Progress Report to the Funding Agencies

For the Period July 1, 1999 to March 31, 2000
Grant R01 LM06594-01
"A Shared Internet Server for Delivering Guidelines"
National Library of Medicine and
US Army Telemedicine and Advanced Technology Research Command (TATRC)

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

Co-Investigators:
Aziz A. Boxwala, MD, PhD
Lucila Ohno-Machado, MD, PhD
Mor Peleg, PhD
Samson Tu, MS

Accomplishments to date:

Since July 1999, Intermed has been working on the development of the third version of the GuideLine Interchange Format, dubbed GLIF 3.0. We have been analyzing the previously published version, GLIF 2.0 (Ohno-Machado, Gennari et al. 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, Shortliffe et al. 2000), EON (Tu, Musen 1999), Prodigy (Sugden, Purves 1999), Proforma (Fox, Rahmanzadeh 1998), and Prestige (Gordon,Veloso 1999), and examining various guidelines from the National Guideline Clearinghouse. Through this analysis, we identified a number of areas in which GLIF 2.0 needed to be extended in order to support fully the encoding of computer-interpretable and executable guidelines.

In November 1999, we compiled a list of functional requirements for a shared guideline representation (Boxwala, Peleg 1999). By the beginning of March 2000, most of these requirements were already supported by the GLIF 3.0 draft technical specification (Peleg, Boxwala 2000) (Peleg, Boxwala 2000).

GLIF 3.0 now supports modeling guidelines at three levels of abstraction: a conceptual flowchart, 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. GLIF 3.0 has three different mechanisms for the management of complexity (Boxwala, Mehta 2000).

The encoding scheme uses a superset of Arden Syntax's logic grammar for specifying logical criteria. It supports a flexible decision model and has an extended action-specification model. GLIF 3.0 has constructs for expressing iteration, events and exceptions, and patient states.

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 have created an RDF schema for GLIF 3.0. This schema specifies the syntax of GLIF 3.0, and RDF files containing GLIF-encoded guidelines can be checked automatically for validity and correctness, when compared against this formal schema using available tools.

We also encoded two ACP-ASIM guidelines using GLIF 3.0 (Peleg, Boxwala 2000). They are (a) Screening for Thyroid Disease, and (b) Management of Patients with Chronic Stable Angina.

Since February, we have been concentrating on the development of domain ontology support (Zeng, Tu 2000), and we estimate that the major part of this task will be completed by the end of April 2000.

On March 3-4, 2000, an invitational workshop on Modeling Clinical Guidelines was hosted by Intermed, supported by conference grants from the NLM and the Army, but also with the sponsorship and participation of the Agency for Healthcare Research and Quality and the Centers for Disease Control. About 80 people attended from academia (from 8 different countries), government agencies, professional organizations (including the ACP-ASIM, which continues as an active collaborator as outlined below), providers, and industry. The participants all recognized the need for a standard sharable computer-interpretable guideline representation. Breakout sessions addressed the following issues:

1) defining the functional requirements for a standard representation of clinical guidelines
2) understanding the different modeling methods for clinical guidelines
3) defining the special modeling needs of clinical trial protocols
4) defining which tools and infrastructure are needed for development and use of computerized guidelines
5) creating a process for standardization of guideline modeling and representation.

Task forces initiated at the workshop are working on these areas.

The workshop helped us to identify certain guideline issues and situations that could not be represented by GLIF 3.0. We are planning to work on these in the future. These issues include: specifying guideline goals, representing negative recommendations and contraindications, and handling uncertainties in data.

