

Computer-Interpretable Clinical Practice Guidelines

Where Are We and where Are We Going?

F. A. Sonnenberg, C. G. Hagerty

Department of Medicine, UMDNJ Robert Wood Johnson Medical School, New Brunswick, New Jersey, USA

Summary

Objectives: To provide a comprehensive overview of computer-interpretable guideline (CIG) systems aimed at non-experts. The overview includes the history of efforts to develop CIGs, features of and relationships among current major CIG systems, current status of standards developments pertinent to CIGs and identification of unsolved problems and needs for future research

Methods: Literature review based on PubMed, AMIA conference proceedings and key references from publications identified. Search terms included *practice guidelines*, *decision support*, *controlled vocabulary* and *medical record systems*. Papers were reviewed by both authors and summarized narratively.

Results: There is a consensus that guideline delivery systems must be integrated with electronic health records (EHRs) to be most effective. Several evolving CIG formalisms have in common, use of a *task network model*. There is currently no dominant CIG system. The major challenge in development of interoperable CIGs, is agreement on a standard controlled vocabulary. Such standards are under development, but not widely used, particularly in commercial EHR systems. The Virtual Medical Record (VMR) concept has been proposed as a standard that would serve as an intermediary between guideline vocabulary and that used in EHR implementation.

Conclusion: CIG systems are in a state of evolution. Standards efforts promise to improve interoperability without compromising innovation. The VMR concept can assist guideline development even before clinical systems routinely adhere to standards. Frontiers for future work include using the principles learned by computer implementation of guidelines to improve the guideline development process and evaluation methods that isolate the effects of specific CIG features. Haux R, Kulikowski C, editors. *IMIA Yearbook of Medical Informatics 2006. Methods Inf Med 2006; 45 Suppl 1: S145-58.*

Keywords

Decision Support Systems, Clinical Practice Guidelines / st [Standards] Programming Languages, Software Design, Vocabulary, Controlled, Medical Records Systems, Computerized

Clinical Practice Guidelines

In its 1992 monograph, the Institute of Medicine defined clinical practice guidelines as “systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances” [1]. Many other terms are used to describe *clinical practice guidelines* including *clinical parameters*, and *practice policies* [2]. For the sake of simplicity, this paper will refer to them simply as *clinical guidelines*. Clinical guidelines have been advocated with increasing frequency to reduce inappropriate care [3], control geographic variations in practice patterns and make more effective use of health care resources. Guidelines are an intrinsic component of disease management programs [4], increasingly applied to improve outcomes for patients with chronic illnesses. Another reason for promoting guidelines is the ever increasing quantity of new knowledge available about health interventions. Guidelines are promoted as a means of providing this knowledge to clinicians in a readily digestible form without requiring that they search through large volumes of published material [5]. Increasingly, adherence to guidelines is considered a measure of quality of care [6] and may be used to award managed care contracts, or determine the reimbursement of practitioners. They are now part of well-established quality measurement

programs and are increasingly becoming part of mandatory quality programs, such as those of the Joint Commission on Accreditation of Healthcare Organizations [7].

Clinical guidelines have been promoted against a background of data showing that there is a significant gap in the application of the evidence base to clinical practice [8-10]. Studies in the U.S. and the Netherlands have shown that about 30–40% of patients do not receive care according to current scientific evidence, and about 20–25% of care provided is not needed or is potentially harmful [11, 12]. More recent studies [10, 13] show that these trends persist although they vary considerably by clinical condition and also by geographic location and provider.

Benefits of Guidelines

There is debate about whether or not guidelines result in improved medical care [14], but there is little question that they are a well-established and increasingly important part of medical practice. Therefore, it is imperative for guidelines to be of the best possible quality. Quality is defined in terms of five objectives specified by Eddy [2]. They must be 1) accurate - producing the intended results 2) accountable - documenting the reasoning behind the recommendations, 3) predictable -the health and financial consequences

should be anticipated, 4) defensible - the guideline will facilitate resolution of conflicts across policies and 5) usable -capable of being employed in practice. A number of formal efforts have been established to evaluate guideline quality [15]. The Agree Collaboration [16] is a European project to evaluate systematically the quality of clinical guidelines. The AGREE instrument consists of 23 key items organized in six domains: 1) scope and purpose 2) stakeholder involvement of guideline development 3) rigor of development 4) clarity and presentation 5) applicability and 6) editorial independence to avoid conflicts of interest. Shiffman et al published the findings from the Conference on Guideline Standardization (known as the "COGS" statement) to define a standard for guideline reporting [17].

The evidence about whether interventions to increase the use of clinical guidelines improve the process of care is unclear. In 1993, Grimshaw and Russell performed a systematic review of rigorous evaluations of 59 clinical guidelines. In nearly all cases significant improvements were observed as a result of the interventions [18]. Using the process of care as a measure of the benefit of clinical guidelines is circular reasoning to a degree because the measures of care are generally those specified in the guidelines. A small minority of these studies examined actual clinical outcomes, but where they did, the majority showed a favorable effect [18]. On the other hand, Worrall et al. found that only a minority of guideline interventions provided statistically significant benefits [19]. In 1999, Shiffman performed a systematic review of 25 studies of computer-based clinical guideline systems [20]. Adherence to guidelines was improved in 14 of 18 studies that evaluated it. Documentation was improved in all 4 studies that reported

it. In 8 studies that reported "clinical outcomes", quasi-clinical outcomes (cost or completion of advanced directives) only 3 showed improvement.

Guideline Dissemination (Implementation)

Most guidelines are issued either as published articles, as specialized monographs or both [21]. In either case, such guidelines often have little impact on clinical practice either because clinicians are unaware of them or because the guideline is not accessible during the provision of care [18, 22, 23]. Tunis et al. showed that physicians were poorly informed about the details of several well-established guidelines yet claimed knowledge of a non-existent guideline presented in the study as a control [24]. Ellrodt et al. showed that only a small part of non-compliance with guidelines resulted from disagreement with them. The remainder resulted from other factors including implementation issues [25]. More recently, studies have shown barriers to adoption of guidelines by physicians, including lack of awareness and unfamiliarity with guidelines, and inertia of previous practice [26, 27]. Moreover, guidelines can suddenly become obsolete in the face of newly published research or the availability of new tests and treatments. For example, the Adult Treatment Panel III (ATP III), the most widely accepted guidelines for management of hypercholesterolemia, were published in 2001 [28]. Since then, and long before a revision of the ATP III guidelines were planned, several major clinical trials resulted in important modifications of the goals set in the original guidelines [29].

A variety of approaches have been used to computerize guidelines. At the most

rudimentary level, some of these are merely electronically readable versions of text guidelines. Guidelines that provide immediate as opposed to delayed feedback to clinicians have been shown to be superior in their effects on clinicians [30]. Furthermore, Weingarten [19] showed that when computer-based reminders stop, compliance with the guideline declines to its baseline [31]. Thus, optimal benefit from guidelines requires that such guidelines be built into existing clinical information systems [30,32]. Guidelines available only "on line" via telecommunication services are, in general, not available during the process of care and given the current state of the art, not linked to patient attributes. Electronic publication potentially may bring the content of guidelines to the point of care, but may not make them easily usable [33]. For example, the National Guideline Clearinghouse (NGCH) [www.guidelines.gov], an online reference source maintained by the Agency for Healthcare Research and Quality (AHRQ) is comprehensive, but guidelines are presented in a variably structured text format. Each guideline is provided with a summary, which summarizes the important recommendations. The summaries are a significant advance over manual methods of reviewing guidelines. However there may be multiple competing guidelines for a single clinical problem, and it may take too long to find a specific guideline for a clinical problem and then to find the portion applicable to a given patient. The full guidelines may be very lengthy documents, and the actual full text may reside on another web site, which may be inaccessible to the clinician without a subscription to a journal or membership in a professional society. The NGCH is a very important reference tool but not ideally suited for use by a busy clinician during clinical encounters.

