Application of the Logical Elements Rule Method for formalization of clinical rules: case study of ACOVE-NLI

Stephanie MEDLOCK1, Saeid ESLAMI1, Dedan OPONDO1, Marjan ASKARI1, Sophia de Rooijb and Ameen ABU-HANNA1

1Department of Medical Informatics, Academic Medical Center, Amsterdam, The Netherlands
2Department of Internal Medicine, Section of Geriatric Medicine, Academic Medical Center, Amsterdam, The Netherlands

Abstract. The Logical Elements Rule Method (LERM) is a step-wise method for formalizing if-then clinical rules. We applied LERM to a set of 40 clinical rules used in pharmacological quality assessment initiatives to assess (1) the amenability of the rules to formalization for decision support application (2) comparing adherence to rules that can and cannot be formalized, and (3) the usefulness of LERM as a tool for this task. Five rules could not be formalized, all due to unclear decision criteria. The adherence to ambiguous, non-formalizable rules was significantly lower than for formalizable ones (p<0.001). We modified LERM with three additions for this task: (a) adding the sub-step of restating the rules in a consistent natural-language grammar before decomposing them into normal form, (b) creating rules to use in lieu of a controlled vocabulary, and (c) adding the requirement that a time frame must be defined for all medications (before hospitalization, current medication, new medication, or discharge medication). Although the clinical rules in this sample are all stated as semi-structured if-then recommendations and are used in quality assessment initiatives, many ambiguities and inconsistencies in the clinical rules were identified by using LERM.

Keywords. Decision support, guideline implementation, formalization, clinical rules

Introduction

Clinical decision support is any computer-based system designed to help people make clinical decisions [1], and has been shown to be one of the most effective means for quality improvement [2]. To provide decision support, clinical knowledge in guidelines, protocols, clinical rules, or other natural-language representations must be formalized to a computer-interpretable form [3].

We have previously developed and published LERM, the Logical Elements Rule Method, for the purpose of formalizing clinical rules [4]. We chose to use the (Assessing Care Of Vulnerable Elders – Netherlands Inpatient) ACOVE-NLI indicators as representative of rigorously-developed, well-established clinical rules suitable to form the basis for clinical decision support. The ACOVE-NLI quality indicators are a
set of clinical rules covering a wide variety of topics relating to medication safety in hospitalized elderly patients [5]. They are rigorously developed in multiple Delphi rounds in two settings [5, 6] and have been used in a manual assessment of the quality of care with high inter-rater agreement (Cohen’s $\kappa=0.8$) [7]. They follow the “if (condition) then (action)” format recommended by Shiffman as a natural language expression of clinical recommendations that facilitates formal knowledge representation [8]. MacLean et al. observed that quality indicators that could be assessed using administrative data tended to have better adherence than indicators that could only be assessed manually [9], thus we hypothesize that clinical rules that can be formalized will have better adherence.

The objective of this study is to assess a set of clinical rules, stated in the form of “if (condition) then (action),” for their amenability to decision support using the Logical Elements Rule Method (LERM), and determine if the ability to formalize the rules is correlated with the adherence. A secondary objective is to critically assess LERM as a tool for the task.

1. Methods

We included all clinical rules from the Avoiding Inappropriate Medication and Continuity and Documentation of Care domains of the ACOVE-NLI set. This gave us 40 rules with a wide range of adherence and covering many aspects of care. The steps of LERM have been described previously [4] and are briefly summarized in Figure 1. We followed each step of LERM for all rules before proceeding to the next step. At each step, we noted ambiguities in the rules and problems with formalization, as well as any problems with using LERM or deviations from the published methods.

![Figure 1. The seven steps of the Logical Elements Rule Method (LERM).](image)
We determined how many rules could be formalized using LERM, could be partly formalized, or could not be formalized. A rule could not be formalized if a problem could not be resolved by consulting the handbook used in the manual assessment or reviewing the relevant literature. A rule could be partly formalized if it contained multiple recommendations, some which could be formalized and some which could not. We hypothesized that the rules which could be formalized would have better adherence than those which could not. We tested this by replicating the methods of MacLean et al., using a t-test to compare the quality score. The quality score is calculated for a set of rules by dividing the total number of times each rule in the set was adhered to by the total number of patients eligible for each rule (fulfilling the criteria in the "if" statement) [9]. We use the number of eligible patients and the adherence for each rule published in the manual assessment [7].

All statistical analyses were performed using R 2.14.1 (The R Foundation for Statistical Computing).

2. Results

2.1. Formalization

Of the 40 rules assessed, 30 could be formalized, 5 could be partly formalized, and 5 could not be formalized. All five of the rules that could not be formalized had unclear decision criteria (e.g. “If the attending physician suspects an elder is depressed...”).

