

# A Systematic and Replicable Approach to Knowledge Formalization

Richard Shiffman, George Michel, Nitu Kashyap, Mark Dixon  
Yale Center for Medical Informatics, New Haven, CT

## ABSTRACT

Transforming knowledge from clinical narrative into executable decision support has been challenging. We describe a 4 stage model and set of tools that support a systematic and replicable formalization process. BRIDGE-Wiz captures knowledge directly from clinical experts and facilitates development of clear, transparent, and implementable recommendations. GLIA (and its electronic version, eGLIA) help to identify obstacles to successful implementation. GEM provides a hierarchical model of guideline document that serves as the backbone for tool development. GEM Cutter facilitates markup and capture of information contained in textual guidelines. EXTRACTOR provides a variety of displays of guideline content. The transition from semiformal to formal executable code is multifaceted requiring careful consideration of local workflow. Proceeding methodically from narrative through semi-structured and semi-formal forms to a formal computable format preserves a record that helps to assure validity of the final product.

## BACKGROUND

Knowledge formalization is a process that takes a natural language knowledge source (such as clinical guidelines and other narrative reports) written for clinician-experts and transforms it into a symbolic representation that allows for automated interpretation by computers. For many organizations, the process of knowledge transformation is ad hoc despite the efforts of many investigators to define a standardized procedure, including prominently (but not limited to) the SAGE [1], Protocure [2], and DeGel [3] projects. Funded by the US Agency for Healthcare Research and Quality, the GLIDES Project (GuideLines Into Decision Support) is a collaboration of guideline developers, disseminators, and implementers working to design, develop, implement, and demonstrate clinical decision support (CDS) using systematic and replicable processes for knowledge transformation and CDS design. Since its beginnings in 2008, the multi-site collaborators have worked to develop a suite of tools to facilitate the acquisition of evidence-based knowledge and deliver it in effective computer-mediated decision support.

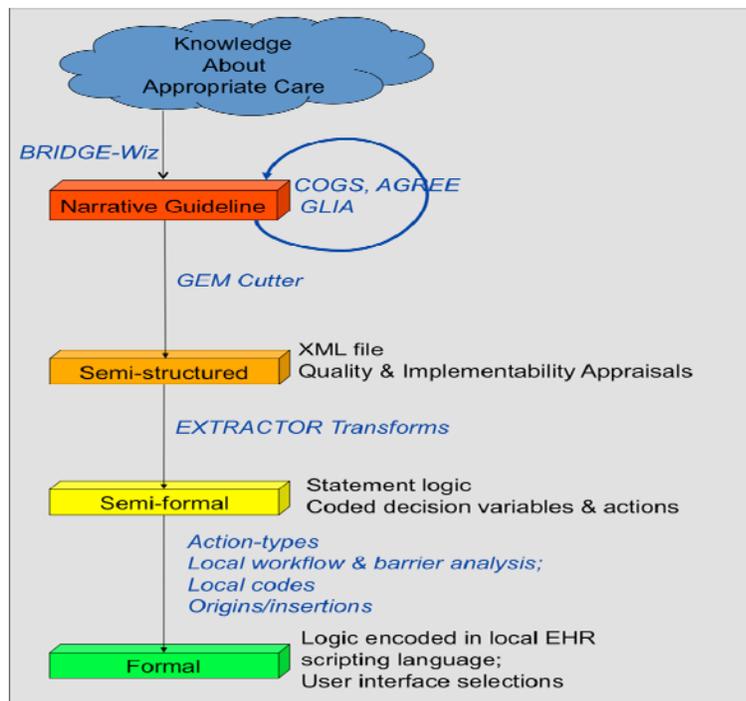


Figure 1. Knowledge formalization is a four-stage process.

GLIDES investigators envision a document-centered knowledge formalization process [4] that proceeds through several steps. (Figure 1):

- Capture of raw knowledge (original research reports, meta-analyses and systematic reviews)
- Synthesis of that knowledge with expert experience and judgments about benefits and harms to yield evidence-based, transparent, *clinical practice guidelines*
- Parsing that knowledge into a *semi-structured XML* representation maintaining clear linkages to the original text
- Transforming the parsed XML into *semi-formal* statement logic with standardized terminology
- Embedding that statement logic into a *formal* encoded form within electronic health record (EHR) systems incorporating locally prescribed meta-information

At each stage, we have devised tools and procedures intended to systematize the process and make it reusable at multiple sites. In this paper, we will describe tools that are available for use at each stage to assist guideline developers and CDS implementers.

