EliXR-TIME: A Temporal Knowledge Representation for Clinical Research Eligibility Criteria

Mary Regina Boland[1], BS, Samson W. Tu[2], MS, Simona Carini[3], MA
Ida Sim[3], MD, PhD, Chunhua Weng[1], PhD
1. **Department of Biomedical Informatics, Columbia University, New York, NY, 10032**
2. **Center for Biomedical Informatics, Stanford University, Stanford, CA, 94305**
3. **Department of General Internal Medicine, UCSF, San Francisco, CA, 94107**

Abstract

Effective clinical text processing requires accurate extraction and representation of temporal expressions. Multiple temporal information extraction models were developed but a similar need for extracting temporal expressions in eligibility criteria (e.g., for eligibility determination) remains. We identified the temporal knowledge representation requirements of eligibility criteria by reviewing 100 temporal criteria. We developed EliXR-TIME, a frame-based representation designed to support semantic annotation for temporal expressions in eligibility criteria by reusing applicable classes from well-known clinical temporal knowledge representations. We used EliXR-TIME to analyze a training set of 50 new temporal eligibility criteria. We evaluated EliXR-TIME using an additional random sample of 20 eligibility criteria with temporal expressions that have no overlap with the training data, yielding 92.7% (76 / 82) inter-coder agreement on sentence chunking and 72% (72 / 100) agreement on semantic annotation. We conclude that this knowledge representation can facilitate semantic annotation of the temporal expressions in eligibility criteria.

1. Introduction

Eligibility criteria are essential to every clinical research study of human subjects. They specify the characteristics of study participants and provide a checklist for screening and recruiting those participants. A computable representation of eligibility criteria can significantly accelerate electronic screening of clinical research study participants and improve research recruitment efficiency.[1] Although 38% of eligibility criteria contain temporal expressions[2], the typical free-text narrative format of these expressions is not amenable to computer processing. A knowledge representation (KR) for temporal expressions is needed to facilitate temporal information extraction from and representation of free-text eligibility criteria and to enable automatic formulation of temporal eligibility queries of electronic patient information.[2,3] Despite a plethora of existing general and clinical temporal KRs,[3-14] particularly for clinical narratives and clinical research protocols, their reusability for clinical research eligibility criteria remains unknown.

This study was designed to reuse existing temporal KRs as appropriate and to adapt or extend them to structure temporal expressions in clinical research eligibility criteria through semantic annotation. We (1) assessed representative temporal KRs for clinical narratives and clinical research protocols and (2) designed a frame-based temporal knowledge representation for temporal expressions in clinical research eligibility criteria called EliXR-TIME, which is sharable on the Protégé (version 3.4.6) platform.[15] This paper presents the design and evaluation results for EliXR-TIME.

2. Method

We implemented a 6-step procedure to model the temporal expressions in eligibility criteria. First, we sampled 100 eligibility criteria from ClinicalTrials.gov[16] to derive the KR requirements. We then surveyed a few representative temporal KRs and compared them with our knowledge representation requirements. On this basis, we reused the applicable top-level semantic types from existing temporal knowledge representations to annotate a training set of 50 criteria with temporal expressions selected from ClinicalTrials.gov.[16] We randomly selected these 50 criteria using both keyword search (i.e. "years", "weeks", "days") to locate eligibility criteria containing temporal expressions and manual review to ensure that the criteria retrieved were not entirely composed of simple temporal expression phrases, e.g., *6 months of chemotherapy*. Also, we removed age criteria, e.g., *6-12 years old*. We manually decomposed these 50 training criteria into sentence segments, labeling each with the initial set of semantic types. We then further annotated each sentence segment into smaller segments through an iterative process until each segment became a semantic type. Throughout this iterative process, we identified the atomic semantic types for each sentence segment and organized these semantic types into hierarchies. To maximize knowledge reuse, we reused class names from previous knowledge representations wherever possible as long as they had the same meaning. We also reexamined the instantiation results for the 50 training criteria and removed rarely used or

confusing classes or attributes and their definitions. We repeated this process until the model stabilized. We called this temporal KR EliXR-TIME. Finally, we evaluated the "fitness for use" of EliXR-TIME by having two human raters independently encode another 20 temporal eligibility criteria obtained from ClinicalTrials.gov in the same manner as the training set.

