Handling Expressiveness and Comprehensibility Requirements in GLIF3

Mor Peleg, Ph.D.[a], Aziz A. Boxwala, M.B.B.S., Ph.D.[b], Samson Tu, MS.[a],
Robert A. Greenes, M.D.,Ph.D.[b], Edward H. Shortliffe, M.D., Ph.D.[c], Vimla L. Patel, Ph.D., D.Sc.[c]

[a] *Stanford Medical Informatics, Stanford University School of Medicine, Stanford, California*
[b] *Decision Systems Group, Brigham & Women's Hospital, Harvard Medical School, Boston, MA*
[c] *Department of Medical Informatics, College of Physicians and Surgeons, Columbia University, New York, NY*

Abstract

Clinical guidelines are aimed at standardizing patient care and improving its quality and cost effectiveness. Guidelines represented in a computer-interpretable (CI) format can be used to provide automatic decision support applied to individual patients during the clinical encounter. The process of creating computer-interpretable guidelines (CIG) removes ambiguities contained in paper-based guidelines, thus making the guideline more comprehensible. For these reasons, CIGs may have a larger impact on clinician behavior than paper-based guidelines. Since much effort goes into creating guidelines in a CI format, it is desirable that different institutions and software systems share them. In a guideline representation workshop hosted by the InterMed Collaboratory in March 2000, the need for a standard representation format for sharable CIGs was recognized. As a first step towards achieving this goal, we proposed a set of functional requirements for sharable CIGs. The requirements encompass the entire life cycle of a CIG: development, implementation, use and maintenance. In this paper we discuss requirements that are important during the development stage of a CIG. We have abstracted the requirements into two groups: expressiveness – the ability to express the knowledge content of different types of guidelines – and comprehensibility – the ability to manage complexity, facilitate coherence, and visualize a guideline model to aid in human comprehension. The Guideline Interchange Format version 3 (GLIF3) is a language for structured representation of CIGs. It is under development to facilitate sharing CIGs among different institutions and systems. We illustrate how GLIF3 meets the specified development requirements.

Keywords:

Clinical practice guideline, knowledge representation

Introduction

Clinical guidelines define recommended strategies for managing health care in specific clinical circumstances. The recommendations are based on evidence from scientific studies, and are created by a consensus process among medical experts. They are intended to minimize practice variation, to improve clinical outcomes, and to reduce inappropriate use of resources. Unfortunately, guidelines have not always been successful at affecting practitioner behavior [1]. Among factors that have been proposed to explain this situation is the lack of usability and comprehensibility of paper-based guidelines [2]. Guideline implementation strategies that provide patient-specific advice automatically at the point-of-care are more likely to increase usability than those in which guidelines are made available in non-patient-care contexts (such as publication in monographs or journals) [3]. CIGs can be used to provide patient-specific recommendations, to support retrospective analysis of quality of care [4], or for educational purposes. Moreover, the process of creating a CIG aids in removing ambiguities that may arise in paper-based guidelines, thus making the guideline more comprehensible.

For guidelines to be computer-interpretable, they must be encoded in a particular model. The model defines the knowledge concepts contained in guidelines, by explicitly specifying their structure and relationships. A number of models for CIGs have been developed. Arden Syntax [5] is a model that is suitable for modeling individual decision rules that generate alerts and reminders. Some examples of methodologies that have been developed for modeling multi-step guidelines are GLIF [6], EON [7], PRODIGY [8], PRO*forma* [9], PATient Workflow Management System (PATMAN) [10], Asbru [11], and PRESTIGE [12]. They differ in many respects that include: (1) the types of guidelines they can represent, (2) the depth to which attributes are structured (i.e., whether attributes are simple or complex data types), (3) the formality of the underlying model, (4) the underlying mathematical model (e.g., state transition graph, rule-based system), and (5) the focus of interest. In

GLIF, the focus of interest is the ability to share guideline encodings. In EON it is the guideline task model, temporal queries, use of domain ontologies, and explanation. Safety is the center of focus in PRO*forma*. In Asbru, it is that of expressing guideline intentions. PRODIGY aims at facilitating knowledge engineering by domain experts by producing a simple, understandable model sufficiently expressive to represent chronic disease management guidelines. Work done by the Medical Informatics group at the University of Pavia emphasizes workflow integration into healthcare organizations.

