Model Formulation ■

GEM: A Proposal for a More Comprehensive Guideline Document Model Using XML

RICHARD N. SHIFFMAN, MD, MCIS, BRYANT T. KARRAS, MD, ABHA AGRAWAL, MD, ROLAND CHEN, MD, LUIS MARENCO, MD, SUJAI NATH, MD

A b s t r a c t   **Objective:** To develop a guideline document model that includes a sufficiently broad set of concepts to be useful throughout the guideline life cycle.

Design: Current guideline document models are limited in that they reflect the specific orientation of the stakeholder who created them; thus, developers and disseminators often provide few constructs for conceptualizing recommendations, while implementers de-emphasize concepts related to establishing guideline validity. The authors developed the Guideline Elements Model (GEM) using XML to better represent the heterogeneous knowledge contained in practice guidelines. Core constructs were derived from the Institute of Medicine's Guideline Appraisal Instrument, the National Guideline Clearinghouse, and the augmented decision table guideline representation. These were supplemented by additional concepts from a literature review.

Results: The GEM hierarchy includes more than 100 elements. Major concepts relate to a guideline's identity, developer, purpose, intended audience, method of development, target population, knowledge components, testing, and review plan. Knowledge components in guideline documents include recommendations (which in turn comprise conditionals and imperatives), definitions, and algorithms.

Conclusion: GEM is more comprehensive than existing models and is expressively adequate to represent the heterogeneous information contained in guidelines. Use of XML contributes to a flexible, comprehensible, shareable, and reusable knowledge representation that is both readable by human beings and processable by computers.

■ **J Am Med Inform Assoc.** 2000;7:488–498.

Over the last decade, clinical practice guidelines have become increasingly important repositories of knowledge about ideal practice. Built on a careful analysis and understanding of research evidence combined with expert consensus, a flood of guidelines are being created in an effort to diminish inappropriate practice,

to improve health outcomes, and to control costs of care.[1]

Sponsoring organizations most often publish practice guidelines as paper-based, prose documents, which sometimes include algorithmic flowcharts. These publications are typically unavailable during clinical consultations. Zielstorff[2] noted that while electronic dissemination can "solve the problem of accessibility to the guideline itself . . . access to knowledge embedded in the guideline can still be problematic." She called for defining guideline knowledge in "standard interchange formats that permit installation in a wide variety of technical infrastructures." Guideline querying, electronic distribution, and decision support systems that implement the guideline's recommendations can all be facilitated when the knowledge contained in the guideline is represented in digital form.

The proposed Guideline Elements Model (GEM) is intended to serve as a document model, i.e., an ideal-

ized abstraction of a guideline document that masks certain details to bring forth others (after Degoulet and Fieschi[3]). By describing concepts pertinent to guideline representation, attributes of those concepts, and relationships among the concepts, GEM promotes translation of natural language guideline documents into a format that can be processed by computers. The framework is intended to be useful to developers, disseminators, implementers, maintainers, and end users of guidelines.

XML (the Extensible Markup Language) offers a powerful technology for representing electronic documents. It allows both computers and human beings to access the information in a document and extract it for reuse or modification.[4] In XML, guideline documents are conceptualized as a hierarchy of elements —basic units of information that store data and define structure by virtue of their position in a tree.[5] Element tags demarcate text and provide labels that characterize the semantic content of the element. Tagging a document does not require programming skill, yet it can create a computer-processable representation of the knowledge contained in a guideline.

An XML document type definition (DTD) models the names of allowable elements and attributes in the document, the content of each element, and the structure of the document (i.e., the order and cardinality of each element). A validating parser can ensure that any tagged document conforms to the DTD. A parser (such as Internet Explorer 5) can read an XML document file and populate a tree in memory, thereby exposing all the elements and attributes to manipulation by an application.[6] That application might, for example, select certain components for presentation to the user (e.g., costs) or interact with facts in a clinical database to provide guideline-based decision support.

Several stakeholders combine their efforts to develop, disseminate, implement, and maintain the knowledge contained in guidelines throughout the guideline life cycle. Not surprisingly, therefore, the models that have been created to represent this knowledge vary, depending on the orientation of the stakeholder.

Models devised by those who build and evaluate practice guidelines tend to emphasize descriptions of the methods applied in guideline development, issues of guideline testing and maintenance, and details about the sponsors, objectives, and intended audiences; however, they often provide few details for conceptualizing recommendations. For example, the National Guideline Clearinghouse (NGC) model contains most of the concepts found in the health services literature, yet it provides only a single slot for "Major Recommendations."[7]

On the other hand, informatics researchers tend to model guideline recommendation components in great detail but often fail to address the concerns of the health services community, such as evidence strength and quality, potential biases of sponsors and developers, and validity checks. For example, although our own augmented decision table model for representing guidelines knowledge recognized the importance of "collateral information," it fell short in defining a comprehensive set of considerations.[8]

One of the most ambitious efforts to build an electronic guideline representation is the Guidelines Interchange Format (GLIF) of the Intermed Collaboratory. GLIF includes constructs for the name of a guideline and its authors, purpose, and eligibility criteria, but it specifies only a vaguely defined component called *didactics* to "provide background or supporting information."[9] The emphasis in GLIF is on detailed specification of guideline recommendations.

