Improving care of vulnerable elders through computerized clinical decision support

Medlock, S.K.

Publication date
2015

Document Version
Final published version

Chapter 5

LERM (Logical Elements Rule Method): a method for assessing and formalizing clinical rules for decision support

Medlock S, Opondo D, Eslami S, Askari M, Wierenga P, de Rooij SE, Abu-Hanna A.

Int J Med Inform. 2011 Apr;80(4):286-95.
doi: 10.1016/j.ijmedinf.2011.01.014

Abstract

Purpose: The aim of this study was to create a step-by-step method for transforming clinical rules for use in decision support, and to validate this method for usability and reliability.
Methods: A sample set of clinical rules was identified from the relevant literature. Using an iterative approach with a focus group of mixed clinical and informatics experts, a method was developed for assessing and formalizing clinical rules. Two assessors then independently applied the method to a separate validation set of rules. Usability was assessed in terms of the time required and the error rate, and reliability was assessed by comparing the results of the two assessors.
Results: The resulting method, called the Logical Elements Rule Method, consists of 7 steps: (1) restate the rule proactively (2) restate the rule as a logical statement (preserving key phrases) (3) assess for conflict between rules (4) identify concepts which are not needed (5) classify concepts as crisp or fuzzy, find crisp definitions corresponding to fuzzy concepts, and extract data elements from crisp concepts (6) identify rules which are related by sharing patients, actions, etc. (7) determine availability of data in local systems. Validation showed that the method was usable with rules from various sources and clinical conditions, and reliable between users provided that the users agree on a terminology and agree on when the rule will be evaluated.
Conclusions: A method is presented to assist in assessing clinical rules for their amenability to decision support, and formalizing the rules for implementation. Validation shows that the method is usable and reliable between users. Use of a terminology increases reliability but also the error rate. The method is useful for future developers of systems which offer decision support based on clinical rules.

Introduction

A clinical guideline is a systematically developed document to assist practitioner and patient decisions about appropriate care [1]. Use of guidelines has the potential to improve patient care, especially if guideline recommendations are provided in the form of clinical decision support [2]. Clinical decision support can be broadly defined as any computer-based system designed to help people make clinical decisions [1].

Much work in clinical decision support in the last decade has been devoted to developing computer-interpretable clinical guidelines [41]. In order to build decision support based on a guideline, the guideline is formalized (transformed from natural language to a logical algorithm) and implemented (using the algorithm to program decision support software which is used in practice). Recent work on formalization has focused on narrative guidelines, which describe a process of care with branching decisions unfolding over time [41].

In recent years, a demand for quality assurance and accountability has led to increased interest in performance indicators and other quality measures. In order for the quality of care to improve as a result of these measures, they must be linked to a process of care [144]. For example, a rule such as "80% of diabetic patients should have an HbA1c below 7.0" could be linked to processes such as: "All diabetic patients should have an annual HbA1c test" and "Patients with values over 7.0 should be rechecked within 2 months." These measure quality and performance at the population level, but in order to improve the quality of care, action is required at the patient level. When quality measures are linked to processes of care, the resulting statements closely resemble what Shiffman [145] called *condition-action* rules. Condition-action rules specify one or at most a few conditions which are linked to a specific action [145], in contrast to narrative guidelines which describe a series of branching or iterative decisions unfolding over time [41]. These quality

measures and condition-action rules are hereafter referred to as "clinical rules."

Narrative guidelines and clinical rules are two ends of a continuum of clinical care standards. Clinical rules represent elementary, isolated care recommendations, while narrative guidelines describe a coherent, unified care process. Clinical rules can be distilled from narrative guidelines [145], although this discards the control flow structure. Guidelines may contain both narrative sections with complex control flow and clinical rules. However, most work in developing computer-interpretable guidelines has focused on the difficult problem of formalizing the time-oriented structure of guidelines [41]. Descriptions of the encoding of clinical rules, for example for encoding in the Arden syntax [146], has left most of the formalization process to the developers.

As a case in point, at our institution a set of 87 clinical rules based on the ACOVE (Assessing Care of Vulnerable Elders) [25] set were used to determine areas for improvement in quality of care [29]. We planned on assessing the rules to determine whether they could be evaluated (that is, determining whether the computer can determine when the condition applies and whether the action should be taken), and formalizing the rules for a proactive clinical decision support system to help clinicians to adhere to these rules. To our knowledge, there is no systematic method described in the literature for assessing and formalizing clinical rules for decision support implementation. Although decision support systems based on clinical rules have been studied [147–154], these studies do not provide much concrete guidance for the process of rule formalization. Williams *et. al.* afford the most detailed description: (1) identifying the data elements needed to evaluate each rule, (2) determining if these are explicitly defined in the local computer systems, and (3) for those which are not explicitly defined, determining whether they can be acquired [154].

