Journal of Biomedical Informatics
doi:10.1006/jbin.2001.1002, available online at http://www.idealibrary.com on IDEAL®

METHODOLOGICAL REVIEW

Methods of Cognitive Analysis to Support the Design and Evaluation of Biomedical Systems: The Case of Clinical Practice Guidelines

Vimla L. Patel,*,† José F. Arocha,† Melissa Diermeier,† Robert A. Greenes,‡ and Edward H. Shortliffe*

**Department of Medical Informatics, College of Physicians and Surgeons, Columbia University New York, New York; †Cognitive Studies in Medicine, Centre for Medical Education, McGill University, Montreal, Quebec, Canada; and ‡Decision Systems Group, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts*

Received January 7, 2001

This article provides a theoretical and methodological framework for the use of cognitive analysis to support the representation of biomedical knowledge and the design of clinical systems, using clinical-practice guidelines (CPGs) as an example. We propose that propositional and semantic analyses, when used as part of the system-development process, can improve the validity, usability, and comprehension of the resulting biomedical applications. The framework we propose is based on a large body of research on the study of how people mentally represent information and subsequently use it for problem solving. This research encompasses many areas of psychology, but the more important ones are the study of memory and the study of comprehension. Of particular relevance is research devoted to investigating the comprehension and memory of language, expressed verbally or in text. In addition, research on how contextual variables affect performance is informative because these psychological processes are influenced by situational variables (e.g., setting, culture). One important factor limiting the acceptance and use of clinical-practice guidelines (CPGs) may be the mismatch between a guideline's recommended actions and the physician–user's mental models of what seems appropriate in a given case. Furthermore, CPGs can be semantically complex, often composed of elaborate collections of prescribed procedures with logical gaps or contradictions that can promote ambiguity and hence frustration on the part of those who attempt to use them. An improved understanding of the semantics and structure of CPGs may help to improve such matching, and ultimately the comprehensibility and usability of CPGs. Cognitive methods of analysis can help guideline designers and system builders throughout the development process, from the conceptual design of a computer-based system to its implementation phases. By studying how guideline creators and developers represent guidelines, both mentally and in text, and how end-users understand and make decisions with such guidelines, we can inform the development of technologies that seek to improve the match between the representations of experts and practitioners. We urge informaticians to recognize the potential relevance of cognitive analysis methods and to begin more extensive experimentation with the their use in biomedical informatics research.

Key Words: cognitive analysis; propositional analysis; semantic networks; comprehension; knowledge representation; clinical-practice guidelines.

1. INTRODUCTION

The capture and use of biomedical knowledge demand a rigorous analytical approach to understanding the cognitive

1532-0464/01 $35.00

processes both of human beings who possess such knowledge and of those who will wish to use the knowledge when it is delivered to them by a computer system. The classical approach to knowledge engineering has involved the interviewing of experts or the analysis of text documents, with the subsequent encoding of knowledge and ideas after they have been expressed in natural language. In this paper we argue that there is an important role for the use of cognitive-analysis methods to support this process of knowledge encoding and the design of systems that will appropriately support the mental models that users bring to their interactions with computers. We illustrate this point with examples drawn from the field of clinical-practice guidelines (CPGs). This is a domain in which important issues arise regarding both the cognitive processes used by experts as they develop and express guidelines and the mental models and constraints that characterize practitioners who may be presented with guideline-based advice by clinical information systems.

The framework we offer is based on a large body of research on the study of how people mentally represent information. This research encompasses many areas of psychology, but the more important ones are the study of memory and the study of comprehension and understanding [1, 2]. Of particular importance to our aim is research devoted to investigating comprehension and memory of text (e.g., written or verbal discourse) [3]. In addition, research on how contextual variables affect performance is informative because these psychological processes are influenced by situational variables (e.g., setting, culture). Research in comprehension has dealt with both of these aspects within the framework of knowledge-based cognition (i.e., study of cognitive processes in content domains, such as medicine, law, sports, chess, and other fields).

The process of representing verbal information begins with thoughts and ideas in the mind. This is an important point of departure because comprehension is dependent on what one already knows [4]. These thoughts are expressed externally using symbols (e.g., natural language) in some physical medium (e.g., an external representation, such as a written text). Because of the role of prior knowledge, the relationship between an internal representation and its external representation is not a simple one-to-one mapping. Typically, the external representation simplifies the intended meaning, since the background knowledge that supports such a representation cannot be expressed (it is potentially infinite). This way, many high-level inferences (e.g., those involving nontrivial interpretations, such as inferring a disease from a cluster of clinical findings), assumptions, and presuppositions are left out of the external representation. In other words, comprehension requires the prior availability of knowledge and beliefs that contextualize the way that external representations are to be interpreted [5, 6]. Furthermore, since prior knowledge varies from one person to another, their representations are likely to be different, even if they possess a common general knowledge background (e.g., medical specialty). The representation that results from combining the information in the external representation and the prior knowledge possessed by the reader is known as the "situation model" [7]. We have been using this framework for investigating various aspects of medical cognition such as diagnostic reasoning, clinical decision making, and clinical guideline understanding [8].

In this paper we describe methods of cognitive evaluation that help us to analyze the understandability and coherence of guidelines. In particular, propositional and semantic analysis methods are presented. These are two tools that have been widely used in cognitive science to evaluate texts. We propose that propositional and semantic analyses, when used as part of the guideline-development process, can be used to improve the usability and comprehension of clinical-practice guidelines by clinicians who are offered guideline-based advice. Furthermore, when used formatively during the process of guideline creation, they can help us to understand the mental processes of experts who create guidelines and, in turn, help us to optimize the design of guideline-authoring environments. Using these methods of analysis, investigations can be carried out that evaluate how difficult the guidelines are to comprehend as well as how to help to assure their practical use for screening, diagnosis, or treatment.

The methods outlined below have origins in cognitive science, a multidisciplinary field involving the study of cognition, drawing on several sciences such as psychology, linguistics, anthropology, computer science, neuroscience, and philosophy. The way that individuals perceive the external world, based on their internal representation of the relevant reality, determines their performance. Thus, internal representations are studied because they are essential in understanding processes in comprehension and problem solving. If two people represent the same information in two different ways, then their conception of the problem, and possibly their solution, will be different. In the case of clinical-practice guidelines, such differences (either among guideline users or authors) may lead to different practice decisions with unintended consequences. We submit that by studying how guideline creators and developers represent guidelines, both internally and externally, and how end-users understand and make decisions with such guidelines, we can inform the development of technologies that seek to improve the match between the representations of the guideline creators and practitioners [9]. In turn, this can help to improve the design and comprehensibility of guidelines.

2. CLINICAL-PRACTICE GUIDELINES

In this paper, we outline methods of data analysis that have proved useful in the study of text comprehension and that we believe can be extended to the study of clinical-practice guidelines. CPGs are aimed at physicians with a wide range of knowledge and experience. A desired result of using guidelines is the adoption of best practices, and decreased variability. However, past research on guidelines has found that they are not widely used [10–14]. Many reasons have been put forth to account for this situation, among them the guidelines' lack of integration into workflow. However, their perceived irrelevance to the actual practice of medicine is also a consideration. Such perception may be due to a mismatch between a guideline's recommended actions and the physician–user's opinion of what should be done for a specific patient in a particular clinical setting. This may be due to disagreement about the recommendations or may result because the focus of the guideline, i.e., the kind of patients for whom it is intended, may not precisely correspond to the actual patient under consideration. Furthermore, CPGs can be semantically complex, often composed of elaborate collections of prescribed procedures with logical gaps or contradictions that can promote ambiguity and hence frustration on the part of those who attempt to use them. Many CPGs also involve the embedding of procedures within procedures or complicated temporal or causal relations among the various steps in a procedure. An understanding of the semantics of CPGs and of the physician's interpretation of the guideline may help to improve such matching, and ultimately the comprehensibility and usability of CPGs.