A clear consensus has emerged that clinical guidelines should be deployed through clinical information systems and made available during clinical encounters. More and more, the literature on the effects of deploying clinical guidelines has focused on computer-based interventions [34-43]. While many of these focus on simple reminder systems, many others represent increasingly complete presentations of entire clinical guidelines. Not all of the studies are positive. Tierney et al. found that a sophisticated electronic health record (EHR) system, did not increase compliance with guidelines for congestive heart failure of the Agency for Health Care Policy and Research (AHCPR, now AHRQ) [44]. Montgomery et al. found that chart reminders were effective in lowering blood pressure but that there was no additional benefit to adding computer-based decision support [45]. Ansari et al. [46] found no benefit from computerized physician reminders to use beta blockers in patients with heart failure, but did find benefit from a nurse-facilitator to carry out the intervention. In Montgomery's study [45], unassisted physicians did a poor job of assessing patients' risk [47]. This was improved by provision of a risk chart but adding computer-based decision support did not provide any additional benefit. There are many reasons why use of guidelines may not be demonstrated to be effective in clinical care:

1. Historically, clinical guidelines have addressed population-based recommendations [48]. Recommendations may not be suitable for specific patients. Moreover, Wetter [49] pointed out that most guidelines are diagnosis oriented and are thus not always well-suited to the more common problem-oriented nature of patient encounters.
2. Guidelines may be logically incomplete; they may not cover all possible clinical scenarios (combinations of patient characteristics) [50].
3. The guidelines themselves may be of poor quality in reflecting the current evidence-base [51]. In part this may reflect the methods used to develop the guidelines [52].
4. Guidelines often employ concepts that require knowledge not contained in the guideline document. In our analysis of the vocabulary requirements of two clinical guidelines [53] we encountered a number of necessary concepts that were not defined within the guideline. For example, the term "high coronary disease risk" is defined in an external reference [54]. A user of the guideline would have to retrieve that reference in order to identify the 7 variables that comprise the concept. Some terms, like the term "end-stage heart disease" which appears in the JNC 7 guideline, and even very mundane terms such as "older patients", "children", and "adolescents" [55] are not defined or referenced in the guidelines. Proper interpretation, therefore, depends on the general medical knowledge of the clinician. Aside from making application of guidelines difficult, missing definitions can result in variability of implementation from one clinician to another. Another example of necessary external knowledge is knowledge about classes of test or treatments. For example, a guideline may specify that a drug from a specific class (e.g. angiotensin converting enzyme (ACE) inhibitors) should be used, but the guideline itself does not specify which drugs belong to each class. A guideline may even make a recommendation such as "choose another drug from the same class" assuming that the clinician will know the alternatives drugs for each class. This knowledge could be provided by the EHR or by the guideline system.

5. Missing knowledge in guidelines may also occur in the form of vague or ambiguous language [56]. Codish and Shiffman [56] distinguish vagueness from ambiguity. Vagueness in guidelines refers to imprecise specification of a recommended action, a decision criterion or the definition of a term. Ambiguity refers to statements which can have multiple interpretations.

Consider the following statement from the JNC VII hypertension guidelines [55]: "[Hypertension in older individuals] In many individuals, lower initial drug doses may be indicated to avoid symptoms; however, standard doses and multiple drugs are needed in the majority of individuals to reach appropriate BP targets". The term "older individuals" is vague because it is not defined anywhere in the document. The statement is ambiguous, because it is not clear whether the recommendation is to start with lower doses or standard doses in older individuals.

Vagueness is often introduced intentionally to reflect a lack of consensus, or imprecision about recommendations [57]. Ambiguity, though possibly deliberate due to lack of consensus, is more likely to be inadvertent. Their resolution poses a challenge for faithful encoding of the original authors intention and the ability to achieve reproducible application of guidelines. The term "clinical judgment" is often used in clinical guidelines to indicate that precise decision criteria are not specified in the guideline and are expected to be supplied externally by the clinician using unspecified means. This suggests that such guidelines were not designed to be applied automatically by computers, since they implicitly assume clinician responsibility for the stated clinical judgment.

Clearly, then, the deployment of clinical guidelines in clinical information

systems is not, by itself, sufficient to result in successful implementation. Despite these mixed results, there is a consensus that clinical guidelines are beneficial and that the most effective way to deploy them is through clinical information systems by making them computer-interpretable. The remainder of this paper will examine the content of clinical guidelines, and will review the array of representation schemes and the processes involved in deploying them in clinical practice.

Comparison of Computer Interpretable Guideline (CIG) Systems

There are a number of existing formalisms for expressing CIG's and these have been the subject of several comprehensive reviews [48, 58, 59]. This paper will not attempt to reproduce those detailed comparisons. Instead, we will summarize the key guideline elements and implementation activities and attempt to indicate where each formalism fits into the progress toward development of guideline implementation standards and more widespread deployment of guideline support in the EHR. In one comparison between CIG systems that use a Task Network Model (TNM) representation, Peleg [48] identified the following common features of CIG models: 1) support for the Task Network Model (TNM); hierarchical decomposition of guidelines into networks of component tasks that unfold over time. This has been referred to as the "task-based paradigm" 2) support for sequential, parallel, cyclical and iterative plans. 3) multiple "entry points" (patient states) into a guideline. 3) clinical actions which include treatments and information gathering (including order-

ing tests). 4) temporal constraints on start times of guideline plan components (but they differ in ability to specify end time and duration). 5) deterministic decision criteria for "switching" among different mutually exclusive alternative plans 6) abstractions: general concepts that encompass a class of specifics, which can be used to characterize temporal, definitional, terminological or scenario concepts. 7) representation of patient data

Arden Syntax [60] is the best established framework for providing decision support in EHR's. Established in 1989, it was recognized in 1992 as a standard by American Society for Testing and Materials (ASTM) [61]. Arden consists of Medical Logic Modules (MLM's) which are collections of rules that invoke specific actions. The relative simplicity of Arden MLM's, allows for integration within medical record systems to provide reminders, and translation into other representations such as the Guideline Interchange Format (GLIF) [62] and Guideline Elements Model (GEM). It has also been demonstrated that GEM-encoded guidelines can be used to generate MLM's [63]. As with any CIG system, clinical terms in MLM's must be implemented and integrated with clinical systems of local institutions. Since the encoding of such terms is embedded directly in the Arden Syntax within curly braces (`{}`), this has come to be known as the "curly braces problem" [59]. While the decision logic of MLM's supports sharing across institutions [64], this need for local implementation greatly increases the effort required to share Arden-specified guidelines. In addition, there is no overall control structure for the execution of MLM's which limits their ability to represent complex plans. Because of these limitations, some authors argue that Arden

should not be considered as a full-fledged CIG system [48].

The EON system [65] is one of the earliest comprehensive CIG systems. Developed at Stanford University, it is a precursor of systems such as GLIF and SAGE and continues to be developed as a research system. Notable features of EON compared to previous systems include use of a general-purpose ontology/framework editing system (Protégé), mapping of terms to a controlled vocabulary and an extensible conceptual framework that supports the modeling of a network of successive decision tasks [48]. These systems are thus called Task Network Models. EON continues to evolve with its descendants such as GLIF, including incorporation of the virtual medical record [66] (VMR) concept (see "Standards" below) which attempts to address the "curly braces problem"

ProFORMA [67], developed at the Advanced Computation Laboratory of Cancer Research in the UK, has a number of distinctive features. It is based on an underlying modeling language, R²L, and a graphical display language Arezzo that is strongly goal-based [48, 58].

The Asbru [68, 69] guideline modeling language was developed jointly by Ben-Gurion University and the Vienna University of Technology. Like ProFORMA, it is goal-based, with goals referred to as "intentions". It is particularly advanced in modeling temporal aspects of guidelines.

GUIDE [70], developed at the University of Pavia in Italy, introduces two distinct features. One is the use of a formal model (Petri nets) to model guideline workflow [58]. The other distinct feature is use of formal decision analytic models for decision making [71]. Adding a decision analytic component to clinical guidelines is one way

to achieve one goal of guideline delivery systems proposed by Eddy; predicting the consequences of alternative actions [2].

GLIF [72-74] is an effort by the Intermed Collaboratory, a collaboration of researchers from Stanford, Harvard and Columbia to develop a sharable CIG representation. GLIF was developed with the same guideline modeling language (Protégé) as EON. The first iterations of GLIF lacked a formal specification of guideline steps and mapping of data elements from the guideline to the EHR [59]. A later version, GLIF3, introduced a more formal object-oriented model and a structured Extensible Markup Language (XML) syntax fully specified by the Resource Description Framework (RDF) [75]. The latest version of the object-oriented model, originally called GEL (Guideline Expression Language) is known as GELLO and is being considered by HL7 as a standard for a query and expression language [76]. GLIF has also adopted a medical data model based on the HL7 Reference Information Model (HL7-RIM) specification which provides support for the virtual medical record (VMR) concept [27]. In addition, GLIF3 has been augmented with GLEE, a guideline execution system [77].