The CIG community, in a workshop that was hosted by the InterMed Collaboratory [13] in March 2000 [14], recognized the need for a sharable standardized representation of clinical guidelines. Sharing encoded guidelines leverages the large effort that is put into encoding. A common format for sharing guidelines will focus the effort of achieving a high quality guideline model and authoring tools and execution applications. As a first step towards achieving such a common format, participants of the guideline-modeling workshop identified a set of functional requirements for sharable CIGs [15]. The requirements encompass the entire life cycle of a CIG, from development, through implementation, to use and maintenance. In this paper we focus on the requirements that are important during the development stage of a CIG: expressiveness – the ability to express the knowledge content of different types of guidelines – and comprehensibility – the ability to manage complexity, facilitate coherence, and to visualize a guideline model so as to aid in human comprehension.

Any robust guideline model should be able to satisfy the set of requirements. GLIF3 is specifically designed for the purpose of sharing guidelines among different institutions and software systems. Its domain ontology allows mapping of guideline terms to standard vocabularies and standard data models of medical concepts. This design is aimed at structuring the way in which guideline encodings are mapped to institutional electronic medical record (EMR) codes. Here we use the GLIF guideline model to illustrate how some of the requirements of a sharable CIG representation format could be handled.

Materials and Methods

Our work is based on the set of functional requirements for a sharable CIG representation format that was raised by participants of the guideline-modeling workshop hosted by InterMed in March 2000 [15]. We begin by abstracting the set of requirements to create an overarching organization that is intended to be easy to follow. We then illustrate how these requirements might be met, using examples from the GLIF3 guideline representation format (hereafter referred to simply as GLIF format).

Results

Many requirements of CIG representation arise during the development life cycle stage. We have abstracted them into two main groups: expressiveness and comprehensibility.

Expressiveness

Expressiveness is the ability to encode the knowledge content of different types of guidelines. Guidelines may be classified according to the clinical domain, the stage of the medical problem and its management (e.g., screening, diagnosis, disease management), multiple or single encounters, setting (e.g., inpatient or outpatient clinic), time frame (emergency, acute, or chronic), and guideline computability (i.e., algorithmic, guiding, or intermediate) [16]. We have used GLIF to encode a variety of guidelines including guidelines for Influenza vaccination [17], management of chronic cough [18], management of stable angina [19], thyroid screening [20], lower back pain [21], heart failure [22], and depression [23], as shown in Table 1.

Table 1 – Classification of GLIF3-encoded Guidelines

Disease/ Condition	Stage of Problem	En- count- ers	setting	Time Frame	Computability
Flu	prevention	1	out	a	algorithmic
Stable Angina	management	n	out	a/c	intermediate
Chronic cough	Diagnosis + management	n	out	a	intermediate
Lower back pain	Diagnosis + management	n	out	a	intermediate
Heart failure	management	n	out	a/c	algorithmic
Depression	management	n	out	a	algorithmic
Thyroid screening	screening	1	out	a	algorithmic

n = many; out = outpatient clinic; time-frame: c̲hronic, a̲cute;

The knowledge content of the different kinds of guidelines can be classified in several ways. One classification takes the point of view of the guideline's structural parts (i.e., recommendations, concept definitions, and algorithms), and the other is a functional classification of the kind of decision-support tasks that a guideline involves.

Definitions, recommendations, and algorithms

To illustrate, the cough guideline defines chronic cough as cough persisting for over 3 weeks. This definition can be represented in GLIF as a *patient-state step*, for which the logical criterion is specified in GLIF's expression syntax. The logical criterion refers to patient data items that are defined by GLIF's domain ontology. The criterion is expressed as: "latest_cough_end_time >= now and lat-

est_cough_start_time < (now – 3 weeks)", where latest_cough_end_time and latest_cough_start_time are time-stamp variables defined elsewhere in the guideline by a Get_Data action specification that links these variables to definitions of patient data in the domain ontology. An example of a data item is shown in Figure 1.