### BRIDGE-Wiz

BRIDGE-Wiz (Building Recommendations in a Developer's Guideline Editor) is a software assistant for use by guideline authors that is intended to improve clarity, transparency, and implementability. The program is useful to a panel of guideline authors that has completed a literature review, but has not yet propounded a guideline's key action statements. BRIDGE-Wiz sequentially leads a panel of guideline developers through a series of questions intended to populate a model of guideline recommendations. The model requires explicit attention to:

- WHEN –i.e., under what circumstances
- WHO –in the guideline's pre-defined Intended Audience
- OUGHT –with what level of obligation
- To Do WHAT
- To WHOM –in the guideline's predefined Target Population
- HOW and
- WHY

What

What type of activity do you propose? EDUCATE/COUNSEL  NOT

Based on the EDUCATE/COUNSEL action type, recommend

select a verb:

Recommend what? diet and exercise

	VERB	WHAT			
	start	metformin as first-line treatment	AND	OR	DEL
AND	recommend	diet and exercise	AND	OR	DEL

Figure 2. Defining recommended actions in BRIDGE-Wiz.

BRIDGE-Wiz makes use of controlled natural language to focus the authors on clear statements of proposed action [5]. A controlled list of transitive verbs is offered based on a classification of proposed action-type (see Figure 2). BRIDGE-Wiz incorporates tests of decidability and executability of the evolving guideline recommendation. Once a clear action statement is defined, the developer panel is next led through documentation of anticipated benefits, risks, harms, and costs that may be anticipated if the

recommendation is followed and judgments about whether there is equilibrium between benefits and harms or a preponderance of one or the other. BRIDGE-Wiz output consists of a natural language recommendation easily translated into IF...THEN format and an evidence profile that summarizes and documents the process of recommendation building [6]. BRIDGE-Wiz has been piloted successfully in 5 guideline development efforts at national professional societies.

## GLIA

The GuideLine Implementability Appraisal is an instrument designed to identify obstacles to successful implementation of guideline recommendations [7]. It focuses on intrinsic impediments to implementability, i.e., those within the purview of the guideline authors. In addition to several items applicable to the guideline as a whole, GLIA explores nine dimensions of implementability that are related to individual guideline recommendations:

- Executability (exactly what to do)
- Decidability (precisely under what conditions, e.g., age, gender, clinical findings, laboratory results, to do something)
- Validity (the degree to which a recommendation reflects the intent of the developer and the strength of evidence)
- Flexibility (the degree to which a recommendation permits interpretation and allows for alternatives in its execution)
- Effect on the process of care (the degree to which a recommendation impacts upon the usual workflow in a typical care setting)
- Measurability (the degree to which the guideline identifies markers or endpoints to track the effects of implementation of a recommendation)
- Novelty/innovation (the degree to which a recommendation proposes behaviors considered unconventional by clinicians or patients)
- Computability (the ease with which a recommendation can be operationalized in an electronic information system; it is only applicable when an electronic implementation is planned)

Recommendations that do not pass criteria for decidability and executability will not be implementable until these deficiencies are addressed. The GLIA instrument was recently revised based on user’s critiques and has been published as GLIA v2.0 (available at GEM.med.yale.edu/glia).

**Project:** Project Test

**Recommendation rec 3:** With term premature rupture of membranes (PROM), labor may be induced at the time of presentation or patients may be observed for up to 24 to 72 hours for the onset of spontaneous labor.

---

**Effect On Process Of Care - the degree to which the recommendation impacts upon the usual workflow of a care setting**

Question	Appraisal	Comments
<b>20</b> Can the recommendation be carried out without substantial disruption in current workflow?	<input type="radio"/> Yes <input checked="" type="radio"/> No <input type="radio"/> NA <input type="radio"/> ? <input type="radio"/> Reset	<div style="border: 1px solid gray; padding: 5px; min-height: 30px;">This is a comment</div>
<b>21</b> Can the recommendation be pilot tested without substantial resource commitment? <i>For example, buying and installing expensive equipment to comply with a recommendation is not easily reversible.</i>	<input checked="" type="radio"/> Yes <input type="radio"/> No <input type="radio"/> NA <input type="radio"/> ? <input type="radio"/> Reset	<div style="border: 1px solid gray; padding: 5px; min-height: 30px;"></div>

Previous Recommendation  
Previous Dimension    Main Menu    Next Dimension  
At last recommendation

Figure 3. eGLIA appraisal screen showing a recommendation and potential responses.