The SAGE (Standards-Based Sharable Active Guideline Environment) Project is a collaboration of six research groups (IDX Systems, University of Nebraska Medical Center, Intermountain Health Care, Apelon, Inc., Stanford Medical Informatics and the Mayo Clinic [78,79]). SAGE was funded in part by The National Institute of Standards and Technology (NIST) Advanced Technology Program. SAGE seeks to establish an infrastructure to permit sharing guidelines in heterogeneous clinical information systems. Key to the SAGE approach is involving vendors and stan-

dards organizations (chiefly HL7) to bridge the gap between guideline logic and implementation in local clinical settings. It is in many ways, an evolutionary successor to all of the above systems, particularly EON and GLIF, and takes a step towards synthesis of standards to permit interfacing to actual clinical information systems.

There are three components to the SAGE information models: a view of patient data, a healthcare organization model and a workflow model. The significance of this arrangement is separating the healthcare enterprise model (generic steps) from the specific workflow model needed in a particular local environment. The architecture consists of *Guideline Recommendation Sets*, (a subset of guideline content tailored to a specific healthcare enterprise), which may be composed as graphs of *Context Nodes*, which specify assumptions (e.g. a physician in an emergency room) about the setting in which care is provided, *Action Nodes*, which specify information system activities carried out in support of the guideline, *Decision Nodes*, and *Routing Nodes*, which control overall execution of a guideline protocol. Routing nodes, analogous to the synchronization steps in GLIF, EON and GUIDE, specify the process by which choices are made. Decision nodes are a particularly interesting addition to guideline representations, in that they allow for the inclusion of "bound models" which may be formal utility-based or neural network decision models to make choices among alternatives [48]. As with most modern systems, SAGE makes use of standard concept codes such as SNOMED-CT and LOINC which are found in the UMLS, and places particular emphasis on incorporating the VMR concept. A notable application of the SAGE methodology

is implementation of hypertension guidelines in the ATHENA project [80, 81] which involved incorporation of SAGE-based decision support into the EHR in use at Veterans Administration clinics [82].

A number of other notable systems have been developed and reported. These include DILEMMA/PRESTIGE [83] from the United Kingdom NHS Information Management Centre, and GASTON from the University of Maastricht [84]. Detailed comparisons of these systems and supporting references may be found at the OpenClinical web site (www.openclinical.org).

At this time, there is no dominant CIG framework and no system that is in widespread clinical use outside of the institution in which it was developed [27]. The major systems are summarized in the Table.

Document-centric Approaches

A different perspective on the representation of guidelines in computer-interpretable form is taken by several groups that consider the original guideline document to be the reference repository of guideline specification. This is called the "document-centric approach" [57] to distinguish it from the "knowledge base-centric approach" taken in many of the above systems. The original guideline document is either marked up, or annotated to produce a more structured format with well-defined semantic elements. This process can be carried out in stages, initially using markup to identify elements and later assigning these elements to specific semantic tags. Markup approaches can help alleviate some of the difficulty in encoding guidelines in specialized computer languages or knowledge-based representations. This

Table 1 Computer-interpretable Guideline Systems

| System | Key references | Developer |
|------------------|------------------------------------|--|
| Arden Syntax | Hripcsak 1991 ^{60, 61} | Arden Homestead Conference |
| | Hripcsak 1993 ⁵⁷ | HL7 |
| | | ASTM |
| GLIF | Ohno-Machado 1998 ⁷² | Intermed Collaboratory |
| | Patel 1998 ⁷³ | Columbia, Stanford, Harvard |
| | Boxwala 2004 (GLIF3) ⁷⁴ | |
| PROforma | Fox 1998 ⁶⁷ | Advanced Computation Laboratory of Cancer Research, UK |
| | Sutton 2003 ¹¹² | |
| Asbru | Shahar 1998 ^{68, 69} | Ben-Gurion University |
| | Shahar 2004 ⁶⁴ | Vienna University of Technology |
| EON/DHARMA | Musen 1996 ⁶⁵ | Stanford University |
| GUIDE/PatMan | Ciccarese-2005 ¹¹³ | University of Pavia, Italy |
| | Quaglioni 2001 ⁷⁰ | |
| PRODIGY | Johnson 2000 ⁹⁸ | University of Newcastle upon Tyne, UK |
| DILEMMA/PRESTIGE | Herbert 1995 ⁸³ | NHS Information Management Centre, Birmingham, UK |

approach offers the possibility that at least, the earlier stages of the transformation can be performed by operators (e.g. clinicians) not skilled in programming. The desired result is transformation of clinical guidelines from a text-based document (which is how guidelines are usually available) into a formal representation that can be used by the execution software [85], while retaining links to the original document. The best-known document-centric approach is the Guideline Elements Model (GEM) [86, 87]. Rather than programming a guideline in computable code, GEM begins with the original narrative guideline text and has an operator use a special-purpose editor (GEM-Cutter; http://gem.med.yale.edu/GEM_Cutter/gem_cutter.htm) to copy text segments into guideline elements which the editor arranges graphically

in a hierarchy. The result is an Extensible Markup Language (XML) document containing the selected guideline excerpts, labeled as elements that are delimited by corresponding markup tags. GEM has more than 100 elements based on guidelines found in the Institute of Medicine's Guideline Appraisal Instrument, the National Guideline Clearinghouse [www.guidelines.gov], and the augmented decision table guideline representation [50]. 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 [87]. Arden MLM's [63] have been encoded auto-

matically from GEM and the elements in the guideline checklist developed by the Conference on Guideline Standardization [17, 88] were derived from this representation. ASTM has adopted GEM as a standard [27]. GEM is considerably richer than other systems in specifying details of the guideline development process (e.g. authors, purpose, intended audience, versions, and the consensus process). However, to date, although it has been shown that GEM decision elements could potentially be mapped to GLIF models, the GEM representation has stopped short of extending to integration with clinical systems [57]. It may thus be thought of more as a guideline specification than a computer-interpretable guideline representation.

Methods such as ActiveGuidelines [89] and the Hypertext Guideline Markup Language (HGML) [90] also begin with the original guideline text document (transformed into HTML) and use an editor to mark up the document by selecting text segments and enclosing them in tags. Unlike GEM, ActiveGuidelines and HGML insert tags directly into the original document, thereby producing an XHTML document which contains the entire original text as well as tags representing statements, variables, conditions and recommendations. The use of XML style sheets and Extensible Stylesheet Language (XSL- <http://www.w3.org/Style/XSL>) translations permit viewing the marked-up guideline in a variety of formats including the original text, the original text with color coding or annotation indicating the semantics of the text (as encoded by the embedded/surrounding markup tags), and purely XML structured formats showing the individual elements displayed hierarchically. An important distinction from many of the other XML systems

is that this approach produces an XHTML document which consists of the original document annotated with additional semantic tags. All of the document-centric approaches are complementary as it is theoretically possible to perform each encoding in such a way that guidelines can be translated back and forth among representations.

Recently, Eriksson et al. [91] described a document-centric approach that annotates documents in the Portable Document Format (PDF) with knowledge representation components. Advantages of this approach include preserving the “look and feel” of the original published guideline document and the fact that the PDF format is well-suited to displaying very large documents in a web browser. Furthermore, by annotating rather than adding markup, problems resulting from overlapping scopes of markup tags from different semantic perspectives are avoided.

Other approaches to guideline markup include the Guideline Markup Tool (GMT) which also produces a formal representation in XML [85] through a successive markup process. The approach has been illustrated primarily to edit Asbru protocols [85], although additional formats such as GEM are possible as well.

Shankar [92] describes a *bridge approach* that links formal knowledge structures of the knowledge base-centric approaches such as GLIF and EON with the natural language style of the guideline text in document-centric approaches. Rather than using explicit links between the two representations as in the other document-centric approaches, this is accomplished using “soft links” that involve queries whose answers are associated with the original text. This approach addresses directly the problem of updating encoded CIG’s when changes are made to the underlying

text documents. Based on the experience that even the markup process involves an encoding bottleneck, several researchers have begun to explore semi-automated identification of concepts in text-based guidelines by identifying candidate condition and recommendations and potentially presenting them to an operator who chooses the best candidates for encoding. This may involve methods such as the identification of concepts from the UMLS metathesaurus [93], or parsing of deontic verbs [94]. By applying natural language processing and machine learning techniques, we hope to improve the accuracy by which the encodings and candidate concepts can be identified and mapped to variables or actions in the EHR (or via a VMR) [53], and thus alleviate some of the burdens of the encoding process.

