rahmanzadeh A. Disseminating medical knowledge: the PROforma approach. *Artificial Intelligence in Medicine* 1998;14:157-181.

Gordon C, Veloso M. Guidelines in Healthcare: the experience of the Prestige project. In: Medical Informatics Europe; 1999; Ljubljana , Slovenia; 1999. p. 733-738.

Grosso WE, Eriksson H, Fergerson R, Gennari JH, Tu SW, Musen MA. Knowledge Modeling at the Millennium (The Design and Evolution of Protege-2000). In: Gains BR, Kremer R, Musen M, editors. The 12th Banff Knowledge Acquisition for Knowledge-Based Systems Workshop.; 1999; Banff, Canada; 1999. p. 7-4-1 to 7-4-36.

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

Irwin RS, Boulet LS, Cloutier MM, Gold PM, Ing AJ, O'byrne P, et al. Managing Cough as a Defense Mechanism and as a Symptom, A Consensus Panel Report of the American College of Chest Physicians. *Chest* 1998;114(2):133S-181S.

Johnson PD, Tu SW, Booth N, Sugden B, Purves IN. Using Scenarios in Chronic Disease Management Guidelines for Primary Care. In: Proc. AMIA Annual Symposium; 2000.

Johnson PD, Tu SW, Musen MA, Purves I. A Virtual Medical Record for Guideline-Based Decision Support. In: submitted to AMIA annual symposium 2001.

Kushniruk AW, Patel, C, Patel VL, Cimino, JJ. 'Televaluation' of clinical information systems: an integrative approach to assessing Web-based systems. *International Journal of Medical Informatics*. 2001;61:45-70.

Musen MA, Tu SW, Das AK, Shahar Y. EON: A component-based approach to automation of protocol-directed therapy. *JAMIA* 1996;3:367-388.

National Institute of Health. The Sixth Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: National Institute of Health; 1997 November. Report No.: 98-4080.

Ohno-Machado L, Gennari JH, Murphy S, Jain NL, Tu SW, Oliver DE, et al. The GuideLine Interchange Format: A Model for Representing Guidelines. *Journal of the American Medical Informatics Association* 1998;5(4):357-372.

Quaglini S, Dazzi L, Gatti L, Stefanelli M, Fassino C, Tondini C. Supporting tools for guideline development and dissemination. *Artif Intell Med* 1998;14(1-2):119-37.

Schadow G, Russler DC, Mead CN, McDonald CJ. Integrating Medical Information and Knowledge in the HL7 RIM. In: Proc. AMIA Annual Symposium 2000; 2000; 2000. p. 764-768.

Shahar Y, Miksch S, Johnson P. An Intention-Based Language for Representing Clinical Guidelines. In: James J. Cimino E, editor. AMIA Annual Fall Symposium; Washington, D.C: Hanley & Belfus; 1996. p. 592-596.

Shiffman RN, Karras BT, Agrawal A, Chen R, Marenco L, Nath S. GEM: a proposal for a More Comprehensive Guideline Document Model Using XML. *JAMIA* 2000;7(5):488-498.

Sugden B, Purves IN, Booth N, Sowerby M. The PRODIGY Project - the Interactive Development of the Release One Model. In: AMIA Symp; 1999; 1999. p. 359-363.

Tu SW, Musen MA. A Flexible Approach to Guideline Modeling. In: AMIA Symp; 1999; p. 420-424.