Characterizing the Dimensions of Clinical Practice Guideline Evolution

Jacques BOUAUD\textsuperscript{a}, Brigitte SÉROUSSI\textsuperscript{b}
\textsuperscript{a}AP-HP, DSI, STIM, Paris, France; INSERM, UMR_S 872, eq. 20, Paris, France
\textsuperscript{b}Université Paris 6, UFR de Médecine, Paris, France; AP-HP, Hôpital Tenon, Département de Santé Publique, Paris, France; APREC, Paris, France

Abstract. The ever growing pace at which medical knowledge is produced requires clinical practice guidelines (CPGs) to be regularly updated. Since clinical decision support systems (CDSSs) are effective means to implement guidelines in routine care, they have to be revised as their knowledge sources evolve. From one version to another, some parts are kept unchanged whereas others are more or less modified. We propose to characterize formally the different dimensions of recommendation evolution in two successive guideline versions from the knowledge modeling perspective. Each atomic recommendation is represented as a rule connecting a clinical condition to recommended action plans. Using subsumption-based comparisons, seven evolution patterns were identified: No change, Action plan refinement, New action plan, Condition refinement, Recommendation refinement, New practice, and Unmatched recommendation. The method has been evaluated on French bladder cancer guidelines in the revisions of 2002 and 2004.

Keywords. Clinical practice guideline updating, clinical decision support system, knowledge base revision, bladder cancer management

Introduction

Significant practice variations are frequently observed and numerous medical errors occur in most medical institutions. Thus, there has been a worldwide interest for clinical practice guidelines (CPGs) expected to improve quality of care by disseminating evidence-based practice and assist practitioners in the state-of-the-art management of patients. However, CPGs are usually produced as textual documents, the dissemination of which is poorly effective in changing the behavior of physicians.

On the contrary, clinical decision support systems (CDSSs) seem to be efficient to influence healthcare professionals in the adoption of CPGs [1]. Defined as any software in which characteristics of individual patients are matched to a computerized knowledge base (KB) for the purpose of generating patient-specific assessments, or recommendations, CDSSs rely on computerized versions of original textual CPGs. However, despite the development of numerous guideline representation formalisms, the translation of textual CPGs into computerized KBs is still difficult and expensive [2]. In addition, KBs are often difficult to compare to original documents.
On the other side, medical knowledge is continuously increasing and the very meaning of “state of the art” or “evidence” are relative and evolving notions. The consequence is that CPGs are necessarily involved in a life-cycle and must be updated regularly by health professional societies or national health agencies in charge of their development. In France, the “Haute Autorité de Santé” (www.has-sante.fr), is responsible of CPG development, dissemination and updating. However, as opposed to living guidelines [3, 4], updating CPGs in France seldom capitalizes on the previous version of the same document, and revised versions of CPGs are often totally new CPGs. As a consequence, when updating guideline-based CDSSs to account for guideline revision, KBs have to be re-developed from the new CPG version.

The aim of this paper is to characterize the differences existing between two successive versions of computerized KBs and propose a formal characterization of knowledge evolution underlying CPG updating. The method has been applied to bladder cancer guidelines developed by the French Association for Urology (“Association Française d’Urologie”) in the revisions of 2002 and 2004.

1. Background

1.1. Document vs. knowledge-base centric approaches to CPG modelling

Most guideline tools focus on CPG documents as the starting point for guideline annotation and KB creation. Document-centric approaches such as the Guideline Element Model (GEM [5]), the Digital electronic Guideline Library (DeGeL [6]), Stepper [7], and the Guideline Markup Tool (GMT [8]) use guideline documents as the knowledge source and operate a gradual process of creating KBs from documents. Most of the time, those markup-based tools use XML-based documents to link textual descriptions with structured guideline models. The relationship between the semi-formal model of marked text parts and the original document is thus preserved. In model-centric approaches, the underlying conceptual model of guidelines is build by domain experts on the basis of their interpretation to clear text ambiguities and complete document incompleteness [2]. In both approaches, the final step is the translation of the formal guideline model into a guideline representation formalism (GLIF, Asbru, PROforma, etc.) and the building of a KB.

On the other hand, SAGE aims at promoting a knowledge-base-centric approach. In order to provide a human-comprehensible view of the guideline knowledge encoded in Protégé ontologies and KBs, the SAGE project developed a method for translating portions of a guideline KB into a legible format. Thus, SAGE uses KBs as the main guideline representation format and generates guideline documents from its models. Eriksson et al. [9] proposed to automate the document production from a computer interpretable guideline KB while combining documents with semantic information from the guideline KB. Although they used the SAGE guideline model, the approach is applicable for KBs that have well-defined semantics.

