Original **Investigation**

JAMIA

Research Paper ■

The GuideLine Interchange Format:

A Model for Representing Guidelines

Lucila Ohno-Machado, MD, PhD, John H. Gennari, PhD,
Shawn N. Murphy, MD, PhD, Nilesh L. Jain, DSc, Samson W. Tu, MS,
Diane E. Oliver, MD, Edward Pattison-Gordon, MS,
Robert A. Greenes, MD, PhD, Edward H. Shortliffe, MD, PhD,
G. Octo Barnett, MD

A b s t r a c t   **Objective:** To allow exchange of clinical practice guidelines among institutions and computer-based applications.

Design: The GuideLine Interchange Format (GLIF) specification consists of the GLIF model and the GLIF syntax. The GLIF model is an object-oriented representation that consists of a set of classes for guideline entities, attributes for those classes, and data types for the attribute values. The GLIF syntax specifies the format of the test file that contains the encoding.

Methods: Researchers from the InterMed Collaboratory at Columbia University, Harvard University (Brigham and Women's Hospital and Massachusetts General Hospital), and Stanford University analyzed four existing guideline systems to derive a set of requirements for guideline representation. The GLIF specification is a consensus representation developed through a brainstorming process. Four clinical guidelines were encoded in GLIF to assess its expressivity and to study the variability that occurs when two people from different sites encode the same guideline.

Results: The encoders reported that GLIF was adequately expressive. A comparison of the encodings revealed substantial variability.

Conclusion: GLIF was sufficient to model the guidelines for the four conditions that were examined. GLIF needs improvement in standard representation of medical concepts, criterion logic, temporal information, and uncertainty.

■ **JAMIA.** 1998;5:357−372.

Recently, market pressures in the health care industry and the trend toward managed care have driven medical organizations to increase productivity and to reduce costs without adversely affecting patient care. One method that has been proposed to achieve these goals is to adopt institutional and national standard practice guidelines that foster effective patient care and reduce both practice variation and inappropriate use of resources.

The substantial work needed to develop good guidelines creates an incentive to make guidelines sharable among different institutions and implementable in computer-based applications. Currently, guideline-sharing is based on the dissemination of relatively unstructured documents. The utilization of these documents for building computer-based applications is indirect. Providing a common format for expressing guidelines is a necessary (but not sufficient) step in the direction of creating shareable guidelines. If guidelines could be encoded in a common representation format and shared among institutions electronically, we would reap four benefits. First, a repository of shareable guidelines would avoid duplication of effort among institutions that wish to use common guidelines. Second, a common electronic format would allow guideline amendments and modifications to be disseminated rapidly. Third, a common format would encourage the development of application tools that help health care practitioners retrieve and use guideline information. Fourth, a formal representation format would encourage guideline authors to be rigorous in the development process, producing guidelines with less ambiguity and fewer errors. Full realization of such benefits requires not only the creation of a common representation format but also the evaluation of a guideline repository, systems that support dissemination of guideline modifications, guideline information-retrieval applications, and guideline authoring tools.

In this article, we discuss the development of a common representation for guidelines. Our group, the InterMed Collaboratory, has developed a format called GuideLine Interchange Format (GLIF) that draws on our experiences with health care guidelines. The InterMed Collaboratory includes informatics workers from Columbia University, Harvard University (Brigham and Women's Hospital and Massachusetts General Hospital), McGill University, and Stanford University. The goal of InterMed is to facilitate collaboration in the development of medical information systems by using cooperative approaches to the development of infrastructure, methods, tools, and resources that can be shard via the Internet.[1,2] An important example of a shared resource is a health care guideline; thus, a primary task of the InterMed Collaboratory has been to develop GLIF.

Our ultimate goal is to develop a standard for representing guidelines that facilitates guideline-sharing across the software tools at different medical institutions that manipulate, analyze, or otherwise compute with an electronic representation of a health care guideline. We refer to such software systems as *guideline applications*. There is a wide range of capabilities among guideline applications in our institutions, and these applications have varying information needs. Nonetheless, we have identified guideline elements that are common to these applications and have developed GLIF as a formal way to specify them. Our hypotheses are that many guideline applications can be written to use GLIF-encoded guidelines and that GLIF will allow these applications to share health care guideline information. In this paper, we propose a common representation, show that it can be used to encode different types of guidelines, and present a study of the encoding process.

We begin by providing a brief discussion of the goal of moving from paper-based guidelines to computer-based guidelines and by presenting a description of currently recognized guideline types and formats. We give a description of the GLIF development process, including an analysis of four precursor guideline systems that contributed to that development, as well as the design of our pilot study in which we encoded four clinical practice guidelines. We report on our experiences with the encoding task and then conclude with a discussion of the limitations of GLIF and future directions.

Background

In 1990, the Institute of Medicine defined *practice guidelines* as "systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances."[3] In 1992, the Institute's Committee on Practice Guidelines clarified this definition further by adopting the following definition of *appropriate care*: "the expected health benefit exceeds the expected negative consequences by a sufficient margin that the care is worth providing."[4]

A 1990 survey by Audet et al.[5] showed that, although guidelines were being widely promoted, organizations were devoting far more attention to guideline development than to guideline implementation, where implementation included dissemination, training, monitoring, and problem identification. This situation still holds, but there is increasing recognition

that computer-based systems offer new opportunities for guideline implementation. The Committee has stated that information and decision-support systems are crucial elements in long-term strategies for promoting the use of guidelines.[4]

Paper-based versus Computer-based Guidelines

Although there are relatively few health care sites where practitioners routinely use computer-based implementations of guidelines, there is good evidence that such computer-based systems can have a positive effect on patient care. In a survey of the literature, Johnston et al.[6] identified 28 clinical trials that assessed the effects of computer-based clinical decision-support systems on clinician performance and patient outcome. The studies showed beneficial effects from systems that provided support for drug-dose determination, intervened with preventive-care reminders, or provided recommendations for the care of active medical problems. Few of the studies showed significant improvements in patient outcome, but the authors concluded that further studies are needed to assess clinical effects on patients.

Not only can computers improve the implementation of guidelines, but the task of translating guidelines into computer-based formats can highlight deficiencies and lead to revisions that make those guidelines more useful.[3] Failure to develop rigorous approaches to guideline construction prompted Shiffman and Greenes[7] to apply decision-table methods to verify guideline completeness and to detect inconsistencies and redundancies. A representational form of guideline knowledge that promotes completeness and minimizes inconsistencies and redundancies is essential if we want to implement and share guidelines for computer-based applications.

Specifying guidelines with sufficient detail for computer use can be difficult because greater precision may be required than is typically found in paper descriptions.[8] Because of author oversight, reliance on "common knowledge," or a reluctance to be specific when scientific evidence is not available, paper-based guidelines may omit details of patient assessment, fail to specify care plans for certain conditions, or contain ambiguous directives regarding what actions are appropriate.[9,10] Conversion of a paper-based guideline to a computer representation tends to reveal these flaws and to require that they be addressed. Tierney et al.[8] found it difficult to incorporate the heart-failure guidelines published by the Agency for Health Care Policy and Research (AHCPR) into their local clinical information system, noting that many terms used in the text-based guideline were defined inadequately ("left ventricular systolic dysfunction"), criteria for

decisions were unclear ("continued symptoms"), patient states were described inadequately ("drug intolerance"), modifiers were vague ("frequently," "recurring"), and other recommendations were vague ("when needed"). These investigators recommended that guidelines be written in a simple if–then–else format, with all parameters defined strictly based on routinely collected clinical data.