In addition to work on the development of GLIF 3.0, Intermed team members have also been working on the development of tools for GLIF 3.0. Harvard is extending their guideline tools (Boxwala, Greenes 1999), and Stanford is building tools that work in the Protégé environment (Grosso, Eriksson 1999). The work done at Harvard so far has been

to design an architecture for guideline tools (authoring tool, guideline browser, execution engine, guideline server) that can be custom-tailored for a variety of applications and integrated into diverse environments. The tools have been partially implemented in Java. At Stanford, Intermed researchers have written a "widget" for Protégé that can translate guidelines, encoded in GLIF 3.0 and created using Protégé, into RDF format. By June 2000, we expect to complete a widget that will import RDF files containing GLIF-encoded guidelines into Protégé, and automatically lay out the flowcharts diagrammatically on the screen.

Another project that has been undertaken by Intermed researchers is the development of a guideline-specification scheme (Bernstam, Ash 2000). The approach is 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.

The McGill research team, in collaboration with Dr. Mottur-Pilson and other representatives from the ACP-ASIM, has been investigating the use of two clinical practice guidelines (those for thyroid screening and diabetes management) at the point of patient care. They have 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 housestaff). The research team has 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. The results have shown that experts, housestaff, 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 provide 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 McGill team has also collaborated with Columbia (J. Cimino and C. Patel) and with Harvard (A. Boxwala) to perform usability analyses of guideline encoding and application in clinical practice (work by Patel and Kushniruk). This work has involved remote tracking of physicians as they access clinical guidelines via a Web-based CPR system. The preliminary results indicate that physicians use guidelines primarily for upgrading their general knowledge. In addition, the technology is in place for remotely collecting video recordings of users as they interact with online guidelines. To date the McGill group has collected usability data over the Web, assessing user interactions with guidelines that are available via ACP-Online. We are currently assessing the relationship between computer-based patient record (CPR) technology and automatic prompts for guideline use during practice. This will also make critical information more accessible to physicians when they need it most. These ongoing laboratory-based and clinical-practice studies have important implications for development of guidelines that are generic as well

as specific for a particular case. They inform us as to how guidelines should be written and encoded to minimize errors and misinterpretations by the eventual end users.

Technical reports 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.

Plan for the next 18 months:

1) Encode 10-15 guidelines in GLIF 3.0. The guidelines will be chosen from different guideline categories, according to the classification scheme mentioned above. In this way, different categories of guidelines will be represented, and thus different guideline features will need to be modeled. This work will also test the applicability of the guideline-classification scheme.
2) Further develop authoring and viewing tools for GLIF 3.0. Tools will be created for checking that guideline specifications are syntactically and semantically correct, insofar as the logical syntax can be parsed correctly.
3) Develop validation tools that will:
   (a) identify reachability problems (i.e., is node B reachable from node A?)
   (b) find deadlock situations, or missing "elses" (e.g., a branch with only one branch destination or a decision with only one next step)
   (c) assess logical inconsistencies or redundancies in criteria (e.g., if the guideline contains a branch criterion that utilizes certain data elements, the checker can make sure that branch criteria are not redundant, and do not result in situations where the logic of the criteria prohibits traversal to certain parts of the guideline)
   (d) analyze decision steps to ensure that alternatives are mutually exclusive, non-redundant or conflicting, nor missing, or at very least to document such limitations.
4) Develop an interpreter that can execute the instructions, ask for input of values for parameters, evaluate the logic conditions, and follow the appropriate guideline paths.
5) Create a guideline server (accessible over the Internet) that can provide support for distributing and sharing encoded versions of guidelines. The server will allow:
   (a) browsing the encoded guidelines using viewing tools (adapted from PROTEGE and GEODE).
   (b) accessing text versions of the guidelines
   (c) computing eligibility criteria for guidelines and execution of some categories of protocols
   (d) downloading guidelines in GLIF format.
6) Compare GLIF 3.0 to the European guideline modeling approaches.

7) Formally study the guideline-development process from writing the background paper for guidelines to developing a formal algorithm and flowchart by ACP-ASIM scientific personnel. This work will be undertaken by the McGill team in collaboration with those from the ACP-ASIM.