What’s in a Guideline?

Definition: The literature on computer-interpretable guidelines (CIG’s) collectively has defined the elements that comprise guidelines. The term “guideline” means different things to different people. Some see guidelines as recommendations for specific contexts, while others view them in terms of flowcharts, graphs or workflows. In common use, a “guideline” refers to both single recommendations for specific patients (e.g. use of ACE inhibitors in diabetics with hypertension or proteinuria) and collections of related guidelines, such as the JNC 7 hypertension guidelines [55] which all address hypertension but in a wide range of clinical circumstances or the ATP III guidelines for management of hypercholesterolemia [28], which also addresses a wide range of clinical circumstances including screening of healthy

patients, those with heart disease and those with other chronic diseases such as diabetes. In keeping with the generally-accepted Institute of Medicine definition, a guideline can be any statement that provides a recommendation for an identifiable group of patients or circumstances.

The following are elements common to all guidelines:

Patient States: CIG’s all require a means of describing a patient. This is variously referred to as “patient states” (GLIF), “scenarios” (EON and PRODIGY) or “contexts”. The term context is particularly appropriate because it coincides with the widely used term “clinical context” which refers to a specific clinical circumstance. Contexts are defined in terms of intrinsic patient variables. The challenge has been that to date, commercial EHR systems have generally used proprietary patient data schemes which require matching of patient variables in an *ad hoc* manner. Effort must be duplicated not only for every system, but frequently even for different sites using the same system. The VMR concept promises to greatly reduce the variability of this process.

While assigning standardized vocabulary to clinical observations on the patient record side and to decision variables on the guideline side is critical to sharable guidelines, there is as yet no agreement on the precise standard vocabulary that should be employed. UMLS, including SNOMED is not ideally suited for this purpose because it was developed primarily for identifying static concepts (e.g. disease states, tests or treatments) and lacks representation of attributes of specific instances (e.g. dates, durations, doses, in relation to specific patients), and the temporal and procedural components of clinical practice. The original HL7 specification is largely a messaging standard for com-

municating results from one clinical information system to another. The HL7 Reference Information Model (RIM) [95] is an attempt to create a shared information model for the EHR and decision support systems [96, 97]. The HL7-RIM is based on the Unified Service Action Model (USAM) which describes all events in terms of an action that was responsible for collecting information about the event [97]. HL7-RIM has been adopted as the patient information model for GLIF [74] and serves as the basis for the virtual medical record (VMR) used by GLIF, EON [65], Prodigy [98] and SAGE.

Execution States: Another aspect of the clinical context for application of a guideline is representation of the state of guideline execution. Thus, a CIG system must “know” not only the state of the patient, but must also know what guideline steps have already been applied, and where in time a clinical encounter fits in relation to previously performed and planned guideline steps. Explicit modeling of this aspect of the context as *execution states* is a component of the DILEMMA/PRESTIGE, PROforma, and Asbru approaches [58]. To some extent execution state could be incorporated implicitly into the triggering conditions in rule-based representations (such as Arden, HGML, or GEM) but this would not have all the advantages of an explicit representation.

Eligibility Criteria: All guidelines have eligibility criteria that determine to whom and in what setting the guideline applies or, in many cases, which of many subguidelines applies to various subsets of a patient population. Thus, the asthma guidelines of the National Heart, Lung And Blood Institute (NHLBI) have separate sections for chronic management, for management of acute exacerbations and for children [99]. Modeling of eligibility includes

explicit representation of inclusion and exclusion criteria as well as exceptions. *Triggers* that indicate to a decision support system that an available guideline applies to a patient at a specific time also must be part of any CIG system.

Goals: Goals are important guideline concepts both in terms of satisfying the overall requirements of a guideline and also as intermediate concepts in the application of complex guidelines. For example, *control of hypertension* is a general goal for the JNC 7 guidelines, but the specific concept *goal systolic blood pressure* is an initially unbound variable in application of the guideline that is instantiated based on specific patient criteria [55]. Another example is the goal low-density lipoprotein (LDL) level in the ATP III cholesterol management guidelines. The specific value of the LDL goal differs depending on several patient characteristics [28]. However, the overall control of portions of guideline execution (e.g. treat-observe-reconsider treatment) can be specified without quantifying the goal. A CIG implementation ideally, should account for these differences and also permit modifying the target goals as new scientific evidence becomes available [29]. Goals may necessitate more complex control structures for treating and observing (waiting steps) and iteration in treat-observe-reconsider cycles.

Classification Schemes: Many guidelines depend on classification schemes and risk calculators. Examples include the guidelines for community acquired pneumonia [100] which includes a Pneumonia Severity Index, the ATP III guidelines [28] which incorporate a predictor of coronary artery disease risk based on the Framingham risk score and the asthma guidelines which incorporate a measure of asthma severity [99]. Some of these classification schemes are purely categorical and others result in a con-

tinuously variable quantitative measure of risk. A CIG implementation should not only represent the variables needed to perform these risk assessments, but should provide an automated mechanism for performing the risk assessment, using it in the guideline decision logic and explaining it to the clinician. As we noted above, decision analytic models offer one means of predicting consequences of diseases or treatments and can thus potentially serve to estimate risk in guideline application.

Decisions: Clinical guidelines typically involve decisions involving choices among alternative strategies or treatments. These choices may involve simple alternatives of treatment (one drug vs. another) or entire alternative guideline pathways. Choices may be unconditional in that they apply to all patients (or at least all who are meet the eligibility criteria for the guideline) whereas others specify alternatives that depend on specific criteria. These criteria may be logic-based and deterministic (e.g. *if a and b then do x*) or may depend on more complex heuristics. In either case, presented in a specific context, there will often be more than one competing choice, for which some method will need to be used for conflict resolution. ProFORMA is particularly strong in representing heuristics based on preferences [48]. GUIDE uses a heuristic based on formal utility theory [48]. Alchemist, developed at Stanford University and Duke University [101] uses a decision analytic model to generate a guideline reflecting the knowledge in the model. Thus, in a sense, Alchemist is the inverse of a CIG system; it creates a guideline from a knowledge representation.

Actions: Guidelines contain actions which consist of treatments, information gathering steps (including ordering of diagnostic tests) and provision

of advice. To represent time delays while requested information is obtained or to observe the results of initial treatment guidelines contain *wait steps* [58] which are represented explicitly, but in different ways in the various CIG representations. EON and PRODIGY differentiate between *actions* (instantaneous actions) and *activities* (plans that are carried out over time) [58]. Guideline actions may be arranged into plans in various ways, although not all systems try to do this. Plans may be *sequential* (one step following another in a fixed order), *parallel* (branching into alternative pathways that may converge at a subsequent step), or *cyclical/iterative*, involving repeated cycles of treat-observe-decide. A plan may contain all of these elements. In Asbru, a plan is decomposable into subplans; when subplans are non-decomposable, they are referred to as *actions* [59].

Standards

Over the past decade, the U.S. Federal Government's National Library of Medicine has made available the Unified Medical Language System (UMLS) [102], which now incorporates many independently developed terminologies such as the Systematized Nomenclature of MEDicine Clinical Terms (SNOMED-CT) [103] and Logical Observation Identifiers Names and Codes (LOINC) [104]. While much ongoing research is being devoted to improving this semantic structure with more formal ontologies, it has formed the basis for a composite standard terminology. Ongoing research is aimed at using this terminology in conjunction with additional standards such as the Reference Information Model (HL7-RIM) to describe the contents of medical records [66, 79].

Arden syntax was the first recognized standard for representation of medical procedural knowledge. It is the most widespread in terms of application, being incorporated into the decision support of clinical systems by several large commercial software vendors. However, Arden syntax has three important limitations that make it unsuitable as a comprehensive guideline formalism. First, implementation of recommendations must be specified uniquely for each application site. Referred to as the "curly braces problem" [96] this problem is not unique to the Arden syntax. Second, Arden syntax rules are modular, and although one Arden rule may invoke another, there is no explicit control structure as there is in the TNM-based guideline representation formalisms [59]. Third, as it stands, the Arden data model is overly simplistic and gives rise to incompatibilities with more expressive data models such as the HL7-RIM [27, 96, 105].