3. METHODS OF ANALYSIS FOR THE REPRESENTATION OF INFORMATION IN BIOMEDICINE

Figure 1 presents the methodological steps that we discuss in this paper. Beginning at the bottom with the thoughts and ideas that an individual brings to a design or problem-solving task, people express those notions using spoken language or written text. In the case of clinical guidelines, the written documents have tended to be the basis for direct manual encoding of the logic into a computer-based representation. As is shown by the dark arrows, however, we describe in this paper a series of analytic techniques, derived from a long history of work in cognitive science and, in particular, psychology. First, we must understand the ways in which thoughts and ideas are reflected in the ways that people write or speak about what they are doing, thinking, or designing. The methods for this kind of work are described in Section 3.1 below. Next, spoken or written text is analyzed to identify and formally encode the propositions that are expressed (Section 3.2). Although such propositions could be directly translated into a computer-based encoding, we further argue for their analysis to form a conceptual representation of the mental model that is reflected in the document or expressed language. Such conceptual representations (Section 3.3) then can be used to provide additional insight and context when computer-based representations are created or when developers seek to assess the adequacy of a developing system as reflected in the thoughts, confusions, or solutions achieved by users.

These methods can be used to analyze both the creation and the use of clinical-practice guidelines since CPGs are composed of thoughts and ideas expressed in text and users can in turn express in language their understanding of such guidelines and how they might apply to a given clinical situation. During analysis, text guidelines or transcribed comments by clinicians can be segmented into clauses so that their propositions are identified and listed. A propositional analysis is used to create a basic semantic model of the text (called the *text-base*). This propositional representation is then used to create a semantic network, which shows the relationships among propositions, including local inferences (i.e., semantic ties between propositions). The semantic network is a conceptual representation of the original text, to which inferences have been added to form a more complete structure and where one can assess the coherence of the text. The process describing how this analysis is carried out is given in detail in the following sections, beginning with a brief presentation of the theoretical framework of text comprehension and how thoughts and ideas in the mind have been represented in comprehension research.

3.1. The Mental Representation of Thoughts and Ideas and the Study of Comprehension and Memory

Comprehension is an essential step in design, problem solving, and decision making, since one needs to have a firm understanding of an issue before effective decisions are made or problems are solved successfully [2]. Although

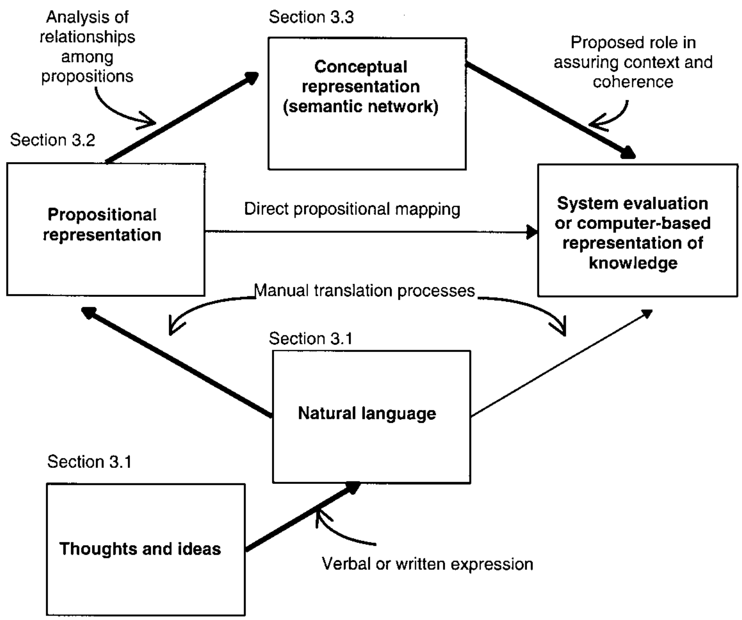

FIG. 1. An outline of the process of translation from internal representations (mental models) into natural and computer-representable languages. We propose the use of cognitive and semantic methods of analysis (thick arrows) to develop a conceptual model of an individual's knowledge or understanding. The thin arrows show typical pathways for developing computer representations, uninformed by cognitive analyses. The proposed model may be used to guide the computer-based representation of knowledge or the evaluation of a system's adequacy as reflected in a user's mental models during problem solving. Details are described in the sections of this article as indicated in the diagram.

some texts, such as CPGs, are designed to provide a plan for action, they cannot be used to bypass this necessary step in medical problem solving, since comprehension is a prerequisite for creation of effective guidelines or successful action by users. Comprehension involves the construction of a mental representation of some aspects of the external world (i.e., an external representation, such as a CPG). Mental representations have been expressed symbolically in various forms, including propositions, production rules, schemas, mental images, and connectionist networks [15]. In the field of comprehension, however, it has been found that propositions are an empirically adequate representational form for investigating text understanding. Text representations are multilayered, in that various levels of discourse can be identified, from the linguistic (e.g., morphology, syntax), to the semantic/propositional, to the conceptual levels [7, 16, 17]. Although the linguistic levels are important aspects of text representation, we deal here only with the psychological aspects relevant to discourse, such as semantic and propositional analysis. Our aim is therefore not to provide a linguistic analysis but to use analysis of the semantics of language as a way to understand cognitive aspects of guideline representation.

In the cognitive processing of written discourse, the reader generates meaning from the text itself by transforming the information in the written text into some semantic form or conceptual message. This interpretation of meaning is "filtered through" the person's prior knowledge about the text message (see Fig. 2). Readers then incorporate this interpreted semantic information into their general store of knowledge. Understanding, then, may be regarded as a process whereby readers attempt to infer the knowledge structure of a writer through the text, based on what they already know. The reader's own knowledge base is used as a source of "data structures" from which inferences are drawn [16]. Thus the mental model developed by an individual who is reading a text is not limited to the information contained in the text itself but is naturally extended to include the larger structure incorporating the person's prior knowledge. Thus, the final product of discourse processing consists of what the text "says" and whatever prior knowledge and experiences the users "attach" to the text. In this sense, the reader constructs a cognitive model of the situation described in the text. It is, then, from the interaction between the text-base and the situation model that the conceptual representation emerges [3]. This representation varies greatly from reader to reader as prior knowledge and experiences differ.

Many processes are involved in the interaction between the text and prior knowledge, including linguistic and cognitive strategies as well as attitudes, interests, plans, and goals. All of these are involved in the generation of inferences as readers attempt to link the content of a text to their personal world knowledge. The types of inferences that concern us here are *knowledge-based inferences* to distinguish them from purely logical inferences. Knowledge-based inferences are generated to fill in the gaps of missing information in the text. Consider a very simple example: "Billy got dressed to go sleighing." In this sentence, the reader will infer that Billy put on warm clothing, not his bathing suit, because the reader knows that one goes sleighing when it is cold outside. Inference generation is done "on-line," as the reader processes the sentence. As each word is processed, activation of related concepts occurs. For example, "sleighing" may activate the concept of "snow," which in turn activates the concept of "cold," which tells us that he would dress in warm clothing.