In contrast to paper presentations, computer-based guideline applications can include additional didactic material, such as images, videos, sounds, simulations, and links to bibliographic databases. Given patient data—obtained either from the user in response to queries or through direct access to an electronic medical record—guideline applications can automatically select the next step in the guideline.[11,12] Some guideline applications can function autonomously (for example, by providing alerts[13,14] or by selecting patients for enrollment in treatment protocols[15]). Lam incorporated practice guidelines in an environment that used computer-based suggestions, as well as reminders, to identify important deviations.[16] Protocols for long-term problems, such as those for cancer chemotherapy[17] and for treatment of AIDS[18,19] and other conditions,[20] have lent themselves to computer-based data collection methodologies, and have spawned knowledge-sharing and knowledge-reuse projects.[21] An important aspect of guidelines is the determination of eligibility of the patient; tools for evaluating eligibility have been developed by Tu et al.[15]

Although there are advantages to using computer-based guideline applications, there are also drawbacks: Such applications often are difficult to develop and validate, have access to limited data, and may not capture all the nuances and uncertainties expressed in natural language. Several groups have experimented with implementing guideline applications, but only a few such applications have been shared among institutions.[22]

Guideline Types and Formats

The 1992 Institute of Medicine report lists five types of guidelines that are categorized by purpose. We use these categories in our discussion of guideline types implemented by existing systems and in the selection of guidelines for evaluation of GLIF. The guideline types, and examples provided by the Institute's report, are as follows:

- *Screening and prevention*: Vaccination for pregnant women who are planning international travel.

- *Diagnosis and prediagnosis management of patients*: Evaluation of chest pain in the emergency room.

- *Indications for use of surgical procedures*: Indications for carotid endarterectomy.

- *Appropriate use of specific technologies and tests as part of clinical care*: Use of autologous or donor blood for transfusions.

- *Guidelines for care of clinical conditions*: Management of patients following coronary-artery bypass graft

Another way to characterize a guideline is by the format in which the knowledge is presented. The Institute's report defines *effective formatting* as ''presenting guidelines in physical arrangements or media that can be readily understood and applied by practitioners, patients, or other intended user groups.''[4] The report presents 16 example guidelines that demonstrate a wide variety of formats, including the most common, narrative text, tables, and flowcharts, as well as graphs, maps, photographs, lists, critical pathways, and if–then–else statements.

Narrative text, tables, and other formats that work well for human processing may not work so well for computer processing. If the computer system is expected to provide patient-specific recommendations in a timely fashion and direct the user's attention to only the most pertinent portions of the guideline, formats that provide greater structure and explicit data representations are required. Thus, effective formatting for computer-based guideline systems must be not only appropriate for the end user but also effective for the intermediate representation used by the computer software to manipulate guideline knowledge and patient data.

Development of GLIF

In April 1996, the Decision Systems Group at the Brigham and Women's Hospital sponsored an InterMed workshop on guideline representation to determine the requirements of a shareable representation and to define a preliminary guideline-interchange format. For us to develop consensus, it was essential that we understand the background and experience of one another. At the workshop, collaborators from the four InterMed sites presented their previous work. Terminology about guideline systems varied markedly, and different groups emphasized different aspects of computer-based guidelines. After gaining a solid understanding of each group's perspective, workshop participants brainstormed to develop a consensus representation.

In the following sections, we describe and analyze the four precursor systems that contributed to the choices

that we made for GLIF, directing our attention to certain features deemed relevant for comparison of systems. Based on this analysis, we derived a set of requirements that is important for a shareable representation. Our prior experience with guideline systems at each institution and our conclusions about requirements led us to the development of GLIF. Finally, we present the GLIF specification and provide a brief example of a GLIF-encoded guideline.

Analysis of the Four Guideline-representation Systems

Topics that we considered relevant to guideline representation in each of the four systems were the following: purpose of the system; examples of guidelines implemented; basic structure and building blocks for storing the knowledge; input and output; representation of sequences of decisions and actions; and representation of eligibility criteria and criteria for making transitions between events. In this section, we give a general description of each system with respect to these features.

Implementation of Medical Logic Modules at Columbia

Medical Logic Modules (MLMs)[23] are software modules, written in the Arden Syntax,[24–26] that, when implemented in a clinical information system,[27] run on databases of patient data and typically provide alerts or reminders for clinicians.[28,29] The purpose of an MLM is to store knowledge required for triggering an action based on data in a patient database. Most MLMs generate messages that are sent to the health care provider. Some messages are alerts that warn the provider of a potentially worrisome patient situation. Other messages provide interpretations of existing patient data or may provide a list of patients who fulfill certain criteria, such as eligibility criteria for clinical trials. Reminders are stored in the electronic patient record so that the clinician can review them during a patient visit.

Examples of guidelines implemented at Columbia are hypokalemia with digoxin therapy (alert), tuberculosis cultures with positive or invalid results (alert), calculation of creatinine clearance (interpretation), and identification of patient records that include admissions but no discharge summaries (quality-assurance tracking). In terms of the Institute of Medicine guideline categories, most MLMs have been designed for screening and prevention. Test interpretation and identification of patients with particular characteristics are not specific guideline categories listed by the Institute but are important examples of guideline types implemented as MLMs.

An MLM contains a *data slot* that specifies the mapping from terms in the MLMs to the corresponding terms used in the local patient database. An *evoke slot* specifies the data element that will cause the logic slot to act. The *logic slot* specifies what needs to be true about the data element for the action to take place. The result of the logical statement is true or false. If it is true, then the action in the *action slot* is executed; if it is false, then no action is taken. In addition to these knowledge slots, there are slots that store documentation about maintenance of the module and references to the literature.

Data input to the system are usually patient data from the database; the output usually is messages to providers. MLMs were originally designed to have a single set of data for input, a single application of criteria logic, and a single set of resulting actions. In more complex guidelines, there may be sequences of actions, each requiring its own set of data, and choices about actions may depend on the results of previous actions and decisions. Sherman et al.[30] studied methods for chaining together multiple MLMs, but such chaining is not well supported by the current Arden Syntax standard.

GEODE-CM at Brigham and Women's Hospital

GEODE-CM is a system that combines guidelines with structured data entry and data retrieval from a clinical database.[31] The purpose of the system is to suggest pertinent data to be collected, decisions to be made, and actions to be taken for a given clinical management state. The developers have emphasized its use for education of medical students, residents, and practicing physicians. GEODE-CM therefore concentrates on presenting recommended processes of care rather than on providing alerts and reminders.

The GEODE-CM screen display for a particular clinical management state initially shows a clinical summary of the information that is currently known about a patient. Subsequently, the display is organized like a SOAP (subjective, objective, assessment, and plan) note. The plan section is composed of actions and transitions. *Actions* include diagnostic tests to be done and therapies to be administered. *Transitions* specify clinical management states to which the patient may go next and criteria that determine whether the patient can reach those states.

An example of an implemented guideline is the diagnostic workup of a breast mass. This guideline belongs to the Institute of Medicine category *diagnosis and prediagnosis management of patients*. Other guidelines that would be appropriate for GEODE-CM are those in the category *guidelines for care of clinical conditions*.

The basic building blocks for storing guideline knowledge in GEODE-CM are nodes that represent clinical management states. A *node* is specified by a unique identifier, a name, pertinent data, actions, warnings, transitions, and didactics. *Didactics* are supplemental material; they include narrative text or links to text, graphics, or video accessible on the World Wide Web.

Input to GEODE-CM includes patient data both retrieved from the database and entered by the user. GEODE-CM differs, therefore, from MLMs in its incorporation of methods to guide data entry by the clinician. The output consists of recommendations for what data to collect, what actions to follow, and what clinical management state to go to next.