## eGLIA

eGLIA is a web-based version of the GuideLine Implementability Appraisal instrument. eGLIA facilitates the appraisal and reporting of potential obstacles to successful guideline implementation. Guideline

appraisers can work efficiently and asynchronously to complete their appraisals online (see Figure 3). The program then helps to focus discussion on areas of appraiser disagreement, thus increasing the efficiency of the appraisal process (see Figure 4). eGLIA provides reports with several levels of detail that may be useful to a guideline development team prior to its final publication or to an implementation team looking to select a guideline or to understand a priori the obstacles to implementation that are likely to be faced.

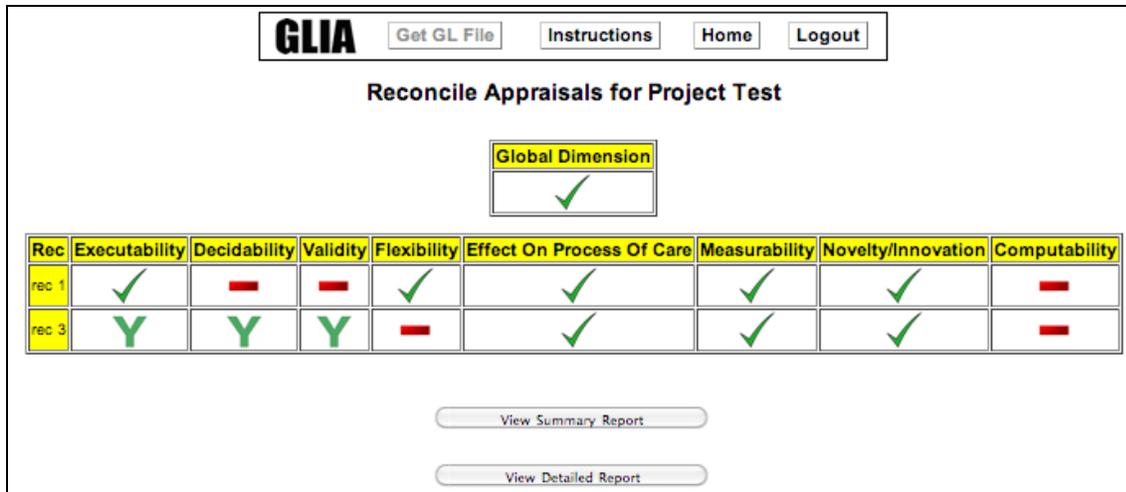


Figure 4. Summary screen for reconciling appraisals where "Y" indicate items on which all appraisers agreed, dashes indicate items on which appraisers disagreed, and checkmarks indicate items on which appraisers initially disagreed but reconciled after discussion. Clicking on cells in this matrix takes the user to a screen that displays details about individual appraisals.

## GEM

The Guideline Elements Model forms the backbone for the formalization process. GEM is an XML-based knowledge model for guideline documents. It incorporates a set of more than 100 tags to categorize guideline content. Initially published as a Document Type Definition (DTD) it was successfully balloted as a standard by ASTM in 2002. In 2006, the model was updated as GEM II, published as an XML schema, and again accepted as a standard (E-2210-06).

GEM is a hierarchical model whose high-level elements include concepts related to a guideline's identity, purpose, developer, method of development, intended audience, target population, knowledge components, testing, revision, and implementation plan. Subtrees of these high level elements specify more detailed characteristics of guideline components. As an XML model, a GEM-ified guideline document can be read by humans but processed in multiple ways by machines (see below).

Knowledge components represent the richest subtree with elements for recommendations, definitions, algorithms, background information, and research agenda. Recommendations—the primary content that distinguishes guidelines from other publications—are further divided into conditionals and imperatives and these elements further subsume subelements relating (inter alia) to a recommendation's decision variables (conditions), actions, reason, evidence quality, and recommendation strength. To allow flexibility of the model, each element can have 0 to many instances. For a graphical overview of the hierarchy, see <http://gem.med.yale.edu/Hierarchy/TopLevel.html>

## GEM Cutter

GEM Cutter is an XML editor specialized for transforming narrative guideline text into a semi-structured GEM file using an intuitive cut and paste metaphor. This tool can be used to mark up a free text guideline and transform it into a semi-structured XML format.