Systematic methods for guideline formalization have been described in the literature, however. These methods suggest that a domain ontology (a formal specification of concepts and relationships) and a control-flow structure (determination of what will execute and in what order) [155] must be specified. Shahar *et.al.* describe a process of semantic markup of the guideline by expert physicians (identifying the domain concepts), cooperative addition of control-flow structure, and then formalization by an expert in knowledge modeling (specification of the ontology) [156]. Svátek and Růžička describe a similar process: (1) input text format (2) course-grained markup (mark parts of the document which describe actions to be taken) (3) fine-grained markup (replace linguistic expressions with formal structures, and resolve ambiguity) (4) modularize the knowledge paths (encapsulate context, abandoning the narrative structure) (5) map to the specific knowledge base, and (6) encode [157].

There is much to be learned from these methods, but several areas require further elaboration. Disambiguating the domain concepts, as suggested in step (3) above, presents a significant barrier for implementation of both clinical rules [154] and guidelines [155]. A flow chart is a natural formal structure for most narrative guidelines, but not for clinical rules, particularly for rule sets which pertain to a heterogeneous patient group (for example, the ACOVE rules, which guide the care of elderly patients). Guidelines must be modularized (step 4 above), but clinical rules are stated as loosely coupled, modular recommendations. As they are integrated into the care process, the developer is faced with the problem of composing the individual rules into a workflow. This is not a problem encountered with narrative guidelines. We felt a method was needed to address these areas, which draws upon and complements the existing work in guideline formalization.

The aim of our investigation was to create a step-by-step method for assessing and formalizing

clinical rules, and to validate this method to assess its usability and reliability. The method we propose is referred to as LERM, the Logical Elements Rule Method.

Methods

In order to create a robust method which could be used with diverse clinical rules, we chose to develop LERM using a sampling of rules from multiple sources. We turned to the literature for studies of decision support systems based on quality measures which were linked to processes or condition-action rules: such as quality indicators, performance indicators, performance measures, standards of care, or clinical rules. We sought studies which implemented decision support systems based on published sets of clinical rules (thus excluding rules which were created specifically for a decision support system). The clinical rules used in these studies, together with the rules which we planned to use in our institution, formed our sample set of clinical rules (listed in Table 5.1). These rules fit our requirements of representing diverse areas of clinical care, from several sources, and in different formats.

Source	**Clinical area**	**Implementation study**
Rand Health	Vulnerable elderly [25]	
Cooperative cardiovascular project	Acute myocardial infarction [158]	Sauaia *et.al* 2000 [147]
Joint Commission	Acute myocardial infarction [159]	Butler *et.al* 2006 [148]
Joint Commission	Heart failure [159]	Butler *et.al* 2006 [148], Niemi *et.al* 2009 [149]
Joint Commission	Pneumonia [159]	Niemi *et.al* 2009 [149]
National Kidney Foundation	Dialysis [160]	Diamond *et.al* 1999 [150]
Arthritis Foundation / Rand Health	Rheumatology [161]	Williams *et.al* 2007 [154]
American Diabetes Association (ADA)	Diabetes [162]	Club Diabete Sicili 2008 [151]

Table 5.1: Sources of Rules

The sets of rules varied widely in size: from 392 rules in the ACOVE set [25] to only 4 rules in the Joint Commission heart failure set [159]. In order to maintain a good mix of rule sources and clinical conditions in our validation set, 2 rules from each set were randomly selected and reserved.

To develop LERM, a focus group was recruited consisting of the primary researchers (SM and DO), two clinical experts (SdR and PW), and one expert in medical informatics (AA). LERM was developed in an iterative process, with the starting assumption that, like guideline formalization, the new method would need to delineate a domain ontology and a control-flow structure [155]. In addition, the requirements of LERM were defined as: (1) assessing clinical rules to determine if

the rules can be implemented as proactive computerized decision support, (2) formalizing rules for implementation of decision support using data from existing clinical information systems (identifying the specific data which are needed so that access to these systems can be prioritized), and (3) a process which can be applied consistently over a large set of rules.