Since different stakeholders who work with a single guideline have different information needs and require different computer applications to support those needs, an ideal electronic model of a guideline should be capable of representing all pertinent aspects of the document. In this work we attempt to bridge the deficiencies in existing designs by proposing a guideline model that includes a more comprehensive set of components. It would be naïve to believe that any model could completely meet the needs of all current and future users. We therefore set a goal of creating a "more comprehensive" model that would represent the most important guideline components and be sufficiently flexible to allow future extensions if needed.

In this work, we describe our approach to developing GEM, followed by a presentation of the model itself. We next illustrate how a variety of guideline models from both the health services and informatics literature map to GEM, and we briefly describe our experience modeling a variety of published guidelines with GEM. We conclude with a discussion of attributes of an ideal guideline document model.

Approach to Model Development

We relied on three primary sources to form the core set of guideline elements. First, for concepts related to guideline development and evaluation, the Institute of Medicine's Provisional Instrument for Assessing Clinical Guidelines provides a detailed set of evaluation criteria for practice guidelines.[10] The purpose of this instrument was to define the attributes of ideal guidelines, to encourage systematic guideline development, and to provide a standardized approach and

structure for the assessment of guideline documents. It consists of 46 questions related to seven guideline attributes: validity, clarity, multidisciplinary process, clinical flexibility, reliability and reproducibility, clinical adaptability, and scheduled review. Pertinent constructs for GEM were extracted from these questions.

Second, for concepts related to guideline dissemination, the NGC provides an online resource for evidence-based guidelines.[7] Sponsored by the U.S. Agency for Healthcare Research and Quality (formerly the Agency for Health Care Policy and Research), the American Medical Association, and the American Association of Health Plans, it provides a schema for classification of the components of guideline documents. This model includes a set of "key attributes" for summarizing each guideline to facilitate search and retrieval of information from the NGC Web site and comparison between guidelines. The NGC also offers several sets of controlled vocabulary constructs to describe concepts in the model.

Third, for concepts related to implementation of guideline recommendations, we apply a set of constructs derived originally from the augmented decision table model.[8] This approach has been used to represent knowledge derived from:

- Guidelines from a variety of sources, including the American Academy of Pediatrics, the American College of Physicians, the Agency for Health Care Policy and Research, the U.S. Preventive Services Task Force, and a managed care organization

- Guidelines that address multiple topics, including **diagnosis** of appendicitis,[11] **risks** of hypercholesterolemia,[12] **indications** for coronary artery bypass graft,[13] and **management** of asthma[14]

- Guidelines that were created using both evidence-based and consensus methodologies

The original model has been substantially enhanced and extended. GEM represents a key infrastructure component for a proposed object-oriented framework for development of computer-based guideline implementations.[15]

Constructs extracted from these sources were supplemented by additional concepts derived from published models. We searched the MEDLINE database (1990 to 1999) using the OVID search engine. The search strategy looked for "practice guidelines" as a subject heading and "guideline" as a text word. The results were combined with "knowledge representation" or "model" or "evaluation."

In another search we looked for papers that had been published in the *Journal of the American Medical Infor-* *matics Association, Methods of Information in Medicine*, or the proceedings of AMIA annual meetings that addressed guidelines. The bibliographies of selected articles were also searched for relevant publications, as were the authors' reference files. Articles were selected that modeled and categorized the content of clinical practice guideline documents. Papers that failed to specify detailed modeling constructs and designs that described models of guideline implementations without describing document models were excluded.

Markup using GEM tags was the natural outgrowth of a system that has been in use in our laboratory since 1995. Members of the Guidelines Review Group at Yale have collaborated with the Committee on Quality Improvement of the American Academy of Pediatrics to review and critique proposed evidence-based guidelines prior to publication and at the time of scheduled review. "Logical analysis" is our name for the cognitive task by which recommendation components are extracted from the natural language text of clinical practice guidelines and specified in a computable format. The first step in logical analysis has been to mark up paper-based documents using colored highlighters to identify and categorize guideline components, such as recommendations, evidence, costs, patient preferences, and clinical options. The GEM hierarchy permits much more detailed categorization than is possible with the physical highlighting system.

Proposed Model

As shown in Figure 1, GEM can be depicted as a directed graph with Guideline Document as the root. The major concepts in the first tier of the GEM hierarchy below the root level are identity, developer, purpose, intended audience, method of development, target population, knowledge components, testing, and review plan. Each of these elements, in turn, comprises one or more additional levels of guideline constructs.