8) Compare the language of representations for guideline development, using both GLIF and formal propositional representations. The former is a declarative representation language whereas the latter is procedural. Preliminary work in this area is already underway between the McGill team and researchers at Stanford (Peleg, Shortliffe), Harvard (Boxwala) and ACP-ASIM (Mottur-Pilson).

9) Evaluate the guideline classification scheme mentioned above.

References

Ash, N., E. Bernstam, et al. (2000). Classifying clinical practice guidelines. Technical Report, Decision Systems Group.

Bernstam, E., N. Ash, et al. (2000). Guideline classification to assist modeling, authoring, implementation and retrieval. Submitted to Proceedings of AMIA Fall Symposium.

Boxwala, A.A., R.A. Greenes, et al. (1999). Architecture for a multipurpose guideline execution engine. Proceedings of AMIA Fall Symposium.

Boxwala, A.A., P. Mehta, et al. (2000). Modeling guidelines using domain-level knowledge representation components. Submitted to Proceedings of AMIA Fall Symposium.

Boxwala, A.A., M. Peleg, et al. (1999). Functional requirements for a shared guideline representation. Submitted to Artifical Intelligence in Medicine.

Clinical Decision Support & Arden Syntax Technical Committee of HL7. (1999). Arden Syntax for Medical Logic Systems, Version 2.0. Draft revision. Health Level Seven. USA**:** 1-119.

Fox, J. and A. Rahmanzadeh (1998). Disseminating medical knowledge: The PROforma approach. Artificial Intelligence in Medicine **14**:157-181.

Gordon, C. & M. Veloso (1999). Guidelines in healthcare: The experience of the Prestige project. Proceedings of Medical Informatics Europe, Ljubljana , Slovenia.

Grosso, W.E., H. Eriksson, et al. (1999). Knowledge modeling at the millennium: The design and evolution of Protege-2000. Proceedings of the12th Banff Knowledge Acquisition for Knowledge-Based Systems Workshop, Banff, Canada.

Ohno-Machado, L., J.H. Gennari, et al. (1998). "The GuideLine Interchange Format: A model for representing guidelines." Journal of the American Medical Informatics Association **5**(4):357-372.

Patel, V.L., D.R. Kaufman, V.G. Allen, E.H. Shortliffe, J.J. Cimino, & R.A. Greenes. (1999) Toward a framework for computer-mediated collaborative design in medical informatics. Methods of Information in Medicine; 38:158-176.

Patel, V.L., V.G. Allen, J.F. Arocha, & E.H. Shortliffe. (1998) Representing a clinical guideline in GLIF: Individual and collaborative expertise. Journal of the American Medical Informatics Association; 5:467-483.

Patel, V.L., J.F. Arocha, & C. Mottur-Pilson. (2000) Representations and use of clinical practice guidelines: Cognitive psychological studies at the point of care. Submitted to Artificial Intelligence in Medicine.

Patel, V.L., A.W. Kushniruk, et al. (2000) Analysis of guideline encoding and application in clinical practice." Technical Report, Centre for Medical Education, McGill University.

Peleg, M., A. Boxwala, et al. (2000). GLIF 3.0 Technical Documentation. http://smi-web.stanford.edu/projects/intermed-web/guidelines/GLIF1.htm

Peleg, M., A. Boxwala, et al. (2000). GLIF 3.0: The evolution of a guideline representation format. Submitted to Proceedings of AMIA Fall Symposium.

Peleg, M., E. H. Shortliffe, et al. (2000). Defining the relationship between GLIF and the Arden Syntax. Technical Report, Stanford Medical Informatics.

Sugden, B., I.N. Purves, et al. (1999). The PRODIGY Project: The interactive development of the Release One model. Proceedings of AMIA Fall Symposium.

Tu, S.W. & M.A. Musen (1999). A flexible approach to guideline modeling. Proceedings of AMIA Fall Symposium.

Zeng, Q., S. Tu, et al. (2000). A three-layer domain ontology for guideline representation and sharing. Submitted to Proceedings of AMIA Fall Symposium.