Medical concept models are important for CIG implementation. However, "a standard medical concept model is currently out of reach because existing vocabularies have not been explicitly designed for clinical decision support and have limitations for such applications" [48]. The VMR (Virtual Medical Record) [66] concept has been introduced as a solution to the diversity of terminology standards in use in clinical information systems and a similar diversity of terminology in CIG representations. In the VMR, a standard set of concepts is mapped to guideline knowledge. Local adaptation to specific EHR systems would require mapping only to the VMR rather than to each different CIG system. Guideline developers could thus accommodate widespread implementation by mapping only to the VMR. The VMR currently is being considered by HL7 as a stan-

dard [78]. This would be an important step toward the development of interoperable guidelines that do not need to know the details of the eventual implementation system.

As noted above, GEM has been endorsed by ASTM as a standard [27]. Because it has de-emphasized implementation, GEM serves more as a guideline specification standard than a CIG standard. It is especially well-suited to considerations of guideline documentation [17] and evaluations of guideline quality [88].

Discussion

This paper assumes the usefulness of clinical practice guidelines and also that computer-interpretable guidelines are the best way to deliver guidelines to clinicians. We do not address in this paper how guidelines are developed; only how they are translated into practice. Although some of the CIG systems have guideline authoring components, we believe that the tools considered here should be thought of primarily as translational tools to bring established guidelines to clinical practice. It is sufficiently difficult to assess whether or not this translation is accurate and functional, without muddying the water by considering what constitutes the best guideline for a specific clinic problem. While some systems, notably GEM do bear on guideline quality, this is a retrospective consideration, after authoring of the guideline. On the other hand, the lessons learned from CIG systems have clarified the shortcoming of existing natural language text representations of clinical guidelines, including ambiguity, vagueness of recommendations and failure to use standard terminology or standard definitions of medical concepts and observations. These lessons should

be fed back to guideline developers. There is a debate about the value of the document-centric approach. Its proponents argue that the original text-based document represents the closest representation to the original intent of the developers. In addition, the original text guidelines typically include background information such as epidemiology and pathophysiology, and cite the evidence on which the guideline recommendations are based. This provides additional support for clinicians who want to know more about the basis for guideline recommendations and may wish to see the evidence for themselves. Preservation of and linking to the original guideline document *as part of the representation scheme* will provide this additional knowledge. In cases of ambiguous or vague recommendations, it also will give clinicians the ability to see how the source document was worded which may clarify the original intent.

In contrast, it has been argued that the document-centric approach leaves a gap between the “idealized document model” and specific guideline implementation in a local clinical environment [106]. Waitman and Miller estimate that 90% of the effort required for successful guideline implementation is local and only the remaining 10% involves “getting the document right” [106]. They argue, correctly that text-based guidelines are geared to average patients and may not suit an individual. Moreover, text-based guidelines (or any standardized formal guideline) fail to take into account local factors such as availability of services. Another valid criticism of a document-centric approach is that maintaining a guideline in multiple formats increases the effort involved in maintaining guidelines as they are modified over time. Encoding by way of a markup approach would have

to be largely redone from scratch when a new version of a guideline is issued. However, this criticism applies also to knowledge base-centric approaches. Approaches that link the text-based representation with the formal knowledge-based representation such as the bridge approach described by Shankar et al. [92] suggest the possibility of both representations being modified together. To this end, it seems desirable to have guideline developers work with knowledge engineers to develop CIG’s collaboratively by simultaneously developing both representations.

While Waitman and Miller’s estimate of the amount of effort required for local adaptation probably is accurate given the current state of the art, it does not seem to anticipate the likely benefits from standardization of EHR’s and the National Health Information Infrastructure (NHII [107]). The promise of these developments is that local effort is likely to decrease substantially over time. Moreover, the guideline development process itself is likely to benefit from increasing awareness of standards development efforts such that deficiencies in guideline documents such as incompleteness, vagueness, ambiguity and lack of attention to standardized vocabulary will be reduced or eliminated. Finally, it is our belief that to a large extent, differences of “local practices” from guideline recommendations are undesirable and can be minimized by application of clinical guidelines.

Computer implementation may result in CIG’s that differ from the intentions of their developers. By enforcing completeness, the implementer may introduce certainty not intended by the guideline developers. The sometimes intentional vagueness of guidelines may result from gaps in the evidence base. Some situations do not permit conclusive recommendations or cannot dis-

tinguish among competing alternatives on any rational basis. We propose that in order to represent vagueness, that CIG representations explicitly incorporate representation of evidence uncertainty. Some clinical contexts will necessarily result in “no recommendation” and these situations should be clearly indicated. Similarly, advice to apply “clinical judgment” should be represented explicitly. More complex decision techniques may be required to resolve these situations. One method is the use of decision analytic models and simulations to predict the consequences of alternative plans and to select among them using a utility-based formalism [108], as in the explicit modeling of decisions incorporated into the SAGE guideline specification [78] using “decision nodes”.

Another argument in favor of formal decision models in guideline deployment is that existing CIG formalisms address overall control of only a single guideline. Applying all applicable guidelines accurately and effectively, may have unintended consequences including burdening patients with multiple treatments, drug interactions and burdensome regimens [109]. Few practice guidelines address the effects of combining their recommendations with those of other guidelines. In addition, many of the recommendations may have little benefit to elderly patients or those with chronic illnesses. While this is an extremely difficult problem, a future goal of guideline implementation systems should include providing methods for estimating the benefits sets of guideline recommendations for a specific patient and prioritizing those that promise greatest benefit and cost-efficiency [110].

The principles learned from efforts to make guidelines computer-interpretable should be applied to future guideline

development. Standardized terminology and nomenclature should be used to ensure that guidelines reflect clinical observations as they are collected in EHR's. Where EHR's use non-standard terminologies, translation mechanisms must be provided to link the EHR terminology with that used in guidelines. Effort must also go into engineering clinical work flow to collect data that is in a useful format for decision support. For example, a guideline that is based on knowing whether a patient has left ventricular hypertrophy [55] will not be supported by an EHR that has only a digitized image of an electrocardiogram (EKG), or even a standard echocardiographic or EKG report stating "left ventricular hypertrophy" unless that observation is stored in a variable that is linked to the guideline logic. This disconnect has been referred to as the "impedance mismatch" between the EHR and decision support [97].

We disagree with the conclusions of Waitman and Miller [106] that most of guideline implementation work should address local adaptation. Clearly, better methods need to be developed for adapting guidelines to local practices, but we believe that working towards national standards for guidelines is useful and will improve quality of guidelines and decrease the amount of work a single institution has to do in order to employ them. Accommodating local variability into a standardized approach seems to us a much more productive endeavor than planning on extensive local customization of implementation. Nevertheless, at this time, we are a long way from having a universal clinical data model that would enable sharable guidelines to be used in multiple institutions without a significant effort at local adaptation. Efforts related to the national health information infrastructure (NHII) [107] will accelerate movement

toward a universal clinical data model that can be adopted as a standard. Indeed, as pointed out by Jenders and Sailors [27], a guideline representation standard is but one component of a comprehensive information infrastructure. The VMR approach referencing a virtual medical record [66] that can be mapped locally to EHR systems acknowledges that for the foreseeable future, clinical information systems will have important proprietary and local differences. It may not yet be time to settle on a single consensus model since various research groups continue to explore the unique features of their own implementations [48]. This approach will permit important research and development work to take place while standards adoption evolves.

A number of issues have not been sufficiently addressed by research efforts to date. One such issue is smoothly linking the updating of clinical guidelines to their computer-based implementation and deployment. Another issue is establishing more formal definitions of what it means to *satisfy* a clinical guideline. Since compliance with guidelines is increasingly used to evaluate and reward clinicians, both the updating and measurement of compliance need to be incorporated into the process of care. This will require more than just sophisticated technology. Both clinicians and the management of healthcare enterprises will have to become active participants in these efforts.

