

Assessing an Ontology for the Representation of Clinical Protocols in Decision Support Systems

Tiago Oliveira, Paulo Novais, and José Neves

Algoritmi Research Centre/Department of Informatics, University of Minho, Braga,
Portugal

{toliveira,pjon,jneves}@di.uminho.pt

Abstract. In order to assess the expressiveness of the CompGuide ontology for Clinical Practice Guidelines, a study was conducted with fourteen students of the Integrated Masters in Biomedical Engineering from the University of Minho in Portugal to whom it was proposed the representation of multiple guidelines according to the ontology. They were then asked to evaluate the ontology through a questionnaire and written reports. Although the results seem promising, there is the need for significant improvements mainly in: the representation of medication prescriptions, the tasks used to retrieve information from the patient, the diversity of actions offered by the ontology, the expressiveness of conditions regarding the state of a patient, and temporal constraints.

Keywords: Clinical Protocols, Ontologies, Clinical Decision Support.

1 Introduction

There are various ways of expressing medical knowledge in Clinical Decision Support Systems (CDSSs) [5], but, among them, decision trees, probabilistic models, and task-network models (TNM) are arguably the most popular [8]. Despite the obvious advantages of each one, task-network models are still preferred over the others because of their representation of clinical guideline knowledge in hierarchical structures containing networks of clinical actions that unfold over time. The main TNMs include Asbru [9], PROforma [2], GLIF3 [1], SAGE [12], and GLARE [11], among others. The first one, Asbru [9], focuses on temporal parameters and offers constructs to define starting points, durations, and ending points of tasks. In addition, it allows the specification of intentions for actions and prescriptions, as well as the expected outcomes. These time-oriented actions, conditions and intentions are represented as patterns which assume the hierarchical structure of plans and sub-plans. As for PROforma [2], it follows a structure somewhat similar to Asbru in the sense that it also resorts to the representation of guidelines as plans. There is a root task to which every plan in a guideline belongs. In turn, a plan has any number of instances of other tasks, from actions to decisions. Its focus is on argumentation in favor or against a decision. GLIF3 [1] was the first model to place its emphasis on the use of standards.

In addition to using the task model just as the other models do, it makes use of terminologies to avoid semantic ambiguity in the definition of clinical terms, and employs the HL7 Reference Information Model (RIM) to ensure that other systems can communicate with a system using GLIF3. As for SAGE [12], a direct evolution of GLIF3, it is considered one of the most complete approaches to Computer-Interpretable Guidelines (CIGs). This model places a high importance on the notion of context. The context coordinates the activation of guideline-based decision support. As such, it has constructs that allow the definition of the conditions in which medical practice takes place, whether they are related to the organizational setting and roles, the patient characteristics, or necessary resources. The procedural guideline logic is represented in an activity graph constrained by the already-mentioned context variables. The GuideLine Acquisition, Representation and Execution (GLARE) [11] model is specialized in the treatment of repeated (periodic) events, which play a major role in clinical therapies. There are other important models featured in comprehensive reviews such as [8] and [6]. However, there is no standard computer-interpretable representation for CPGs, and many of the existing are criticized for lack of expressiveness.

Ontologies are explicit representations of the concepts from a domain. They are the basic construction units of the semantic web and their objective is to allow applications to process the content of information, rather than just presenting it. The Web Ontology Language (OWL) [3] is a standard proposed by the World Wide Web Consortium (W3C) designed for facilitating machine interpretation. Ontologies have not been widely explored in the representation of the procedural logic in clinical protocols, yet, they provide an ideal support for this knowledge. Thus, the objective is exploring OWL as the underlying language for CIGs and use it to develop a CIG ontology with the intention of building a sufficiently expressive representation that would be capable of accommodating knowledge from different types of guidelines.

This work presents a preliminary study designed to assess the CompGuide ontology for clinical protocols. As such, the organization of the paper is as follows. Section two provides a brief description of the ontology with its main primitive classes and properties. Section three describes the materials and methods for the study. Section four presents the results and their discussion. Finally, in section five conclusions and future work considerations are provided.