Which concepts are activated depends on the level of knowledge that the reader possesses of the domain to which the text relates. Memory researchers have identified two memory systems that are used to explain how comprehension takes place: working memory (WM, sometimes also referred to as short-term memory) and long-term memory (LTM). WM stores the information about the current text segment in a rapidly accessible form (e.g., "getting dressed," "sleighing"). Without WM, comprehension would be impossible. Long-term memory includes everything a person knows and remembers; i.e., their prior knowledge (e.g., "sleighing is done in snow," "snow is cold"). Researchers have also identified some conceptual structures in LTM that serve to represent the text in a most efficient way and which develop as the reader accumulates experience with the domain. These are termed *retrieval structures* [18] and they serve to assist in the interpretation of incoming information. If the reader has a high level of expertise in the text's domain of discourse, then retrieval structures are used and comprehension and memory are greatly improved. However, if the reader does not possess sufficient knowledge of the domain, then the retrieval structures will not be available and text comprehension and memory will not be as efficient [19]. For example, for an inexperienced physician, data elements that may seem like a collection of independent signs and symptoms for a patient may cue an expert endocrinologist to generate the concept of hyperthyroidism. In this case, the pattern of signs and symptoms is linked, in the expert's knowledge base, to the generated concept, which then serves to guide further investigation of the patient's condition. In contrast, the inexperienced physician must generate a number of concepts (hypotheses about the patient's condition) that may be irrelevant to the diagnosis. Although the actual patient information available to both is the same, the inferences drawn from such information are radically different.

Given that interpretation depends on knowledge and that the aim of guidelines is to decrease variation in clinical practice (regardless of a user's background clinical knowledge), analyses of CPGs and the ways in which they are interpreted may be an important addition to the guideline-development process [8]. This can help to identify sources of incorrect inferences (e.g., inferences that are based upon ambiguous text or are affected adversely by inadequate prior knowledge). The techniques of propositional and semantic analysis can be used to identify ambiguous areas in the text that lead to such misunderstanding [20]. CPGs must be

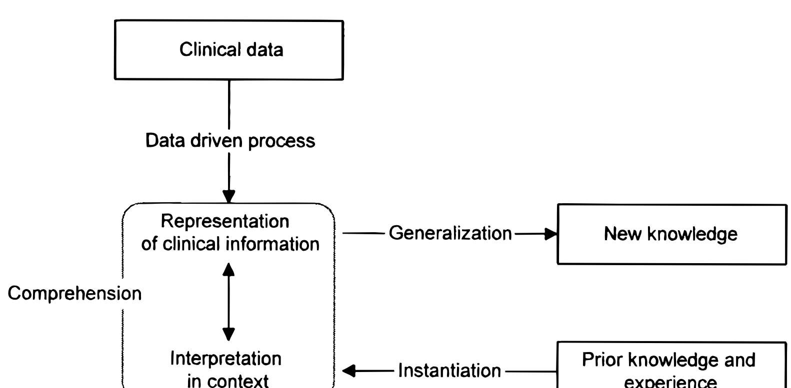

FIG. 2. Schematic representation of the comprehension process, using the interpretation of clinical data as a motivating example. Clinical data are interpreted in the light of prior knowledge and in a particular context. This process may produce new knowledge when generalized to other situations.

designed so that they are both flexible enough to be applicable in the real world, across different levels of expertise, and sufficiently explicit to ensure that correct inferences are made in most cases.

3.2. Propositional Analysis: Capturing the Semantics of Language

Propositional analysis is a method of semantic analysis developed in cognitive science [1], for representing linguistic information, which we have used extensively in the domain of biomedicine [19]. In particular, propositional analysis is a formal method for investigating representation of meaning in memory. Propositions are assumed to be the units of thought and their analysis provides a technique for identifying and classifying such units. The usefulness of the notion of propositions is that propositional representations provide ways of explicitly representing ideas and their interrelationships.

3.2.1. Basic notions in propositional analysis. A proposition is an idea underlying a piece of text (where a text can be a phrase, a sentence, a paragraph, or a story, written or transcribed from speech), which corresponds to the basic unit of the mental representation of symbolic information in human memory. Propositional analysis is used to capture the detailed semantic structure of symbolic discourse (e.g., text, dialogue, and diagrams) as understood by users.

In its simplest form, a proposition is composed of two concepts and a relation between the concepts. For instance, the sentence "A breast lump was revealed" may be analyzed as one proposition relating the concepts of "breast," "lump," and "reveal." Propositional analysis, however, does not end with the identification of the concepts and their relations. It also involves the categorization of the concepts and relations in the text. In the example above, "lump" is the object of the action and "breast" is a location. One can go further in this analysis by identifying the agent and the goals of the action and the instrument used. For simplicity, however, let us limit the analysis of the following examples to the explicitly stated concepts. Thus, a simplified propositional representation of the sentence may be written as

1.1 Reveal OBJ: lump; LOC: breast; NUM: (one); TNS: past,

where the number 1.1 in the first column represents the proposition number, the second column, "Reveal," is the head element or *predicate*, and the third column is called the *argument*. Notice that the propositional representation is always in the present tense (use of "reveal" instead of "revealed"), while tense information is given in the argument.

Table 1 presents an example of propositional analysis of a portion of a physician's patient-recall protocol [20] (i.e., a post-hoc construction by a physician after he has seen a patient or read about a patient's case). The sentence analyzed is: "The patient complained of some shortness of breath when he tried to climb two flights of stairs in his apartment." This sentence contains two clauses, namely: *The patient complained of some shortness of breath* and *when he tried to climb two flights of stairs in his apartment*. The first phrase contains two propositions and the second one contains seven propositions.

As described above, the first column represents the proposition number (1.1, 1.2, and so on), the second column is the head element or predicate (e.g., "Complain"), and the third column is the argument (e.g., AGT:Patient, THM:1.2, = TEM:_, TNS:PAST;). The first proposition (1.1) represents the fact that the patient (AGT) complained about some *theme* (THM:1.2), to be described in the next proposition (1.2). The category TEM (temporal) is left blank because there is no specific information about when the patient made the complaint. TNS:PAST (past tense) indicates that the complaint was in the past. The second proposition presents the theme of the complaint (indicated in proposition 1.1 as "THM:1.2"), which is shortness of breath. The shortness is coded as an attribute (ATT) of the breathing, and DEG representing the degree of the shortness of breath ("some").

The analysis of the second phrase is somewhat more complex. The first proposition expresses that the patient (AGT or agent of an action) tried to do something (i.e., achieve a goal, at this point unspecified). This goal (climb stairs) is presented in the second proposition. It is not clear whether the goal was achieved. The propositions numbered 2.3 and 2.4 convey information about the stairs (that they are located in the apartment and that it was two flights of stairs that the patient tried to climb). Proposition 2.5 represents that it was the patient's own apartment; and proposition 2.6 represents the *temporal equivalence* between the shortness of breath and the attempt to climb the stairs. The last proposition (the temporal equivalence between shortness of breath and climbing two flights of stairs)—although not directly expressed in the text—is nonetheless very important. It provides information that is likely to be interpreted by various physicians in different ways, or even to be missed by an inexperienced physician or a trainee.