Each successive version of LERM was tested with a sample of rules from the development set, and the results discussed at the focus group meetings. The questions discussed in the focus group meetings were: (1) Is the clinical intent of the rules in that iteration maintained? (2) Is the mapping of phrases from one terminology to another correct? (3) Are there inconsistencies between or within rules or other problems which were revealed? (4) Could these problems be identified earlier, and would it be useful to do so? The method was revised on the basis of these results and the process repeated.

Validation was performed to assess both the inter-user reliability and usability of the method. Usability was assessed in terms of the time required to assess the rules and the error rate. Errors were defined as any inconsistency between the formalized rule and the original rule (for example, omitting a concept), or between the identified data elements and the terminology (for example, failure to locate a concept in the terminology which maps to the concept in the rule). Errors were located by comparing the results from the two assessors. Reliability was measured by assessing agreement between the two sets of results for each step of the method, and noting points of divergence and re-convergence. Each assessor independently formalized the rules, then the results were discussed together. Agreement and Cohen's κ were calculated in R. It was anticipated that use of a standard terminology would affect agreement, thus the assessment was performed first without a standard terminology, and then the relevant steps repeated using SNOMED-CT (Systematized Nomenclature of Medicine - Clinical Terms) [163] as the target terminology.

Results

LERM was created as a step-by-step method in order to systematize rule transformation and facilitate its contemporaneous application to many clinical rules. Although the process is presented linearly, there are several places where steps may be carried out in parallel. In the illustration of the method in Figure 5.1, these places are indicated by branches. Clinical rules are formulated for a particular purpose, and a clinical expert is needed to ensure that this intent is maintained as the rule is formalized. Side-by-side cooperation of a clinical expert and an informatics knowledge expert may be the best approach [156], but points are noted where clinical expert involvement is essential.

LERM is best described by illustrating its steps using concrete examples. Table 5.2 gives examples of clinical rules from the various rule sets, including the rules used in the examples below.

1. Determine whether the rule can be proactively operationalized.

   In order to provide proactive decision support for a rule, the rule must be stated such that it can be proactively operationationalized. That is, it must be possible to use the rule to make a decision for an individual patient before it is too late to carry out the recommended care. For example, outcome-oriented indicators will need to be restated in terms of a care process which can be improved [144].

   > **Example 1.1**: (from the Joint Commission AMI set) [158] Eligible [patients]: All patients with confirmed AMI. Criterion [for adherence]: Received aspirin during hospitalization. Exclusions: [*list of exclusions*].

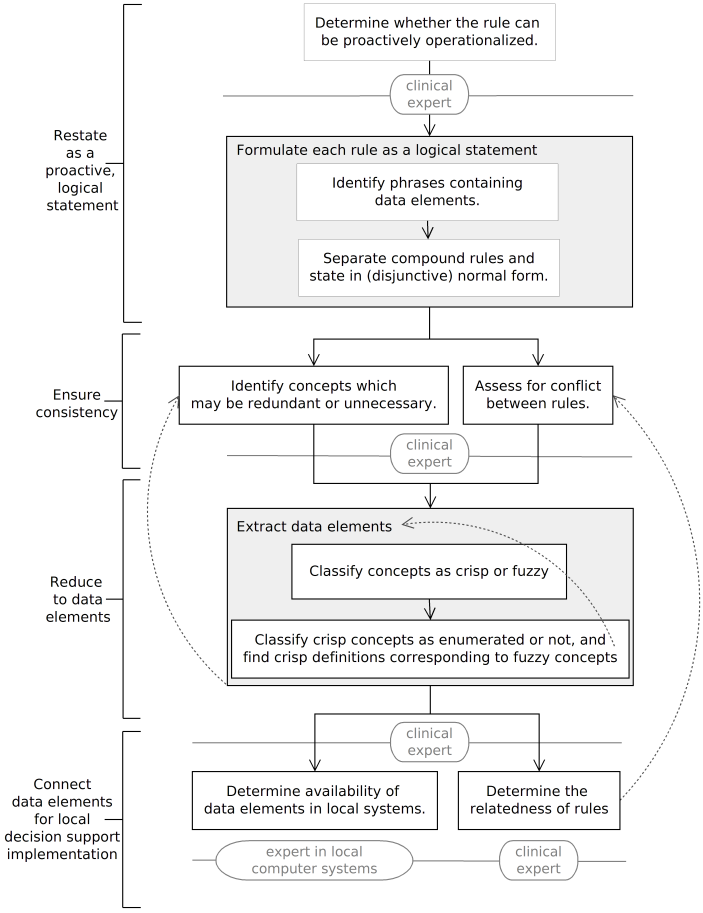

Figure 5.1: LERM: the Logical Elements Rule Method. Although the method is presented linearly in the text, in practice some steps may be done in parallel, as shown here. Some steps, such as extracting data elements or checking for conflicts between rules, may need to be repeated with the results of later steps as input.

