Designing Ontology-based Patterns for the Representation of the Time-Relevant Eligibility Criteria of Clinical Protocols

Christopher L. Crowe[1], Cui Tao, PhD[2]
[1]University of Virginia, Charlottesville, VA; [2]University of Texas School of Biomedical Informatics, Houston, TX

Abstract

The amount of time and money required to screen patients for clinical trial and guideline eligibility presents the need for an automated screening process to streamline clinical trial enrollment and guideline implementation. This paper introduces an ontology-based approach for defining a set of patterns that can be used to represent various types of time-relevant eligibility criteria that may appear in clinical protocols. With a focus only on temporal requirements, we examined the criteria of 600 protocols and extracted a set of 37 representative time-relevant eligibility criteria. 16 patterns were designed to represent these criteria. Using a test set of an additional 100 protocols, it was found that these 16 patterns could sufficiently represent 98.5% of the time-relevant criteria. After the time-relevant criteria are modeled by these patterns, it will allow the potential to (1) use natural language processing algorithms to automatically extract temporal constraints from criteria; and (2) develop computer rules and queries to automate the processing of the criteria.

Introduction

By testing the efficiency and credibility of emergent treatment options, clinical research plays a vital role in the advancement of medicine. However, these advancements are often delayed as studies frequently take longer than expected or are terminated due to the slow accrual of eligible patients[1]. Clinical trials require meticulous screening to ensure the study's outcome is not invalidated due to unintentional inclusion of patients who do not satisfy all eligibility criteria. As a result, such slow accrual rates may be attributed to the up to 1,554 hours of manual review required for completion of evaluations for a single study and up to $336.48 required to evaluate each patient[2, 3]. Likewise, while clinical guidelines present the best procedures for given conditions, physicians must first manually determine which guideline is applicable to each scenario, thus wasting valuable time that could otherwise be spent treating the patient[4].

Previous attempts to solve this issue have shown that the vast amount of patients' medical histories found in the electronic health records (EHRs) can be used to simplify the screening process by reducing the amount of manual work required[5-7]. For example, it has been shown that using computer-based systems to deal with clinical guidelines can lead to advancements in supporting physicians in the diagnosis and treatment of diseases[4]. In effect, the switch from human- to computer-based selection allows researchers to conduct screenings at greater speeds and with lower costs. Additionally, allowing clinical researchers to query the EHRs and automatically retrieve the total number of patients who meet all or partial eligibility criteria while their trials are still in the preliminary stages of planning may help them predict the feasibility of their study prior to its initiation[1]. With the promise of maximizing enrollment in clinical trials and decreasing the number of slow accrual terminations, the incentive to develop a tool that can efficiently use the EHRs to determine clinical trial eligibility is evident.

In recent years, several methods have been developed to evaluate data from the EHRs and much progress has been made. While some attempts have been successful and were specifically designed to analyze clinical trial eligibility criteria, there is still room for improvement. As an example, endeavors such as the E-Screening method created for the NIH-sponsored ACCORD clinical trial and the CTA system focused only on filtering out patients who did not meet basic criteria such as age or diagnosis requirements[5, 6]. These approaches helped to decrease the total number of patients who were screened manually, but failed to automate the entire process because more complex criteria, such as temporal requirements, were not covered and had to be evaluated manually. Also of significance is the EliXR-TIME temporal knowledge representation[8] and related works[9-11], which were successful in classifying the various types of temporal expressions and making progress towards the facilitation of temporal information extraction from free-text criteria. Alternatively, works that pertained to the analysis of EHR data but were not explicitly linked to clinical trial eligibility criteria are relevant but may require further investigation because these works have a broad focus that needs to be refined to be applied in this domain. Specifically, the overall aims of the eMERGE network deal with the general concept of phenotyping[12], and in recent works, methods for designing

phenotyping patterns for the eMERGE network have been developed[13]. These patterns are mainly designed for phenotyping algorithm development and do not specifically focus on the time aspects.