Components of GEM are defined as XML elements. Elements have distinct names and are delimited with start and end tags, e.g.,

⟨**Title**⟩ Diabetic Nephropathy ⟨**Title**⟩

Elements may contain other elements, they may store text, or they may be empty. Elements may appear as often as required. Most elements store information that is presented literally in the guideline text itself, e.g., release date, name of sponsoring organization,
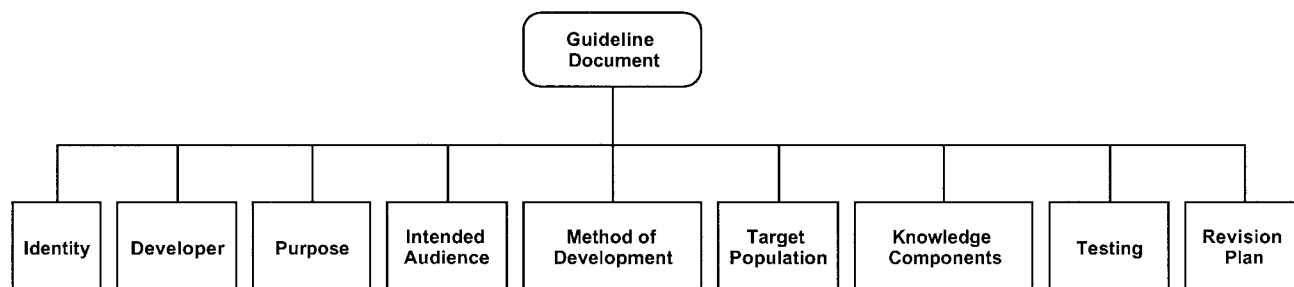
Figure 1 High-level concepts in the Guideline Elements Model.

and recommendation text. A small number of meta-level tags provide information **about** the guideline, which has been interpreted, e.g., *developer.type*. To indicate whether an element's content is explicitly stated in the guideline document or was inferred by the person who performed the markup, each element has an attribute called "source." The source attribute can take values of "explicit," "inferred," or in some cases "NGC" (to indicate that the NGC structured vocabulary is used).

GEM has been proposed as an ASTM E31.25 standard representation for guideline documents. Following ASTM and HL7 conventions, element names in GEM are formatted in lower case and words are separated by periods. In this report, italics indicate specific elements (e.g., *title, decision.variable*). The complete GEM hierarchy, definitions for all elements, a GEM template, the document type definition, and the schema can be viewed at http://ycmi.med.yale.edu/GEM. In the next sections, we describe the major elements of the GEM hierarchy.

Identity

Information that identifies a particular guideline document and describes it in general terms is clustered in the identity construct. The *identity* element includes the guideline's complete *title*, a *citation* that references its publication, its *release date*, its *availability* (in *electronic* and *print* formats), and a person or organization that can be contacted for further information. The *status* element indicates whether the guideline has been updated or revised. Since many current guidelines are released as packages that may include patient education materials, foreign language versions, quick reference guides, and technical reports, a construct for *companion.documents* is included. An entry stored in the *adaptation* element indicates whether the guideline has been adapted from another publication.

Developer

The organization responsible for development of the guideline is identified and described. A *developer.type*

element (e.g., medical specialty society, federal government agency, managed care organization) provides a structured description of the guideline's sponsor. The formal name of the committee within the developing organization as well as its members' names and individual or committee expertise are represented. In addition, sources of financial support for the guideline's development, the names of organizations that have endorsed the guideline, and reference to other organizations' guidelines on the same topic are included.

Purpose

Purpose elements describe the main health practices, services, or technologies addressed by the guideline and reasons for the guideline's development. Guideline *category* classifies the major focus of the guideline, e.g., diagnosis, treatment, or prevention.

The *rationale* for guideline development (e.g., evidence of inappropriate practice, wide practice variation) is subtly different from the *objective* of the guideline (e.g., to increase use of a particular test, to diminish inappropriate use of a therapy), and either (or both) may be described. The *health.outcome* element stores the specific health outcomes or performance measures that the guideline is intended to affect. The *available.option* element describes the principal alternative preventive, diagnostic, or therapeutic interventions that are available. *Exception* refers to factors that may permit an exception to be made in applying the guidelines, including home and family situation and constraints on the health care delivery system. Strategies, performance measures, and plans for implementing the recommendations may be stored in the implementation.strategy element.

Intended Audience

The *intended.audience* element refers to the health care providers whose behavior the guideline is intended to influence. It includes both *professional.group* and *care.setting* constructs, indicating where a guideline recommendation may be applicable, e.g., office, inten-

sive care unit, or a particular health maintenance organization. The *clinical.speciality* element applies the NGC structured vocabulary to categorize the intended users.

Method of Development

The validity of a guideline's recommendations is closely tied to concepts incorporated in *development. method*. Evidence-based guideline development processes relate recommendations directly to the scientific evidence that supports them. Such constructs are clearly important to developers and implementers and to end users of guideline recommendations as well, as they decide whether the recommendations should influence their behavior.

The *description.evidence.collection* element refers to approaches taken by the guideline developers to identify and retrieve scientific evidence. The *method.evidence. collection* element stores an NGC structured construct; *number.source.documents* refers to the number of documents identified during evidence collection. *Evidence.time.period* refers to the publication dates of the evidence. *Method.evidence.grading* stores criteria used to gauge the quality of information from different sources and may include a formal *rating.scheme. Method.evidence.combination* refers to formal methods of synthesis used to develop summary measures that reflect the strength of scientific evidence, e.g., meta-analysis, decision analysis, or formal group judgment techniques.