Sequences of decisions and actions are important in GEODE-CM. There are specific *eligibility criteria* for entering the guideline initially, a start node, and additional nodes linked by transitions. *Transitions* specify *transition criteria* for going to another particular state. Both eligibility criteria and transition criteria are intended to be logical statements in conjunctive normal form. A clinical management state may contain more than one action if the order of performing the actions does not matter.

MBTA at Massachusetts General Hospital

MBTA is an architecture for building large knowledge-based medical systems, tailored especially to providing clinical reminders and practice guidelines.[32] It does not specify a particular syntax, and it is not designed only for specifying computer-based clinical guidelines. Instead, it is a general method for managing procedural modules and data objects. Workers have applied this method to health care guidelines and built a system that provides recommendations to clinicians for particular patients. MBTA separates guideline knowledge and logic from the electronic patient-record system.[33] A client, such as a clinical workstation or a World Wide Web browser, connects to and interacts with the MBTA guideline server.

Examples of guidelines implemented are treatment of urinary incontinence, screening for and treatment of hypercholesterolemia, management of low-back pain, and selection of patients for influenza vaccinations. Using the Institute of Medicine categorization, these guidelines would be examples of *screening and prevention* (influenza vaccination, hypercholesterolemia) and *guidelines for care of clinical conditions* (low-back pain, urinary incontinence, and hypercholesterolemia).

The basic components of the MBTA system are modules, objects, and explainers. An *MBTA module* is the procedural portion, similar to a procedure in a traditional programming language. However, instead of

passing in parameters, MBTA enables modules to use objects for input and output. An *MBTA object* is like a C++ object but also provides temporal representation and the ability to support automated vocabulary mapping. An object is a data structure that has a unique name and a set of data fields. Each field has a name and a value or list of values. For example, a cholesterol object might have four data fields that are relevant to a patient who is concerned about cholesterol: age, diseases, gender, and cholesterol level.

An *MBTA explainer* is a special type of module that uses previously created groups of objects to produce an explanation in narrative text about those objects. The explanation about a particular object is stored separately from that object, thereby making it possible for the system to generate different explanatory descriptions for the same object.

Input to the MBTA system comes from the clinical database; patient data are packaged as MBTA objects. Output from the MBTA system is in the form of explanations.

Sequences of decisions and actions are represented in the procedural code of MBTA modules. Eligibility criteria and transition criteria are not specifically labeled as such, but criteria that serve the same functions are embedded in the procedural code of the modules.

EON at Stanford

EON is a component-based architecture for building systems that provide decision support for guideline-based care.[34] The components of EON that have been implemented identify guidelines that are appropriate for a given patient, track a patient's progress through a guideline, recommend appropriate tests and therapies for a given patient at a given time, and compare a clinician's management plan with the plan recommended by the guideline.

An example of a guideline implemented in EON is a clinical-trial protocol for breast cancer treatment. Clinical-trial protocols belong to the IOM category of *guideline for care of a clinical condition*.

Basic knowledge structures included in the EON representation are protocol steps, intervention states, eligibility criteria, conditions, and revision rules that change the current intervention being considered. *Protocol steps* are intervention steps that specify an intervention, assessment steps that specify data to collect to assess or diagnose a patient, or a combination of intervention and assessment steps. *Intervention states* may be active, completed, aborted, or suspended to indicate how far along an intervention is. Such states are particularly important for treatment protocols in which, for example, a round of outpatient chemother-

apy may be aborted or suspended if a patient does not tolerate the chemotherapeutic agents. *Revision rules* can change the current state of an intervention or modify properties such as the dose of a prescribed drug. As in GEODE-CM, *eligibility criteria* are represented explicitly. *Conditions* contain the conditional logic for progressing from one protocol step to another and for making modifications in interventions.

Input to EON includes both patient data obtained from a clinical database and data acquired at run time from the user. The output is recommendations for actions to be taken by the clinician.

To represent sequences of decisions and actions in EON, guideline authors specify eligibility criteria for entering the guideline, a start step, and one or more steps that may be taken following each step. Often there are conditions specified that must hold true for a particular next step to be taken. In other cases, a choice is not specified by the guideline, with the expectation that user preferences will dictate which next step to follow (although user preferences are not formally represented). In either case, a selection must be made regarding which step to take next. If there are multiple possible next steps, the guideline author can indicate whether one of those steps, some of those steps, or all of those steps must take place.

Other components of guideline management that are important in EON are a controlled vocabulary, a temporal database manager,[35] and decomposition of steps into finer-grained steps. Patient data elements mentioned by the guideline must be mapped to concepts defined in the controlled vocabulary. The temporal database manager can recognize temporal patterns in data stored in a database, such as a situation in which a patient has a low platelet count 7 to 14 days after receiving the second course of a chemotherapy regimen. A collection of fine-grained steps, such as a set of drug prescriptions, can be aggregated to a composite step, such as a chemotherapy regimen.

Requirements for a Shareable Guideline Representation

Analysis of the four existing systems revealed a number of commonalities and several ideas that were unique to given systems. Clearly, a wide variety of terminologies were used by the different research groups, and our goal was to acquire an understanding of which terms represented features that were common across systems and which represented features that were unique. For example, a *node* in GEODE-CM represents a clinical management state, which is similar to but not equivalent to a *step* that represents an action in EON. A node in GEODE-CM is followed by

one or more *transitions*, each of which has *transition criteria*. A step in EON is followed by a *selection* of one or more *conditions*. Again, these constructs are similar but not exactly equivalent. Certain features, such as a temporal database manager or specific links to the World Wide Web, are unique to particular systems.

Creating a GLIF specification required moving beyond differences in local terminology to determine desired functionality and to choose a mutually acceptable terminology for guideline structural elements. From our discussions, we developed a set of representational requirements for the GLIF model.

Representation of Complex Guidelines with Branching Logic

All four systems share the goal of representing knowledge about guideline steps, with the intent of providing clinical recommendations given patient data. We all wanted to be able to represent complex guidelines with branching logic, including guidelines that recommend diagnostic workups, therapies for particular conditions, and screening and prevention.

GEODE-CM and EON emphasize multiple steps linked together with branching possibilities. MLMs in the current standard Arden Syntax tend to be used for simpler guidelines with a single combination of a logical statement and an action. However, if MLMs were linked together, they could potentially represent more complex guidelines. MBTA can also represent complex guidelines, but the representation of choices and actions is less explicit. We concluded that for a shareable representation, it is important to represent explicitly the steps that specify actions or choices in the guideline. Each of the four systems could be mapped to such a representation.

Representation of Patient Data Elements

The decision-support system of implemented MLMs running at Columbia–Presbyterian Medical Center is triggered by data in the patient database used for actual patient care. The other three systems are research prototypes that do not currently interact with real patient data but are intended to be triggered by stored patient data or by data requested as input from the user. MLMs have a data slot for specifying which data the program must retrieve from the database and how those data elements map to the clinical parameters in the logic slot; GEODE-CM has a section for pertinent data, which may refer either to data to collect or to data already stored; MBTA has objects that contain data from the database packaged in a particular manner; and EON has data elements that conform to a controlled vocabulary linked to the system. We certainly need data structures that link the guideline to

the data elements in the database or to data elements that are entered by the user. However, for the purposes of this initial version of GLIF, we chose to avoid the problem of controlled vocabulary, because there is currently no universal clinical vocabulary standard for the electronic medical record.

Representation of Eligibility Criteria

GEODE-CM and EON both have constructs for representing eligibility criteria. MLMs have only one logic statement that serves as the criterion for performing a designated action, and this statement can be viewed as the eligibility criteria for a single-action guideline. However, eligibility criteria are intended to be used specifically for determining patient entrance into a complex guideline; thus, this type of criteria is set apart as distinct from all others within the guideline. In a shareable representation, we propose that eligibility criteria for entrance into the guideline should be labeled as such.