3. Results

3.1 Knowledge Requirements for Eligibility Criteria

We annotated 100 eligibility criteria from ClinicalTrials.gov as a bag of UMLS-recognizable concepts, including temporal concepts.[17] We then manually reviewed their semantic patterns, yielding a set of 11 KR requirements (Table 1).

Table 1. Knowledge Representation (KR) requirements for eligibility criteria

KR Requirement	Example
Relative events	*Must initiate hormonal therapy including ovarian suppression at least 4 weeks prior to initiation of vaccinations*
Relative time interval	*Past six months*
Temporal patterns	*Every 5 hours per day*
Comparison operators	*More than, less than, equal to*
Temporal relations	*After, before, during, within, onset, until.*
Conjunction	*And, or*
Logical temporal expression	*Evidence of active substance abuse during 30 days prior to entry into trial*
Arithmetic temporal expression	*30 days prior to entry into trial*
Combinatory temporal expressions that modify a single Event or Anchor	*More than 28 days but within 12 weeks prior to enrollment*
Recursive or hierarchical representation of complex temporal expressions	*Chronic administration (more than 14 days) of systemic high dose immunosuppressant drugs during a period starting from six months prior to administration of the vaccine and ending at study conclusion*
Intrinsic duration or frequency of events	*Chronic administration (defined as more than 14 days)*

3.2 Comparison of Temporal Expressions in Clinical Texts

Next, we compared temporal expressions among different types of clinical texts (e.g., in-patient clinical narratives, clinical trial study calendar, and clinical trial protocols) along two dimensions.

The first dimension is representation granularity as measured by the size of a representation unit. For example, "*within 6 months of past surgery*" can be represented as a duration ("*6 months*") associated with an event ("*past surgery*") in a model with coarse granularity. However, to more precisely represent the meaning of this temporal constraint, the anchor point ("*surgery*") should be related to the duration ("*6 months*") by a time lag or direction modifier ("*past-before*").   Because it is composed of smaller representation units, we consider the second representation more granular.

The second dimension is reference time. Temporal expressions in different clinical texts assume different contexts and implicit reference times. For example, clinical narratives use the context of patient care activities that are relative to observational times or episodic time, such as *from start of pneumonia* and *from hospital discharge*. An example temporal expression in clinical narrative is *today the patient's toe hurt*,[11] where the date or time is relative to the observation and documentation time. Such expressions are common in clinical notes because they are observations by individual care providers. In contrast, they are not found in clinical trial protocol eligibility criteria because these are generic instructions for clinical researchers.[11] Clinical trial protocols use the context of research activities that are relative to protocol starting time (e.g., enrollment, visit 1). In a clinical trial study calendar, time is usually relative to the date of consent or randomization. In eligibility criteria, the implicit reference time is often the time of eligibility determination, which may be different from time of enrollment and usually different from time of first visit. Event-dependent temporal expressions that refer to a research event as an anchor[11] are important in eligibility criteria because they constitute the bulk of the temporal expressions. An example is *at least 4 weeks prior to initiation of vaccinations*. The duration *at least 4 weeks* is relative to the event *initiation of vaccinations* via the temporal relation *before.*

3.3 Identification of Applicable Temporal Entities

Temporal KRs have been primarily developed for processing clinical narratives and clinical trial study calendars. We selected two representative temporal KRs for these texts to analyze their generalizability to clinical research eligibility criteria. We found that the clinical narrative Temporal Constraint Structure (TCS) was most similar to our KR requirements while the generic markup language for temporal expressions (TimeML) was only partially applicable.

The TCS for extracting temporal information from clinical narratives consists of 10 fields: event_point, anchor, anchor_point, anchor_modifier, relation, time_unit, quantity, direction, interval_operator, and vagueness.[11,13] All are necessary but insufficient to represent temporal expressions in eligibility criteria because the TCS does not represent recurrent temporal patterns explicitly.[11] Another related technology is the recently developed Clinical Narrative Temporal Relation Ontology (CNTRO 1.0).[7,8] The TCS was better suited to our representational requirements (Table 1) because it represents the event and anchor separately. Other KRs, such as that developed by Weng et al.[9], the Epoch model[6], and the temporal representation of the Knowledge-Based Temporal Abstraction theory, rely heavily on time-stamped data for representation of events thus were not considered for adoption.