GEM Cutter's main display screen can be divided into 3 columns (see Figure 5). In the leftmost column, users import the text of a guideline in rtf, html, plain text, or pdf formats. The middle panel contains an expandable tree view of the Guideline Elements Model. The rightmost panel includes an area in which text can be edited and a panel in which the program displays the definition of selected elements. Users select guideline text, identify where it belongs in GEM, and click the right arrow, at which point the text appears in the tree view and populates the appropriate element in the XML file. Other buttons allow for replication and deletion of elements and subtrees.

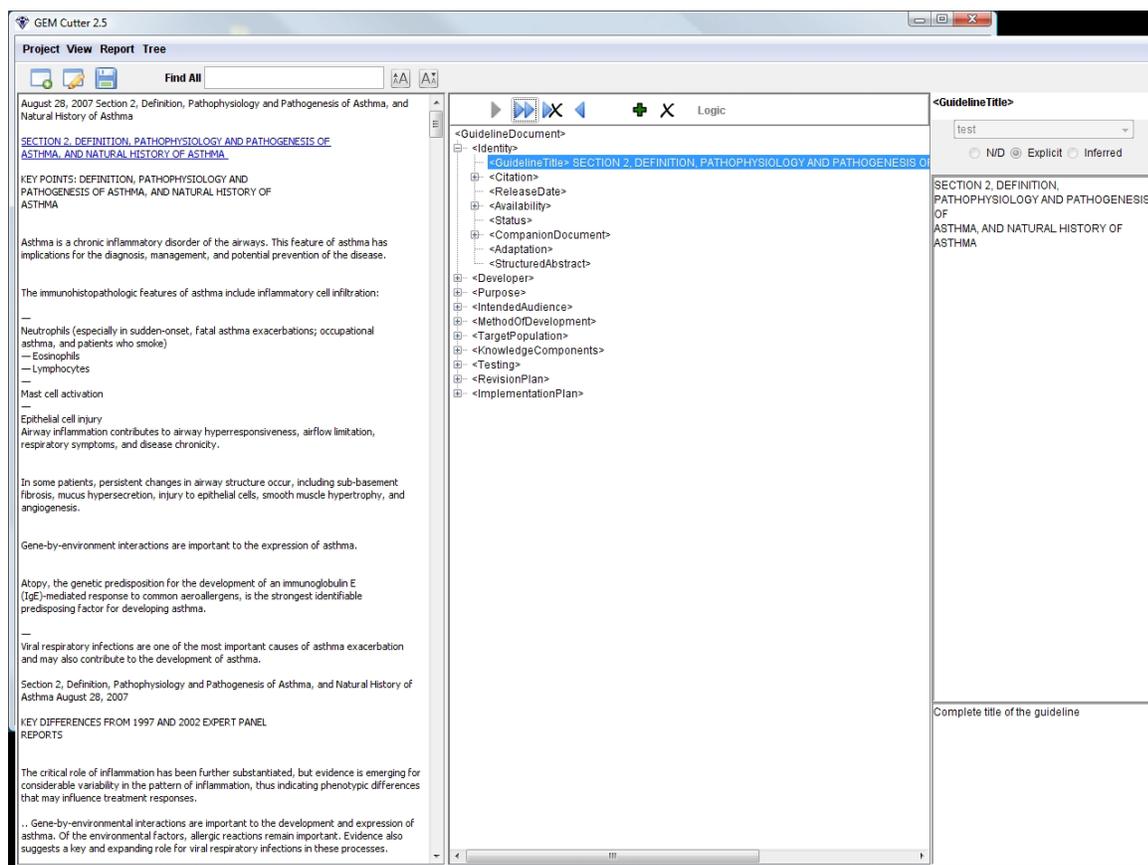


Figure 5. GEM Cutter's 3 panels.

GEM Cutter permits users to visualize an XML file as it is being created. In addition, GEM Cutter can dynamically generate a number of reports. Users can select a GEM-COGS report in which the guideline text that satisfies guideline quality criteria determined by the Conference On Guideline Standardization are displayed [8]. Finally, GEM Cutter can display EXTRACTOR reports that display a variety of views of the guideline logic and relevant knowledge components (see below).