The field needs more rigorous evaluations of various aspects of CIG representations. Such evaluations should be designed to distinguish clearly the effects of the guidelines themselves from the effects of computer implementation in general and from specific implementation features. In order to do this, evaluations will have to use carefully designed controls in order to isolate the

changes being studied. Thus, in order to evaluate whether a particular format of computer-based guideline presentation provides additional benefit over previous formats, a study should compare two CIG implementations, one with the feature and one without. We offer as a starting point, the identification of several steps in the guideline implementation process that should be important foci of evaluation.

1) Guideline Encoding Process

- The ease of and time required for guideline encoding.
- Skill level required – Can encoding be carried out by non-technical clinical personnel or by technical, non-clinical personnel?
- Reproducibility of encoding – Will multiple encoders produce substantially the same formal representation.
- Accuracy of encoding. – Does the encoding process faithfully capture the content of the original guideline?
- Comparative features of encoding – Which features of CIG's produce greatest improvement in compliance – e.g. Does explicit representation of overall plan structure result in better compliance with guidelines? Do embedded formal decision methods improve compliance?
- Maintenance: How easily can CIG's be updated when the original guideline changes?

2) Implementation

- Incorporation of guideline into clinical information system. How effective are guidelines built in to the EHR vs. supplied by a separate system?
- Matching of variables between clinical guidelines and clinical information systems. To what extent is the EHR able to capture variables needed to execute the guideline? To what extent is actual workflow capturing those variables?

3) Evaluation of the CIG System

- What is the best form for CIG output? Does a simple list of recommendations, a set of recommendations with available detailed explanations, or automatic construction of test and treatment orders achieve the best results?
- Accuracy of the output. Do the outputs of the CIG accurately reflect the knowledge in the original guideline for all possible clinical scenarios? To this end, we recommend establishment of a set of *reference cases*, guidelines for which a correct set of outputs is agreed upon by a group of clinical experts. The output of an experimental system can then be compared to the reference set. The reference cases should cover a range from simple guidelines (e.g. when to perform mammograms) to complex disease-management guidelines (e.g. asthma management).

In the future, it will be important to carry out comparative evaluations of alternative CIG systems, as most studies to date have compared new systems with traditional non-system approaches. Finally, we encourage continued efforts to assess the quality of published clinical guidelines according to accepted criteria such as the COGS statement [17] or the AGREE instrument [16] and by newer criteria that address implementability of guidelines [111]. It is to be hoped that in a manner similar to that in which the conduct and reporting of clinical trials in the medical literature has been reshaped by rigorous standards for reporting, that the quality and implementability of clinical guideline development and publication will be improved.

References

1. Field MJ, Lohr KN, Institute of MCcCP, Guidelines. Guidelines for clinical practice from development to use. Washington, D.C: National Academy Press; 1992.
2. Eddy DM. Clinical decision making: from theory to practice. Practice policies—guidelines for methods. JAMA Apr 4 1990; 263(13): 1839-41.
3. Leape LL, Park RE, Solomon DH, Chassin MR, Koseoff J, Brook RH. Does inappropriate use explain small-area variations in the use of health care services? JAMA Feb 2 1990; 263(5): 669-72.
4. Ellrodt G, Cook DJ, Lee J, Cho M, Hunt D, Weingarten S. Evidence-based disease management. JAMA Nov 26 1997; 278(20): 1687-92.
5. Woolf SH. Practice guidelines: a new reality in medicine. I. Recent developments. Arch Intern Med 1990; 150(9): 1811.
6. Epstein A. Performance reports on quality—prototypes, problems, and prospects. N Engl J Med 1995; 333(1): 57.
7. Joint Commission on Accreditation of Healthcare Organizations: Specification Manual for National Hospital Quality Measures (2005); 2005.
8. Bodenheimer T. The American health care system—the movement for improved quality in health care. N Engl J Med Feb 11 1999; 340(6): 488-92.
9. Julian DG. Translation of clinical trials into clinical practice. J Intern Med Mar 2004; 255(3): 309-16.
10. McGlynn EA, Asch SM, Adams J, et al. The quality of health care delivered to adults in the United States. N Engl J Med 2003; 348(26): 2635-45.
11. Schuster MA, McGlynn EA, Brook RH. How good is the quality of health care in the United States? Milbank Q 1998; 76(4): 517-63, 509.
12. Grol R. Successes and failures in the implementation of evidence-based guidelines for clinical practice. Med Care Aug 2001; 39(8 Suppl 2): II46-54.
13. Schuster MA, McGlynn EA, Brook RH. How good is the quality of health care in the United States? 1998. Milbank Q 2005; 83(4): 843-95.
14. Brook RH. Practice guidelines and practicing medicine. Are they compatible? JAMA 1989; 262(21): 3027.
15. Graham ID, Calder LA, Hebert PC, Carter AO, Tetroe JM. A comparison of clinical practice guideline appraisal instruments. Int J Technol Assess Health Care Autumn 2000; 16(4): 1024-38.
16. Cluzeau F. Appraisal of Guidelines for Research & Evaluation; The AGREE Instrument. Vol 2006; 2001.
17. Shiffman RN, Shekelle P, Overhage JM, Slutsky J, Grimshaw J, Deshpande AM. Standardized reporting of clinical practice guidelines: a proposal from the Conference on Guideline Standardization. Ann Intern Med Sep 16 2003; 139(6): 493-8.
18. Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. Lancet Nov 27 1993; 342(8883): 1317-22.
19. Worrall G, Chaulk P, Freake D. The effects of clinical practice guidelines on patient outcomes in primary care: a systematic review. CMAJ 1997; 156(12): 1705-12.
20. Shiffman RN, Liaw Y, Brandt CA, Corb GJ. Computer-based guideline implementation systems: a systematic review of functionality and effectiveness. J Am Med Inform Assoc 1999; 6(2): 104-14.
21. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. BMJ Feb 20 1999; 318(7182): 527-30.
22. Andrews EJ, Redmond HP. A review of clinical guidelines. Br J Surg Aug 2004; 91(8): 956-64.
23. Lomas J, Anderson GM, Domnick-Pierre K, Vayda E, Enkin MW, Hannah WJ. Do practice guidelines guide practice? The effect of a consensus statement on the practice of physicians. N Engl J Med Nov 9 1989; 321(19): 1306-11.
24. Tunis SR, Hayward RS, Wilson MC, Rubin HR, Bass EB, Johnston M, et al. Internists' attitudes about clinical practice guidelines. Ann Intern Med Jun 1 1994; 120(11): 956-63.
25. Ellrodt AG, Conner L, Riedinger M, Weingarten S. Measuring and improving physician compliance with clinical practice guidelines. A controlled interventional trial. Ann Intern Med 1995; 122(4): 277-82.
26. Cabana MD, Rand CS, Powe NR, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. JAMA 1999; 282(15): 1458-65.
27. Jenders RA, Sailors RM. Convergence on a standard for representing clinical guidelines: work in health level seven. Medinfo 2004; 11(Pt 1): 130-4.
28. Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). JAMA 2001; 285(19): 2486-97.
29. Grundy SM, Cleeman JI, Merz CN, Brewer HB Jr, Clark LT, Hunninghake DB, et al. Implications of recent clinical trials for the National Cholesterol Education Program Adult Treatment Panel III guidelines. Circulation 2004; 110(2): 227-39.
30. Tierney WM, Hui SL, McDonald CJ. Delayed feedback of physician performance versus immediate reminders to perform preventive care. Effects on physician compliance. Med Care Aug 1986; 24(8): 659-66.
31. Weingarten SR, Riedinger MS, Conner L, et al. Practice guidelines and reminders to reduce duration of hospital stay for patients with chest pain. An interventional trial. Ann Intern Med Feb 15 1994; 120(4): 257-63.
32. Safran C, Rind DM, Davis RB, et al. Guidelines for management of HIV infection with computer-based patient's record. Lancet Aug 5 1995; 346(8971): 341-6.
33. Zielstorff RD. Online practice guidelines: issues, obstacles, and future prospects. J Am Med Inform Assoc May-Jun 1998; 5(3): 227-36.
34. Subramanian U, Fihn SD, Weinberger M, Plue L, Smith FE, Udris EM, et al. A controlled trial of including symptom data in computer-based care suggestions for managing patients with chronic heart failure. Am J Med Mar 15 2004; 116(6): 375-84.
35. Vissers MC, Biert J, van der Linden CJ, Hasman A. Effects of a supportive protocol processing system (ProtoVIEW) on clinical behaviour of residents in the accident and emergency department. Comput Methods Programs Biomed 1996; 49(2): 177-84.
36. Rossi RA, Every NR. A computerized intervention to decrease the use of calcium channel blockers in hypertension. J Gen Intern Med 1997; 12(11): 672-8.
37. Schriger DL, Baraff LJ, Rogers WH, Cretin S.