We also analyzed TimeML,[18] which defines four major entities as part of its temporal specification language: EVENT, TIMEX3, SIGNAL, and LINK. There are seven Event types: Occurrence, State, Reporting, I-Action, I-State, Aspectual, and Perception. The type I-State is used for intentional states, such as *feel, love, hope, believe,* and *suspect*.[19] The temporal information contained within such statements is needed for a complete temporal specification language. However, this level of detail is not necessary for representing the temporal expressions in eligibility criteria; therefore, we only adopted the general event semantic type from TimeML. The anchor semantic type was separated from the event type because most events in eligibility criteria are relative to anchors. This distinction of event and anchor is a difference between our knowledge representation requirements and the features offered by TimeML.[18]

3.4 The EliXR-TIME Knowledge Representation

This research is part of the **Eli**gibility Criteria E**x**traction and **R**epresentation (EliXR) project;[20] therefore, our knowledge representation was named EliXR-TIME. It is designed to be an interval-based model where every object (e.g., event or anchor) is an interval. TCS[11] and TimeML[18] are also interval-based KRs but depend on time-stamped information. Because EliXR-TIME is also made available as a fr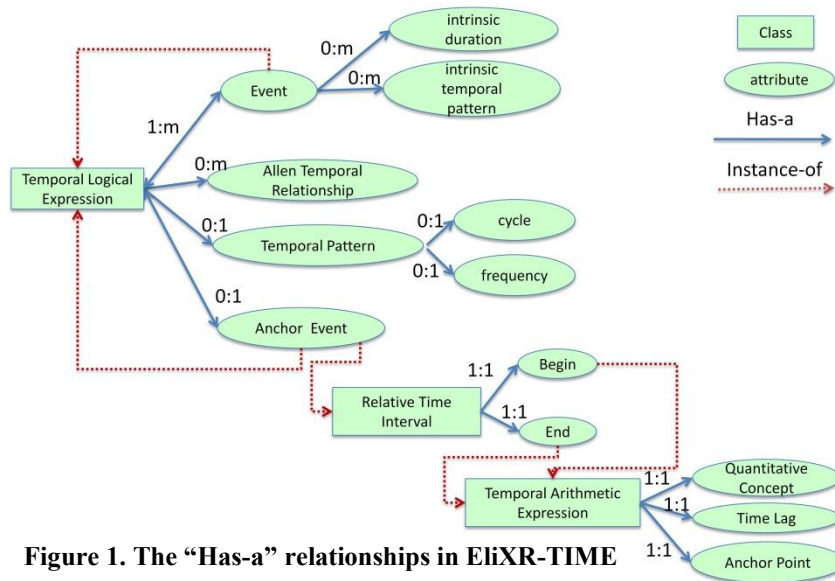ame-based knowledge representation to support semantic annotation, its construction and usage is closely coupled with natural language processing (NLP) considerations for structuring free-text eligibility criteria. Table 2 shows the definitions for the classes and attributes. (Appendix Table 1, accessible online at http://people.dbmi.columbia.edu/~chw7007/2012CRI_Appendix.htm, shows a comparison among related classes in EliXR-TIME, TCS, and TimeML.[18]) The EliXR-TIME Allen Temporal Relation class (Table 2) uses Allen's formalism for Interval Algebra (13 relationships) in order to represent the various types of temporal relations found in eligibility

Figure 1. The "Has-a" relationships in EliXR-TIME

criteria.[21,22] Figure 1 shows the "has-a" hierarchy for EliXR-TIME.

Importantly, each criterion at the top-level must contain a Temporal Logical Expression (TLE) that returns a Boolean value because each eligibility criterion is a statement that evaluates to true or false. Within this TLE, other imbedded temporal expressions can exist, such as another TLE, a Temporal Arithmetic Expression (TAE), or an event. Each TLE contains the following slots: event, Allen temporal relation, temporal pattern, and anchor. Each event can have an intrinsic duration and an intrinsic temporal pattern. Each pattern can have a cycle and a frequency specification. An anchor can be either a temporal expression (logical or arithmetic) or a relative time interval (e.g., *the past 6 months*). Figure 2 illustrates a common temporal constituent breakdown at the class level with the instantiation shown in *italics*. This example contains a top-level TLE with a relative time interval functioning semantically as the anchor and a TAE as the start of that interval.