Figure 2 shows how the following recommendation from the cough guideline is represented in GLIF: "four-view sinus radiographs should be ordered prior to beginning therapy for postnasal drip syndrome (PNDS) when cough is productive (Grade II-2)". The strength of evidence (Grade II-2 in this example) is an attribute of the guideline step class in the GLIF model.

Figure 1 – A medical action specification of a chest X-ray showing a medical task data item
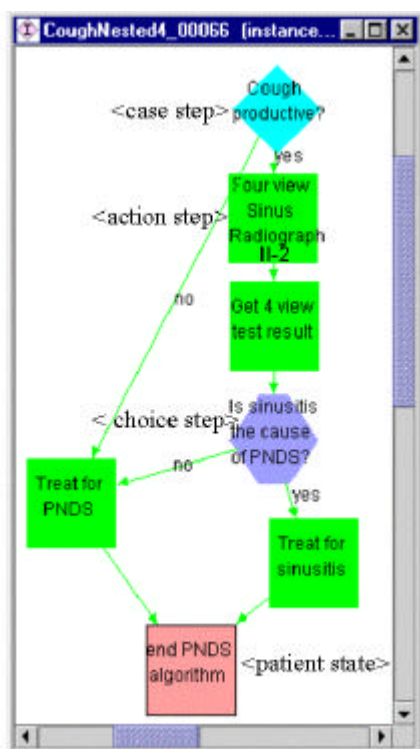

Figure 2 – An algorithm representing evaluation and treatment of PNDS as the cause of cough

Representing decision-support guideline tasks

Another way to classify the knowledge content of the different kinds of guidelines is a functional classification of the kind of decision-support tasks that a guideline involves [7]. The tasks are: making decisions, setting goals, specifying work to be performed, data interpretation, and generating alerts or reminders. In this section, we illustrate how the GLIF model can specify these tasks. GLIF differentiates between two types of decision steps: case and choice. *Case steps* represent decisions that can be automated by directly evaluating logical criteria based on data items from the EMR. *Choice steps* represent choices that should be made by the user since they are either safety-critical or require knowledge that is not specified by the guideline. To aid in the choice, rule-in and rule-out criteria may be specified for each alternative. Examples of case and choice steps are shown in Figure 3. The decision criteria follow a three-valued semantics of true, false, or unknown, which is more appropriate for medical data then Boolean semantics.
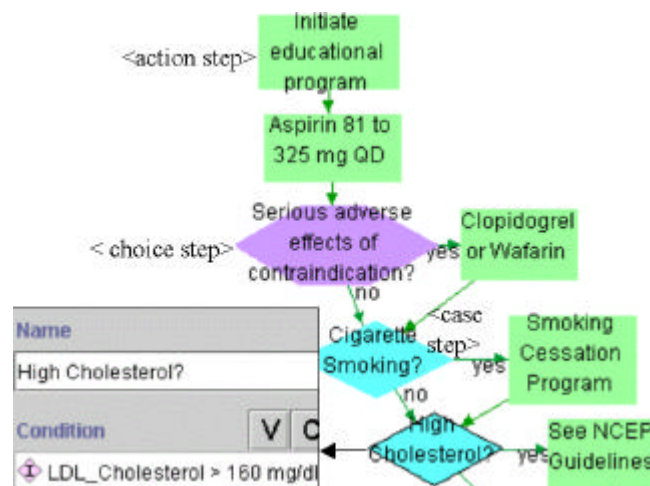

Figure 3 – Two levels of representation of a guideline

Goals can be specified in GLIF as text strings that define the intention of a guideline or sub-guideline. Goals are represented in a much more systematic way by Asbru [11].

Work to be performed is specified in a GLIF guideline's algorithm by *action steps*. Each action step includes an action specification(s) that is either programming-oriented (e.g., data retrieval from the EMR, calling of a sub-guideline, sending a message to the user) or medically oriented. Medically oriented action specifications refer to medical concept codes from standard vocabularies and to standard medical data frameworks, which in the default case is HL-7's Unified Service Action Model (USAM) [24], but can alternatively be defined by the guideline authors. The medical data frameworks classify medical data into a top-level hierarchy, and specify their attributes and relationships. For example, the USAM model specifies 4 top-level classes that are used by GLIF: the act class, which is used to represent patient data, and 3 of its subclasses – observation, medication and (medical) procedure. An ex-

ample of a medical action specification that models a chest X-ray order is shown in Figure 1.