2 Developed Ontology

The CompGuide ontology was initially presented in [7]. The ontology provides a task network model representation for clinical guidelines in OWL. In order to fulfill that purpose, it follows a logic in which complex information elements are represented as individuals with multiple object properties connecting them to other individuals, and simple information that cannot be further decomposed is represented using data properties. However, simple information that is reusable and will most likely be needed across different guidelines is represented as class individuals as well. In this regard the representation is similar to a linked list

of procedures. As such, a CPG is represented as an instance of the *ClinicalPracticeGuideline* class. Individuals from this class have a set of data and object properties that enable the representation of descriptive and administrative guideline information such as the name of the guideline, its general description, date of creation and last update, version, clinical specialty, category, intended users, and target population. An example of the initial definition of a guideline is given in Fig. 1.

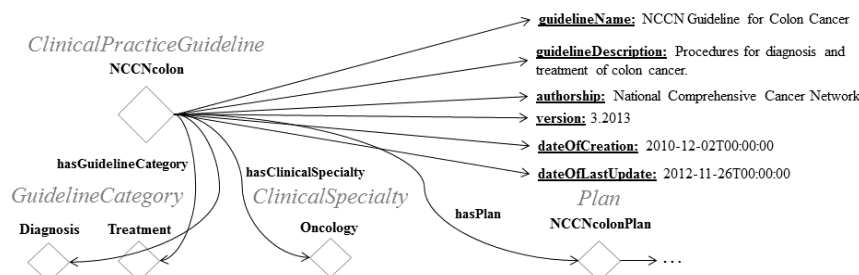


Fig. 1. Initial definition of a National Comprehensive Cancer Network guideline for the treatment of colon cancer in the CompGuide ontology.

Every guideline is connected to an individual of the class *Plan* which, in turn, is connected to other individuals that represent basic tasks. The procedural logic and workflow of clinical tasks is represented using three basic classes: *Action*, *Decision* and *Question*. The objective here is to create a recommendation plan that contains references to specific types of tasks. The *Action* class expresses a procedure that should be carried out by a health care professional. There are several subtypes of actions in the ontology that specify their nature with more detail. The *Decision* class is used to make assertions about the state of the patient, to infer new information from the existing one. The most obvious example of such a task is clinical diagnosis. The *Question* task is used to get information about the symptoms of a patient, to register information from the observations of the physician, and to store results from clinical exams. This type of task gathers all the information necessary for the execution of the clinical algorithm. Through object properties, it is possible to define the sequence of execution of tasks or if they should be executed simultaneously or concurrently.

3 Materials and Methods

The objective of the preliminary study was to assess the expressivity of the CompGuide ontology in four important aspects of CPGs, namely the representation of administrative information, the construction of workflow procedures, the definition of temporal constraints, and the definition of clinical constraints. These are considered the fundamental aspects of CIG representation and the pre-requisites of a good CIG model [8]. For that purpose, 14 students from the fourth year of the Integrated Masters in Biomedical Engineering, branch in

Medical Informatics, from the University of Minho, in Braga, Portugal, aged between 22 and 23 years old, were selected. They had no prior knowledge of OWL, and received training in both OWL and Protégé [10] for a total of six hours distributed by three two-hour sessions. After the training, they were taught about the structure, classes, and properties of CompGuide in a two-hour session.

Then, the students were asked to do an assignment which consisted in the representation of a CPG in the referred ontology using Protégé. They were handed a .owl file containing the definition of the ontology which they should fill in by adding the necessary elements. The set of CPGs used in the assignment is showed in Table 1. They were randomly distributed among the students. As much as possible, one tried that each guideline included multiple categories, namely diagnosis, evaluation, treatment, and management. The assignment had the duration of one month, by the end of which the students were asked to fill in a questionnaire which consisted of sixteen statements regarding the expressiveness of the model in the four above-mentioned aspects. The statements used in the questionnaire complete the general statement: "The CompGuide ontology allowed the representation of:". Statements 1-9 were about the construction of workflow procedures, statements 10-12 were related with the definition of clinical constraints, statements 13-15 were devised to assess the definition of temporal constraints, and, finally, statement 16 was about the representation of administrative information. The set of statements can be consulted in Figure 2. The answers were provided in a five point Likert rating scale [4] (*1-strongly disagree, 2-disagree, 3-neutral, 4-agree, 5-strongly agree*). It was also asked that the students handed a ten-page report describing their principal difficulties and observations while performing the task.

Table 1. List of the guidelines that were used in the study, featuring their name, organization and the number of people assigned to their representation.