## EXTRACTOR

EXTRACTOR is a suite of XSL (eXtensible Stylesheet Language) transforms that permit a GEM file to be manipulated and displayed in a variety of ways. Perhaps most valuable is a view of guideline knowledge components organized by recommendation and showing title, guideline narrative text, an IF...THEN statement of the guideline recommendation with each condition (decision variable) and each action parsed. In addition, fields are available to display the guideline text that justifies the recommendation (reason), the aggregate evidence quality supporting the recommendation and the guideline authors' view of the recommendation strength. This semi-formal view of the recommendation is ready for coding of the recommendation components into selected controlled vocabularies.

## RECOMMENDATION: STATEMENT 1. DIAGNOSIS

**Conditional:** Clinicians should diagnose hoarseness (dysphonia) in a patient with altered voice quality, pitch, loudness, or vocal effort that impairs communication or reduces voice-related QOL.

**Decision Variable:** patient with altered voice quality  
**Decision Variable:** patient with altered pitch  
**Decision Variable:** patient with altered loudness  
**Decision Variable:** patient with altered vocal effort  
**Decision Variable:** impairs communication  
**Decision Variable:** reduces voice-related quality of life (QOL).

**Action:** Clinicians should diagnose hoarseness (dysphonia)  
**Benefit:** Identify patients who may benefit from treatment or from further investigation to identify underlying conditions that may be serious, promote prompt recognition and treatment, and discourage the perception of hoarseness as a trivial condition that does not warrant attention  
**Risk/Harm:** Potential anxiety related to diagnosis

**Evidence Quality:** Aggregate evidence quality: Grade C, observational studies for symptoms with one systematic review of QOL in voice disorders and two systematic reviews on medication side effectsC: Observational studies (case-control and cohort design)

**Recommendation Strength:** Recommendation based on observational studies with a preponderance of benefit over harm.Recommendation: A recommendation means the benefits exceed the harms (or that the harms exceed the benefits, in the case of a negative recommendation), but the quality of evidence is not as strong (Grade B or C\*). In some clearly identified circumstances, recommendations may be made based on lesser evidence when high-quality evidence is impossible to obtain and the anticipated benefits outweigh the harms.Clinicians should also generally follow a recommendation, but should remain alert to new information and sensitive to patient preferences.

**Logic:** If (patient with altered voice quality OR patient with altered pitch OR patient with altered loudness OR patient with altered vocal effort ) AND (impairs communication OR reduces voice-related quality of life (QOL). ) Then Clinicians should diagnose hoarseness (dysphonia)

Figure 6. EXTRACTOR view of a recommendation.

Another EXTRACTOR view of guideline recommendations simply lists decision variables or actions divorced from the supporting recommendation text. In the absence of context, ambiguity, vagueness and underspecification are amplified providing guideline authors and implementers an opportunity to clarify the statements. These views also offer an automated summary of concepts to be encoded and prescribed activities for which implementations may be planned.

### From Semiformal to Formal Representation

Transitioning from recommendations expressed in statement logic—with conditions and actions encoded in structured vocabularies—to functional decision support is a complex, multifaceted process. Although several groups have offered guidance regarding keys to successful implementation including [9-11], no single set of best practices defines the best path through implementation.

A cornerstone of decision support design is to involve end-users in the development of tools and systems they will use. As Elson has noted, just as in real estate the key is “location, location, location, in decision support design the key is “workflow, workflow, workflow” [12]. Systems that do not accommodate or effectively reengineer workflow are destined to fail.

Therefore, there is a tension between the implementation considerations that can be specified centrally and those that must be devised at a local level. Central specification of all aspects of CDS design is unlikely to reflect and adapt to local requirements. Ultimately, the CDS must be operationalized in the EHR system in use for a particular set of users. Often, the final design requires accommodation to limitations in those systems [14]. Osheroff et al. have produced a handbook of valuable worksheets for developing metadata to support implementation activities [11].

Local workflow and barrier analysis is necessary to demonstrate decision support “origins,” i.e., when in the course of clinical care, each decision variable is likely to have been instantiated, and “insertions,” i.e., when in the course of clinical care it is appropriate for the decision support to appear. Similar considerations will also dictate to whom the decision support should be addressed.

In addition to local workflow considerations, CDS implementers must also integrate their knowledge transformation tools and processes with local systems development and maintenance methodologies associated with the EHR and other clinical systems. Creating a tightly-knit CDS project management methodology and systems development lifecycle that does not integrate well with the broader conventions in place for managing the EHR is unlikely to be adopted or work effectively.