Representation of a Starting Point

We agreed that there needs to be a specific starting point to the guideline. Although we believed that it would be beneficial to permit entry to the guideline at any point, we concluded that we would limit entry to a single point in the initial version of GLIF. A starting point is not relevant to MLMs that follow the current Arden Syntax standard, because MLMs do not occur in series but instead fire independently. However, it is possible to assign priorities to indicate which MLMs should fire before others. MBTA modules do have implicit starting points but do not label those points explicitly. GEODE-CM and EON both identify a starting point—a clinical management state in GEODE-CM and a protocol step in EON—and we concluded that a starting point would be necessary for the shareable representation.

Representation of Actions

GEODE-CM has actions, EON has interventions, MLMs have action slots, and MBTA permits representation of actions within the code of MBTA modules. Furthermore, guidelines can specify a considerable amount of detailed information about an action. A therapeutic intervention, for example, might specify drug, amount, route of administration, and frequency. It is essential that a guideline representation provide for the representation of these details. In GLIF, we chose to have an action entity that contains only a single action.

Representation of Options That Follow an Action

One node in GEODE-CM can lead to multiple transitions, and a protocol step in EON can lead to a se-

lection of multiple possibilities. Given the goal of representing branching logic, it was clear that our shareable representation also needed a way to represent multiple options that follow an action. EON developers have found the constructs *one-of*, *some-of*, and *all-of* to be useful, and we adopted this approach as well. If multiple actions must be performed concurrently, there must be a mechanism for representing the point at which the multiple paths converge. The mechanism used in the EON system makes use of a construct called a *synchronization step*. We have adopted this same approach for GLIF.

Representation of Criteria for Proceeding

Every action that may occur in the guideline must be preceded by criteria for proceeding to that action. All four systems have ways of representing such criteria: transition criteria in GEODE-CM, predicates in EON, the logic slot in MLMs, and code embedded in MBTA modules. These systems permit true or false values for each criterion. We considered the additional possibility that there would be situations in which it would be desirable to specify that a subset of n criteria must be satisfied before proceeding, and we included this option in the shareable representation. Developers of the four systems intend that the logical statement for a criterion take the form of a predicate–logic sentence, but none of those systems enforce a particular structure to logical statements. We recognize the need for greater structure to represent criteria, but in the initial version of GLIF we leave these statements as narrative text.

Decomposition of Actions

For complex guidelines it is useful to be able to decompose actions into smaller actions. The EON sys-

tem includes this ability because it was designed for complex clinical-trial protocols. We believe that a strategy for decomposing actions is important for any complex guideline that we might want to encode in our shareable representation.

Links to Guideline Resources

MLMs include a *citations slot* and a *links slot* that specify relevant citations in the literature and links to other supportive information, respectively. GEODE-CM provides the option of referencing supplemental material in association with every node, action, transition, and pertinent datum. MBTA emphasizes the importance of explaining the reasoning behind recommendations to users, and through the explainers can refer to or summarize published guideline information. We concluded that providing links to text, citations, and World Wide Web resources is crucial for the shareable representation.

GLIF Specification

The GLIF specification consists of the GLIF model and the GLIF syntax. The *GLIF model* consists of a set of classes for guideline entities, attributes of those classes, and data types for the attribute values. A particular guideline encoded in GLIF is an instance of the general guideline model. Figure 1 depicts the classes and attributes in the model. (See the GLIF Web site[36] for a full specification of the GLIF model.)

To encode a guideline according to the GLIF model, authors express the guideline knowledge using *GLIF syntax*. The GLIF syntax specifies the format of the text file that contains the encoding. Although we do not present a formal specification of the syntax here, we present, in Figure 2, sample text from a guideline encoded using the syntax.
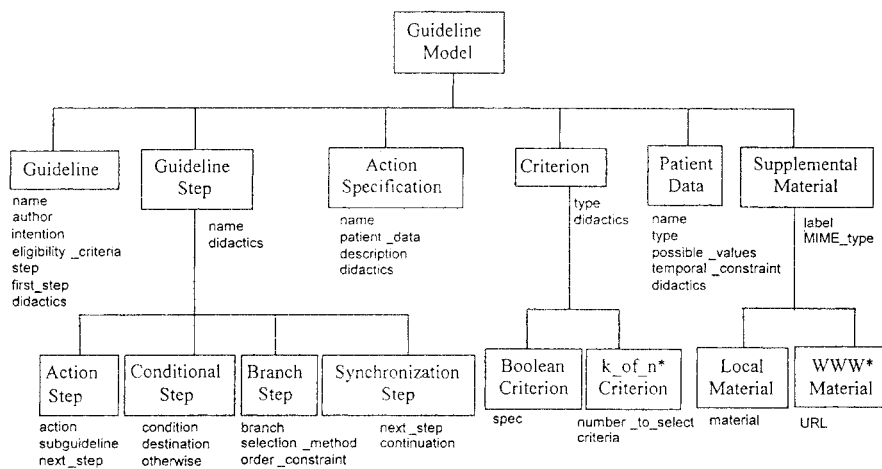

Figure 1 GLIF classes and attributes. Classes (shown in boxes) are arranged in a hierarchy. Each class has its own attributes (shown in a list) as well as the attributes it inherits from the class above. NOTE: k_of_n indicates k criteria from a total of n criteria; WWW, World Wide Web.

GLIF Model

The GLIF guideline object has a name, a list of authors, a characterization of the guideline's intention, a specification of the patient-eligibility criteria, an unordered list of all steps in the guideline, an indication of the starting step in the guideline, and a list of supporting didactic material.

The *guideline intention* is a characterization of the purpose of the guideline, currently encoded as narrative text. *Eligibility criteria*, encoded as as set of *criterion* objects, are conditions that must be true before a guideline can be applied to a particular patient. *Didactics* provide background or supporting information; such information may be stored locally as text (including bibliographic citations) or provided by reference to a uniform resource locator (URL) that designates a site on the World Wide Web. Other classes in GLIF also allow optional inclusion of external didactic information.

The guideline specification consists of a collection of steps that are linked together in a directed graph. There are four types of *guideline steps*: action steps, conditional steps, branch steps, and synchronization steps.

Action steps specify clinical actions that are to be performed in the patient-care process. An action step may name a *subguideline*, which provides greater detail for the action. Each action step contains exactly one action specification and one pointer to the next step in the guideline. An *action specification* has a name, a list of patient data, a description of the action in narrative text, and an optional list of associated supporting didactic materials. If the action involves the collection of patient data, such data are specified as a set of data elements associated with that particular action. A *patient data element* is defined by a name, a data type, an optional list of acceptable values, a temporal constraint on how recent a value must be to be considered valid, and a list of supporting didactic materials. If the action is purely therapeutic and involves no data collection, then the set of patient data is empty.