In this paper, we introduce our efforts on representing time-relevant criteria using an ontology-based approach. The importance of the temporal dimension of criteria can be seen from the fact that 38% of criteria contain temporal expressions, and yet these expressions are found in a free-text narrative format that is typically not amenable to computer processing due to their multi-dimensional complexities[8, 14]. Whereas non-temporal criteria may be met by a simple logical statement as to whether or not a condition (i.e. living in a specific city, being diagnosed with a disease, etc.) is satisfied, temporal criteria must first be broken down into components that can be related to one another. Although many research projects have focused on temporal relation modeling on EHR data[8, 13-16], our solution was tailored to the nuances of clinical trial eligibility criteria by examining only time-relevant criteria and designing generalized patterns to represent temporal constraints based on a review of these criteria. Furthermore, our solution is supported by an ontology, the Clinical Narrative Temporal Relation Ontology (CNTRO), which allows the patterns to be represented with respect to CNTRO[17]. This approach can enable formal representations of the criteria, semantic alignment to other domain ontologies, and direct leverage of semantic web querying technologies. Here we proposed the common basic patterns to represent time-related criteria. Using these common basic patterns, more complex patterns can be composed to represent different time-related criteria with respect to the CNTRO ontology. Based on these patterns, (1) natural language processing algorithms can be implemented to automatically extract the temporal constraints from the criteria; and (2) computer rules and queries can be composed to enable automatic processing of complex temporal eligibility criteria. Eventually, these advances can be used to create a streamlined, automated process for clinical trial patient screening, thus minimizing the time and cost required to conduct clinical research.

Methods

Selection of Representative Temporal Eligibility Criteria

A group of representative eligibility criteria was compiled using criteria from real clinical trials and guidelines that were randomly selected from the protocols stored in the clinicaltrials.gov and guideline.gov databases[18, 19]. Each protocol was manually reviewed by a single reviewer, and all criteria that contained time-relevant components were extracted. To be considered time-relevant, a criterion had to make use of at least one of the time classes defined by the EliXR-TIME model[8]: fixed duration, comparative duration, range duration, relative time interval, frequency constraint, temporal arithmetic expression, temporal logical expression, or Allen temporal relation[20]. After the initial list was made, the criteria were reviewed once again by the same reviewer and all criteria that were the same type of criteria as another were removed from the list. For example, criteria such as "at least 30 days since prior transfusion" and "at least 7 days since prior chemotherapy" were deemed similar because they followed the same relational pattern, so only one of the two was added to the list of representative criteria[17].

Parsing the Criteria Segments

Each criterion was deconstructed into basic temporal segments using a combination of one or more of the previously mentioned time classes defined by the EliXR-TIME model. For example, given "No corticosteroids used during the trial unless started at least 8 weeks prior to beginning of study," it was noted that this criteria was composed of 4 different segments: an atomic event (corticosteroids used; beginning of study), temporal logical expression (no corticosteroids used during trial), comparative duration (at least), and temporal arithmetic expression (8 weeks prior to beginning of study).

Identification of Temporal Patterns

The parsed segments were then examined and the general relationships used to link multiple segments together were noted. A set of patterns was then developed so that each criterion could be expressed by using one or more of the patterns. After defining the initial patterns, the criteria were reexamined and the coverage of the patterns was tested by determining if each criterion could be fully represented using the patterns. If it was found that no combination of the patterns was sufficient for a given criterion, the temporal relations used in that criterion were reassessed and a new pattern was designed. This process was repeated until the patterns were fully capable of representing all of the criteria in the training pool.

Results

In total, the eligibility criteria of approximately 400 trials and 200 guidelines were manually reviewed. The process of ensuring that each criterion in the representative set was unique and non-redundant led to a total of 37 sample criteria.

The iterative review process for the temporal pattern identification resulted in the definition of 16 patterns. A list of the patterns and example criteria from the representative set is shown below (Table 1).