Decision support can be offered in a wide variety of formats—not simply as alerts and reminders. These formats include displays of relevant information arranged so as to facilitate appropriate care, calculators of various forms that manipulate text and numbers to simplify and improve the accuracy of computation, order sets, documentation templates that prompt the user to collect appropriate information, algorithms, and infobuttons that provide context-sensitive information when requested.

In developing a decision support intervention, our group has found that classification of the action-type(s) prescribed is useful. Recognizing that there is a finite set of activities called for by guideline recommendations [13], categorizing them can lead to a consideration of a pattern of beneficial services associated with that action. For example, recommendations to “test” (from a simple blood test to an organ biopsy) might be associated with:

- Display of test indications
- Presentation of test options or alternatives
- Display of recent pertinent test results
- Display of test costs
- Assistance with test scheduling
- Providing information about specimen collection and transport
- Interpretation aids
- Result norms based on age, race, and gender
- A tickler follow up system
- Assistance with test coding (CPT, LOINC, etc).

Similar considerations assist in the design of relevant systems implementing the other action types.

Evaluation is vital component of implementation that is often overlooked and underfunded. Kawamoto [10] found that CDS improved clinical practice in 68% of trials. Since CDS is not invariably successful, knowing what works, what doesn’t work, and why is essential.

## **Discussion**

We describe a knowledge formalization model and offer a set of software tools to facilitate transformation of knowledge based on evidence in a systematic and replicable manner. Guideline developers, disseminators and implementers have used these tools to convert clinical knowledge into formats that are amenable to manipulation by computers and ultimately to influence care processes. Proceeding methodically from narrative through semi-structured and semi-formal forms to a formal computable format preserves a record that helps to assure validity of the final product.

The tools described are available from the GEM website at <http://ycmi.med.yale.edu/GEM>

## **Acknowledgements**

We are grateful for the support of the Agency for Healthcare Research and Quality for its support through HHS 290200810011 and the National Library of Medicine for its support through 2R01LM007199.

### References

1. Tu SW, Campbell JR, Glasgow J, Nyman MA, McClure R, McClay J, et al. The SAGE Guideline Model: achievements and overview. *J Am Med Inform Assoc* 2007;14(5):589-98.
2. Balsler M, Coltell O, Croonenborg Jv, Harmelen Fv, Jovell A, Lucas P, et al. Protocure: supporting the development of medical protocols through formal methods. *Stud Health Tech Inf* 2004;101:103-7.
3. Hatsek A, Young O, Shalom E, Shahar Y. DeGeL: a clinical-guidelines library and automated guideline-support tools. *Stud Health Technol Inform* 2008;139:203-12.
4. Svatek V, Ruzicka M. Step-by-step mark-up of medical guideline documents. *Int J Med Informatics* 2003;70:329-335.
5. Shiffman RN, Michel G, Krauthammer M, Fuchs NE, Kaljurand K, Kuhn T. Writing clinical practice guidelines in controlled natural language. In: Fuchs NE, editor. *Proceedings Controlled Natural Language 2009*: Springer; 2010. p. 265-280.
6. Rosenfeld R, Shiffman R. Clinical practice guideline development manual: a quality-driven approach for translating evidence into action. *Otolaryngology—Head and Neck Surgery* 2009;140:S1-S43.
7. Shiffman RN, Dixon J, Brandt C, Essaihi A, Hsiao A, Michel G, et al. The GuideLine Implementability Appraisal: development of an instrument to identify obstacles to guideline implementation. submitted 2004.
8. 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* 2003;139:493-498.
9. Bates DW, Kuperman GJ, Wang S, Gandhi T, Kittler A, Volk L, et al. Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. *J Am Med Inform Assoc* 2003;10(6):523-30.
10. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *BMJ* 2005;330:765-72.
11. Osheroff JA, Pifer EA, Teich JM, Sittig DF, Jenders RA. Improving outcomes with clinical decision support: an implementer's guide. Chicago, IL: HIMSS; 2005.
12. Elson RB, Connelly DP. Computerized decision support systems in primary care. *Primary Care* 1995;22:365-84.
13. Essaihi A, Michel G, Shiffman RN. Comprehensive categorization of guideline recommendations: creating an action palette for implementers. In: Musen MA, editor. *Proc AMIA 2003*; 2003; Washington, DC: Hanley and Belfus; 2003.
14. Wright A, Sittig DF, Ash JS, Sharma S, Pang JE, Middleton B. Clinical decision support capabilities of commercially-available clinical information systems. *J Am Med Inform Assoc* 2009;16(5):637-44.