Conditional steps direct flow from one guideline step to another. A conditional step may link any guideline step to any other guideline step. A conditional step contains a condition, or *criterion*, which is a logical statement that may be evaluated as true or false. If the condition is true, then control flow goes to the step specified by the *destination* attribute. Alternatively, if the criterion is false, control flow goes to the step specified by the *otherwise* attribute. No transition is specified if available data do not allow evaluation of the criterion.

```
Guideline Example
{
    name = "Guideline for Vaccine X";
    authors = SEQUENCE 1 { "Mary Doe, MD"; };
    eligibility_criteria = NULL;
    intention = "Decide whether to recommend the Generic vaccine
                and at what dosage";
    steps =
        SEQUENCE 8
        {   (Branch_Step 1);         (Action_Step 1);        (Action_Step 2);
            (Synchronization_Step 1);  (Conditional_Step 1);   (Conditional_Step 2);
            (Action_Step 3);         (Action_Step 4);
        };
    first_step = (Branch_Step 1);
    didactics =
        SEQUENCE 1
        {
            Supplemental_Material 1
            {   label = "critique";
                MIME_TYPE = "text/plain";
                material = "Published guideline does not contain explicit eligibility criteria.";
            };
        };
}
//_____
Branch_Step 1
{
    name = "collect data"
    branches =
        SEQUENCE 2
        {   (Action_Step 1);
            (Action_Step 2);
        };
    selection_method = all_of;
    order_constraint = any_order;
    didactics = SEQUENCE 0 {};
}
//_____
Action_Step 1
{
    name = "Get occupation";
    action = . . . (see 37)
}
//_____
Action_Step 2
{
    name = "Get age";
    action = . . . (see 37)
}
//_____
Synchronization_Step 1
{
    name = "wait until data collected";
    next_step = (Conditional_Step 1);
    continuation = wait_for_all;
    didactics = SEQUENCE 0 {};
}
//_____
Conditional_Step 1
{
    name = "Age < 12?";
    condition =
        Boolean_Criterion 1
        {   type = k_two_valued;
            spec = "AGE < 12";
            didactics = SEQUENCE 0 {};
        };
    destination = (Action_Step 3);
    otherwise = (Conditional_Step 2)
    didactics = SEQUENCE 0 {};
}
//_____
Action_Step 3
{
    name = "Pediatric dosage";
    action =
        Action_Spec 3
        {   name = "Administer pediatric dosage",
            intension = "Administer pediatric vaccine dosage";
            patient_data = SEQUENCE 0 {};
            description = "inject 1 cc of vaccine X into muscle";
            didactics = SEQUENCE 0 {};
        };
    subguideline = NULL;
    next_step = NULL;
    didactics = SEQUENCE 0 {};
}
```

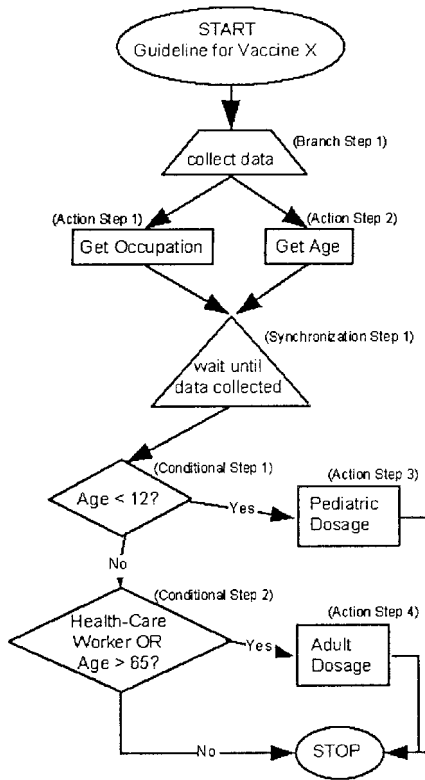**Figure 2** A portion of a GLIF encoding of a simple vaccine guideline.

Figure 3 Graphic representation of the simple vaccine guideline.

Branch steps direct flow to multiple guideline steps. The author specifies whether all, some, or only one of these steps must take place. In addition, the author specifies whether the steps are to occur in a specified order or in parallel.

Synchronization steps are used in conjunction with branch steps. When a branch step is followed by multiple guideline steps, the flow of control must eventually converge in a single step. Each branch may lead to a series of steps, resulting in a set of branching paths. The step at which the paths converge is the synchronization step. The direct predecessors of a synchronization step can be any type of guideline step; at some previous point, however, there must have been a branch step. When the flow of control reaches the synchronization step, a *continuation* attribute specifies whether all the preceding steps must have been completed before control can move to the next step or whether just one of the branches needs to have been completed before control can move on.

Sample GLIF-encoded Guideline

To demonstrate the encoding of a guideline in GLIF, we give a simple example, using a hypothetic text-based guideline for administration of a vaccine:

Patients at high risk should receive the vaccine. Patients at high risk include health care workers, patients older than 65 years, and children under 12 years. Children under 12 years should receive the pediatric dosage. Other patients for whom the vaccine is indicated should receive the adult dosage.

A portion of the GLIF-encoded version is shown in Figure 2. (The full guideline can be found at the GLIF Web site.[37]) A graphic representation of the entire guideline is shown in Figure 3.

The sample guideline relies on data about a patient's age and occupation to determine recommendations for administering the vaccine. The text-based version of the guideline does not specify the order in which age and occupation data should be collected. Therefore, the encoded version of the guideline uses a branch step, which leads to two action steps for collecting these data. The branch step says that the two actions can be performed in any order. Following collection of age and occupation data, the flow of control leads to a synchronization step, where control waits until both action steps have been completed before moving on.

The guideline then uses a conditional step to determine whether the patient is under 12 years of age. If so, the next step is an action step, in which the pediatric dose is administered. Otherwise, control passes to a conditional step to determine whether the patient is a health care worker or is older than 65 years. If one of these characteristics holds, then the conditional step indicates, by the destination attribute, that the next step is an action step, in which the adult dose is administered. If neither criterion is met, the guideline finishes.

Evaluation of Encoding Guidelines in GLIF

We performed a pilot study to assess the expressivity of GLIF in the representation of four selected clinical guidelines and to measure the variability that occurs

Table 1 ■

Guidelines Encoded in the GLIF Evaluation

Guideline	Encoders	Reviewer
Breast-mass workup guideline	BWH, Stanford	Columbia
Breast-cancer treatment protocol	BWH, Stanford	MGH
Cholesterol screening and management	BWH, MGH	Columbia
Influenza-vaccine recommendations	Stanford, MGH	BWH

NOTE: BWH indicates Brigham and Women's Hospital; MGH, Massachusetts General Hospital.

when two encoders encode the same text-based guideline. The selected guidelines were for influenza vaccination,[38] cholesterol screening and management,[39] breast-mass workup,[40] and breast cancer treatment protocol.[41] We chose these four guidelines because at least one InterMed site had experience with each guideline and because each guideline is of a different type. Using the Institute of Medicine categorization, we included one guideline whose purpose is *screening and prevention* (influenza vaccination), one guideline intended for *screening and prevention* that also can be categorized as a simple *guideline for care of a clinical condition* (cholesterol management), one guideline that reflects *diagnosis and prediagnosis management* of patients (breast-mass workup), and one relatively complex *guideline for care of a clinical condition* (breast cancer treatment protocol).

For each guideline, two researchers from different InterMed sites independently encoded the guideline from the same text-based source. For the breast-mass workup and the breast cancer treatment protocol, a flow chart was also available for use by encoders. None of these guidelines was originally authored for direct use in a computer-based application. After the two encoders assigned to each guideline had completed the encoding task, a third researcher from a different site reviewed the pair of encodings. Table 1 shows the test guidelines, the sites responsible for encoding, and the sites responsible for reviewing.

The encoders found the encoding task relatively straightforward, although sometimes laborious and time consuming.* None of the encoders reported any problems in performing this task, and each thought that GLIF was adequately expressive. However, this may be simply because all encoders were members of the InterMed group that developed GLIF and thus shared an understanding of how the classes were intended to be used. In contrast, a comparison of the encodings created for the same guideline by two different encoders revealed substantial variability.