Table 2. The class hierarchy and definitions for EliXR-TIME

| Class	Subclass	Definition	Examples
Event	Atomic Event	An occurrence specified by a noun-phrase representing a time interval. Both of the semantic annotation labels event and anchor can be instances of this class	*Expected survival, active inflammation, had a baby*
	Combinatory Events	A group of events related via a conjunction or disjunction	*History of alcohol or drug abuse*
Duration	Fixed	A duration whose upper and lower limits are both clearly stated and do not form a range but a period	*Jan. 20th, 2006, 3 weeks, 2 years, 11:00 am*
	Comparative	A duration with one clearly specified upper or lower limit coupled with a comparison operator	*Less than 6 weeks*
	Range	A duration bounded on both sides by upper and lower limits (maximum, minimum)	*More than 28 days but within 12 weeks*
Relative Time Interval	-------	An interval with a clearly defined begin point and end point that is relative to some medical occurrence	*A period starting from 6 months* [begin point] *prior to administration of the vaccine* [end point]
Allen Temporal Relation	-------	The 13 Allen temporal relations	*After, before, during, equals, meets, finishes, overlap, starts*
Frequency Constraint	-------	A temporal pattern characterized by a recurring event of a specified duration. The name is based in part on work taking place on a Generalized ERGO Annotation	*Greater than three stools per day* [recurring event] *for greater than 7 days* [duration]
Temporal Arithmetic Expression	-------	An arithmetic expression that returns an instantaneous interval (similar to a time-point) using a calculation. *After* and *before* represent a time lag (+ or -) rather than an Allen temporal relation	*4 weeks before study treatment; 6 months prior to administration of the vaccine*
Temporal Logical Expression	-------	A temporal comparison expression that returns a Boolean result. It uses Allen's temporal relations where *after* and *before* are comparison operators and not time lags	*Administration or planned administration of immunoglobulins and/or any blood products **during** a period*
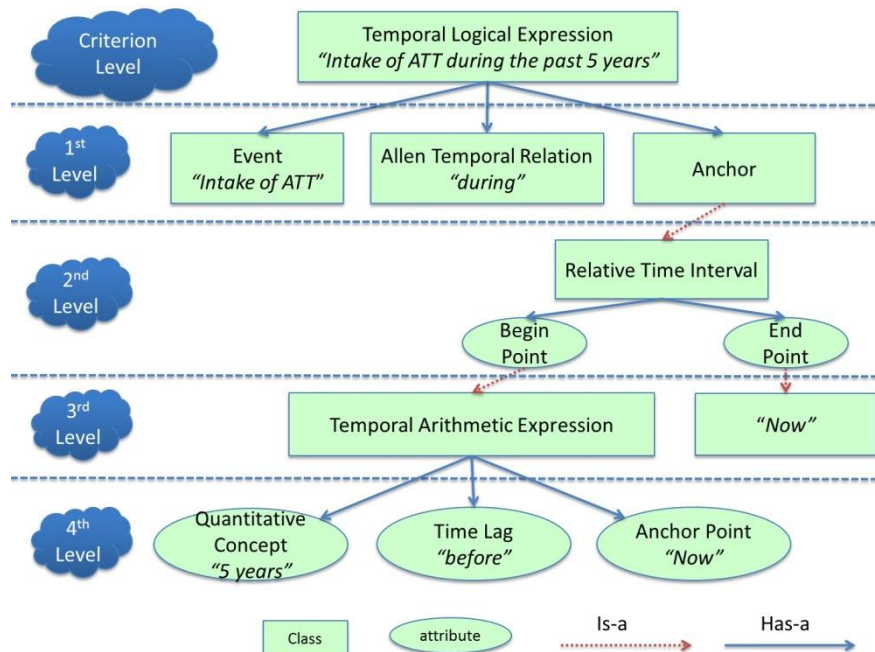
Figure 2. The hierarchical annotation of an eligibility criterion using EliXR-TIME

The BNF syntax for EliXR-TIME is shown below.

Temporal Logical Expression: \<name\>\<event\>$^+$[\<event-conjunction\>]\<temporal-relation\>$^+$ [\<temporal-pattern\>]\<anchor\>
event: \<Atomic_Event\> | \<Combinatory_Event\> | \<Temporal_Logical_Expression\>|\<Temporal_Arithmetic_Expression\>
event-conjunction: \<"AND"\> | \<"OR"\>
temporal-relation: \<Allen_Temporal_Relation\>
temporal-pattern: \<Frequency_Constraint\>
anchor: \<Event\> | \<Temporal_Arithmetic_Expression\> | \<Relative_Time_Interval\> | \<Duration\>

Temporal Arithmetic Expression: \<name\>\<quantitative_concept\>\<time_lag\>\<anchor_point\>
quantitative_concept: \<Duration\>
time_lag: \<"BEFORE_-"\> | \<"AFTER_+"\>
anchor_point: \<Atomic_Event\> | \<Combinatory_Event\> | \<Temporal_Logical_Expression\>