Clinical Practice Guideline	Organization	People Assigned
Clinical Practice Guidelines in Oncology - Colon Cancer	National Comprehensive Cancer Network	2
Clinical Practice Guidelines in Oncology - Rectal Cancer	National Comprehensive Cancer Network	2
Clinical Practice Guidelines in Oncology - Distress	National Comprehensive Cancer Network	2
Clinical Practice Guidelines in Oncology - Palliative Care	National Comprehensive Cancer Network	2
Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults	National Heart Lung and Blood Institute	1
Diagnosing and Managing Asthma	National Heart Lung and Blood Institute	1
Diagnosis, Evaluation and Management of von Willebrand Disease	National Heart Lung and Blood Institute	1
Diagnosis and Treatment of Respiratory Illness in Children and Adults	Institute for Clinical Systems Improvement	1
Diagnosis and Management of Diabetes	Institute for Clinical Systems Improvement	1
Diagnosis and Treatment of Ischemic Stroke	Institute for Clinical Systems Improvement	1

The process resulted in the diverging stacked bar chart in Figure 2. The chart presents the total percentage of agreement (calculated as *agree + strongly agree*), the total percentage of disagreement (calculated as *disagree + strongly disagree*), and the percentage of participants who were neutral (equal to the percentage of the *neutral* category), for each statement, in order to show the central tendency in each item.

4 Results and Discussion

By consulting the chart of Figure 2, and specifically items 1 to 9 which refer to the representation of different procedures and tasks in a workflow, it is possible to verify that, for each item in this group, the level of agreement is at least equal or above 50%. Indeed, the item about medication prescriptions (item 1) is the one that has the lowest agreement, the highest percentage in the *neutral* category (43%), and the only one that has percentage in the *strongly disagree* category (7%). This is indicative that the representation of medication prescriptions may have issues. In fact, in the reports the participants mentioned that the representation of medication prescriptions was impractical at times. In the ontology, a medication has to be defined as the subtype of an *Action* individual, and one *Action* can only have one prescription. However, in several of the represented guidelines what were perceived as single actions included the administration of more than one drug, requiring the representation of medication schemes as several parallel actions instead of a single action with a clear objective. Another criticism to the representation of medication prescriptions was that it was mandatory to define an active ingredient, dosage, pharmaceutical form, and posology for a drug, but in certain guidelines these elements were not available. Both items 1 and 2 seem to correspond to the requirements of guideline representation as they have high percentages of agreement. Item 4 also has a high percentage of agreement, but it is, among the nine items, the one that has the highest percentage of disagreement (14%). This may be due to some limitations of the *Question* class such as the absence of a description data property where it would be possible to provide a detailed description of the information that the task aims to obtain. The participants considered that the way in which the ontology is designed allows the representation of series of questions, decisions and actions, which mimics the organization of the algorithms of clinical protocols. This is evident in the high levels of agreement of items 5 to 9. Overall, the organization of the procedural logic of the guideline and the grouping of tasks in plans was considered to be advantageous, mainly because this helps the delimitation of different diagnoses, treatments, and realities. The item that refers to this grouping of tasks, item 6, has an agreement of 100%. As a whole, and given the topics presented in the questionnaire, it can be considered that the participants widely agreed that the CompGuide ontology could effectively be used to represent the guidelines in question. Nonetheless, there were concerns expressed in the reports that the available subtypes of actions (namely medication, clinical exam, observation and procedure) would not cover all the possible actions that clinical protocols may have. Many CPGs have knowledge encoded as index tables which are necessary in order to calculate health indexes which, in turn, are later used in decision making. This type of knowledge could not be represented, which is another aspect to improve. On the other hand, the participants reported that, by following the design pattern of the ontology, they were able to find redundant elements in the guideline algorithms which did not trigger any kind of event or have any consequence further ahead in the clinical process. This means that the structure of the ontology is at the same time intuitive and can