In Table 2, we show the number of each type of GLIF object used in each encoding. These data demonstrate that it is possible for two people to encode the same guideline and to make different modeling decisions about when to use action steps, conditional steps, branch and synchronization steps, and patient data

Table 2 ■

Number of Objects in GLIF from Two Different Encodings

	Encoding 1	Encoding 2
Breast-mass Workup Guideline		
Guideline	1	3
Steps		
Action step	19	20
Action specification	19	13
Conditional step	19	12
Branch/synchronization steps	1	6
Criteria		
Boolean	19	12
k_of_n	0	0
Patient_data	11	15
Breast Cancer Treatment Protocol		
Guideline	1	4
Steps		
Action step	31	26
Action specification	30	11
Conditional step	13	17
Branch/synchronization steps	12	13
Criteria		
Boolean	14	14
k_of_n	0	0
Patient_data	5	21
Cholesterol Guideline		
Guideline	1	5
Steps		
Action step	33	41
Action specification	12	28
Conditional step	15	29
Branch/synchronization steps	6	0
Criteria		
Boolean	12	26
k_of_n	3	0
Patient_data	13	11
Influenza Vaccine Recommendations		
Guideline	1	3
Steps	40	38
Action step	19	20
Action specification	19	13
Conditional step	19	12
Branch/synchronization steps	1	6
Criteria	19	12
Boolean	19	12
k_of_n	0	0
Patient_data	11	15

NOTE: k_of_n indicates *k* criteria from a total of *n* criteria.

elements. We sought to understand sources of variation by analyzing the resultant encodings to find reasons for such differences.

One type of variability was due to differences in interpretation of the sequence in which data elements should be collected. For example, published guidelines often recommend that specific actions be performed if certain values for particular data elements

---

*The encoding of a guideline from natural language text to GLIF is a cognitive task and could have different results depending on the background knowledge and perspective of the encoder. Recognizing this problem, we have studied the process using cognitive-evaluation techniques and provide a detailed description of this process and resulting lessons elsewhere.[42]

are obtained. Those guidelines may not specify, however, in what order the data should be collected. We found examples in which certain encoders decided to organize data collection in parallel, using branch and synchronization steps, whereas other encoders used their background medical knowledge to propose a more realistic collection of information (e.g., history items first, followed by physical examination, followed by laboratory tests).

Frequently, the encoded guideline varied in its level of detail. For example, for the influenza-vaccination guideline, the concept "no contraindications to influenza vaccine" was specified in one encoding whereas "no sensitivity to eggs" and "no history of hypersensitivity to the vaccine" were specified in the other encoding. Similarly, in the breast-mass workup guideline, the concept "risk factors for breast cancer" was specified in one encoding whereas "history of fibrocystic disease" and "late first pregnancy" were specified in the other encoding. In the breast-cancer treatment protocol, one encoding asked "OK to start chemotherapy?" whereas the other asked specific questions such as "serum creatinine < 2 mg/dl?" and "SGOT < 1.2 times normal?"

The number of conditional steps varied depending on whether criteria were specified as atomic sentences or as composite sentences. We define an *atomic sentence*, in this context, as a logical sentence in which only one data element is checked against one value. We define a *composite sentence* as a logical sentence that is the conjunction or disjunction of atomic sentences. Typically, a composite sentence checks the values of multiple data elements. For example, one encoder used a series of three conditional steps in which each step tested one data element ("age > 3," "age < 18," and "on aspirin") whereas the other encoder used a single conditional step that stated several tests in the same criterion ("3 < age < 18 and on aspirin"). This difference resulted in a different number of criteria and conditional steps for the same guideline.

There were significant differences in choices made in the specification of data elements. Sometimes terms could be viewed as synonyms, as for example "health care worker" or "health care provider." In other cases, the differences were more extreme. For example, a data element in one encoding of the breast cancer treatment guideline was "bone-marrow aspirate and biopsy" and a data element in the other encoding was "bone-marrow cell-morphology statistics." It is not clear whether these entities are the same or whether the possible values of these data elements would be the same.

Some differences occurred because of omission of guideline information by an encoder. In the breast-mass workup guideline, for example, the data element "menstrual status" was not used in any criterion in once encoding of the guideline; this omission would have resulted in an incorrect recommendation for a patient who was menstruating at the time of visit. In the influenza guideline, the data elements "HIV status" and "pregnancy" were not used in criteria for administering the vaccination in one of the encodings, and the two encodings of this guideline were not equivalent. In the cholesterol guideline, both encodings included the data element "LDL cholesterol," but one encoding did not include the criterion "LDL > 220," thereby neglecting to give recommendations for high LDL values.

In summary, we found that it was possible—if not inevitable—that two encoders would encode the same guideline in two different ways. Sources of variation included differences in the order in which data elements are collected, differences in the level of detail represented, differences in the use of atomic sentences or composite sentences in criteria, differences in the specification of data elements, and omissions due to human error. We believe that it is not necessary to guarantee uniqueness of encoding with a shareable guideline-representation format. However, the results of this small study do point out the need for authors of text-based guidelines to specify the order of steps when order is important and to be specific about the level of detail required. In this experiment, the encoders who had a medical background were more likely to enhance what was written in the text-based guideline when they thought that it was necessary to provide a clear computer-based guideline than were encoders who did not have a medical background. (See also Patel et al.[42])

Differences in the use of atomic sentences and composite sentences would be less common if criteria were represented in a more structured form. The current version of GLIF uses narrative text to represent logical criteria, but future versions could have additional classes that explicitly represent conjunction, disjunction, negation, and specific operators such as *is equal to*, *is greater than*, and *is less than*.

Differences in the selection and naming of data elements are a fundamental problem for sharing guidelines that act on clinical data in patient databases. If a standard clinical vocabulary and a standard data model for clinical data existed, encoders of guidelines would be more likely to encode guidelines in the same way. Thus, the development of shareable guidelines requires more than a common representation such as GLIF and is closely tied to the development of standards for the representation of patient data.

Finally, omissions due to human error are clearly undesirable. Ideally, the encoders of guidelines would also be the domain experts who create the guidelines, in which case error due to lack of knowledge would not be a problem. However, for domain experts to encode guidelines, we need good authoring tools that allow the experts to concentrate on the guideline itself rather than on the syntax of its formal representation. Even for encoders who are familiar with the encoding format, authoring tools can be beneficial because they can identify unintended gaps or inconsistencies in the guideline logic, and they can potentially make the encoding process less time-consuming. In this experiment, most of the encoders used word processors to write GLIF code, but in a few circumstances they used early versions of authoring tools created by Stanford University[43] and by the Brigham and Women's Hosptial.[31] The development of sophisticated authoring tools will be an important aspect of future work on computer-based guidelines.

Discussion and Future Directions

In addition to MLMs, GEODE-CM, MBTA, and EON, other projects are confronting the problem of guideline representation. For example, skeletal-plan instantiation (SPIN), developed by Uckun,[44] is used for protocol-based treatment planning, plan execution, and execution monitoring. SPIN uses skeletal-plan refinement, in a manner similar to EON.[34] It also has a syntax for specifying the flow control of various protocol steps, and an execution model that monitors mismatches among intended and expected effects of the guideline actions and observations. SPIN is limited to dealing with a subset of guidelines—namely, treatment protocols—whereas GLIF is designed to handle a greater variety of guideline types.

ASBRU is an intention-based guideline-execution language, developed by Shahar et al.,[25] that represents a guideline as a set of plans.[45] In ASBRU, each plan has a name and five components: preferences, intentions, conditions, effects, and a plan body that describes the actions to be executed. ASBRU's characteristic feature is that the system couples each plan with a rich set of intentions that reasoning modules can use to compare intelligently a clinician's management plan with the plan suggested by a guideline. ASBRU's language for flow control is similar to that of GLIF.