Specification.harm.benefit describes qualitatively the anticipated benefits, potential risks, or adverse consequences associated with implementing the guideline recommendations, while *quantification.harm.benefit* stores mathematical models and numeric estimates.

The *role.value.judgment* element stores information related to whose values were applied in determining the relative desirability of a health practice. For example, guidelines that optimize health care from the point of view of the individual patient, the payor, and society may well differ. Likewise, the *role.patient.preference* element stores information about how preferences were applied to determine guideline policies.

Target Population

Target.population refers to the group of persons who are the subject of the guideline recommendations. The *eligibility* element may include criteria—the *inclusion.criterion* and *exclusion.criterion*—that determine the specific portion of the target population for which recommendations are applicable. For example, a guideline on managing otitis media in young chil-

dren defines the inclusion criteria as "age 1 through 3 years" and "otherwise healthy except for otitis media with effusion." Exclusion criteria are specified as "craniofacial or neurologic abnormalities" and "sensory deficits."[16] The NGC specifies sex and age ranges for categorization of the target population.

Testing

The *external.review* element stores information about the findings of persons and groups outside the sponsoring organization, who have reviewed recommendations. The *pilot.testing* element stores text that refers to testing of the guideline's recommendations in clinical settings.

Revision Plan

The *scheduled.review* and *expiration* elements store review and expiration dates that help determine the validity of the recommendations in light of new evidence.

Knowledge Components

Knowledge components store and categorize the expert knowledge that is the salient feature of clinical practice guidelines. We have classified *knowledge. components* into three high-level classes—*recommendation*, *definition*, and *algorithm*—because the sub-elements of each of these call for different approaches to processing (Figure 2). Each knowledge component and its sub-tree in the GEM hierarchy are discussed below.

Recommendation

Recommendations are the unique components that distinguish guidelines from other clinical publications; recommendations are intended to influence practitioners' behavior. When recommendations are analyzed into atomic concepts (and perhaps encoded in a structured vocabulary), they can be executed by a computer's logic.

Recommendations can be categorized as *conditional* or *imperative* statements. While conditional statements clearly delineate the situations in which they apply, imperatives are broadly applicable to the target population and do not impose constraints on their pertinence.

Conditional recommendations can be described in rules that take the form:

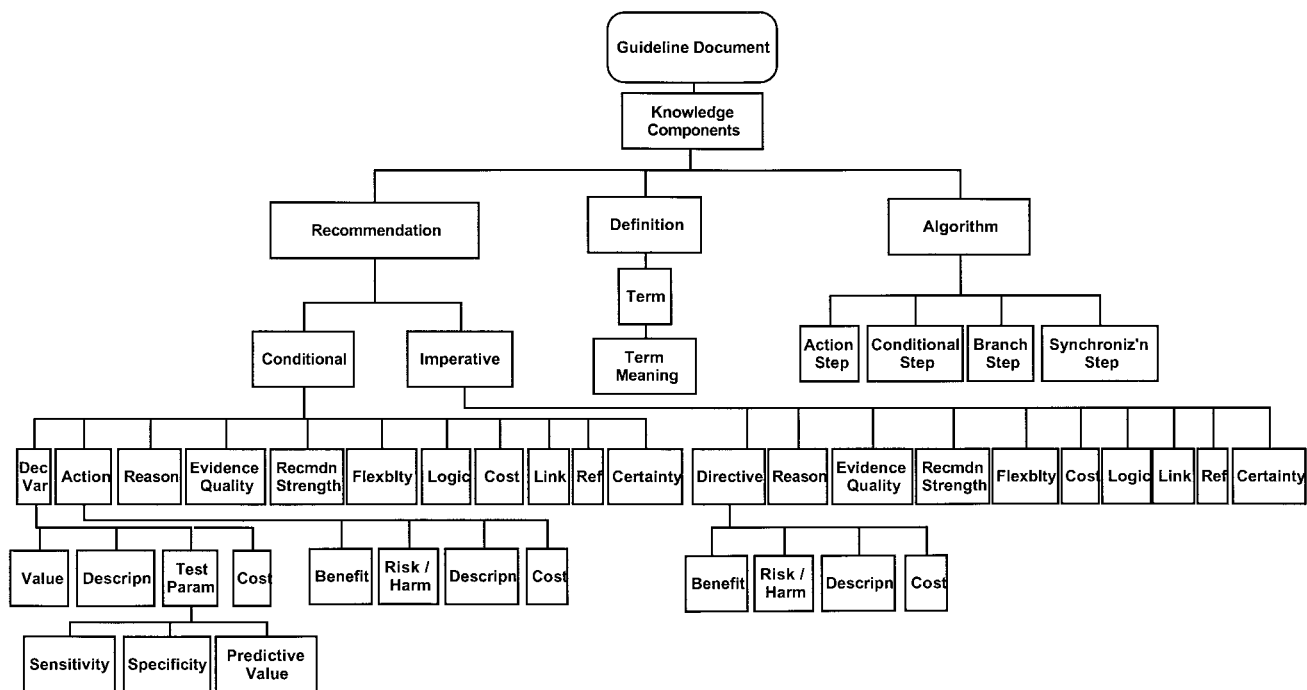> If CONDITION then ACTION(s)
> {because REASON(s)}

Figure 2 Detailed model of the knowledge components hierarchy.