help to identify points in which the integrity of guidelines are compromised. The representation of clinical constraints is central to the ontology. Through trigger conditions, pre-conditions, and expected outcomes it is possible to respond to changes in the state of the patient and control the execution of tasks. From the levels of agreement of items 10, 11, and 12, of which the lowest is 79%, it can be said that the representation primitives for these elements fulfilled, for the most part, their role. As that may be, the participants mentioned that there were some obstacles to the definition of conditions. One of them was that conditions did not allow the representation of intervals for a value of a clinical parameter. It was possible to use inequality constraints, but to define an upper and lower bound for a clinical parameter it would be necessary to create two separate conditions. This situation requires extra work from the guideline encoder and may introduce errors in the encoding. The items referring to temporal restrictions, namely items 13 to 15, have low agreement when compared to the majority of the other items in the chart. The agreement that the CompGuide ontology allowed the representation of the duration of clinical constraints was 71%. 29% of the participants answered in the neutral category. As a matter of fact, the participants observed that, while it was possible to define how long a task should last, the expressive power of the ontology was limited in this regard. It was not possible to define intervals of duration for tasks with minimum and maximum values. However, this type of information element occurred very often in the guidelines. Meanwhile, item 14, which concerns the repetition of clinical tasks, gathered only 43% of agreement, and 50% of the participants answered in the *neutral* category. Despite recognizing the usefulness of the ontology element that enables the definition of the number of times that a task should be executed, the participants believed that a crucial element was missing, and that was conditional repetitions, i.e., the possibility of stating that a task should be repeated if a the state of a patient does not improve. This was also an observation made within the scope of item 15. Finally, item 16 got 93% of agreement, which seems to convey that the ontology elements responsible for the representation of administrative information such as authoring, name of the guideline, its general description, date of creation, and so forth, are fulfilling their target function.

On a final note, representation formats such as CompGuide have to also be capable of representing situations in which the decision is left to the health care professional. Occasionally, the elements for a decision may not be all present in the description provided by a CPG, and in such situations guidelines may recommend that health care professionals follow their best judgment according to the available evidence.

5 Conclusions and Future Work

Although there was no access to an entire statistical population of interest, given the time-consuming nature of the survey, the study still provides useful hints for the development of the CompGuide ontology. Essentially, one may consider that the guidelines used in the survey were accurately represented according to the ontology, despite the need for certain adaptations, which did not affect the logic

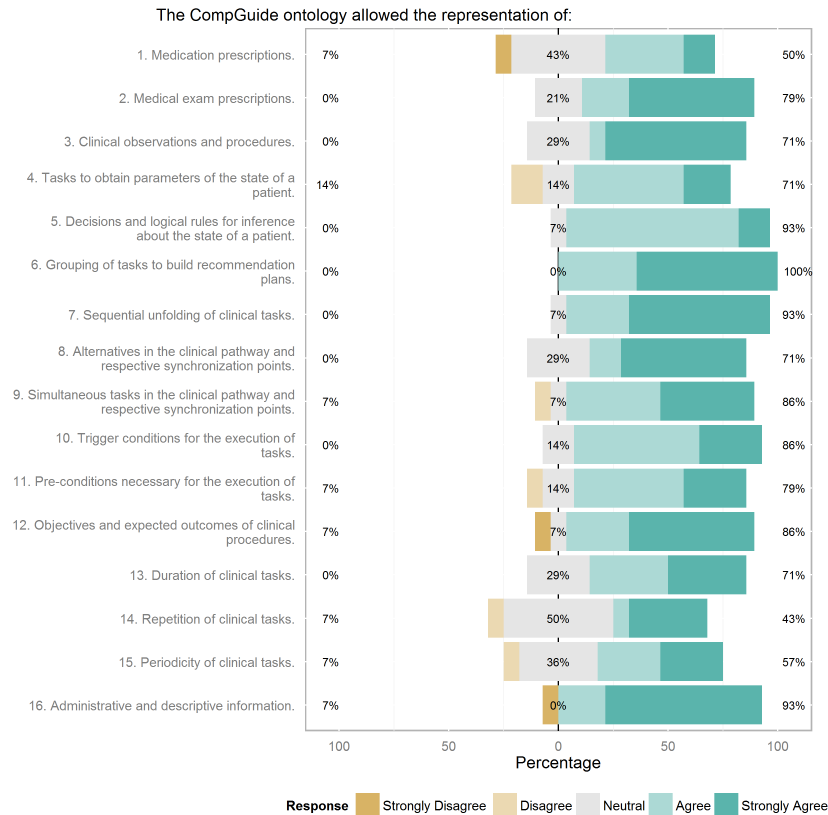


Fig. 2. Diverging stacked bar chart showing the results of the questionnaire to assess the expressiveness of the CompGuide ontology.

of the clinical process represented in CPGs. Nonetheless, there is a need for significant improvements, mainly in: the representation of medication prescriptions, the tasks used to retrieve information from the patient, the diversity of actions offered by the ontology, the expressiveness of conditions regarding the state of a patient, and temporal constraints as a whole.