The PRESTIGE project in Europe[46] builds on the earlier DILEMMA generic protocol model (DGPM). In DGPM (as in GLIF), a protocol contains actions, transitions, and criteria for transitions. DGPM and GLIF share a hierarchic decomposition structure and a distinction between the prescriptive specification (pro-

tocol flow) and the actions. In DGPM, just as in GLIF, the implementation of a protocol step results in an action. Both methods use transition and transition criteria to define guidelines. However, DGPM models the states of an action. For example, an action may be in the state of being requested, rejected, started, suspended, abandoned, completed, and so on. DGPM depicts sequences of action states, not actions, whereas GLIF depicts sequences of actions. For example, in DGPM, a transition criterion may be defined between abandoning the administration of a drug and requesting a laboratory test. GLIF does not include that level of detail, which is instead incumbent on the application that utilized the shared guideline.

In this paper, we described our collaborative development of a guideline representation and our evaluation of the expressiveness of this representation. Recognizing that a standard representation of guideline logic had not yet evolved, the InterMed Collaboratory sought to identify features that were common and universally important by analyzing four existing systems and working toward a consensus model. This guideline representation needed to fulfill at least the requirements of the InterMed sites and potentially the requirements represented by guideline implementations at other institutions. Our evaluation of GLIF found that the representation was sufficient to model the clinical algorithms of the guidelines and protocols that were examined. On the other hand, there was considerable variability in how encoders in each institution used GLIF to model the same guidelines. We found the contributing factors to the variability to be ambiguities in the original guidelines, differences in medical expertise among encoders, differences in the level of detail at which guidelines were modeled, lack of a well-defined criteria language and common vocabulary, and errors and omissions.

We have tried to limit our requirements for further development as much as possible, but GLIF needs improvement in the following areas:

- *Representation of medical concepts*: Computer-based guidelines contain information about medical concepts, such as diseases, symptoms, physical findings, surgical procedures, and laboratory tests. Such concepts are essential for the expression of eligibility criteria, conditions, actions, and patient data in a GLIF-encoded guideline and for the representation and use of patient-specific data with a particular guideline. Unfortunately, different institutions do not always use the same clinical vocabulary or coding system for concept representation. In the absence of a standard clinical vocabulary, the shareability of computer-based guidelines that act on pa-

tient data is limited. Therefore, our ability to share GLIF-encoded guidelines depends on the evolution of vocabulary standards.

■ *Representation of criterion logic*: We currently represent conditional expressions (e.g., in eligibility criteria or in conditional steps) as strings. The guideline author writes such a criterion in narrative text using any vocabulary for clinical concepts and any other natural-language terms to create a readable clause. For example, a condition might be *patient is over 65 years old or patient has a chronic pulmonary disease, such as asthma, emphysema, or chronic bronchitis*. However, we do not have natural-language processors that can parse and make sense out of such clauses. Thus, a formal syntax for the representation of conditional expressions is essential. Existing standards include KIF, which supports first order predicate logic, and Arden Syntax, which supports Boolean expressions as well as simple temporal conditions.[25]

■ *Representation of temporal information*: Complex expressions that show temporal trends, such as *rapidly decreasing platelet counts*, are expressed as narrative text. Temporal expressions have been of great interest to the medical informatics community.[47–50] Shahar and Musen[51] have studied in depth the problem of detecting temporal patterns, such as direction and rate of change of laboratory test results. Recently, he and his colleagues[52] have developed a rich notation for specifying temporal intervals, which allows multiple time lines and uncertainty in starting time, ending time, and duration. Their work addresses the need to express temporal information in clinical guidelines. These models could provide a starting point for representing complex temporal expressions in GLIF.

■ *Representation of uncertainty*: There are several types of uncertainty that affect patient data and guidelines. First, the individual who collected the patient data may be uncertain about the truth of a given medical statement. For example, a clinician might believe, on the basis of a patient history, that the patient may have had a myocardial infarction in the past, but in the absence of concrete evidence, the clinician cannot be certain. Second, data required by a guideline may be impossible to obtain. For example, a clinician may document that the family history is unknown because a patient was adopted and does not know who his family is. Third, a data value may be uncertain because it is missing or has not yet been collected. A guideline may recognize the difference between data that have not yet been collected (prompting a particular action to obtain the data) and data that cannot be collected, at which

point the guideline must continue without it. At present, GLIF cannot recognize and handle these kinds of uncertainty about patient data.

These four limitations in GLIF should be addressed by future work. Other tasks to undertake are the creation of an Internet-accessible guideline server; of guideline-authoring tools that permit graphic display of guideline information, guideline construction, and automatic encoding; and of guideline-application programs that are based on the shareable representation and interact directly with clinical data stored in local patient databases.

Creation of GLIF was possible because the InterMed project permitted workers at different locations to recognize common goals in computer-based guideline research and to collaborate despite differences in local computing infrastructure in the health care delivery environment and differences in institutional demands. The GLIF specification is the product of a team effort, in which striving for consensus was an important part of the task. The consensus-building process demonstrated that a common understanding of functionality and terms used for guideline structural elements was necessary for the different groups to reach agreement on a guideline-representation specification. We evaluated our model and the GLIF encoding process in a preliminary study. We found the syntax to be adequately expressive but recognized that different individuals produced different encodings as a result of different modeling choices, different representations of criteria given the use of narrative text in the current version of GLIF, and selection of different terminology for data elements in the absence of standards for clinical vocabulary and data models. We are enhancing the GLIF representation as described here; GLIF is therefore a model in evolution. We hope that this work on GLIF will encourage the development of standards for representing clinical guidelines for use in computer-based systems.

References ■

1. Shortliffe EH, Barnett GO, Cimino JJ, Greenes RA, Huff SM, Patel VL. Collaborative medical informatics research using the Internet and the World Wide Web. Proc 20th Annu Symp Comput Appl Med Care. 1996:125–9.
2. Shortliffe EH, Patel VL, Cimino JJ, Barnett GO, Greenes RA. A study of collaboration among medical informatics research laboratories. Artif Intell Med. 1998;12:97–123.
3. Field MJ, Lohr KN (eds). Clinical Practice Guidelines: Directions for a New Program. Institute of Medicine. Wash-

ington, DC: National Academy Press, 1990.

4. Field MJ, Lohr KN (eds). Guidelines for Clinical Practice: From Development to Use. Washington, DC: National Academy Press, 1992.

5. Audet A, Greenfield S, Field M. Medical practice guidelines: current activities and future directions. Ann Intern Med. 1990;113:709–14.

6. Johnston ME, Langton KB, Haynes RB, Mathieu A. Effects of computer-based clinical decision support systems on clinician performance and patient outcome. Ann Intern Med. 1994;120:135–42.

7. Shiffman RN, Greenes RA. Improving clinical guidelines with logic and decision table techniques: application to hepatitis immunization recommendations. Med Decis Making. 1994;14:245–54.

8. Tierney WM, Overhage JM, Takesue BY, et al. Computerizing guidelines to improve care and patient outcomes: the example of heart failure. J Am Med Inform Assoc. 1995;2:316–22.

9. McDonald CJ, Overhage JM. Guidelines you can follow and trust: an ideal and an example. JAMA. 1994;271:872–3.

10. Musen MA, Rohn JA, Fagan LM, Shortliffe EH. Knowledge engineering for a clinical trial advice system: uncovering errors in protocol specification. Bull Cancer. 1987;74:291–6.

11. Liem EB, Obeid JS, Shareck EP, Sato L, Greenes RA. Representation of clinical practice guidelines through an interactive World-Wide-Web interface. Proc 19th Annu Symp Comput Appl Med Care. 1995:223–7.