Table 1. Temporal patterns and example criteria from the representative criteria. Many criteria are composed of multiple patterns, the specific segment that exemplifies the pattern is italicized.

PATTERN	EXAMPLE
Event (X) before/after fixed time instant (Y)	No *blood transfusions before 1990*
Event (X) before/after event (Y)	No *allergic reaction after taking drug of similar composition to EZN-2968*
Event (X) before/after start of interval (Y)	Patient must be a cancer survivor (defined as *cancer before start of trial*)
Event (X) before/after end of interval (Y)	No *disease progression* or *relapse after completion of high-dose chemotherapy*
Negation	*No* treatment until 4 weeks after the end of glucocorticoids treatment for CLL unless ≤ 10 mg of prednisolone /day
Exception	No ureteral obstruction before start of trial *unless* stent or nephrostomy tube has been placed
Compare Number of Occurrences	Veterans who report *at least 2* of the following 3 symptoms that began in 1990 or thereafter: fatigue, musculoskeletal pain, psychological symptoms
Exact Time Offset	Co-enrollment in therapeutic protocol allowed if begun at least 30 days following the week 20 immunization (*week 20 immunization is exactly 20 weeks after start of trial*)
At Least/Most Time Offset	*> 45 days* but *< 6 months* from completion of last treatment
Order of Events (X → Y → Z)	Test positive on 2 out of 3 *consecutive* samples
Compare Duration	Patients may not have received any cancer therapies *< 4 weeks or 5 half-lives (whichever is shorter)* of initiating study
Cumulative duration of events	No receipt of more *than 7 days (cumulative)* of prior antiretroviral therapy at any time prior to study entry, with the exception of zidovudine
Interval/event (X) has some occurrence at the same time as interval (Y) (i.e. Contain, During, Equal, Finish, Start)	No *concurrent therapeutic anticancer agents*
Event (X) repeated exactly at frequency (Y)	Those not performing moderate levels of activity for *exactly 30 minutes daily* on at least 5 days of the week
Event (X) repeated at least/most at frequency (Y)	No concurrent sodium fluoride at daily doses ≥ *5 mg per day*
Interval with start (X) and end (Y)	Stable weight - variation of less than 5 kg *over 3 months prior to screening*

It should be noted that there were fewer patterns defined than there were sample criteria because any number of patterns could be combined to create a number of unique criteria that exceeds the total number of patterns. For example, the criterion "test positive on 2 out of 3 consecutive tests" only requires the use of the pattern for *the order of events* to ensure each of the tests were consecutive. Similarly, the criterion "> 45 days but < 6 months from completion of last treatment" uses only the pattern for *a minimum or maximum time offset*. However, another criterion, "must have a 50% increase in PSA which is sustained for 3 consecutive observations obtained at least 1 week apart from each other," makes use of the pattern for *the order of events* and the pattern for *a minimum time offset*. In this example, 3 unique types of criteria were represented using patterns from a pool of only 2 patterns. This principle can be further applied to explain how 37 unique criteria can be formed using combinations of patterns from a pool of only 16 patterns.

Use of Patterns

For each defined pattern, we can use computer rules and queries to represent it. This can facilitate automatic executions of the criteria matching the patterns. In this paper, we aligned the patterns we developed to the CNTRO ontology. We then represented the patterns with respect to CNTRO using the Web Ontology Language Description Logic (OWL DL) and/or the Semantic Web Rule Language (SWRL)[21, 22]. The implementation of these patterns in

CNTRO will present the possibility of submitting queries to select all patients who have a medical history that satisfies the temporal patterns required for a given criterion. That is, once the rule has been defined, a reasoner can be used to return a list of all patients who meet the eligibility criteria contained in the rule. Examples of the OWL DL and SWRL representations of a few of the patterns can be seen below (Table 2).

Table 2. Sample OWL DL and SWRL implementations of select patterns.