3.2.2. Identificationof inferences. Propositional and semantic analyses provide information on the coherence and comprehensibility of text. Through these analyses, verbal

TABLE 1

Example of Propositional Analysis of the Phrase[a]

1. The patient complained of some shortness of breath	
1.1. Complain	AGT:Patient, THM:1.2, = TEM:_, TNS:PAST;
1.2. Breath	= ATT:shortness, DEG:some;
2. when he tried to climb two flights of stairs in his apartment.	
2.1. Try	AGT:he (patient), GOAL:(to)2.2, = TEM:2.6, TNS:PAST;
2.2. Climb	AGT:(he), OBJ:2.3;
2.3. Stairs	= LOC(in)apartment, PRT:2.4;
2.4. Flights	= NUM:two;
2.5. (POSS)	PAT:he (patient), OBJ:apartment;
2.6. EQUIV:TEM (when)	[1.2], [2.1]

[a] "The patient complained of some shortness of breath when he tried to climb two flights of stairs in his apartment."

protocols may be examined for the generation of inferences and the directionality of reasoning (i.e., whether from data to hypothesis generation or from hypotheses to data collection). In this section, we explain how this semantic information can be derived from a propositional analysis, extending beyond simple analysis of syntax.

We previously mentioned the importance of inferences in discourse comprehension. A subject's transcribed protocol can be analyzed for inferences when they are asked to read a text and are then asked to recall what they have read. By specifying both the propositional structure of the text and of the "recall" that the subject provides, one can determine which parts of the subject's recall are straightforward recapitulations of the text and which are modifications or inferences. Inferences are reflected as "propositional transformations made on a text."

The scoring of the subject's protocol involves marking as a recall every item in the protocol that corresponds exactly to the message base as defined in the original text. Similarly, whenever there are transformations made by subjects when compared with the message base, such changes are scored as inferences.

For example, consider the following sentence from a cancer guideline:

"If hemoptysis or persistent cough, then examine tracheal bronchial tree."

The physician's recall protocol reads as follows:

"If hemoptysis or persistent cough, then examine for malignancy in lungs."

The propositional analysis of the two segments is presented in Table 2.

In comparing these two analyses, proposition 1.1 is identical to proposition R1.1, and 1.2 is identical to R1.2. Therefore, both of these propositions would be coded as pure "recall." However, proposition R1.3 would be coded as an inference because there is a change in the consequent slot of the proposition as compared to 1.3. Finally, R1.4 would also be coded as an inference because it involves a replacement of the concept "tracheal bronchial tree" with the more general concept "lungs." Transformations of the text (e.g., from examine *tracheal bronchial tree* to examine *malignancy in lungs*) are made on the basis of the subject's prior knowledge. This physician was able to make the inference because she knew that examination of the tracheal bronchial tree was motivated by a search for malignancy in the lungs.

The generation of inferences is linked to the understandability and the coherence of a text. One can argue that coherence makes a text easier to understand by constraining the amount and types of inferences that are made during interpretation. In turn, the amount of prior knowledge is also related to inference generation and comprehensibility.

3.2.3. Study of coherence through propositional analysis. The notion of text *coherence* is associated with both

TABLE 2

Propositional Analysis of a Text Segment and Its Recall by a Physician

Text segment: "If hemoptysis or persistent cough, then examine tracheal bronchial tree"		
1.1.	COND: (if)	[hemoptysis, 1.2]; [1.3];
1.2.	Cough	ATT: persistent;
1.3.	Examine	OBJ:[1.4];
1.4.	Tree	ATT: tracheal; ATT: bronchial;
Recall segment: "If hemoptysis or persistent cough, then examine for malignancy in lungs."		
R1.1.	COND: (if)	[hemoptysis, 1.2]; [1.3];
R1.2.	Cough	ATT: persistent;
R1.3.	Examine	THM: [(for) 1.4];
R1.4.	Malignancy	LOC: (in) lungs;

the nature of the text itself and the prior knowledge possessed by a reader. Some texts are complex or difficult to understand as a result of the number of ideas that are presented in a given section, such as in a sentence or a paragraph. Propositional analysis allows us to identify various types of propositions that serve to indicate the level of coherence of the text, based on its degree of *connectedness*. Some propositions express single ideas, whereas others express ideas that in turn reference other concepts. Texts with more connected ideas are generally more coherent. Thus coherence is defined for our purposes as a measure of the degree of connectedness among propositions provided explicitly in text. We can categorize propositions into three types based on the degree to which they provide coherence to a text: *single*, *embedding*, and *linking*.

- *Single* propositions express only one idea. These propositions are self-contained (i.e., they do not refer to other propositions). For example, the sentence "She lost weight" is represented as a simple proposition:

1.1 Lose PAT: She; ATT:Weight; TEM:_, TNS:past;

- *Embedding* propositions are identified by the presence, in the argument, of one or more concepts or proposition numbers that refer to other propositions. An example of an embedding proposition is found in the analysis of the sentence "She lost weight, resulting in an improvement of the control of her blood sugar." This sentence would be analyzed as two propositions: "She lost weight" and "improvement in control of her blood sugar," plus the link between these two propositions by a resultive relation (i.e., one proposition *results* from the other). Since the argument of proposition 1.1 contains proposition 1.2, this is coded as embedding. Similarly coded is proposition 1.2 because it contains proposition 1.3 in its argument.

1.1 Lose PAT: She; ATT:Weigh, RSLT:1.2, TNS:past;
1.2 Improve ACT: Control; THM:1.3;
1.3 Sugar LOC:Blood;

- *Linking* propositions contain a relation (i.e., a term in the propositional grammar, such as causal, temporal, or logical relations) as a predicate, and at least two propositions or concepts as part of the argument. An example is given by the sentence "She lost weight relative to her premorbid state." In this case, the two propositions, "She lost weight" and "premorbid state" are connected by a temporal order relationship, as follows:"

1.1 Lose PAT: She; ATT:Weight, TEM:_, TNS:Past;
1.2 State PAT:(her); ATT:Premorbid, TEM_;
1.3 ORD:TEM; relative to [1.2], [1.1];

Typically, most propositions in a simple text are single propositions. One's comprehension of a text is related to the extent to which its propositions require linking to other propositions, or require making inferences that interrelate different parts of the text. When individuals attempting to comprehend a text have background expertise in the area, evidence suggests that an increase in the number of embedding and linking propositions results in a decrease in the number of inferences needed to interpret the text. Since inferences introduce variability in interpretation, a text with fewer of these types of propositions will be more variable in its interpretation than a text with more of them. Interestingly, it will also be less easily recalled and understood. Studies of text comprehension have shown that the quality of understanding increases as the coherence of the text increases [21, 22]. This is because embedding and linking propositions bridge various parts of the text, making more salient the inferences that need to be made to capture the intended meaning.

For example, we used propositional analysis to characterize a major CPG developed by the American College of Physicians–American Society of Internal Medicine (ACP–ASIM), dealing with depression [23]. Table 3 shows the frequency of propositions for each of the three categories that we have described. Because this analysis indicates that there are numerous embedding and linking propositions (close to 40%), we can predict that the text is reasonably coherent and relatively easier to understand than a text with a lower percentage of such types of propositions, as comprehension research demonstrates [24]. Coherence reduces the probability of errors in understanding by providing the reader with retrieval structures that index different parts of the text and help with memorization.