A condition, in turn, is specified by one or more combinations of a *decision.variable* and its value linked by comparison operators, e.g., ⟨*decision.variable*⟩platelet count⟨/*decision.variable*⟩⟨*value*⟩less than 50,000⟨/*value*⟩. In many cases, the value of a decision variable is not explicitly stated in guideline text but is implied to be true or present.

Fulfillment of the condition triggers at least one guideline-specified *action*. *Reason* elements explain why the action has been triggered. The *evidence.quality* that led the guideline developers to call for a particular recommendation and the *strength* they attach to a particular recommendation are also tagged. The *flexibility* element describes optional conditions or actions that relate to a particular rule and are often recognizable by the presence of "or" statements in the guideline text. Defining a condition and executing an action often entail an economic burden that can be described in cost elements associated with an individual *decision.variable* or *action* or with the higher-level *conditional*. Information about the relationships between recommendations is stored in the link element. Such links might define a temporal sequence or a part–whole relationship or relate one part of the hierarchy to another. A *reference* slot can be used to store citations to specific evidence that supports a particular recommendation. The *logic* element summarily stores the Boolean connectives that link component decision variables and actions; for example:

> IF *decision.variable* 1 AND *decision.variable* 2
> THEN *action* 1 OR *action* 2.

At deeper levels of the conditional tree, elements store information that describes in detail individual decision variables and actions. Specific elements define quantitative test parameters for individual decision variables (sensitivity, specificity, predictive.value) and benefits and risks or harms associated with individual actions.

In contrast to conditional recommendations, imperative recommendations present broadly applicable directives (which parallel the actions in a conditional recommendation), e.g.,

> A major aspect of initial treatment should consist of lifestyle modifications, such as weight loss, reduction of salt and alcohol intake, and exercise. . . .[17]

> The laboratory must use a screening procedure that will detect sickle hemoglobin in the newborn. . . . Test results must be reported in understandable language that includes the identified phenotype, diagnostic possibilities, and sources where additional information may be obtained. The laboratory also should inform the infant's mother of the screening result, unless prohibited by law.[18]

Imperatives often include terms such as "require," "must," and "should" but do not contain conditional

text (e.g., "if," "when," "whenever") that would limit their applicability to specified circumstances. With the exception of *decision.variable* elements (which exist only in the conditional sub-tree), most of the deeper-level elements of the knowledge components hierarchy are similarly applicable to both imperative and conditional statements.

Definition

A *definition* element stores important guideline terminology as well as the meaning of the terms. For example, a guideline that advises on appropriate diagnostic testing for children with febrile seizures includes a careful definition of "simple febrile seizure."[19] A Centers for Disease Control guideline on hepatitis B immunization recommends more intense immunoprophylaxis for infants of "high risk mothers," a high-level concept defined to include intravenous drug abusers and women with sexually transmitted diseases during pregnancy or pre-existing liver disease.[20] Indeed, the American Academy of Neurology has issued a guideline that is expressed as a set of case definitions—rather than recommendations—for HIV-associated neurologic disease.[21]

Algorithm

Many (but not all) guidelines include an algorithm that is graphically represented in flowcharts. This describes a temporal sequence of activities and the branching decision logic that implement the guideline's recommendations. In GEM, a flowchart can be included en bloc as an *algorithm* element, or it can be broken down into its component parts.

The GLIF specification consists of a collection of "guideline steps," which are linked in a directed graph.[9] The GEM *algorithm* hierarchy includes elements derived from the GLIF steps model: *action.step*, which specifies a clinical action that is to be performed in the patient-care process; *conditional.step*, which directs flow from one guideline step to another on the basis of evaluation of a criterion; *branch.step*, which directs flow in alternate directions; and *synchronization.step*, which represents a convergence of other steps.

Mapping

Although considerable research has been focused on representing guideline knowledge, no single model has gained wide acceptance. We selected, from our literature review, a sample of existing models that we think are representative of a range of guideline document models, to explore our hypothesis that current models are limited in their comprehensiveness.

We attempted to map GEM elements to constructs described in these published models. Each publication was reviewed by at least two authors to establish whether a particular concept was described in the publication. Conflicts were resolved by discussion among the authors. In general, we tended to be liberal in these mappings, because we recognized that the published specifications of models might be incomplete.