Atomic_Event: \<name\>\<intrinsic-duration\>{\<intrinsic-duration\>}[\<conjunction-between-durations\>]\<intrinsic-temporal-pattern\>{\<intrinsic-temporal-pattern\>}[\<conjunction-between-patterns\>]
intrinsic-duration: \<Duration\>
conjunction-between-durations: \<"AND"\> | \<"OR"\>
intrinsic-temporal-pattern: \<Frequency_Constraint\>
conjunction-between-patterns: \<"AND"\> | \<"OR"\>

Combinatory_Event :
\<name\>\<atomic_event\>$^+$\<conjunction_between_events\>[\<intrinsic_duration\>][{\<intrinsic_duration\>}][\<conjunction-between-durations\>][\<intrinsic-temporal-pattern\>][{\<intrinsic-temporal-pattern\>}][\<conjunction-between-patterns\>]
atomic_event: \<Atomic_Event\> | \<Temporal_Logical_Expression\>
conjunction_between_events: \<"AND"\> | \<"OR"\>

Relative_Time_Interval : \<name\>\<begin_point\>\<end_point\>
begin_point: \<Temporal_Arithmetic_Expression | Event | Fixed\>
end_point: \<Temporal_Arithmetic_Expression | Event | Fixed\>

Frequency_Constraint: \<name\>\<recurrence_cycle\>\<lasting_duration\>
recurrence_cycle: \<Duration\>
lasting_duration: \<Duration\>

Duration: \<Comparative\> | \<Fixed\> | \<Range\>
Comparative: \<name\>\<comparative_relationship\>\<base_interval\>
comparative_relationship: \< "at_least" | "at_most" | "equal_to" | "less_than" | "more_than"\>
base_interval: \<Fixed\>

Fixed: \<Date\> | \<Num_Unit\> | \<Time\>
Date: \<name\>\<date_value\>
date_value: \<string\>
Num_Unit: \<name\>\<number\>\<time_unit\>
number: \<Integer\>
time_unit: \<"Day" | "Month" | "Year" | "Week" | "Hour"\>
Time: \<name\>\<time_value\>
time_value: \<string\>
Range: \<name\>\<max_duration\>\<min_duration\>
max_duration: \<Duration\>
min_duration: \<Duration\>

Allen_Temporal_Relation: \<After\> | \<Before\> | \<During\> | \<During_inverse\> | \<Equals\> | \<Finishes\> | \<Finished_by\> | \<Meets\> | \<Met_by\> | \<Overlaps\> | \<Overlap_inverse\> | \<Starts\> | \<Started_by\>
After: \< "after"\> | \< "since" \>
Before: \< "before"\> | \<"prior_to"\> | \<"preceding"\> | \< "past"\> | \<"prior"\> | \<"last"\> | \<"ago"\>
During: \< "during"\> | \<"within"\>

During_inverse: < "during_inverse">
Equals: < "equal_to">
Finishes: < "ends_with"> | <"until"> | < "ending">
Finished_by: < "finished_by">
Meets: < "meets">
Met_by: < "met_by">
Overlaps: < "overlaps">
Overlap_inverse: < "overlap_inverse">
Starts: < "begins_with"> | < "onset"> | < "starting">
Started_by: <"started_by">

3.5 Descriptive Statistics

Table 3 shows the distribution of the semantic annotation labels among the sentence constituents in the 50 training criteria. A TLE usually contains three annotation labels (temporal patterns are rare): event, relation, and anchor. A TAE contains time lag, quantitative concept, and anchor point. The classes used to describe a TLE's annotation labels are highly variable (Table 3), whereas a TAE's quantitative concept varies only among the Duration subclasses. Sixty-nine percent of sentence segments labeled as *event* are atomic events and 28% are another TLE. Eighty-seven percent of sentence segments labeled as *anchor* are relative time intervals and 11% are another event. Ninety-two percent of relationships between an anchor and an event are "during", 6% are "before", and the remaining 2% are "after". Most sentence segments labeled as quantitative concept are of fixed duration while 19% are comparative duration. This demonstrates that one semantic annotation label, corresponding to the natural language text, can evaluate to multiple EliXR-TIME classes. Also, we found only three Allen temporal relations in the training corpus. These mappings are essential for proper extraction and representation of the information contained within each criterion.