3.3. *Semantic Networks and the Structure of Concepts*

Because a propositional representation is a list, it is difficult to visualize the whole structure of someone's understanding of a text. Hence the next step in the analytical process we propose (Fig. 1) is creation of a semantic network

TABLE 3
Frequency of Propositional Types as Indicators of Inferential Requirements in a CPG for Depression [23]

	Frequency	
Type of proposition	Total number	Percentage
Single	286	61.6
Embedding	58	12.5
Linking	120	25.9

that provides a picture of the whole conceptual representation at any desired level of detail. Semantic networks are graphs consisting of a non-empty set of nodes and a set of links connecting such nodes [25, 26]. Nodes may represent clinical findings, hypotheses, or steps in a procedure, whereas links may represent directed connections between nodes. One can think of semantic networks as being complementary to propositional representations, making it possible to visualize the relational structure of the propositions in their totality. Within this structure, propositions that describe attribute information form the nodes of the network (these are the semantic structures), and those that describe relational information form the links (the logical structures). Therefore the content of the network is defined by the nodes, and the structure of the network is defined by the links. Thus, the semantic network conveys two types of information: conceptual (i.e., the concepts used) and structural (i.e., how the concepts relate to one another). For this reason, it is sometimes necessary to generate ideas that convey or presuppose other more specific ideas. We represent these higher order ideas by enclosing various concepts in one box.

Figure 3 presents an example of a semantic network that includes a higher level concept in such a box. The figure illustrates the visual properties of semantic network representations that allow us to investigate two aspects of cognitive activity. First, we can investigate global and local text *coherence* based on the connectivity of the semantic graph. Second, networks provide a relatively precise means for characterizing the *directionality of reasoning*, as is described below. In addition, semantic network representations can allow us to analyze temporal relations as is described in the next section.

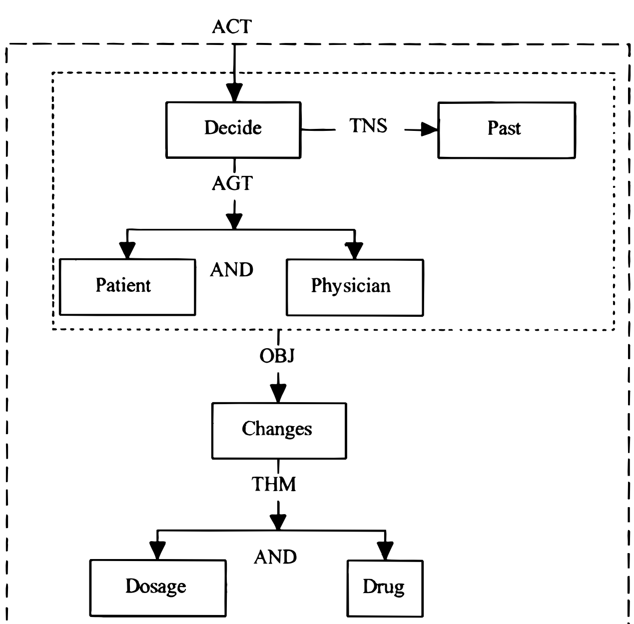

FIG. 3. Example of semantic network analysis of the propositions "The patient and the physician decided on changes in the dosage and the drug." This semantic network captures the overall coherence of the text. The high coherence of this sentence is reflected in the complete connectedness of the nodes and links in the diagram.

3.3.1. Temporal relations. Temporal relationships among actions or events are also captured in propositional analysis with the help of propositional types, as we have seen in previous sections of this paper. Propositions reflecting temporal relationships can be represented as procedures within the semantic network. In general, such temporal relations can be expressed directly in the text (e.g., the patient felt dizzy before collapsing), or they can be inferred from the tacit temporality present in the text, in which case propositional analysis links different parts of the textual information. The temporal nature of procedures is often assumed, rather than explicitly mentioned, in a CPG and thus must be inferred by a guideline user.

Since the user's situation model allows him to fill gaps and sometimes to reorganize information, the identification of places in the guideline where temporal sequences are meaningful may help in avoiding possible misinterpretations. Such variations in processing temporal relations are only one example of the way in which differences among experts and less expert physicians, as they process medical information, suggest that they would benefit from different types of guidelines.

Consider, for example, experiments in which we studied the interpretation of text guidelines by individuals with different backgrounds [9]. Figure 4 shows sample networks comparing a physician's representation of the temporal sequence within a portion of a text guideline (left side) with that of a computer scientist who was also asked to interpret the same guideline. Because the physician knew the clinical domain, he relied on high-level retrieval structures, tending to skip steps in processing medical information (e.g., constructing a procedure for managing a patient), whereas the computer scientist, although familiar with biomedical topics, was more likely to specify every step in the procedure. Thus, because their situation models were at different levels of generality, the same information given to both led to different representations. Such difference may also occur among individual physicians with varying backgrounds and perspectives, leading to different interpretations and actions when presented with the same clinical scenario and the identical

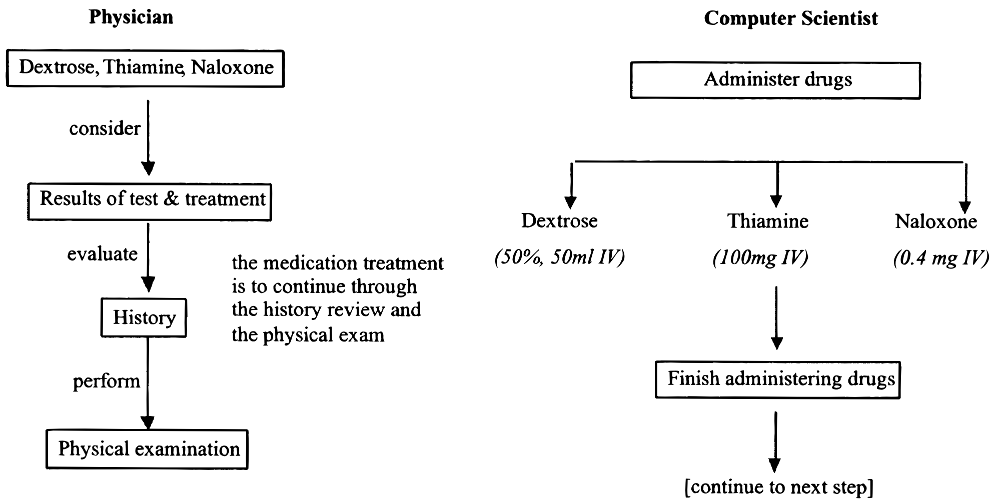

FIG. 4. Semantic network representation of the temporal sequence of a guideline as assessed by a physician and a computer scientist, showing the differences between the two individuals, reflecting their prior domain knowledge [9].

clinical guidelines. Sensitivity to this issue is crucial in the development of systems for representing guideline knowledge and presenting advice to clinicians based on CPGs.