PATTERN	DL/SWRL
Event (X) before/after fixed instant (Y)	Before: X(?e1), Patient(?p1), hasEvent(?p1, ?e1), hasNormalizedTime(?e1, ?t1), lessThan(?t1, Y) → ValidInstant(?e1)
	After: X(?e1), Patient(?p1), hasEvent(?p1, ?e1), hasNormalizedTime(?e1, ?t1), greaterThan(?t1, Y) → ValidInstant(?e1)
	ValidPatient = hasEvent some ValidInstant
Event (X) before/after start of interval (Y)	Y(?i1), hasStartTime(?i1, ?t1) → ValidStart(?t1)
	ValidInstant = X and before/after some ValidStart
	ValidPatient = hasEvent some ValidInstant
Exception	ValidPatient = (hasEvent some ValidInstant) or (not hasEvent some ValidInstant and hasEvent some ValidException)
Interval/Event (X) has some occurrence at the same time as interval (Y)	ValidInstant = X and some overlap some Y
	ValidPatient = hasEvent some ValidInstant

Discussion

Evaluation

An additional 50 trials and 50 guidelines were selected from clinicaltrials.gov and guidelines.gov to create a test set of 100 protocols. Within these 100 protocols there were 1,206 eligibility criteria, 408 of which were time-relevant criteria. Using the 408 time-relevant criteria, the coverage of the 16 patterns was evaluated by determining if the criteria could be represented using the patterns. The test indicated that the patterns were sufficient for 402 out of the 408 time-relevant criteria, for a coverage of 98.5%. Of the criteria that were not covered, a notable example includes criteria which indicated a patient must have a time event that recurs at a given frequency but allowed for variation as long as the end result is equivalent to the specified pattern. For example, one trial required that patients consume no more than 30 mg of prednisone per day or an equivalent variation. Clearly, a pattern could be used to express the requirement for a maximum consumption of a 30 mg pill one time every 24 hours. However, since the criteria allows the maximum to be defined by any other variation, another type of pattern is needed to represent that the criteria may also be satisfied by taking a 15 mg pill twice a day, a 10 mg pill three times a day, or any other combination that totals to 30 mg per day.

Limitations

It is important to note that the EHRs generally do not contain information explicitly stating the conditions a patient does not have. This, along with OWL's open world assumption, presents a problem in representing patterns, such as the *exception* pattern, which makes use of a *not hasEvent* statement. We can handle the negation or exception patterns by locating all the patients that are eligible for the condition in the negation or exception patterns, then exclude these patients from the eligible cohort. For example, for the criterion "no treatment in four weeks before trial", we will find all the patients who have had at least a treatment in four weeks before the trial, then exclude them from the trial cohort.

Additionally, it was found that CNTRO cannot handle one type of temporal information that appeared in the representative criteria. That is, there is no CNTRO class to represent a time duration that is characterized by a given number of chemical half-lives of a therapeutic drug (e.g. "Patients may not have received any cancer therapies < 4 weeks or 5 half-lives (whichever is shorter) of initiating study")[18]. Another type of information that would present issues for CNTRO were those that referred to a point in time as a negative number to represent the number of days the event occurs before another event. For example, it could be said that the patient may not have any surgeries after day -5, 5 days before the beginning of the trial begins. Although these types of criteria rarely appeared in the approximately 700 trials and guidelines that were reviewed (600 in the training pool and 100 in the test set), their addition is necessary to extend coverage to all types of criteria.

Conclusion and Future Work

Based on time-relevant criteria in clinical trials and guidelines, we developed a set of patterns that can be used to represent eligibility criteria. The 16 defined patterns were found to sufficiently cover 98.5% of the time-relevant eligibility criteria in the test set of 100 clinical protocols. The patterns presented in this paper are fairly comprehensive.

There are several future directions we would like to pursue. First, the defined patterns will be reviewed and evaluated by more domain experts to ensure a more scalable coverage. Second, we will explore how to use natural language processing technologies to automatically classify eligible criteria and clinical guidelines to these patterns and extract the important segments. Finally, we will use semantic-web-based technologies to make the criteria automatically executable.