To represent "health services" models, we selected the Institute of Medicine's Appraisal Instrument,[10] the NGC model,[7] a proposal for structured abstracts of scientific papers that describe clinical practice guidelines,[22] and a recent evaluation of guideline quality.[23] Informatics models of guideline documents include the Arden Syntax, an international standard for encoding logic in decision support systems[24]; GLIF, a knowledge representation intended to facilitate guideline sharing[9]; a relational model that captures both structured guideline content and procedural logic[25]; PRESTIGE, a generic approach to representation of guideline knowledge in the European Community[26]; a Web-based guideline dissemination system available nationally on the Kaiser Permanente intranet[27,28]; and a model based on augmented decision tables.[8]

In Figure 3, we show the high-level concepts from GEM that are represented in these selected models. In general, health services models were more explicit with regard to developer, purpose, intended audience, and method of development concepts than were informatics models, as indicated by the density of the gray bars in those segments. On the other hand, although every model had a construct for "recommendation," the informatics models tended to atomize knowledge components into more detailed constructs than did the health services models. The display is somewhat deceptive, because each element is given equal visual weighting. Nonetheless, it is clear that many informatics models lack constructs for encoding knowledge about guideline development methodology and validity assessment. Likewise, the health services models under-specify features that facilitate implementation of recommendations.

Guideline Markup with GEM

The usability and expressive adequacy of GEM were tested by applying it to a selection of guidelines. The authors modeled published guidelines as GEM documents from a variety of disciplines that represented areas of their expertise. Only practice guidelines sponsored by national organizations were modeled, although we believe that the process should be appli-

	Health Services				Informatics						GEM
	NGC	IOM	Informative Abstracts	Shaneyfelt	Arden Syntax	GLIF	Relational	Prestige	CPG DTD	ADT (Decision Tables)	
Identity											
Title											
Citation											
Release Date											
Availability											
Contact											
Status											
Companion Document											
Adaptation											
Developer											
Developer Name											
Committee Name											
Funding											
Endorser											
Comparable Guideline											
Purpose											
Health Practices											
Category											
Rationale											
Objective											
Available Options											
Implementation Strategy											
Health Outcomes											
Exceptions											
Intended Audience											
Care Setting											
Clinician Users											
Method of Development											
Evidence Collection											
Evidence Time Period											
Evidence Grading											
Combining Evidence											
Specification of Harm/Benefit											
Quantification of Harm/Benefit											
Value Judgment											
Patient Preference											
Qualifying Statement											
Cost Analysis											
Target Population											
Eligibility Criteria											
Population Description											
Knowledge Components											
Recommendation											
Conditional (decision variable)											
Action											
Logic											
Reason											
Strength of Recommendation											
Evidence Quality											
Cost											
Certainty											
Algorithm											
Definition											
Testing											
External Review											
Pilot Testing											
Revision											
Expiration Date											
Scheduled Review											

Figure 3 GEM constructs represented in a variety of guideline models. A high-level concept was considered to be present if it or any of its subordinate concepts was described. See text for sources.

```
<guideline.document>
    <identity>
        <title>The Diagnosis, Treatment, and Evaluation of the
        Initial Urinary Tract Infection in Febrile Infants and Young
        Children</title>
            <citation>American Academy of Pediatrics. Practice
            Parameter: The diagnosis, treatment, and evaluation of
            the initial urinary tract infection in febrile infants and
            young children. Pediatrics 1999;103:843-52.
                <guideline.length>9 pages</guideline.length>
            </citation>
            <release.date>1999</release.date>
            <availability>
                <electronic>http://www.pediatrics.org/cgi/content/
                full/103/4/e54</electronic>
                <print>Pediatrics(Journal)</print>
                <contact/>
            </availability>
            <status/>
            <companion.document>A detailed description of the
             methods by which the parameter was derived is
            available in a technical report from the American
            Academy of Pediatrics.
                <patient.resource/>
            </companion.document>
            <adaptation/>
    <identity>
```

F i g u r e 4 Identity elements from a guideline depicted hierarchically as a GEM document. Opening and closing tags are shown in boldface. *Contact, status,* and *patient.resource* are empty. The complete markup of this guideline may be viewed at http://ycmi.med.yale. edu/GEM.

cable to local guidelines. No effort was made in this study to model critical pathways or clinical trial protocols, which may require additional elements. Electronic versions of the guidelines were marked up using Microsoft XML Notepad 1.5, by copying pertinent text and pasting it into an empty GEM Schema template. The guidelines can be viewed using Internet Explorer 5.0 (or later versions) at http://ycmi.med.yale. edu/GEM. An example of a portion of a GEM document is shown in Figure 4.

As might be expected, there was substantial variation in the use of GEM elements from guideline to guideline. No two guidelines—whether produced by the same organization on different topics or produced by different organizations and covering the same health condition—are constructed identically. We noted that a number of tags were not used to model any of the documents.

Moreover, as has been noted in the evaluation of GLIF encoding,[9,29] there was considerable variation in the way modelers analyzed the guidelines. GEM offers

flexibility with respect to the granularity at which individual elements are encoded, and this flexibility was exploited. This effect was not formally studied and will be the subject of future research.

Discussion

We propose a document model for practice guidelines that can store and organize the heterogeneous information they contain. Although the elements identified in this work could be added to most existing health services and informatics models, GEM describes concepts and knowledge more comprehensively than do other current models.