> As it is stated, this rule can only be evaluated after the patient has been discharged. A proactive restatement of this rule might be:
>
> "All patients with confirmed AMI should receive aspirin within 1 day of admission, and daily during hospitalization, except [*list of exclusions*]."

The rule may need to be made more explicit in order to be actionable, but the intent of the rule should be preserved. Since a clinical understanding of the rule is required, a clinical expert should be involved in this step.

2. Formulate each rule as a logical statement.

   Like guidelines, rules need to be transformed from natural to formal language. This step combines elements of steps 2 an d 3 from Svátek and Růžička (markup and replacing linguistic structures with formal structures) [157]. A potential caveat of this transformation is that the formal version must be medically valid, but clinicians may have difficulty understanding formal restatements of the rules. Thus this is broken into smaller steps, and

ACOVE [25] (Assessing Care of Vulnerable Elders): IF a vulnerable elder is prescribed an ongoing medication for a chronic medical condition, THEN there should be a documentation of response to therapy
ACOVE [25]: IF a vulnerable elder requires analgesia, THEN meperidine should not be prescribed
Cooperative Cardiovascular Project [158]: Criterion: Received aspirin during hospitalization. Eligible: All patients with confirmed AMI (acute myocardial infarction). Exclusions: [*list of exclusions*]
Joint Commission AMI [159]: Description: Acute myocardial infarction (AMI) patients who are prescribed aspirin at hospital discharge. Numerator Statement: AMI patients who are prescribed aspirin at hospital discharge. Denominator Statement: AMI patients. Excluded Populations: [*list of exclusions*]
Joint Commission pneumonia [159]: Description: Pneumonia patients transferred or admitted to the ICU (intensive care unit) within 24 hours of hospital arrival, who had blood cultures performed within 24 hours prior to or 24 hours after hospital arrival. Numerator Statement: Number of pneumonia patients transferred or admitted to the ICU within 24 hours of hospital arrival who had blood cultures performed within 24 hours prior to or 24 hours after arrival at the hospital. Denominator Statement: Pneumonia ICU patients 18 years of age and older. Excluded Populations: none
National Kidney Foundation [160]: The delivered dose of hemodialysis should be measured at regular intervals no less than monthly
Arthritis Foundation Quality Indicators [161]: If a patient has RA (rheumatoid arthritis) and is being treated with a DMARD (disease-modifying antirheumatic drug) and reports worsening of symptoms over a 6-month period of time and there is evidence of active disease, then one of the following should be done: increase DMARD dose, change DMARD, add an additional DMARD or, start or increase dose of glucocorticoids
ADA (American Diabetes Association) Standards of Medical Care in Diabetes [162]: Statin therapy should be added to lifestyle therapy, regardless of baseline lipid levels, for diabetic patients with overt CVD (cardiovascular disease) or without CVD who are over the age of 40 and have one or more other CVD risk factors

Table 5.2: Examples of clinical rules. Clinical rules are independent statements which link one or a few conditions to a conclusion.

the original vocabulary and structures such as negations and chronological relationships are kept intact.

(a) Identify phrases containing data elements.

Williams *et. al.* [154] use the term *data elements* to describe the units of clinical information used by the decision support system. In some cases a single phrase may later be broken down into multiple data elements, but the goal at this stage is to identify domain concepts (shown in the example as underlined phrases) without changing the language used.

> **Example 2.1**: (from the ADA set) [162] Statin therapy should be added to lifestyle therapy, regardless of baseline lipid levels, for diabetic patients with overt CVD or without CVD who are over the age of 40 and have one or more other CVD risk factors.

(b) Separate compound rules. Restate the rules in disjunctive (or conjunctive) normal

form.

Here the basic control-flow structure is defined by defining which condition will trigger which action. Rules often contain compound statements (joined by "and", "or", etc.). There are two reasons to break the rules down into their simplest parts. One is that it makes it easier to later program the rules in a decision support system. The other is that if only part of the data is available, part of the rule may still be implementable, which may provide useful decision support.