3.3.2. Directionality of reasoning. It is important to analyze the directionality of reasoning because this gives us information about the actual thinking process being used to interpret patient information. Data-driven reasoning corresponds to an oriented path from data to hypothesis. Data-driven rules are identified whenever a physician, for instance, generates a hypothesis from the findings in a case. Hypothesis-driven rules correspond to an oriented path from a hypothesis to data collection. Pure data-driven reasoning refers to a network where all paths are oriented from data to hypotheses. Pure hypothesis-driven reasoning refers to a network where all paths are oriented from hypotheses to data-collection activities.

Data represented in semantic networks are analyzed for the nature and types of concepts and relationships generated by the subjects as they solve a problem. In the context of CPGs, clinicians can be asked to solve clinical problems both with and without the availability of written guidelines or, for system evaluations, computer-presented guidelines. In such settings the analysis allows for the identification and comparison of the procedures specified in the guideline with those generated by the subjects, either on their own or when attempting to be advised by the guideline. The nodes and links are identified and then coded as either data (attributes that appear in the original text) or hypotheses (attributes that do not appear in the text and have been generated by the subject). Then, using propositional analysis, those propositions that involve directionality are identified as described below. These are mainly conditional relationships and causal relationships. The path of reasoning (data-driven or hypothesis-driven) can be determined from the semantic network and this, in turn, can help us to understand how users of CPGs may be confused or misguided, rather than aided, as they attempt to provide quality care. In turn, systems for delivering guidelines benefit from the insights derived from such experiments in that we can learn how to avoid modes of presentation or guidance that confuse or complicate rather than assist by providing clarity.

Analysis of the directionality of reasoning is carried out when the subject is engaged in one of several tasks, such as think-aloud, recall, or explanation. In our work in biomedical cognition, a subject is typically asked to recall clinical information, to solve a problem while thinking aloud, or to explain the underlying pathophysiology of a clinical problem [19, 20]. Recent experiments have mimicked such studies while asking clinicians to make decisions while reviewing CPGs that are pertinent to their case [8]. In such experiments the resulting protocols are analyzed for data-driven or hypothesis-driven reasoning strategies by representing the protocol in propositional form from which a semantic network is generated as previously described. Figure 5 presents as an example the semantic network from a physician's think-aloud protocol as he solves a diagnostic problem. The physician describes a patient with a number of signs and symptoms (the data) that immediately suggest cardiovascular disease (the hypothesis), in a data-driven fashion. This hypothesis, in turn, suggests what to do (e.g., what additional tests to perform, and what patient data to collect), based upon predictions that can be made about likely scenarios. In our example, the patient also presents with a history of extreme fatigue and sleepiness, which the physician cannot interpret within the hypothesis of cardiovascular disease alone. In this case, the physician must generate an additional hypothesis to explain this anomalous finding [27].

Such mixing of data- and hypothesis-directed reasoning is common and to be expected in any problem-solving situation [17]. However, experienced individuals will often leap quickly to hypotheses that would be slower in coming (and typically would require more data-collection and data-driven analysis) for individuals who are less experienced in an area. This difference may, in turn, encourage those with experience to make many assumptions based on their familiarity with the domain, and may in turn result in their skimming, or even ignoring, sections of a guideline that is offered to them. Others might follow the guideline's logic more closely. Understanding this mix, and how different individuals will respond to and interpret the same guideline, can have important implications for how a computer-based delivery environment should be custom-tailored, how the guideline knowledge should be represented and conveyed, and how much flexibility should be offered in the modes of communication with the user.

3.3.3. Representing guideline information in computer-based language Consider now our use of similar analytic techniques in experiments that studied system developers with different backgrounds as they sought to encode a text CPG in a computer-readable format [9]. This corresponds to studying individuals as they convert text guidelines directly to computer code as shown by the thin arrow at the bottom right of Fig. 1. We studied individuals as they encoded CPGs into GLIF (GuideLine Interchange Format), a language designed to capture clinical guidelines so that they may be interpreted by computers in clinical settings, ideally integrated into health-care information systems [28, 29]. Computer-based guidelines are important because they can provide an automated method for delivering patient-specific recommendations to practitioners at the point of care. Without such capability, the only alternative for applying guidelines is the unrealistic expectation that clinicians will read the guidelines published in journals, remember their advice, and apply it subsequently in their practice. While authoring of guidelines directly in GLIF-compatible formats is a desirable goal, most existing guidelines have been generated originally as paper-based and must be translated into GLIF.

What is the process involved in translating natural language guidelines into the GLIF format? This question was addressed in our study [9] and we found differences among encoders similar to those described earlier when we mentioned differences in temporal assumptions between physicians and computer scientists who were interpreting guidelines. The results indicate that there is a great variability in the way guidelines are translated into GLIF. Variability in strategies used by the subjects was found to be dependent on the degree of prior experience and knowledge of the domain. The representations developed by physicians used

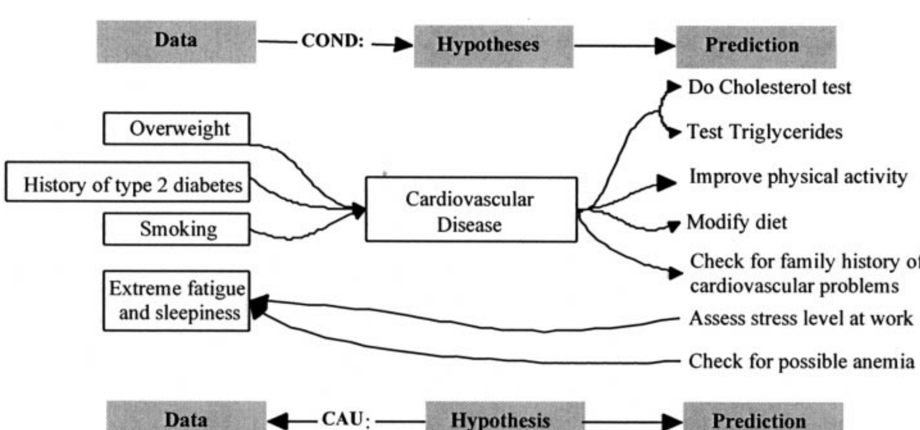

FIG. 5. Patterns of data-driven and hypothesis-driven reasoning to account for patient information. The left-hand side of the figure represents data from a patient, the middle box represents the substantive hypothesis generated (a retrieval structure linked to the data in the physician's memory), and the right-hand side of the figure represents the hypotheses for diagnostic procedure. Often although not always, data-driven reasoning is marked as conditional, while hypothesis-driven reasoning is coded as causal.

additional information and organization not explicitly stated in the guidelines, reflecting the physicians' understanding of the underlying pathophysiology. The computer scientists developed more literal representations of the guideline, and when they made additions, these mostly related to specifications mandated by the logic of language. However, collaboration between physicians and computer scientists resulted in more consistent representations where both domain-specific knowledge of medicine and generic knowledge of guideline structures were integrated [9]. Understanding such differences among situation models, and the way in which the same information given to various individuals can lead to different representations, can assist developers when they seek to anticipate problems in the design of systems for CPG delivery. A possible solution is to develop processes whereby participants with different backgrounds and expertise collaborate in the guideline-development and translation process, sharing and clarifying concepts and issues based on their knowledge of, and experience in, their domains of expertise.