1.2. CPG updating

In document-centric approaches, when a guideline document is updated, updating the corresponding KB is well-handled in the case of “living guidelines” [3, 4]. Indeed,
living guidelines are incrementally built from previous versions and highlight the information that was changed: modifications of every revision are annotated, time-stamped and marked by arrows. In this case, differences between two successive versions of the document are easily recognizable. Unchanged parts of the text correspond to already modelled parts in the KB and are thus inherited. Only new textual parts have to be modelled.

In the case of “non-living guidelines”, even using natural language processing, identifying what was changed in the new textual version is not easy and the new guideline-based KB has often to be built from scratch. In addition, the development of a new version of non-living guidelines usually occurs every 4 to 5 years, or according to the availability of new evidence impacting clinical practice. This process may take up to 2 years, so that CPGs can be out of date as they are published. Thus, KBs may be more frequently updated on an ongoing basis as new evidence becomes available.

We propose a method to compare two KBs built from two successive revisions of the same CPGs. Beyond highlighting the differences between the two KBs, the method provides a formal characterization of practice guideline knowledge evolution.

2. Method

2.1. Notations and identity

CPGs can be considered as a set of atomic recommendations, noted \( \{R_i\} \). Each recommendation \( R_i \) is characterized by a pair \((S_i,P_i)\) such as \( S_i \Rightarrow P_i \), denoting that an action plan \( P_i \) is recommended in the clinical situation \( S_i \). A clinical situation \( S_i \) is defined by a set of instantiated decision variables: \( S_i = \{c_{i,n}\}_n \). An action plan \( P_i \) is defined by an ordered sequence of medical actions: \( P_i = (a_{i,m})_i \).

The basic comparison between 2 recommendations is based on identity: \( R_i \) and \( R_j \) are identical if and only if their clinical situations are identical (same sets of decision variables) and their action plans are identical (same sets of actions in the same order): \( R_i = R_j \iff (S_i = S_j) \land (P_i = P_j) \)

2.2. Similarity and subsumption

Even when not identical, recommendations may share some commonality. We propose to use subsumption functions to compare recommendations at a more abstract level.

In “ontological” abstraction, abstract concepts have to be identified and a mapping, noted \( abst \), between basic elements and their abstract class is built. To avoid the problem of over-generalisation and, for practical reasons, we limited the abstraction by specifying for each concept only one abstract concept relevant for the application. Abstraction is performed on decision variables and on actions. For instance, \( abst(geicbitin) = abst(cysplatin) = chemotherapy \) so that gemcitabin and cysplatin are both considered as chemotherapies, but are not comparable to any radiotherapy or surgery. Action plan similarity, noted \( sim \), is a boolean function taking 2 action plans as arguments. We consider that 2 action plans are similar if they involve the same actions. Since action plans are sequences, it consists in not considering order, and similarity is therefore defined as set identity: \( sim(P_i,P_j) \iff (P_i \subseteq P_j) \land (P_j \subseteq P_i) \).

Clinical situations can be compared using a structure-based subsumption function, noted \( subsum \), taking \( S_i \) and \( S_j \) as arguments, and returning true when \( S_j \) is more specific
than $S_i$, typically by involving additional decision variables. A simple subsumption function is then to consider set inclusion: $\text{subsum}(S_i, S_j) \iff (S_i \subseteq S_j)$.

Similarity and subsumption functions can be enhanced using abstraction on their arguments. We introduced the two functions $\text{Sim}$ and $\text{Subsum}$ defined as follows:

$$\text{Sim}(P_i, P_j) = \text{sim}(\text{abst}(P_i), \text{abst}(P_j))$$
$$\text{Subsum}(S_i, S_j) = \text{subsum}(\text{abst}(S_i), \text{abst}(S_j))$$

2.3. Evolution patterns

Using the notations previously introduced and the comparison algorithm described in figure 1 between 2 recommendations, assuming that $R_j$ is an evolution of $R_i$, we formally obtain seven evolution patterns that are qualitatively described as follows:

1. **No change**: recommendation $R_i$ remains unchanged; $R_i$ and $R_j$ are identical.

2. **Action plan refinement**: both clinical situations are identical, but their action plans, though different, share common properties according to the plan similarity function, e.g. $S_i^{2002} = S_j^{2002} = \{\text{no complete tumor resection, invasive tumor, non metastatic disease, N2, no-contraindication to chemotherapy}\}$, $P_i^{2002} = \{\text{chemotherapy}\}$ and $P_j^{2004} = \{\text{adjuvant chemotherapy}\}$.

3. **New action plan**: both clinical situations are identical, but their action plans are not considered similar. $R_j$ recommends a new action plan in the same situation e.g. $S_i^{2002} = S_j^{2004} = \{\text{no complete tumor resection, invasive tumor, non metastatic disease, N2, no-contraindication to chemotherapy}\}$, $P_i^{2002} = \{\text{radiotherapy}\}$ and $P_j^{2004} = \{\text{follow-up}\}$.