> **Example 2.2**: The rule in Example 2.1 can be broken into independent parts:
>
> (1) diabetes AND overt CVD → statin therapy AND lifestyle therapy
>
> (2) diabetes AND (NOT CVD) AND over age 40 AND CVD risk factor [other than diabetes] → statin therapy AND lifestyle therapy

3. Assess for conflict between rules.

   Simply restating the rules in a uniform grammar may reveal inconsistencies within a set of rules which were not noticed by the developers of the rule set, or inconsistencies between sets of rules. Although it is mentioned as an early step in the process, further formalization may reveal other conflicts. Performing this as an early step avoids the potential caveat of failing to recognize a conflict until later, leading to additional work. Conflicts within and between guidelines are well recognized [155], and clinical rules are susceptible to this problem as well.

4. Identify concepts which may be redundant or otherwise unnecessary.

   A potential caveat of implementing clinical decision support is the work involved in extracting patient data from proprietary databases, or taxing the user with additional data entry. It is useful to determine the minimum data needed to provide decision support for the rule. If all data are available, redundant concepts can be used to check one another. There are two categories of phrases which are dealt with in this step: phrases which can be excluded without changing the meaning of the rule (Example 4.1); and phrases which are not redundant, but are not needed to interpret the rule in current clinical practice (Example 4.2).

   > **Example 4.1**: (from the ACOVE set) [25] IF a vulnerable elder is prescribed an ongoing medication for a chronic medical condition, THEN there should be a documentation of response to therapy.
   >
   > The vast majority of ongoing medications are prescribed for an ongoing (chronic) condition. The phrase "for a chronic medical condition" can probably be omitted without changing the meaning of the rule, and would greatly simplify use of this rule for decision support.
   >
   > **Example 4.2**: (from the ACOVE set) [25] IF a vulnerable elder requires analgesia, THEN meperidine should not be prescribed.
   >
   > There are no indications for meperidine other than analgesia (pain relief). However, the indications for meperidine could change as practice changes. The rule can be implemented by assuming that meperidine is always contraindicated in elderly patients, but this implementation should be checked regularly to ensure that there are no new indications for meperidine which would justify its use.

   At this stage, the rules have been broken into their simplest parts and restated in a consistent logical grammar, and apparent redundancies and conflicts annotated. The following

steps entail translating the phrases used by the authors of the clinical rules to the language which will be used inside the decision support system. Thus it is important at this stage to consult with a clinical expert to ensure that the rules, as stated in their new form, are medically valid and retain the intent of the original rules.

5. Extract data elements.

   Data elements are the units of clinical information which will be used by the software which provides decision support. The level of granularity which is "elemental" for a particular set of rules will depend on the clinical context where it is deployed. The decision support system may use its own terminology [156], in which case the data elements are the units of information at the level of granularity of this terminology.

   (a) Classify phrases as crisp or fuzzy.

   Disambiguating clinical rules can present a significant barrier to implementation [154]. The terms crisp and fuzzy are drawn from set theory. A set of objects is a fuzzy set if it is possible to have degrees of membership in that set. Likewise, a set is crisp if it is possible to unambiguously determine whether something is or is not a member of the set [164]. For the purposes of evaluating a clinical rule, the concepts in the rule are crisp if it is possible to unambiguously determine whether the rule applies and has been followed, based on mapping patient data to the concepts in the rule. Failure to recognize this ambiguity can lead to the caveat of needing to better define the concepts late in implementation. Preserving the original fuzzy terms (i.e. letting a clinician decide whether the rule applies) can undermine the intent of the decision support system in improving care.

   Using the same rule as in Example 2.1 and 2.2:

   Example 5.1:
   crisp: diabetes, statin therapy, age > 40
   fuzzy: overt CVD, CVD risk factor, lifestyle therapy

   (b) Classify crisp concepts as enumerated or not enumerated, and where possible, find crisp definitions corresponding to fuzzy concepts.(see Figure 5.2). A crisp

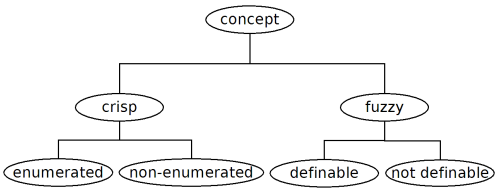

Figure 5.2: Classification of crisp and fuzzy concepts: A concept is crisp if it can be unambiguously determined whether something is included in the concept or not. A concept is enumerated if all relevant examples are listed. A concept is fuzzy if it is possible to have degrees of membership in that concept. It may or may not be possible to agree on a crisp definition of a fuzzy concept for the limited purpose of evaluating the rule.

concept is enumerated if all known examples are listed [164]. In this case, a concept is considered enumerated if all examples which are relevant to evaluation of the rule are listed. If a terminology is used, a concept is enumerated if it maps to

exactly one concept in the target terminology. If a terminology is not used, then this classification is subjective.