4. DISCUSSION

In this paper we have provided a theoretical and methodological framework for the use of cognitive analysis to support biomedical knowledge representation and the design of clinical systems. Illustrative examples have been drawn from the development and use of clinical-practice guidelines. The theoretical ideas are borrowed from psychology, more specifically from the area of text comprehension (knowledge representation/interpretation and the role of prior knowledge) and its relationship to problem solving (knowledge use for arriving at solutions or for making decisions). We have proposed the use of closely related analytical methods or tools, namely propositional and semantic analysis. These are formal methods of data analysis, which, together with other complementary methods, can provide us with powerful evaluation tools. We stress that the techniques described here are not the only methods that have value in the study of knowledge representation or system design. We have attempted, rather, to show the complementary role that they can play when used by medical informaticians in the design and implementation of biomedical systems. There are a host of other psychological and cognitive issues that are not encompassed by assessment of comprehension. In the case of CPGs, research needs to be carried out that looks at how they are developed and used at the point of care, mapping the whole process from the time CPGs are created to the time they are used in actual clinical settings.

Most evaluations of CPGs have been based on how easily and safely practitioners can use them in their daily practice. When the outcome of the guideline use is positive, as expected, we accept this guideline as being successful even though we do not know the precise aspect of the guideline or its implementation that made the guideline useful and effective. In the absence of this knowledge, we are unable to make sure that we are duplicating our success when we undertake subsequent development of guidelines or systems for their delivery.

Similarly, when the guidelines are not used effectively by practitioners, there are many things that could have gone wrong. We typically try to use post-hoc "trouble-shooting" techniques to figure out the source of errors or nonacceptance, an important process in that it gives us some insight into the problem. However, given that there are complex processes involved, it is very difficult to be precise about the nature of the problem in "after the fact" analyses. In addition, such analyses are subject to hindsight bias.

We have proposed, rather, a formative, proactive analysis of the processes in guideline development and use—one that is based on theories from cognition and comprehension. Detailed analyses of the description of the CPG development process can also help us to contextualize and interpret the errors that arise during use. The identification and categorization of the errors can be fed back into the guideline-development cycle. Such feedback of information can allow designers either to improve the clarity and accuracy of the information to be delivered or to tailor it to the characteristics of specific users (e.g., according to their levels of expertise). This cognitive approach gives us a methodological tool that allows us to compare how the meaning intended by the guideline designer is interpreted and understood by the users. Furthermore, we can identify points of misunderstanding and errors that can have consequences for the appropriate implementation of CPGs.

Figure 6 presents a diagram of the dynamic process of guideline comprehension and use. It demonstrates how both the interpretation and application of the guideline are affected by many factors, and how the entire process is always related to the particular clinical goal the reader has in mind.

We close by summarizing some ways in which cognitive methods of analysis could be used fruitfully to improve CPG generation, implementation, and use.

1. CPGs are typically developed by a number of highly qualified experts in the domain of the guideline. It has been demonstrated that experts typically do not need the level

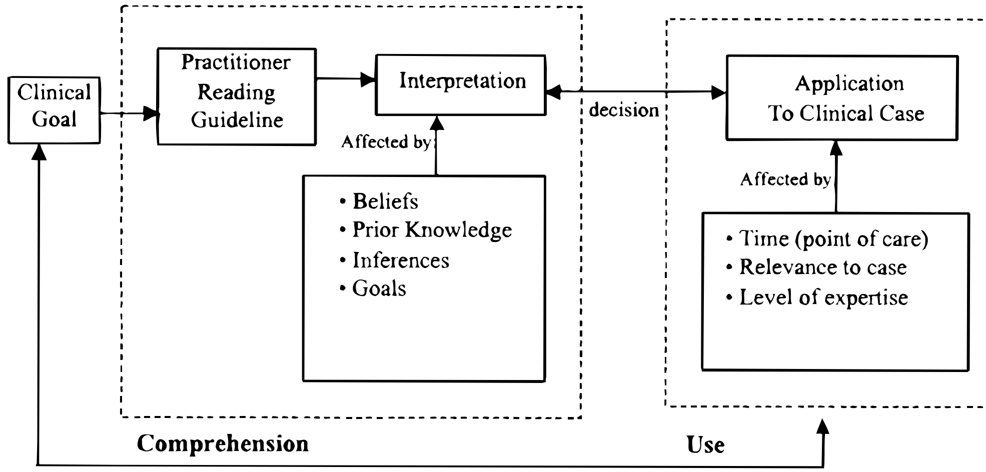

FIG. 6. The process of guideline comprehension and use. Arrows indicate directionality, where two-way arrows represent mutual influence.

of informational detail that less-than-expert people require. Sometimes guidelines developed by experts are not understood well by nonexperts. For this reason, text versions of CPGs may require a large number of inferences for their proper interpretation. These inferences provide coherence to the representation and may be crucial in the proper understanding of the guideline. However, it is essential that the user make the correct inferences if correct actions are to be taken. The identification of embedding and linking propositions in a text guideline may be beneficial for the development of more explicit guidelines and for identifying areas of potential confusion.

2. Propositional analysis can also be employed to assess the use of working memory in interpreting a text. The size of the working memory differs as a function of expertise. As expertise increases, there is a greater use of intermediate constructs (e.g., retrieval structures), which in medicine are concepts that refer to underlying disease processes and which summarize information at lower levels of aggregation (such as the number of clinical signs and symptoms). For instance, an expert physician while observing a patient may interpret a case as *hypometabolic state*. In contrast, a novice interpreting the same case may observe a series of disconnected signs and symptoms, such as *feeling cold and tired*, *dry and brittle hair*, *dry and itchy skin*, and *constipation*. The expert is using intermediate constructs to put these symptoms together into the diagnostic category of hypometabolic state. This process requires the use of working memory. The novice cannot generate these intermediate constructs and therefore cannot efficiently use working memory to retrieve the relevant information. Using results from propositional analysis, guideline information can be tailored to a particular target audience, thereby reducing the excessive use of working memory and the possibility of incorrect inferences.

3. Using techniques of data collection such as think-aloud protocols, and subsequently using our analytical techniques, we can seek to understand the thoughts and ideas of guideline creators as they design CPGs. We know that some written guidelines have problems with inconsistencies and logical gaps, as well as inappropriate assumptions about the backgrounds and expertise of readers. By investigating the underlying thought patterns in addition to their text representation by CPG creators, we can better understand the guideline-authoring process and, in turn, assist in the development of robust authoring environments. The goal, of course, is to help to assure that future guidelines are created without the kinds of flaws and mismatches that characterize too many of today's CPGs.

4. When trying to encode clinical guidelines for use in automated delivery systems, developers would benefit from insights into the dependence of the CPG logic on linking or embedding propositions. The variability in interpreting CPGs can be lessened by converting complex propositions to sets of single propositions, easing the need for generating inferences by the reader. To tune the information to different levels of user expertise, it may be possible to make this information optionally provided. For example, by browsing through the CPG on the computer (either in text or graphical form), an expert physician may be offered a means to skip through this information (and therefore make the necessary inferences himself or herself), whereas the inexperienced physician may prefer to inspect the information in detail.