Table 3. Distribution of semantic annotation mappings in the 50 training criteria

Semantic Annotation Label	Protégé Class	Percent (N)
Event	Atomic Event	69 (87)
	Combinatory Event	2 (87)
	Temporal Logical Expression	28 (87)
	Temporal Arithmetic Expression	1 (53)
Anchor	Relative Time Interval	87 (53)
	Event	11 (53)
	Temporal Arithmetic Expression	1 (53)
Allen temporal relation	During	92 (53)
	Before	6 (53)
	After	1 (53)
Quantitative concept	Duration, Fixed	81 (54)
	Duration, Comparative	19 (54)
Anchor point	Event	100 (54)

3.6 Evaluation

3.6.1 Sentence Segmentation

Two raters independently annotated 20 additional temporal eligibility criteria to evaluate the suitableness of EliXR-TIME for semantic annotation. Before testing the coverage of EliXR-TIME, each rater was acquainted with rules (see Appendix) regarding the usage of EliXR-TIME and the training set of 50 instantiated criteria. During the evaluation, each criterion was first chunked into sentence segments. For example, segmenting the criterion *Laboratory confirmed influenza disease within 6 months* yields (1) *Laboratory confirmed influenza disease,* (2) *Within,* and (3) *6 months.* Our measurements include inter-rater agreement for sentence segments generation (or sentence chunking) and semantic annotation labeling for the generated sentence segments. One rater (CW) generated 79 sentence segments and the other (MB) generated 80. The union set included 82 segments containing 100 temporal constituents. Inter-rater agreement for sentence segmentation was 92.7% (76 / 82). Four criteria contained six segmentation discrepancies and we analyzed the reasons for the discrepancies (Appendix Table 2).

Difficulties in interpreting implied information resulted in two sentence chunking discrepancies. One rater (MB) failed to represent the implied duration of *currently* in the criterion *Patients currently on stable ART (anti-retroviral therapy) for at least 12 weeks, who need to change their ARV regimen because it is currently failing, with a viral load of > 1000 copies/mL.* Because of the modifier "currently*", the TLE should have been event = *on stable ART,* Allen temporal relation = *during,* and anchor = *now.* Some of the differences in sentence segmentation resulted

from different interpretations of the criterion. For the criterion *Willingness to have blood stored for up to 10 years for use in additional assays to evaluate immune responses to influenza or the alphavirus vector if such assays become available*, one rater (MB) broke the sentence into three segments: 1. *Willingness to have blood stored,* 2. *During*, and 3. *Interval now up until 10 years from* now. The other rater (CW) broke the sentence into two segments: 1. *Blood stored* and 2. *Up to 10 years.* The meaning of these two segmentations is different. The first (MB) represents the "up to 10 year" interval as starting today, independent of when the blood is stored; for instance, if the blood is stored 2 years from now, then the duration would be only 8 years (10 – 2 years). In other words, this rater believed that the phrase *willingness to have blood stored for up to 10 years* referred to the patient's willingness starting now and lasting up to 10 years from now. The other rater (CW) was correct because blood cam be stored for no more than 10 years.

3.6.2 Semantic annotation

Each criterion sentence segment was instantiated into the Protégé-based EliXR-TIME. Exact agreement of temporal constituents was 72.0% (72 / 100). Of the 28 temporal constituents involved in a discrepancy (Table 4), 28.6% consisted of semantically equivalent modeling differences. As an example such a difference for representing the segment *at least 14 days*, one rater (CW) set the length_comparison_operator = ">=" and the duration = "*14 days*" (a Fixed Duration); whereas, the other (MB) set the duration = "*at least 14 days"* (a Comparative Duration). We removed slots that were redundant and that caused the raters to produce syntactically different but semantically equivalent representations of the same sentence segment (Appendix Table 3).