> **Example 5.2**: Diabetes may be considered enumerated if, in the logic of the decision support system, it is a single concept. More typically, the term "diabetes" would be considered crisp but not enumerated, because diabetes is commonly divided into types (I and II) and sometimes etiological subtypes.

In order to create a domain ontology, crisp definitions will need to be found which correspond to fuzzy concepts. The crisp definition may not encompass the whole of the fuzzy concept, but should encompass the part of the concept needed to evaluate the rule correctly. The same phrase may be defined differently for different rules. In the simplest case, the supplementary material of the rule set may provide a crisp definition. In other cases, the concept may be redefined for the limited domain of the rule. Ambiguity can represent a lack of evidence or consensus [155] and thus it may not be possible to agree on an unambiguous definition. A working definition may be reached by local consensus, but this should be updated regularly as new evidence emerges, and reassessed if the software is deployed in a new clinical setting.

> **Example 5.3**: "CVD risk factors" are listed in the supplementary text: dyslipidemia, hypertension, smoking, a positive family history of premature coronary disease, and the presence of micro- or macroalbuminuria. These terms will in turn need to be classified as crisp or fuzzy and defined as necessary. By contrast, "lifestyle therapy" is not defined in the text. A review of the evidence or expert consensus may be needed to agree on a definition of "lifestyle therapy" for the purpose of evaluating this rule.

The result of these steps is a restatement of the rule in terms of data elements:

> **Example 5.4**: diabetes [diabetes type I, diabetes type II] AND overt CVD [AMI, peripheral artery occlusive disease...] $\rightarrow$ statin therapy [atorvastatin, fluvastatin, lovastatin...] AND lifestyle therapy [weight counseling, stop-smoking...]
> OR
> diabetes [diabetes type I, diabetes type II] AND age > 40 AND CVD risk factors [dislipidemia [TC > 240, LDL > 160...]...] $\rightarrow$ statin therapy [atorvastatin, fluvastatin, lovastatin...] AND lifestyle therapy [weight counseling, stop-smoking...]

A clinical expert will need to advise on the definition of fuzzy concepts, and confirm that the rules are valid and maintain the intent of the original rule with their new vocabulary. A clinical expert can also advise on whether there are exceptions to the rule or other revisions needed in order to accurately evaluate the rule. These exceptions may also need to be specified in terms of data elements. If the clinical experts cannot agree on a crisp definition for all concepts in the rule, then it may not be possible to implement decision support based on the rule.

6. Determine relatedness of rules.

   By this stage the rules are defined in a crisp language and uniform grammar. The relationships between the domain concepts can now be mapped. Rules may be related if they share data elements, share a patient population, or require the same action from the clinician. Recognizing these convergence points and encoding them into the control-flow

structure avoids potential caveats such as disjointed workflow or multiple alerts for the same patient. If multiple rules result in the same recommendation, the clinician should still receive only one message with an appropriate explanation.

> **Example 6.1**: It can now be seen that the AMI rule from Example 1.1 overlaps with the diabetes rule from Example 2.1, in that they will both apply to patients with AMI. Other rules in the Joint Commission AMI set suggest discharge medications for the AMI patient, including aspirin. The rules should be linked to ensure that the patient gets a continuous, appropriate aspirin prescription, and that the clinician gets a single list of recommended discharge medications.

7. Determine availability of data elements.

   To this point, analysis has been independent of the local data systems. In this step, the results of the above analysis are applied to the local setting. The analysis described above allows quantification of the importance of a particular data element to the interpretation of the rule set. Thus, if that data element is not recorded, there can be a discussion with the clinicians as to whether separate entry of that data item is worth their time (Example 7.1).

   > **Example 7.1**: In a hypothetical hospital, all data elements for the diabetes rule in the example above are available except lifestyle therapy. After discussion with clinicians, a tab is added to the electronic patient record where lifestyle therapy can be documented.

Results of validation

LERM was validated with a set of 16 rules which were reserved for validation prior to developing the method. The results of this assessment are summarized in Table 5.3.

Without use of a terminology, the process took about 4 hours for one participant and about 6 for the other. Variation was noted in the proactive rephrasing of the rule. Upon discussion, the differences in results stemmed from different visions of *when* a rule would be evaluated. For example, one of the rules states that AMI patients should receive aspirin within 24 hours of arrival. One assessor envisioned a consulting system (suggests aspirin in the list of orders upon arrival), the other envisioned a critiquing system (issues a reminder if there is no aspirin after 24 hours). Other than differences resulting from this early departure, there was little divergence in the subsequent steps, except as anticipated in step 5. Errors were rare, with 7 errors noted (although one error was systematic, with the same concept omitted 3 times by each assessor). In all cases the source of error was omission of a phrase in a complex rule while restating it in conjunctive normal form. This experience suggests that the results of this step should be carefully checked for errors, or performed by two persons and the results compared.