12. Cimino JJ, Socratous SA, Clayton PD. Automated guidelines implemented via the World Wide Web interface. Proc 19th Annu Symp Comput Appl Med Care. 1995:941.

13. Jenders RA, Hripcsak G, Sideli RV, et al. Medical decision support: experience with the Arden Syntax at the Columbia–Presbyterian Medical Center. Proc 19th Annu Symp Comput Appl Med Care. 1995:169–73.

14. Barnett GO, Winickoff R, Dorsey JL, Morgan MM, Lurie RS. Quality assurance through automated monitoring and concurrent feedback using a computer-based medical information system. Med Care. 1978;16:962–70.

15. Tu SW, Kemper CA, Lane NM, Carlson RW, Musan MA. A methodology for determining patient's eligibility for clinical trials. Methods Inf Med. 1993;32:317–25.

16. Lam SH. Implementation and evaluation of practice guidelines. Proc 17th Annu Symp Comput Appl Med Care. 1993:253–7.

17. Shortliffe EH. Update on ONCOCIN: a chemotherapy advisor for clinical oncology. Med Inform. 1986;11:19–21.

18. Carlson RW, Tu SW, Lane NM, et al. Computer-based screening of patients with HIV/AIDS for clinical-trial eligibility. Online J Curr Clin Trials [online journal]. March 28, 1995; doc 179.

19. Safran C, Rind DM. The development of knowledge-based medical records for clinicians caring for patients with HIV infection. In: Lun KC, Degoulet P, Piemme TE, Rienhoff O (eds). Medinfo. 1992:495–500.

20. Henderson SE. Computerized clinical protocols in an intensive care unit. Proc 13th Annu Symp Comput Appl Med Care. 1989:284–8.

21. Musen MA. Dimensions of knowledge sharing and reuse. Comput Biomed Res. 1992;25:435–67.

22. Pryor TA, Hripcsak G. Sharing MLM's: an experiment between Columbia–Presbyterian and LDS Hospital. Proc 17th Annu Symp Comput Appl Med Care. 1993:399–403.

23. Hripcsak G. Writing Arden Syntax medical logic modules. Comput Biol Med. 1994;24:331–63.

24. Clayton PD, Pryor TA, Wigertz OB, Hripcsak G. Issues and structures for sharing knowledge among decision-making systems: The 1989 Arden Homestead Retreat. Proc 13th Annu Symp Comput Appl Med Care. 1989:116–21.

25. American Society for Testing and Materials. E1460 Standard Specification for Defining and Sharing Modular Health Knowledge Bases (Arden Syntax for Medical Logic Modules). 14.01 ed. Philadelphia, Pa.: ASTM, 1992:539–87.

26. Hripcsak G, Ludemann P, Pryor TA, Wigertz OB, Clayton PD. Rationale for the Arden Syntax. Comput Biomed Res. 1994;27:291–324.

27. Hripcsak G, Cimino JJ, Johnson SB, Clayton PD. The Columbia–Presbyterian Medical Center decision-support system as a model for implementing the Arden Syntax. Proc 15th Annu Symp Comput Appl Med Care. 1991:248–52.

28. Hripcsak G, Clayton PD, Jenders RA, Cimino JJ, Johnson SB. Design of a clinical event monitor. Comput Biomed Res. 1996;29:194–221.

29. Barrows RC Jr, Allen BA, Smith KC, Arni VV, Sherman E. A decision-supported outpatient practice system. Proc 20th Annu Symp Comput Appl Med Care. 1996:792–6.

30. Sherman EH, Hripcsak G, Starren J, Jenders RA, Clayton P. Using intermediate states to improve the ability of the Arden Syntax to implement care plans and reuse knowledge. J Am Med Inform Assoc. 1995;1:238–42.

31. Stoufflet PE, Ohno-Machado L, Deibel SRA, Lee D, Greenes RA. GEODE-CM: a state-transition framework for clinical management. Proc 20th Annu Symp Comput Appl Med Care. 1996:924.

32. Barnes M, Barnett GO. An architecture for a distributed guideline server. Proc 19th Annu Symp Comput Appl Med Care. 1995:233–7.

33. Barnett GO. The application of computer-based medical record systems in ambulatory practice. N Engl J Med. 1984;310:1643–50.

34. Musen MA, Tu SW, Das AK, Shahar Y. EON: a component-based approach to automation of protocol-directed therapy. J Am Med Inform Assoc. 1996;2:367–88.

35. Das AK, Shahar Y, Tu SW, Musen MA. A temporal-abstraction mediator for protocol-based decision-support systems. Proc 19th Annu Symp Comput Appl Med Care. 1995:320–4.

36. The GLIF guideline page. GLIF Web site. Available at: http://dsg.harvard.edu/public/GLIF/pubs/JAMIA98_AppendixA.html.

37. The GLIF guideline page. GLIF Web site. Available at: http://dsg.harvard.edu/public/GLIF/pubs/JAMIA98_AppendixB.html.

38. Centers for Disease Control and Prevention. Prevention and control of influenza: recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR Morb Mortal Wkly Rep. April 21, 1995;44(RR-3):1–22.

39. National Cholesterol Education Program. Summary of the second report of the NCEP expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (Adult Treatment Panel II). JAMA. 1993;269:3015–23.

40. Borten M. Breast mass. In: Friedman EA, Borten M, Chapin M (eds). Gynecological Decision Making. 2nd ed. Philadelphia, PA.: BC Decker, 1988:72–3.

41. Eastern Cooperative Oncology Group. Protocol EST 2190 [computer program], PDQ version. 1994.

42. Patel VL, Allen VG, Arocha JF, Shortliffe EH. Representing clinical guidelines in GLIF: individual and collaborative expertise. J Am Med Inform Assoc, in press.

43. Musen MA, Gennari JH, Eriksson H, Tu SW, Puerta AR. PROTEGE II: Computer support for development of intelligent systems from libraries of components. In: Greenes RA, Peterson HE, Protti DJ (eds). Medinfo. 1995:766–70.

44. Uckun S. Instantiating and monitoring treatment protocols. Proc 18th Annu Symp Comput Appl Med Care. 1994:689–93.

45. Shahar Y, Miksch S, Johnson P. An intention-based language for representing clinical guidelines. Proc 20th Annu Symp Comput Appl Med Care. 1996:592–6.

46. Herbert SI. Informatics for care protocols and guidelines: towards a European knowledge model. In: Gordon C, Christensen JP (eds). Health Telematics for Clinical Guidelines and Protocols. Amsterdam, The Netherlands: IOS Press, 1995:27–42.

47. Kohane IS. Temporal reasoning in medical expert systems. In: Salamon R, Blum B, Jorgensen M (eds). Medinfo. 1984:170–4.

48. Kahn MG. Modeling time in medical decision-support pro-

grams. Med Decis Making. 1994;11:249–64.

49. Das AK, Musen MA. A temporal query system for protocol-directed decision support. Methods Inf Med. 1994;33:358–70.

50. Dolin RH. Modeling the temporal complexities of symptoms. J Am Med Inform Assoc. 1995;2:323–31.

51. Shahar Y, Musen MA. Knowledge-based temporal abstraction in clinical domains. Artif Intell Med. 1996;8:267–98.

52. Miksch S, Shahar Y, Johnson P. ASBRU: a task-specific, intention-based, and time-oriented language for representing skeletal plans. In: Motta E, Harmelen FV, Pierret-Golbreich C, Filby I, Wijngaards N (eds). Proceedings of the 7th Workshop on Knowledge Engineering: Methods and Languages. Milton Keynes, England: The Open University, 1997:9.1–9.20.