An ideal guideline knowledge model should be:

- **Comprehensive**, i.e., capable of expressing all the knowledge contained in the guideline. Existing health services models of guidelines are inadequate for expressing the complexity of knowledge components in sufficient detail to facilitate electronic translations. On the other hand, existing informatics models are insufficient to model constructs that express and support guideline validity. Lack of confidence in the validity of guideline recommendations may ultimately limit end-user adherence.[30]

- **Expressively adequate** to convey the complexities and nuances of clinical medicine while remaining **informationally equivalent** to the original guideline.[29] Most tagged elements in GEM store the actual language of the guideline, thereby remaining true to the original. Moreover, GEM does not require recommendation knowledge to be structured in a temporal sequence, an often artificial transformation necessary for algorithmic representations.

- **Flexible**, i.e., a useful model must be able to deal with the variety and complexity of guidelines.[31] The representation should permit modeling at high and low levels of granularity, so that guidelines can be interpreted at different levels of abstraction.[29] GEM allows markup using high-level tags, or deeper analysis using elements lower in the hierarchy. In addition, the open XML document model can be modified easily, if necessary, to accommodate missing semantic constructs.

- **Comprehensible**, i.e., it should match the stakeholders' normal problem-solving language and allow domain experts to describe their knowledge with little effort. GEM markup does not require knowledge of programming. The markup process parallels physical highlighting of a document and should be learned easily by nonprogrammers.

- ■ **Shareable** across institutions. The use of XML for knowledge representation and markup provides unparalleled cross-platform compatibility.

- ■ **Reusable** across all phases of the guideline life cycle.

GEM markup can be used as a first step in translating paper-based, narrative guidelines into formats that can be processed electronically. Developers can use a GEM-encoded document as a set of empty slots to be filled to create a high-quality guideline; e.g., a fully tagged document could facilitate decision table verification of guideline logic,[32] or an XSL-formatted template can help automate the extraction of components that indicate methodological quality.[23] Disseminators can use XML's Web capabilities to publish guidelines. Implementers can use the tagged data for assistance in encoding recommendations, understanding terminology, and even direct execution. For example, a conditional recommendation's decision variables could be automatically extracted from a GEM document and used to label a data collection control, while the potential values are used to name radio buttons. End users can select various aspects of interest from GEM-encoded documents (e.g., the quality of evidence that supports individual recommendations, the costs of interventions) or compare guidelines on the same topic from different sources.

Use of XML for representation of GEM offers a number of advantages.[33] The self-descriptive capability of XML improves searching for, indexing, and locating information. Moreover, the open XML standard facilitates development of tools for document processing. XML is an intrinsic part of the Web, with presentation and parsing capabilities built into Web browsers. Software to process XML documents is expected to become ubiquitous and inexpensive. Over their lifetime, documents represented in XML can be used and reused in a multitude of ways, including (most likely) some that have not yet been invented.

GEM has several limitations. The model is simply an abstraction of the guideline document and, as such, must rely on extrinsic systems to apply it in ways that are useful. GEM does little to resolve the ambiguities that are present in many guidelines. It can, however, faithfully present them to a user for resolution. Use of a system that forces developers to define recommendations as if-then-else statements might help avoid introduction of ambiguous statements.[34] Although GEM extends the work of multiple researchers, this model is probably not comprehensive. Additional elements, attributes, and relationships may be necessary to adequately encode guidelines, depending on the needs

of stakeholders. The XML representation can be updated easily to accommodate these needs. Since the model currently incorporates more than 100 elements, effective markup with GEM will require training and practice to achieve optimal results.

Next steps in our work with GEM will involve refining the model and building and evaluating tools that facilitate activities throughout the guideline life cycle. We are working to create parsing and editing tools, specifically designed for guideline markup, that will promote consistent encoding. We envision Web-based tools for guideline developers that will allow them to collaborate effectively without face-to-face meetings —a major source of guideline development expense. Another goal of this project is to create clinical decision support tools automatedly from guideline documents stored in GEM format. Boxwala et al.[35] have described an architecture for a guideline execution engine using ActiveX, which operates from guidelines encoded in GLIF. We plan to apply ASP (active server pages) technology to dynamically configure intranet Web pages from XML documents.

GEM is intended to meet the needs of a wide variety of stakeholders in the guidelines initiative. More comprehensive models are necessary to describe fully the heterogeneous knowledge contained in clinical practice guidelines. The extensibility and computability of XML make it ideal for guideline document representation. We offer this model as an open, extensible framework and welcome contributions from others working in this area.

References ■

1. Woolf SH. Practice guidelines: a new reality in medicine. Arch Intern Med. 1990;150:1811–8.
2. Zielstorff RD. Online practice guidelines: issues, obstacles and future prospects. J Am Med Informatics Assoc. 1998;5: 227–36.
3. Degoulet P, Fieschi M. Introduction to Clinical Informatics. New York: Springer-Verlag, 1997.
4. Laurent SS. XML: A Primer, 2nd ed. Foster City, Calif: M&T Books, 1999.
5. Ceponkus A, Hoodbhoy F. Applied XML: A Toolkit for Programmers. New York: Wiley, 1999.
6. Marchal B. XML by Example. Indianapolis, Ind: QUE, 2000.
7. National Guideline Clearinghouse Web site. Available at www.guideline.gov. Accessed Dec 31, 1999.
8. Shiffman RN. Representation of clinical practice guidelines in conventional and augmented decision tables. J Am Med Inform Assoc. 1997;4(5):382–93.
9. Ohno-Machado L, Gennari JH, Murphy SN, et al. The guideline interchange format: a model for representing guidelines. J Am Med Inform Assoc. 1998;5:357–72.
10. Institute of Medicine. Guidelines for Clinical Practice: From Development to Use. Washington, DC: National Academy Press, 1992.