LERM was used to identify ambiguities and inconsistencies in the rules. The number of problems found at each step, the number of rules affected, and an example of a problem found in each step is given in Table 1.

Table 1. Results of assessing the clinical rules from the Avoiding Inappropriate Medication and Continuity and Documentation of Care domains of the ACOVE-NLI set.

<table>
<thead>
<tr>
<th>step</th>
<th>number of problems found</th>
<th>number of rules affected</th>
<th>example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2a</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2b</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>5a</td>
<td>48</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>5b</td>
<td>23</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>All unambiguous concepts can be recorded, but missing data is common, especially in diagnoses.</td>
<td>All unambiguous concepts can be recorded, but missing data is common, especially in diagnoses.</td>
<td></td>
</tr>
</tbody>
</table>
It was hypothesized that rules which could be formalized would have better adherence than those that could not be formalized. The quality score calculated from the set of thirty rules that could be formalized was 37.3%; from the set of five rules that could not be formalized it was 9.3% (p<0.001).

2.2. Assessment of LERM

We extended LERM with some additions for this task. Step 2, breaking the rules into statements in normal form, seemed to naturally divide into three sub-steps rather than two. After identifying key phrases, we restated the rules so that they all used the same grammar before splitting them into subparts. For example, some rules are stated, “If A then B unless C,” and were restated as “If A and not C then B.” We had already reported that using a standard terminology improves reliability, but our decision support platform does not use a standard terminology. We addressed this by creating a set of definitions for use with these rules. For example, the phrase “unless contraindicated” was considered ambiguous unless specific contraindications were named. Finally, we required a time frame to be defined for all medications: pre-hospital, current in-hospital medication, new in-hospital medication, and/or discharge medication. This allows expression of a different behavior if a contraindicated medication is listed as a pre-hospital medication versus newly prescribed in the hospital.

3. Discussion

This heterogeneous set of rules was assessed for its amenability to decision support using LERM. Of the 40 rules assessed, 75% could be formalized. Although these rules followed the “if-then” format proposed by Shiffman to facilitate formalization, there were still many ambiguities and other problems. Rules that could not be formalized had a significantly lower adherence than rules that could be formalized. We modified LERM for this task by adding the sub-step of restating the rules in a consistent natural-language grammar before decomposing them, creating rules to use in absence of a controlled vocabulary, and requiring a time frame to be defined for all medications (before hospitalization, current medication, new medication, or discharge medication).

The main limitation of this study is that the rules were assessed by a single researcher, who was also one of the developers of LERM. Applying the theory of the rules to users other than the researchers is a subject for future research. Another limitation is that only five rules in this set could not be formalized. Therefore, although statistically significant, our conclusions about the differences in adherence between rules that could and could not be formalized should be considered preliminary.

Boxwala et al. suggest four layers of structuring in the process of formalizing a clinical document for decision support: unstructured (narrative), semi-structured (recommendations), structured (computable), and executable [3]. In this classification, the ACOVE-NLI clinical rules were already semi-structured, and LERM was used to transition to the structured layer. For measuring the quality score, further formalization to an executable form could be accomplished by combining LERM with the recently published CLIF method for formalizing quality indicators into a query language [10]. As there is not yet a standard executable decision support language, to provide decision support the final transformation will be to the executable format native to the decision support tools in which it is implemented [3]. LERM implies a standardization of this
form of clinical decision making which, in the setting of a trial, has the ability to bring LERM closer to the bedside and improve the quality of care through a stronger understanding of its application.

Consistent with MacLean et al., we found that the adherence to rules that could be formalized was significantly higher than those that could not not [9], though the reason for this difference is not yet known. It may be that basic decision support already exists for simpler rules, or that unambiguous recommendations are easier to remember. Rules which are easier to formalize may also be easier to assess manually, thus may have higher inter-rater reliability and less between-study variation. The difference in adherence between rules which can and cannot be formalized should be considered as a potential confounder when comparing quality scores measured by different methods.

Despite the rigor with which the ACOVE-NLI indicators were developed, and their high inter-rater reliability in manual quality assessment, we still found many ambiguities. The problems identified here are likely to occur in other clinical rules and recommendations. LERM may prove useful as a guide to the developers of clinical rules and quality indicators, to create clearer clinical recommendations in the future.

In conclusion, we were able to formalize 75% of this clinically heterogeneous set of clinical rules using LERM, with minor additions to the previously described method. Although the clinical rules are stated as simple if-then recommendations, many ambiguities and inconsistencies were identified by using LERM. Indicators that could be formalized were more likely to be adhered to than those that could not. Although our results are preliminary, this potential difference should be considered when comparing adherence as assessed by manual chart review with automated assessment.

References