Table 4. Distribution of annotation discrepancies among the semantic labels

Types of representation discrepancies	Frequency (%, N = 28)
Different syntactically but equivalent semantically	28.6
Implied past tense	21.4
Discrepancies caused by differences in segmentation	17.9
Implied duration	10.7
Error in rater's understanding of criterion's meaning	10.7
Event vs. Relative Time Interval to represent *treatment period*	3.6
Incorrect Allen temporal relation (*before* vs. *finishes*)	3.6
Relative Time Interval vs. Range Duration for *1-2 weeks*	3.6

Differences in representing the implied past tense accounted for 21.4% of modeling discrepancies. For the sentence segment *within 6 months*, there is an implied past tense and therefore it should be represented as *during* the *past 6 months*, where the *past 6 months* is a relative time interval. One rater (MB) represented *within 6 months* as *during 6 months*, with *6 months* as an instance of Fixed Duration. While this type of discrepancy is not caused by EliXR-TIME, it illustrates the semantic complexity of eligibility criteria and the difficulties in inferring an implied context even among human annotators. Another 10.7% of the discrepancies were the result of missing an implied duration. For example, *during now* had to be logically inferred from one criterion and in another *within 2 weeks* had to be inferred even though the sentence only stated *2 weeks*. In total, 32.1% of the discrepancies resulted from errors in understanding implied information either past tense or duration.

3.6.3 Temporal expression not represented by EliXR-TIME

When performing the evaluation we encountered only one temporal expression that was not handled by EliXR-TIME. The criterion was *Treatment with an investigational drug within 4 weeks or 5 half-lives, whichever is longer, before first study dose.* The temporal expression *5 half-lives* refers to the chemical half-life of the investigational drug and is a temporal period specific for a particular drug or medication. EliXR-TIME does not support this type of temporal expression.

4. Discussion

4.1 Comparison to TCS, TimeML, and ERGO

No existing KR for temporal expressions met our requirements. We developed EliXR-TIME based upon reusing existing temporal KRs and named the classes and slots based on their functional and semantic types. We provide a further comparison of EliXR-TIME to related temporal schemas or ontologies below.

The TCS did not facilitate recursive temporal patterns in clinical text.[11] In EliXR-TIME, we have a Frequency Constraint class that supports expressions such as *daily or 5 days per week.* For a clinical note, the TCS evaluated not only temporal information in the note, e.g., *6 days before a visit*, but the date and time of the note itself. However, in eligibility criteria, the actual time-stamped date of *visit* will not be known until each individual patient's

record is queried. Therefore, *visit* must be explicitly defined in order to perform these short time-interval calculations; this was handled by our TAE class.

TimeML is a highly detailed temporal specification modeling language. EliXR-TIME differs from TimeML by emphasizing the integrity of a semantic unit. For example, a sentence segment from TimeBank 1.2 is "*at least 30 days before closing the purchase*". The TimeML annotation does not generally represent it as a single semantic unit but groups "*at least 30 days*" into a TIMEX3 expression, while "*closing*" and "*purchase*" are labeled with type EVENT; "*before*" would be represented as a SIGNAL and a TLINK would link a temporal relation to the EVENT. In EliXR-TIME, we label the entire phrase as a single TAE with quantitative_concept = *at least 30 days*, time_lag = *before*, and anchor_point = *closing purchase*. This allows the construction of a hierarchical representation of the entire phrase. In contrast, TimeML only permits annotations of individual terms connected by SIGNALs or LINKs and not annotations of an entire phrase. Furthermore, TimeML is designed to capture temporal information related to types of statements, such as those that use intention verbs like *feel, love, hope, believe,* and *suspect.* However, these types of statements are not relevant for eligibility criteria because inclusion and exclusion criteria state conditions, diseases, actions, etc. that a prospective patient has had in the past. In EliXR-TIME, we sought to focus only on representing the semantic types that were necessary for annotating the information contained within the text. We omitted unnecessary semantic types in an attempt to balance expressiveness and tractability for knowledge representation.

Another related project is the Eligibility Rule Grammar and Ontology (ERGO)-annotation.[2] ERGO-annotation[2] does not represent temporal information in eligibility criteria, though its extension, a Generalized ERGO Annotation, is under development that does include temporal constraints on the main noun phrase of an eligibility criterion. EliXR-TIME differs from ERGO in that EliXR-TIME focuses on defining the semantic types (e.g., events and anchors) for sentence segments in temporal eligibility criteria and the combination patterns, or frame-based templates, of these semantic types. In contrast, ERGO defines the constraint types logically, which often requires intelligent translation or mapping from sentence segments to constraint types.