11. Shiffman RN, Greenes RA. Use of augmented decision tables to convert probabilistic data into clinical algorithms for the diagnosis of appendicitis. Proc 15th Annu Symp Comput Appl Med Care. 1991:686–90.

12. Shiffman RN, Greenes RA. Rule set reduction using augmented decision table and semantic subsumption techniques: application to cholesterol guidelines. Proc 16th Annu Symp Comput Appl Med Care. 1992:339–43.

13. Shiffman RN, Leape LL, Greenes RA. Translation of appropriateness criteria into practice guidelines: application of decision table techniques to the RAND criteria for coronary artery bypass graft. Proc 17th Annu Symp Comput Appl Med Care. 1993:248–52.

14. Shiffman RN, Freudigman KA, Brandt CA, Liaw Y, Navedo DD. A guideline implementation system using handheld computers for office management of asthma: effects on adherence and patient outcomes. Pediatrics. 2000;105:767–73.

15. Corb GC, Liaw Y, Brandt CA, Shiffman RN. An object-oriented framework for the development of computer-based guideline implementations. Methods Inf Med. 1999;38:148–53.

16. Stool S, Berg A, Berman S, et al. Otitis Media with Effusion in Young Children. Agency for Health Care Policy and Research Clinical Practice Guideline 12. Washington, DC: AHCPR, 1994.

17. American Diabetes Association. Diabetic nephropathy. Diabetes Care. 1999;22(suppl 1):S66–S73.

18. Sickle Cell Disease Guideline Panel, U.S. Department of Health and Human Services. Sickle Cell Disease: Screening, Diagnosis, Management, and Counseling in Newborns and Infants. Agency for Health Care Policy and Research Clinical Practice Guideline 6. Washington, DC: AHCPR, 1993.

19. American Academy of Pediatrics Provisional Committee on Quality Improvement. Practice parameter: the neurodiagnostic evaluation of the child with a first simple febrile seizure. Pediatrics. 1996;97:769–72.

20. Centers for Disease Control. Hepatitis B virus: a comprehensive strategy for eliminating transmission in the United States through universal childhood vaccination: recommendations of the Immunization Practices Advisory Committee (ACIP). MMWR. 1991;40(RR-13):1–25.

21. Working Group of the American Academy of Neurology (AAN) AIDS Task Force. Nomenclature and Research Case Definitions for Neurological Manifestations of Human Immunodeficiency Virus Type-1 (HIV-1) Infection. AAN Practice Handbook. St. Paul, Minn: AAN, 1997:41–55.

22. Hayward RSA, Wilson MC, Tunis SR, Bass EB, Rubin HR, Haynes RB. More informative abstracts of articles describing clinical practice guidelines. Ann Intern Med. 1993;118: 731–7.

23. Shaneyfelt TM, Mayo-Smith MF, Rothwangl J. Are guidelines following guidelines? The methodological quality of clinical practice guidelines in the peer-reviewed medical literature. JAMA. 1999;281:1900–5.

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

25. Lobach DF, Gadd CS, Hales JW. Structuring clinical practice guidelines in a relational database model for decision support on the internet. AMIA Annu Fall Symp. 1997:158–62.

26. Gordon C, Herbert I, Johnson P. Knowledge representation and clinical practice guidelines: the dilemma and prestige projects. In: Medical Informatics Europe '96: Human Facets in Information Technologies. Amsterdam: IOS, 1996:511–5.

27. Dolin RH, Alschuler L, Biron PV, et al. Clinical practice guidelines on the Internet: a structured, scalable approach. MD Comput. 1999;16:60–4.

28. Lukoff JY, Dolin RH, McKinley CS, Fuller LM, Biron PV. Validation of an XML-based process to automatically Web-enable clinical practice guidelines: experience with the smoking cessation guideline. AMIA Annu Symp. 1999:311–4.

29. Patel VL, Allen VG, Arocha JF, Shortliffe EH. Representing clinical guidelines in GLIF: individual and collaborative expertise. J Am Med Inform Assoc. 1998;5:467–83.

30. Weingarten S. Practice guidelines and prediction rules should be subject to careful clinical testing. JAMA. 1997; 277(24):1977–8.

31. Tu SW, Musen MA. A flexible approach to guideline modeling. AMIA Annu Symp. 1999:420–4.

32. Shiffman RN, Greenes RA. Improving clinical guidelines with logic and decision table techniques: application to hepatitis immunization recommendations. Med Decis Making. 1994;14(3):245–54.

33. McGrath S. XML by Example: Building E-Commerce Applications. Upper Saddle River, NJ: Prentice-Hall, 1998.

34. 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.

35. Boxwala AA, Greenes RA, Deibel SRA. Architecture for a multipurpose guideline execution engine. AMIA Annu Symp. 1999:701–5.