4. **Condition refinement**: clinical situation $S_j$ is more specific than $S_i$, i.e. $R_j$ is more restrictive than $R_i$, but the same action plan is recommended in both situations e.g. $S_i^{2002} = \{\text{no complete tumor resection, superficial tumor, prior TUR, no-contraindication to cystectomy}\}$, $S_j^{2004} = \{\text{no complete tumor resection, superficial tumor, prior TUR, T4b, no-contraindication to chemotherapy}\}$, $P_i^{2002} = P_j^{2004} = \{\text{new TUR}\}$.

5. **Recommendation refinement**: $S_j$ is more specific than $S_i$ and their action plans are similar (according to the plan similarity function). In this case, refinements occur at both clinical situation and action plan levels.

6. **New practice**: $S_j$ is more specific than $S_i$, but their action plans are not similar.

7. **Unmatched recommendation**: $R_i$ and $R_j$ don’t have anything in common.

![Figure 1: Pseudo code for identifying evolution patterns from R_i to R_j](image-url)
2.4. Comparison of recommendation sets

When comparing the two recommendation sets \( \{R_i^v\} \) and \( \{R_j^{v+1}\} \), which correspond respectively to the version \( v \) of a knowledge base and its following version \( v+1 \), each \( R_i^v \) is compared to each \( R_j^{v+1} \) with the above algorithm (Fig. 1) to determine which evolution pattern is relevant. At the end of the process, each recommendation \( R_j^{v+1} \) that always fell in the Unmatched recommendation case is considered New recommendation. Similarly, recommendations \( R_i^v \) that never matched any recommendation of the \( v+1 \) set are considered Obsolete recommendations.

3. Results

The method has been applied to bladder cancer. Textual 2002 CPGs and their 2004 revision were priorly formalized as decision trees. Each decision tree is structured to represent patient profiles clinicians may theoretically encounter as sequences of parameters (making the difference between clinical descriptors and therapeutic “coordinates” to locate the step the patient actually reached in the guideline-based care process) associated with the appropriate recommendations [10]. Structured KBs were then expanded as recommendations sets. The comparison of the 2 KBs is reported in table 1. In the 2004 version, many decision variables and actions were newly introduced, some were already in the prior version, while others were not used anymore. There were 577 recommendations in 2002 and 1,081 in 2004; only 47 were identical.

Table 1: Comparison of decision variables, actions and recommendations in 2002 and 2004 CPG versions

<table>
<thead>
<tr>
<th></th>
<th>2002-specific</th>
<th>Identical</th>
<th>2004-specific</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision variables</td>
<td>8</td>
<td>32</td>
<td>23</td>
</tr>
<tr>
<td>Actions</td>
<td>9</td>
<td>27</td>
<td>22</td>
</tr>
<tr>
<td>Recommendations</td>
<td>530</td>
<td>47</td>
<td>1034</td>
</tr>
</tbody>
</table>

The detailed comparison between 2002 and 2004 recommendation sets was performed. For each decision variable and action, an abstract concept was associated to implement the \( abst \) function. Then, the algorithm described in the previous section was run. Table 2 reports the distribution of the distinct evolution patterns we identified in the 2004 update of the guidelines.

4. Discussion and conclusion

The evolution of bladder cancer CPGs between 2002 and 2004 mostly concerns the extension of CPG coverage. It is noticeable that only less than 5% of the recommendations remained unchanged and that two third of them were considered new. Thus, in this particular update, and within the proposed framework, less than one third of recommendations might be derived from the previous version. Compared to the initial version, nearly 60% of the 2002 recommendations were considered obsolete.
Table 2: Distribution of recommendation evolution patterns from the 2004 version point of view

<table>
<thead>
<tr>
<th>Pattern</th>
<th>(n)</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obsolete recommendation (2002)</td>
<td>339</td>
<td>58.8%</td>
</tr>
<tr>
<td>(2002-basis : total = 577)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No change</td>
<td>47</td>
<td>4.3%</td>
</tr>
<tr>
<td>(2004-basis : total = 1,081)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Action plan refinement</td>
<td>3</td>
<td>0.3%</td>
</tr>
<tr>
<td>New action plan</td>
<td>40</td>
<td>3.7%</td>
</tr>
<tr>
<td>Condition refinement</td>
<td>49</td>
<td>4.5%</td>
</tr>
<tr>
<td>Recommendation refinement</td>
<td>39</td>
<td>3.6%</td>
</tr>
<tr>
<td>New practice</td>
<td>180</td>
<td>16.7%</td>
</tr>
<tr>
<td>New recommendation</td>
<td>723</td>
<td>66.9%</td>
</tr>
<tr>
<td>(2004-basis : total = 1,081)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The framework proposed to compare two successive versions of KBs can be used in knowledge-base-centric approaches. Indeed, it could provide health professional societies or national health agencies in charge of CPGs development with candidate documents produced by updated KBs. Another use could be to build “living KBs” where differences between two successive versions of KBs are easily recognizable and the types of the differences automatically identified.

5. References