- Implementation of clinical guidelines using a computer charting system. Effect on the initial care of health care workers exposed to body fluids. *JAMA* Nov 19 1997; 278(19): 1585-90.
38. Rollman BL, Gilbert T, Lowe HJ, Kapoor WN, Schulberg HC. The electronic medical record: its role in disseminating depression guidelines in primary care practice. *Int J Psychiatry Med* 1999; 29(3): 267-86.
 39. Cannon DS, Allen SN. A comparison of the effects of computer and manual reminders on compliance with a mental health clinical practice guideline. *J Am Med Inform Assoc* 2000; 7(2): 196-203.
 40. Schriger DL, Baraff LJ, Buller K, et al. Implementation of clinical guidelines via a computer charting system: effect on the care of febrile children less than three years of age. *J Am Med Inform Assoc* Mar-Apr 2000; 7(2): 186-95.
 41. Rollman BL, Hanusa BH, Gilbert T, Lowe HJ, Kapoor WN, Schulberg HC. The electronic medical record. A randomized trial of its impact on primary care physicians' initial management of major depression [corrected]. *Arch Intern Med* Jan 22 2001; 161(2): 189-97.
 42. Ornstein SM. Translating research into practice using electronic medical records the PPRNet-TRIP project: primary and secondary prevention of coronary heart disease and stroke. *Top Health Inf Manage* Nov 2001; 22(2): 52-8.
 43. Maviglia SM, Teich JM, Fiskio J, Bates DW. Using an electronic medical record to identify opportunities to improve compliance with cholesterol guidelines. *J Gen Intern Med* Aug 2001; 16(8): 531-7.
 44. Tierney WM, Overhage JM, Murray MD, Harris LE, Zhou XH, Eckert GJ, et al. Effects of computerized guidelines for managing heart disease in primary care. *J Gen Intern Med* 2003; 18(12): 967-76.
 45. Montgomery AA, Fahey T, Peters TJ, MacIntosh C, Sharp DJ. Evaluation of computer based clinical decision support system and risk chart for management of hypertension in primary care: randomised controlled trial. *BMJ* Mar 11 2000; 320(7236): 686-90.
 46. Ansari M, Shlipak MG, Heidenreich PA, et al. Improving guideline adherence: a randomized trial evaluating strategies to increase beta-blocker use in heart failure. *Circulation* Jun 10 2003; 107(22): 2799-804.
 47. Montgomery AA, Peters TJ, Fahey T. Reasons physicians do not follow clinical practice guidelines. *JAMA* 2000; 283(13): 1685; author reply 1686.
 48. Peleg M, Tu S, Bury J, Ciccarese P, Fox J, Greenes RA, et al. Comparing computer-interpretable guideline models: a case-study approach. *J Am Med Inform Assoc* Jan-Feb 2003; 10(1): 52-68.
 49. Wetter T. Lessons learnt from bringing knowledge-based decision support into routine use. *Artif Intell Med* Mar 2002; 24(3): 195-203.
 50. Shiffman RN. Representation of clinical practice guidelines in conventional and augmented decision tables. *J Am Med Inform Assoc* Sep-Oct 1997; 4(5): 382-93.
 51. 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* May 26 1999; 281(20): 1900-5.
 52. Fretheim A, Williams JW, Jr., Oxman AD, Herrin J. The relation between methods and recommendations in clinical practice guidelines for hypertension and hyperlipidemia. *J Fam Pract* 2002; 51(11): 963-8.
 53. Sonnenberg F, Hagerty CG, Acharya J, Pickens D, Kulikowski C. Vocabulary Requirements for Implementing Clinical Guidelines in an Electronic Medical Record: A Case Study. *Proc AMIA Annu Fall Symp* 2005: 709-13.
 54. Major outcomes in high-risk hypertensive patients randomized to angiotensin-converting enzyme inhibitor or calcium channel blocker vs diuretic: The Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT). *Jama*. Dec 18 2002;288(23):2981-2997.
 55. Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL Jr, et al. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. *JAMA* 2003; 289(19): 2560-72.
 56. Codish S, Shiffman RN. A Model of Ambiguity and Vagueness in Clinical Practice Guideline Recommendations. *Proc AMIA Annu Fall Symp* 2005: 146-50.
 57. Shiffman RN, Michel G, Essaihi A, Thornquist E. Bridging the guideline implementation gap: a systematic, document-centered approach to guideline implementation. *J Am Med Inform Assoc* Sep-Oct 2004; 11(5): 418-26.
 58. Wang D, Peleg M, Tu SW, Boxwala AA, Greenes RA, Patel VL, et al. Representation primitives, process models and patient data in computer-interpretable clinical practice guidelines: a literature review of guideline representation models. *Int J Med Inform* Dec 18 2002; 68(1-3): 59-70.
 59. de Clercq PA, Blom JA, Korsten HH, Hasman A. Approaches for creating computer-interpretable guidelines that facilitate decision support. *Artif Intell Med* May 2004; 31(1): 1-27.
 60. 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 Annu Symp Comput Appl Med Care*. 1991: 248-52.
 61. Hripcsak G, Wigertz OB, Kahn MG, Clayton PD, Pryor TA. ASTM E31.15 on health knowledge representation: the Arden Syntax. *Stud Health Technol Inform* 1993; 6: 105-12.
 62. Wang D. Translating Arden MLMs into GLIF guidelines—a case study of hyperkalemia patient screening. *Stud Health Technol Inform* 2004; 101: 177-81.
 63. Agrawal A, Shiffman RN. Using GEM-encoded guidelines to generate medical logic modules. *Proc AMIA Symp* 2001: 7-11.
 64. Pryor TA, Hripcsak G. Sharing MLM's: an experiment between Columbia-Presbyterian and LDS Hospital. *Proc Annu Symp Comput Appl Med Care* 1993: 399.
 65. Musen MA, Tu SW, Das AK, Shahar Y. EON: a component-based approach to automation of protocol-directed therapy. *J Am Med Inform Assoc* Nov-Dec 1996; 3(6): 367-88.
 66. Johnson PD, Tu SW, Musen MA, Purves I. A virtual medical record for guideline-based decision support. *Proc AMIA Symp* 2001: 294-8.
 67. Fox J, Johns N, Lyons C, Rahmzadeh A, Thomson R, Wilson P. PROforma: a general technology for clinical decision support systems. *Comput Methods Programs Biomed* Sep 1997; 54(1-2): 59-67.
 68. Shahar Y, Miksch S, Johnson P. The Asgaard project: a task-specific framework for the application and critiquing of time-oriented clinical guidelines. *Artif Intell Med Sep-Oct 1998; 14(1-2): 29-51.*
 69. Shahar Y, Young O, Shalom E, Galperin M, Mayaffit A, Moskovitch R, et al. A framework for a distributed, hybrid, multiple-ontology clinical-guideline library, and automated guideline-support tools. *J Biomed Inform* 2004; 37(5): 325-44.
 70. Quaglini S, Stefanelli M, Lanzola G, Caporusso V, Panzarasa S. Flexible guideline-based patient careflow systems. *Artif Intell Med* Apr 2001; 22(1): 65-80.
 71. Quaglini S, Stefanelli M, Cavallini A, Micieli G, Fassino C, Mossa C. Guideline-based careflow systems. *Artif Intell Med*. Aug 2000; 20(1): 5-22.
 72. Ohno-Machado L, Gennari JH, Murphy SN, Jain NL, Tu SW, Oliver DE, et al. The guideline interchange format: a model for representing guidelines. *J Am Med Inform Assoc* 1998; 5(4): 357.
 73. Patel VL, Allen VG, Arocha JF, Shortliffe EH. Representing clinical guidelines in GLIF: individual and collaborative expertise. *J Am Med Inform Assoc* Sep-Oct 1998; 5(5): 467-83.
 74. Boxwala AA, Peleg M, Tu S, Ogunyemi O, Zeng QT, Wang D, Patel VL, et al. GLIF3: a representation format for sharable computer-interpretable clinical practice guidelines. *J Biomed Inform* Jun 2004; 37(3): 147-61.
 75. Brickley DaGRV. Resource Description Framework (RDF) Schema Specification 1.0. W3C Candidate Recommendation. Boston: World Wide Web Consortium; 2000.
 76. Sordo M, Ogunyemi O, Boxwala AA, Greenes RA. GELLO: an object-oriented query and expression language for clinical decision support. *AMIA Annu Symp Proc* 2003: 1012.
 77. Wang D, Shortliffe EH. GLEE—a model-driven execution system for computer-based implementation of clinical practice guidelines. *Proc AMIA Symp* 2002: 855-9.
 78. Campbell J, Tu S, Mansfield JG, et al. The SAGE Guideline Model: A Knowledge Representation Framework for Encoding Interoperable Clinical Practice Guidelines SMI Technical Report 2003-0962; 2003.
 79. Tu SW, Campbell JR, Musen MA. SAGE Guideline Modeling: Motivation and Methodology Paper presented at: Computer-Based Support for Clinical Guidelines and Protocols: Proceedings of the Symposium on Computerized Guidelines and Protocols; 2004.
 80. Goldstein MK, Hoffman BB, Coleman RW, Musen MA, Tu SW, Advani A, et al. Implementing clinical practice guidelines while taking account of changing evidence: ATHENA DSS, an easily modifiable decision-support system for managing hypertension in primary care. *Proc AMIA Symp* 2000: 300-4.
 81. Goldstein MK, Coleman RW, Tu SW, Shankar RD, O'Connor MJ, Musen MA, et al. Translating