Step 5 (concerning crisp and fuzzy concepts) was repeated using SNOMED-CT. For this purpose, a concept was considered crisp if it could be completely represented using SNOMED-CT concepts, and enumerated if it was represented in SNOMED-CT by a concept or a composition of concepts which had no child-concepts. The 87 phrases were represented by a minimum of 114 SNOMED-CT concepts. (SNOMED-CT concepts are not disjoint [165], thus there can be more than one way to represent a concept. The simplest representation was preferred.) As predicted, inter-user reliability improved considerably. However, this came at the cost of more time (about 4 additional hours) and an increased error rate: 17 errors were noted in this step, including 3 systematic errors. Most were due to the difficulty of determining whether a concept was truly absent from

Step	Result	Agreement
Determine whether rule can be proactively operationalized	12/16 rules need to be rephrased	agreement on which rules need to be changed, differences in how they were operationalized in 5/12 rules
Phrases containing data elements		minor variation, except those caused by step 1
Separate compound rules and restate in normal form	4/16 rules were compound	minor variation, except those caused by step 1.
Assess conflict between rules	potential conflict between 4/16 rules	same conflicts noted by both assessors
Check for unnecessary phrases	no unnecessary phrases	assessors agreed
Data elements	100 elements according to one assessor, 83 elements according to the other	13/100 and 2/83 represented unique concepts, the others were the same or minor variations
Crisp/fuzzy concepts	without terminology: of 67 classified the same: 60 crisp and 7 fuzzy	without terminology: 77.0% agreement ($\kappa = 0.315, p < 0.001$)
	with terminology: of 82 classified the same: 72 crisp, 10 fuzzy.	with terminology: 87.4% agreement ($\kappa = 0.760, p < 0.001$)
Crisp-enumerated / fuzzy-defined concepts	without terminology: 52 classified the same: 37 enumerated, 8 non-enumerated, 3 defined, 4 fuzzy-undefined.	without terminology: 58.6% agreement ($\kappa = 0.322, p < 0.001$)
	with terminology: of 76 classified the same: 14 enumerated, 52 non-enumerated, 3 defined, 7 fuzzy-undefined.	with terminology: 87.4% agreement ($\kappa = 0.763, p < 0.001$)
Relatedness of rules	each rule related to at least 1 other rule by sharing data; 4 share patients, 4 share recommendations	agreed on which rules shared data, populations, and recommendations

Table 5.3: Results of the validation of LERM.

SNOMED-CT. The assessors noted concepts which seemed to be as crisp as other SNOMED-CT concepts, but were omitted by chance from the terminology. For example, SNOMED-CT contains the concept "on admission" but not "at discharge." In total 15 such concepts were noted by both assessors as possible omissions, and subsequently classified as crisp concepts.

Discussion

To our knowledge, LERM is the first method for formalizing clinical rules for use in decision support, focusing on condition-action clinical recommendations rather than time-oriented guidelines. We present a step-by-step method for formalizing clinical rules, which has been validated with a sample of rules from diverse sources and clinical domains.

The most important limitation of this investigation is the inherent subjective nature of development through inductive methods. It is possible that a different team, using the same or different rules, would arrive at a different set of steps. That said, each step in the method can be linked to a potential caveat in developing the decision support system. Another limitation is the relatively small number of rules which were reserved for validation, leaving open the possibility that the randomly selected rules are not representative of all the rules in that set. Even so, validation allowed an assessment of patterns in the reliability, error rate, and time cost of formalization using the method. Formalization is recognized as a time- and labor-intensive process [155], but most studies do not report the time required for formalization, though one group reported a total implementation time of 7 weeks per rule [153].

Another well-recognized problem is disambiguation of natural language text [155]. Ultimately, concepts are sufficiently defined when they can be mapped to elements of the patient record and the resulting electronic representation of the rule generates correct results. However, directly mapping to a specific patient record would be detrimental to another important goal for decision support systems: interoperability [146, 166]. Thus LERM does not bind the formalization to a particular set of patient data until the last step, using the terms crisp, fuzzy, and enumerated to help define ambiguity for the intermediate steps.