Data interpretation is handled by GLIF in the following way. Definitions of concepts are based on standard vocabularies. Concept relationships (e.g., Penicillin is an antibiotic) can be represented as well. Therefore, a criterion that asks if the patient is on antibiotics is evaluated as true if the patient is taking penicillin. Sending alerts and reminders is modeled by the message-type action specification of GLIF.

Comprehensibility

Comprehension involves the construction of a mental representation of some aspects of the external world (i.e., an external representation, such as a clinical guideline). In order for guidelines to standardize care, their comprehension should result in a mental model that matches the guidelines' recommended procedures [2]. The comprehensibility of a guideline model is important for guideline authors who want to follow easily the guideline specification that they are generating, as well as for users who want to follow the guideline logic. Comprehensibility entails visualization and readability, complexity management, and coherence facilitation.

Visualization and readability

GLIF has a layered representation that is designed to separate knowledge abstractions from the formal view and from implementation-dependent details. The topmost layer, the conceptual flowchart level, is useful for human readers who want to have a conceptual view of the guideline in which the guideline flow is easy to follow. The formal computable level is needed to support automatic guideline interpretation and verification of logical correctness. The implementation-dependent layer has not yet been developed. Figure 3 shows an example of the education and risk factor modification algorithm of a guideline for stable angina that was created by the Protégé authoring tool [25]. On the right side of the figure, the conceptual flowchart is shown. Each of the guideline steps can be expanded to show its computable specification, as illustrated for the "High cholesterol?" case step, on the left side of the figure.

Complexity management

CIGs can be very complex in terms of the magnitude of knowledge, the amount of embedded and linked propositions within a single sentence and the temporal flow of the logic [2]. In addition, guideline models can contain a large number of constructs. This may result in very complex guideline encodings. GLIF provides three constructs to manage complexity: nesting, macros and views.

Nesting allows a guideline author to represent a top-level view of the guideline's algorithm, and recursively to show the details of action and decision steps using sub-guidelines. For example, the entire sub-guideline algorithm shown in Figure 3 can be collapsed into a single action

(Education and risk factor modification) in a higher-level guideline. Similarly, the details of the "Smoking Cessation Program" and "See NCEP Guideline" action steps are given by sub-guideline.

Macros are special kinds of steps that represent high-level knowledge components that can be mapped to underlying flowcharts of primitive GLIF steps (i.e., action, decision, patient state, branch and synchronization steps). For example, an Arden macro in GLIF contains evoke, condition and action slots, that specify the information that is necessary to instantiate the sequence of a case step followed by an action step. Macro steps let the guideline author represent the guideline in terms of high-level constructs instead of using primitive GLIF steps. This helps to control the complexity of the guideline and to aid in its understanding.

Views are filters that collapse segments of the guideline into a default view that is customized to a given kind of user (e.g., general practitioner, nurse) or setting (e.g., inpatient vs. outpatient clinic). The user can always see more or less detail than what is shown to him by default.

Coherence facilitation

Coherence reduces the probability of errors in understanding by providing the reader with ways to make associations among different parts of the guideline [2]. Coherence makes a guideline easier to understand by constraining the amount and types of inferences that are made during interpretation. As was mentioned earlier, GLIF supports expression of different structural parts of guidelines, including definitions, recommendations, and algorithms. It provides a highly structured model that represents a guideline as a network of guideline steps containing references to patient data, decision criteria, and action specifications. GLIF in turn defines the inter-relations among these different objects in an explicit and coherent way, which eases inference. In addition, GLIF guidelines and guideline steps can reference supporting material that may facilitate understanding of the reasons and evidence behind the recommendations. The supporting material can be inline text, URLs, or keywords. The purpose of the supporting material is further specified as *evidence*, *comment*, or *indexing*.

Discussion

We have presented requirements of a sharable CIG model that are important for the development stage of a guideline. They were part of a set of requirements that were proposed by participants of a workshop entitled "Towards representations for sharable guidelines", that gathered 80 stakeholders from the CIG community [15].