5. Our proposed methods of cognitive analysis, used in tandem with computer formalisms, could result in more

effective systems for guideline design and development. Duff and Casey [30] suggest that computer-based tools can provide strategies that support the access, communication, and evaluation of guidelines through three functions: (1) knowledge browsing (i.e., accessing knowledge from a knowledge base), (2) messaging (i.e., using computational tools to exchange information), and (3) counting (i.e., generalizing and analyzing data about the impact of the guideline on clinical practice). Using complementary cognitive methods of analysis can prove to be effective in providing realizable solutions for the automation of CPGs. For instance, Shiffman [31] has proposed a rule-based guideline mark-up language that is complementary to propositional analysis. The advantage of the mark-up language is that it would facilitate implementation by virtue of being easy to translate into computer-based form. He proposes that the recommendations be written in a simple *if–then* format (similar to a conditional relation in the propositional structure), which would simplify guideline evaluation for correctness, completeness, and clarity. Guideline recommendations could be added or taken away without affecting the structure of the knowledge base because each rule is an individual chunk. However, this could be complemented by coherence measures, such as the use of embedding and linking propositions. Other formalisms have also been proposed that attempt to provide a general framework for the representation, modeling, authoring, and implementation of clinical guidelines [29, 32]. We are currently involved in the process of providing a cognitive methodology to complement such efforts [8].

In summary, methods of cognitive analysis, and in particular those related to comprehension (propositional and semantic analysis), offer the potential to complement other techniques that support the design and implementation of biomedical knowledge-based systems. We have illustrated these points in the context of clinical-practice guidelines and have cited some of our early studies that apply these methods to CPG authoring, encoding, implementation, and use. There are many more studies required before the full utility of these methods can be demonstrated, but now is the time to recognize their potential relevance and to begin more extensive experimentation with the use of cognitive analysis in biomedical informatics research.

ACKNOWLEDGMENTS

This work was supported in part by National Library of Medicine Grant LM06594, awarded in cooperation with the Department of the Army and the Agency for Healthcare Research and Quality, and in part by an award from the American College of Physicians–American Society of Internal Medicine (ACP–ASIM). The studies were conducted at McGill University.

REFERENCES

1. Kintsch W. The role of knowledge in discourse comprehension: a construction–integration model. Psychol Rev 1988; 95(2): 163–182.
2. Kintsch W. Comprehension: a paradigm for cognition. Cambridge: Cambridge Univ Press; 1998.
3. Patel VL, Arocha JF. Cognitive models of clinical reasoning and conceptual representation. Methods Inform Med 1995; 34(1):1–10.
4. Bransford JD, Johnson MK. Contextual prerequisites for understanding: some investigations of comprehension and recall. J Verbal Learning Verbal Behav 1971; 11:717–726.
5. Evans DA, Gadd CS. Managing coherence and context in medical problem-solving discourse. In: Evans DA, Patel VL, editors. Cognitive science in medicine: Biomedical modeling. Cambridge, MA: MIT Press, 1989; 211–255.
6. Patel VL, Evans DA, Kaufman DR. Cognitive framework for doctor–patient interaction. In: Evans DA, Patel VL, editors. Cognitive science in medicine: Biomedical modeling. Cambridge, MA: MIT Press; 1989; 253–308.
7. van Dijk TA, Kintsch W. Strategies of discourse comprehension New York: Academic Press, 1983.
8. Patel VL, Arocha JF, Mottur-Pilson C, How J. Cognitive psychological studies of representation and use of clinical practice guidelines. Submitted for publication, 2000.
9. Patel VL, Allen VG, Arocha JF, Shortliffe EH. Representing clinical guidelines in GLIF: Individual and collaborative expertise. J Am Med Informatics Assoc 1998; 5:467–483.
10. Tunis SR, Hayward RSA, Wilson MC, et al. Internists' attitudes about clinical practice guidelines. Ann Intern Med. 1994; 120(11):956–963.
11. Lomas J, Anderson GM, Dominic-Pierre K, Vaida D, Enkin MW. Do practice guidelines guide practice? The effect of a consensus statement on the practice of physicians. N Engl J Med 1989; 321:1306–1011.
12. McAlister FA, Teo KK, Lewanczuk RZ, Wells G, Montague TJ. Contemporary practice patterns on the management of newly diagnosed hypertension. Can Med Assoc J 1997; 157(1):23–30.
13. Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: A systematic review of rigorous evaluations. Lancet 1993; 342(8883):1317–1322.
14. Wilson MC, Hayward RSA, Tunis SR, Bass EB, Guyatt G. Users' guides to the medical literature, VIII. How to use clinical practice guidelines, B. What are the recommendations and will they help you in caring for your patients? JAMA 1995; 274(20):1630–1632.
15. Zhang J, Norman DA. Representations in distributed cognitive tasks. Cogn Sci 1994; 18:87–122.

16. Frederiksen CH. Representing logical and semantic structure of knowledge acquired from discourse. Cogn Psych 1975; 7:371–458.
17. Groen GJ, Patel VL. The relationship between comprehension and reasoning in medical expertise. In: Chi M, Glaser R, Farr M, editors. The nature of expertise. Hillsdale, NJ: Lawrence Erlbaum, 1988; 287–310.
18. Ericsson KA, Kintsch W. Long-term working memory. Psych Rev 1995; 102(2):211–245.
19. Patel VL, Arocha JF, Kaufman DR. Diagnostic reasoning and medical expertise. In: Medin DL, editor. The psychology of learning and motivation: advances in research and theory. Vol. 31. San Diego, CA: Academic Press; 1994; 187–252.
20. Patel VL, Groen GJ. Knowledge-based solution strategies in medical reasoning. Cogn Sci 1986; 10:91–166.
21. Kintsch W, Vipond D. Reading comprehension and readability in educational practice and psychological theory. In: Nilsson LG, editor. Perspectives on memory research. Hillsdale, NJ: Erlbaum; 1979.
22. Britton B K, Gulgoz S. Using Kintsch's computational model to improve instructional text: effects of repairing inference calls on recall and cognitive structures. J Educ Psych 1991; 83:329–345.
23. Snow V, Lascher S, Mottur-Pilson C. Pharmacological treatment of acute major depression and dysthymia. Ann Intern Med 2000; 132(9):738–742.
24. Miller JR, Kintsch W. Readability and recall of short prose passages: a theoretical analysis. J Exp Psych Hum Learn Memory 1980; 6:335–354.
25. Sowa JF. Conceptional structures: information processes in mind and machine. Reading, MA: Addison-Wesley, 1983.
26. Sowa JF. Principles of semantic network explanations in the representation of knowledge. San Mateo, CA: Morgan-Kaufmann, 1991.
27. Patel VL, Groen GJ, Norman GR. Reasoning and instruction in medical curricula. Cogn Instruct 1993; 10(4):335–378.
28. Ohno-Machado L, Gennari JH, Murphy SN, et al. The guideline internet format: A model for representing guidelines. JAMIA 1998; 5(4):357–372.
29. Peleg M, Boxwala A, Ogunyemi O, et al. GLIF3: The evolution of a guideline representation format. Proceedings of the AMIA annual symposium. Los Angeles CA; 2000.
30. Duff L, Casey A. Implementing Clinical Guidelines: How can informatics help? J Am Med Inform Assoc 1998; 5:225–226.
31. Shiffman R N. Representation of clinical practice guidelines in conventional and augmented decision tables. J Am Med Inform Assoc 1997; 4:382–393.
32. Bernstam E, Ash N, Peleg M, et al. Guideline classification to assist modeling, authoring, implementation and retrieval. Proceedings of the AMIA annual symposium. Los Angeles, CA; 2000.