The terms crisp and fuzzy are borrowed from set theory. The sets in this context are the set of conditions in which the rule should be evaluated, and then dividing those into disjoint sets where support should be offered or not (meaning the rule is "decidable" [167]). Thus "diabetes" in this context refers not to the disease, but to patients to which rules about diabetes should be applied. The onset of a disease such as diabetes is gradual, and the diagnosis is not always clear, but it is the task of doctors to disambiguate the patient's state. While there is no defined moment when a patient becomes diabetic, there is a defined moment when the patient is diagnosed with diabetes, and this is when rules about diabetes apply. Thus, although "diabetes" is fuzzy, the diagnosis of diabetes is a crisp concept.

If a terminology is used, terms can be unambiguously classified as enumerated or not. If no terminology is used, then classification into enumerated or not is subjective. This may be sufficient if the subjective definition is used consistently. If terms are not sufficiently specified, the work of specifying them is offloaded to the last step of the method, where the terms are mapped to actual patient data.

LERM suggests that fuzzy concepts should be mapped to crisp concepts. Implicit in this is the assumption that rules intend to make specific clinical recommendations for a specific, definable patient population. An alternative is to use fuzzy logic, where uncertainty is quantified in terms of degrees of truth. The response of the decision support system could then be changed based on the truth values. In cases where fuzzy concepts were used intentionally by the rule developers to allow for clinical judgment, this may be the best alternative. However, the impact of introducing this uncertainty into clinical decision support either directly (by including the level of certainty in the message to the user) or indirectly (by introducing a percentage of inaccurate messages) is beyond the scope of this investigation.

In order to implement a decision support system, knowledge of medicine and its intricacies must be combined with knowledge of computer reasoning to create a program which is formally valid and medically useful. This requires the cooperation of both clinical experts and experts in medical informatics [156]. To facilitate this cooperation, we chose to separate the steps of changing the grammar to a logical form and changing the vocabulary to data elements. Tools such as DeGEL (Digital electronic Guideline Library) [156], GEM (Guideline Elements Model) [168] or DELT/A (Document Exploration and Linking Tool / Add-ons) [169] could assist in maintenance and tracking of these changes.

After transformation, clinical rules are suitable for implementation as Medical Logic Modules (MLMs) [146], with data elements in the data slot (which are mapped to a local database). Rules which have been recomposed into clinical paths can be implemented using guideline implementation software such as Asbru, GLIF (GuideLine Interchange Format), PROforma, and Gaston [41, 155]. Composing the rules into paths is reserved as a final step, as connections between rules may not be apparent until the rules are fully specified.

The primary audience for LERM are others who wish to implement decision support based on clinical rules. We have developed an approach which has proven robust and reliable. Although intended for human users, a systematic approach such as LERM may also be informative for development of automated formalization [170]. In addition to its primary audience, LERM may also prove useful for improving the rules themselves. Guidelines are easier to follow when they are clearly specified [145, 155]. Similarly, clinical rules which use crisp terms, or acknowledge candidly when the evidence is not sufficient and clinical judgment is required, would likely be easier for clinicians to follow as well as easier to formalize. Often, clinical rules are developed in order to audit quality and provide performance feedback. Automated evaluation of rules to assess adherence could allow for continuous auditing and more frequent feedback. By omitting the first step of the method, LERM could also be applied to this task.

Thus far, this method has only been tested *in vitro*. Its use in developing decision support for use in a clinical trial is currently under investigation. The steps described here end with assessing the availability of data. As many hospitals are in transition from paper to electronic records, the result of this assessment will often be that key data elements are not available or are incomplete. Others have leveraged existing data to infer more information about the patient than is directly recorded in the system [149], but the impact of such inferences on the quality of decision support which can be offered is not yet known. Adapting the type of support offered to the quality of underlying data may affect the quality of support which is offered, and, in turn, the quality of care.

In summary, the Logical Elements Rule Method (LERM) is presented to assist in assessing clinical rules for their amenability to decision support, and formalizing the rules for implementation. The method was validated with a sample of clinical rules from diverse sources pertaining to a variety of clinical conditions. Validation showed that the method was robust and reliable between assessors, provided that it is agreed in advance what terminology will be used and when the rule will be evaluated. Use of a terminology increased inter-user reliability but also increased the error rate. Formalization by two persons, at least in step 5 (extraction of data elements), is recommended to ensure reliability. We envision an important role for LERM in this era of increasing attention to performance indicators and quality assurance. LERM can assist in transforming an indicator into an improvement in the everyday practice of medicine.