The requirements encompass the entire life cycle of CIG: development, implementation, use and maintenance. In this paper we have focused on requirements that are important during the development life cycle stage of a CIG model, abstracting requirements into two groups: those that define the expressive power of the guideline model, and those that define the ability of the guideline encoding to be compre-

hended easily. Both categories are important for sharing CIGs. An expressive model is necessary so that the guideline model can represent all knowledge components and decision-support tasks of different guidelines. Comprehensibility is desirable for guideline authors and users who want to follow the guideline's logic.

The other life cycle stages of a CIG, implementation, use and maintenance entail other requirements, such as the ability to adapt the guideline to site-specific constraints and to integrate it into local environments, to provide version control, and to support different usage modes and user interfaces. GLIF is still a developing methodology. We do not have experience yet in the implementation and use of GLIF encoded guidelines in medical institutions.

We believe that any robust guideline representation format should support the functional requirements presented in this paper. We have illustrated how these requirements are met by GLIF.

Acknowledgments

References

[1] Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PA, and Rubin HR. Why don't physicians follow clinical practice guidelines? A framework for improvement. JAMA 1999; 282(15):1458-65.

[2] Patel VL, Arocha JF, Diermeier M. Cognitive Analysis Methods to Support Implementation of Clinical Practice Guidelines. Submitted to the Journal of biomedical Informatics 2000.

[3] Shea S, DuMouchel W, Bahamonde L. A Meta-analysis of 16 Randomized Controlled Trials to Evaluate Computer-based Clinical Reminder Systems for Preventative Care in the Ambulatory Setting. JAMIA 1996;3(6):399-409.

[4] Advani A, Lo K, Shahar Y. Intention-Based Critiquing of Guideline-Oriented Medical Care. In: Proc. AMIA Annual Symposium; 1998; p. 483-487.

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

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

[7] Tu SW, Musen MA. From Guideline Modeling to Guideline Execution: Defining Guideline-BasedDecision-Support Services. In: AMIA Annual Symposium; 2000; p. 863-867.

[8] Johnson PD, Tu S, Booth N, Sugden B, Purves IN. Using Scenarios in Chronic Disease Management Guidelines for Primary Care. In: Proc AMIA Symp 2000; p. 389-393.

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

[10] Dazzi L, Fassino C, Saracco R, Quaglini S, Stefanelli M. A Patient Workflow Management System Built on Guidelines. In: Proc AMIA Annu Fall Symp; 1997; p. 146-150.

[11] Shahar Y, Miksch S, Johnson P. The Asgaard Project: A Task-Specific Framework for the Application and Critiquing of Time-Oriented Clinical Guidelines. Artificial Intelligence in Medicine 1998;14:29-51.

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

[13] Shortliffe EH, Patel VL, Cimino JJ, Barnett GO, Greenes RA. A study of collaboration among medical informatics research laboratories. Artificial Intelligence in Medicine 1998;12(2):97-123.

[14] InterMed. Guideline Workshop. In: www.GLIF.org/workshop, Nov. 2000.

[15] InterMed. Guideline Discussion Groups. In: www.GLIF.org/discussion/ , Nov. 2000.

[16] Bernstam E, Ash N, Peleg M, Tu S, Boxwala AA, Mork P, et al. Guideline classification to assist modeling, authoring, implementation and retrieval. In: Proc. AMIA Annual Symp 2000; p. 66-70.

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

[18] American College of Chest Physicians. Managing Cough as a Defense Mechanism and as a Symptom, Chest 1998;114(2):133S-181S.

[19] ACC/AHA/ACP-ASIM. Guidelines for the Management of Patients with chronic Stable Angina. J Am Col Cardiol 1999;33:2092-2197.

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

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

[22] 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.

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

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

[25] 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; p. 7-4-1 to 7-4-36.

Address for correspondence

Mor Peleg, Ph.D., Stanford Medical Informatics, Medical School
Office Building x-208, 251 Campus Drive
Stanford, CA 94305-5479
Phone: (650) 723-7711 Fax: (650) 725-7944
peleg@smi.stanford.edu