References

1. Eriko S, Satoshi T, Keiichi Y, Motohiko S, Kenya Y, Masayuki Y. The correlation between the number of eligible patients in routine clinical practice and the low recruitment level in clinical trials: a retrospective study using electronic medical records. *Trials*. January 10, 2013;14(1):2-19.
2. Penberthy L, Dahman B, Petkov V, Deshazo J. Effort Required in Eligibility Screening for Clinical Trials. *Journal Of Oncology Practice*. November 2012;8(6):365-370.
3. Penberthy L, Brown R, Puma F, Dahman B. Automated matching software for clinical trials eligibility: measuring efficiency and flexibility. *Contemporary Clinical Trials*. 2010;31:207-217.
4. Anselma L, Terenziani P, Montani S, Bottrighi A. Automatic treatment of temporal issues in clinical guidelines in the GLARE system. *Studies in Health Technology and Informatics*. 2007;129(Pt 2):935-940.
5. Thadani S, Weng C, Wajngurt D. Electronic screening improves efficiency in clinical trial recruitment. *Journal of the American Medical Informatics Association*. December 2009;16(6):869-873.
6. Embi P, Jain A, Harris C. Development of an electronic health record-based clinical trial alert system to enhance recruitment at the point of care. *AMIA Annual Symposium Proceedings.*2005;231-235.
7. Weng C, Kahn M, Gennari J. Temporal knowledge representation for scheduling tasks in clinical trial protocols. *AMIA Annual Symposium Proceedings*. 2002;879-883.
8. Boland MR, Tu SW, Carini S, Sim I, Weng C. EliXR-TIME: a temporal knowledge representation for clinical research eligibility. *AMIA Joint Summits on Translational Science Proceedings*. 2012;71-80.
9. Pustejovsky P, et al. TimeML: robust specification of event and temporal expressions in text. *AAI Spring Symposium on New Directions in Question-Answering*. 2003;28-34.
10. Hripcsak G, Zhou L, Johnson S. Modeling electronic discharge summaries as a simple temporal constraint satisfaction problem. *Journal of the American Medical Informatics Association*. 2005;55-63.
11. Weng C, Tu S, Richesson R. Formal representations of eligibility criteria: a literature review. *Journal of Biomedical Informatics*. 2010;43(3):451-467.
12. Shahar Y. A framework for knowledge-based temporal abstraction. *Artificial Intelligence*. 1997;90:79-133.
13. Rasmussen L, et al. Design patterns for the development of electronic health-record driven phenotype extraction algorithms. *Journal of Biomedical Informatics*. 2014.
14. Luo Z, Johnson S, Weng C. Extracting temporal constraints from clinical research eligibility criteria using conditional random fields. *AMIA Annual Symposium Proceedings*. 2011; 843-852.
15. McCarty C, Chisholm R, Wolf W. The eMERGE network: a consortium of biorepositories linked to electronic medical records data for conducting genomic studies. *BCM Medical Genomics*. 2011;4:13.
16. Weng C, Wu X, Johnson S. EliXR: an approach to eligibility criteria extraction and representation. *Journal of the American Medical Informatics Association*. 2011;i116-i124.
17. Tao C, Solbrig H, Chute C. CNTRO 2.0: a harmonized semantic web ontology for temporal relation inferencing in clinical narratives. *AMIA Joint Summits on Translational Science Proceedings*. 2011;64-68.
18. http://www.clinicaltrials.gov
19. http://www.guideline.gov
20. Allen, J.F., An interval-based representation of temporal knowledge, *Proceedings International Joint Conference on AI*. 1981, 221-226.
21. OWL web ontology language reference. http://www.w3.org/TR/owl-ref.
22. SWRL: a semantic web rule language combining OWL and RuleML. http://www.w3.org/Submission/SWRL