These promising results may be due, in part, to the nature of the sample used in the survey. Since all of the participants were students of medical informatics, one can say that they have a clearer understanding than most about the role played by technology as a support for medical knowledge, and, although they were not familiar with OWL or Protégé, they already knew about similar models and principles. Moreover, the study should have included a broader range of CPGs in terms of origin and clinical specialty in order to expose the participants to a wider diversity of clinical situations. In future surveys, and once the issues identified are addressed, these aspects should be taken into consideration. Future work also includes completing the ontology proposed in this work with the international standards and data models proposed by HL7.

Acknowledgements

This work is part-funded by ERDF - European Regional Development Fund through the COMPETE Programme (operational programme for competitiveness) and by National Funds through the FCT Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) within project FCOMP-01-0124-FEDER-028980 (PTDC/EEI-SII/1386/ 2012). The work of Tiago Oliveira is supported a FCT grant with the reference SFRH/BD/85291/ 2012.



References

1. Boxwala, A.a., Peleg, M., Tu, S., Ogunyemi, O., Zeng, Q.T., Wang, D., Patel, V.L., Greenes, R.a., Shortliffe, E.H.: GLIF3: a representation format for sharable computer-interpretable clinical practice guidelines. *Journal of biomedical informatics* 37(3), 147–61 (2004)
2. Fox, J., Ma, R.T.: Decision Support for Health Care : the PROforma Evidence Base. *Informatics in Primary Care* 14(1), 49–54 (2006)
3. Group, O.W.: OWL 2 Web Ontology Language Document Overview. Tech. Rep. October, W3C (2009), <http://www.w3.org/TR/owl2-overview/>
4. Jamieson, S.: Likert scales: how to (ab)use them. *Medical Education* 38(12), 1217–1218 (2004), <http://dx.doi.org/10.1111/j.1365-2929.2004.02012.x>
5. Musen, M., Shahar, Y., Shortliffe, E.: Clinical decision-support systems. In: Shortliffe, E., Cimino, J. (eds.) *Biomedical Informatics*, pp. 698–736. Health Informatics, Springer New York (2006)
6. Oliveira, T., Novais, P., Neves, J.: Development and implementation of clinical guidelines: An artificial intelligence perspective. *Artificial Intelligence Review* 42(4), 999–1027 (2014)
7. Oliveira, T., Novais, P., Neves, J.: Representation of Clinical Practice Guideline Components in OWL. In: Pérez, J.B., Hermoso, R., Moreno, M.N., Rodríguez, J.M.C., Hirsch, B., Mathieu, P., Campbell, A., Suarez-Figueroa, M.C., Ortega, A., Adam, E., Navarro, E. (eds.) *Trends in Practical Applications of Agents and Multiagent Systems SE - 10, Advances in Intelligent Systems and Computing*, vol. 221, pp. 77–85. Springer International Publishing (2013)
8. Peleg, M.: Computer-interpretable clinical guidelines: A methodological review. *Journal of biomedical informatics* 46(4), 744–63 (2013)
9. Seyfang, A., Miksch, S., Marcos, M.: Combining diagnosis and treatment using ASBRU. *International journal of medical informatics* 68(1-3), 49–57 (Dec 2002)
10. Stanford Center for Biomedical Informatics Research: Protege Wiki (2014), http://protegewiki.stanford.edu/wiki/Main_Page
11. Terenziani, P., Montani, S., Bottrighi, A., Torchio, M., Molino, G., Correndo, G.: The GLARE approach to clinical guidelines: main features. *Studies in health technology and informatics* 101(3), 162–6 (2004)
12. Tu, S.W., Campbell, J.R., Glasgow, J., Nyman, M.A., McClure, R., McClay, J., Parker, C., Hrabak, K.M., Berg, D., Weida, T., Others: The SAGE Guideline Model: achievements and overview. *Journal of the American Medical Informatics Association* 14(5), 589 (2007)