- research into practice: organizational issues in implementing automated decision support for hypertension in three medical centers. *J Am Med Inform Assoc* Sep-Oct 2004; 11(5): 368-76.
82. Tu SM MA, Shankara R, Campbell J, Hrabak K, McClayb J, Huffc SM et al. Modeling Guidelines for Integration into Clinical Workflow. *Proc Medinfo* 2004:174-8.
 83. Herbert SI, Gordon CJ, Jackson-Smale A, Salis JL. Protocols for clinical care. *Comput Methods Programs Biomed* Sep-Oct 1995; 48(1-2): 21-6.
 84. De Clercq PA, Blom JA, Hasman A, Korsten HH. GASTON: an architecture for the acquisition and execution of clinical guideline-application tasks. *Med Inform Internet Med.* Oct-Dec 2000; 25(4): 247-63.
 85. Votruba P, Miksch S, Kosara R. Facilitating knowledge maintenance of clinical guidelines and protocols. *Medinfo* 2004; 11(Pt 1): 57-61.
 86. Shiffman RN, Agrawal A, Deshpande AM, Gershkovich P. An approach to guideline implementation with GEM. *Medinfo.* 2001; 10(Pt 1): 271-5.
 87. Shiffman RN, Karras BT, Agrawal A, Chen R, Marengo L, Nath S. GEM: a proposal for a more comprehensive guideline document model using XML. *J Am Med Inform Assoc* 2000; 7(5): 488-98.
 88. Shiffman RN, Michel G. Toward improved guideline quality: using the COGS statement with GEM. *Medinfo* 2004; 11(Pt 1): 159-63.
 89. Tang PC, Young CY. ActiveGuidelines: integrating Web-based guidelines with computer-based patient records. *Proc AMIA Symp* 2000: 843-7.
 90. Hagerty CG, Pickens D, Kulikowski C, Sonnenberg F. HGML: a hypertext guideline markup language. *Proc AMIA Symp* 2000: 325-9.
 91. Eriksson HT, SW; Musen, M. Semantic Clinical Guideline Documents. *Proc AMIA Annu Fall Symp* 2005: 236-40.
 92. Shankar RD, Tu SW, Martins SB, Fagan LM, Goldstein MK, Musen MA. Integration of textual guideline documents with formal guideline knowledge bases. *Proc AMIA Symp* 2001: 617-21.
 93. Hagerty CC J, Pickens DS, Kulikowski CA, Sonnenberg FA. Semi-Automated Encoding of Guidelines. *Proc Medinfo* 2004: 1625.
 94. Georg G, Jaulent M. An Environment for Document Engineering of Clinical Guidelines. *AMIA Annu Symp Proc* 2005: 276-80.
 95. HL HL 7 Reference Information Model. Web Page] http://www.hl7.org/Library/data-model/RIM/modelpage_mem.htm. Accessed 1/12/06.
 96. Jenders RA, Sujansky W, Broverman CA, Chadwick M. Towards improved knowledge sharing: assessment of the HL7 Reference Information Model to support medical logic module queries. *Proc AMIA Annu Fall Symp* 1997: 308-12.
 97. Schadow G, Russler DC, Mead CN, McDonald CJ. Integrating medical information and knowledge in the HL7 RIM. *Proc AMIA Symp* 2000: 764-8.
 98. Johnson PD, Tu S, Booth N, Sugden B, Purves IN. Using scenarios in chronic disease management guidelines for primary care. *Proc AMIA Symp* 2000: 389-93.
 99. National Heart LaBI. Expert Panel Report 2: guidelines for the diagnosis and management of asthma. SOURCE(S): U.S. Department of Health and Human Services; 1997.
 100. Mandell LA, Bartlett JG, Dowell SF, File TM, Jr., Musher DM, Whitney C. Update of practice guidelines for the management of community-acquired pneumonia in immunocompetent adults. *Clin Infect Dis* Dec 1 2003; 37(11): 1405-33.
 101. Sanders GD, Nease RF, Jr., Owens DK. Publishing web-based guidelines using interactive decision models. *J Eval Clin Pract* May 2001; 7(2): 175-89.
 102. Lindberg DA, Humphreys BL, McCray AT. The Unified Medical Language System. *Methods Inf Med* 1993; 32(4): 281.
 103. Wang AY, Sable JH, Spackman KA. The SNOMED clinical terms development process: refinement and analysis of content. *Proc AMIA Symp* 2002: 845-9.
 104. McDonald CJ, Huff SM, Suico JG, et al. LOINC, a universal standard for identifying laboratory observations: a 5-year update. *Clin Chem* Apr 2003; 49(4): 624-33.
 105. Peleg M, Boxwala AA, Bernstam E, Tu S, Greenes RA, Shortliffe EH. Sharable representation of clinical guidelines in GLIF: relationship to the Arden Syntax. *J Biomed Inform* Jun 2001; 34(3): 170-81.
 106. Waitman LR, Miller RA. Pragmatics of implementing guidelines on the front lines. *J Am Med Inform Assoc* Sep-Oct 2004; 11(5): 436-8.
 107. Stead WW, Kelly BJ, Kolodner RM. Achievable steps toward building a National Health Information infrastructure in the United States. *J Am Med Inform Assoc* Mar-Apr 2005; 12(2): 113-20.
 108. Matchar DB, McCormick Ka MSRA. Application of Decision Analysis to Guideline Development. *Clinical Practice Guideline Development: Methodology Perspectives*; Agency for Health Care Policy and Research Publication No. 95-00009: U.S. Department of Health and Human Services; 1995.
 109. Boyd CM, Darer J, Boulton C, Fried LP, Boulton L, Wu AW. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. *JAMA* Aug 10 2005; 294(6): 716-24.
 110. O'Connor PJ. Adding value to evidence-based clinical guidelines. *JAMA* Aug 10 2005; 294(6):741-3.
 111. Shiffman RN, Dixon J, Brandt C, Essaihi A, Hsiao A, Michel G, et al. The GuideLine Implementability Appraisal (GLIA): development of an instrument to identify obstacles to guideline implementation. *BMC Med Inform Decis Mak* Jul 27 2005; 5: 23.
 112. Sutton DR, Fox J. The syntax and semantics of the PROforma guideline modeling language. *J Am Med Inform Assoc* 2003; 10(5): 433-43.
 113. Ciccarese P, Caffi E, Quaglini S, Stefanelli M. Architectures and tools for innovative Health Information Systems: the Guide Project. *Int J Med Inform* Aug 2005; 74(7-8): 553-62.

Correspondence to:
 Frank A. Sonnenberg
 Division of General and Internal Medicine
 Robert Wood Johnson Medical School
 Clinical Academic Building
 125 Paterson St.
 New Brunswick, NJ 08901-1977
 USA
 E-mail: sonnenbe@umdnj.edu