4.2 Limitations

This study has two limitations. First, we only instantiated a relatively small sample of eligibility criteria containing temporal expressions, with a data set of 70 temporal eligibility criteria: 50 in the training set and 20 in the evaluation set. However, each of the 70 criteria was unique, i.e., non-redundant, and together represented a variety of temporal expression constructions. We are confident that EliXR-TIME represents most of the temporal expressions found in eligibility criteria. Second, the two raters who instantiated the test criteria were also the developers; therefore, the general usability and reliability across raters independent of the development team remains to be proven. However, a separate study successfully demonstrated the potential of using this model to develop a conditional random fields algorithm to automatically extract and annotate temporal expressions from eligibility criteria.[17]

4.3 Future Work

Our future work involves using EliXR-TIME to use information extraction tools to automatically chunk and annotate semantic segments in temporal expressions for a large eligibility criteria corpus.

We identified four research challenges for future study.

First, expressions containing medically specific temporal information can be implicit. For example, the inherent meaning of *cancer in remission* contains temporal information including the idea that a diagnosis of cancer was made in the past, treatment (e.g., surgery, chemotherapy, radiation) was performed, negative lab test results indicating absence of cancer have been received, and finally a certain time period has elapsed since a positive lab test result. This type of expression contains important temporal information and should be represented as a past event relative to "*Now*", where *now* is the time of inference.

Second, for many criteria, translating English into logic can be daunting because a word such as *during* can be mapped to multiple Allen temporal relations. Therefore, we did not construct EliXR-TIME to be limited to one rigidly structured temporal representation per criterion but rather to be sufficiently flexible to allow multiple temporal relations to exist between an event and an anchor. For example, the criterion *Administration or planned administration of immunoglobulins and / or blood products during a period starting from 3 months prior to administration of the vaccine and ending at study conclusion* can be represented by five Allen relations: *equals, during, finishes, starts,* and *during inverse* (Figure 3).
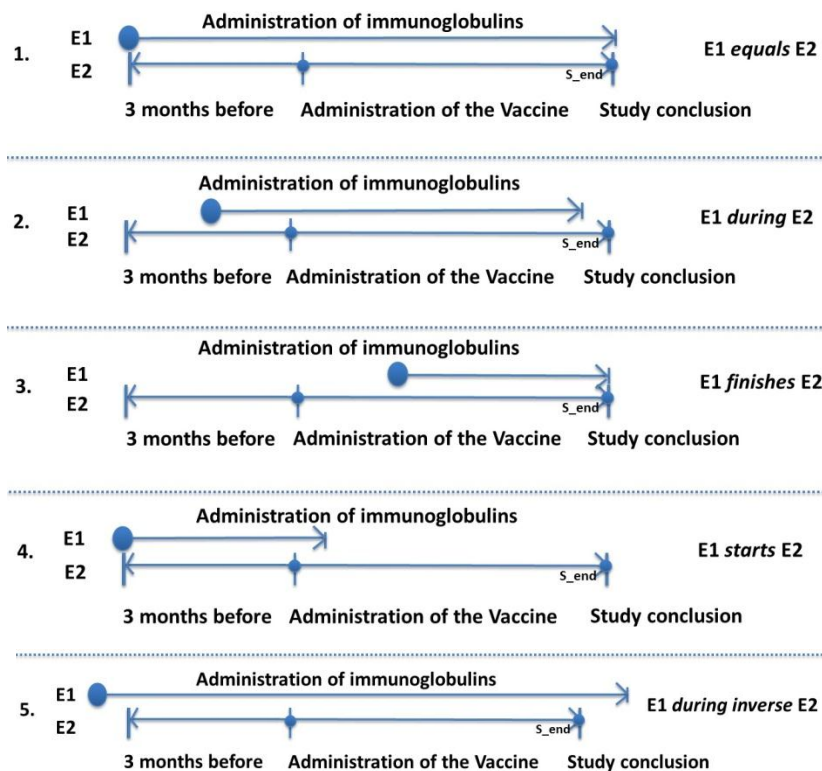
Figure 3. Multiple valid Allen temporal relations for the criterion:
Administration or planned administration of immunoglobulins and / or blood products during a period starting from 3 months prior to administration of the vaccine and ending at study conclusion

Third, mapping between criterion sentence constituents, on one hand, and EliXR-TIME classes and attributes on the other can still be challenging. The difficulty lies in mapping and accurately representing the words and phrases in criterion text to a structure that represents their appropriate semantic meaning. For instance, the criterion *Laboratory confirmed influenza disease within 6 months* can be broken down into a TLE containing an event = *Laboratory confirmed influenza disease,* Allen temporal relation = *during,* and an anchor = *the past 6 months.* This criterion presents two challenges: 1. *within* must be translated into the Allen temporal relation *during* and 2. the past tense must be inferred so that *6 months* is represented as *the past 6 months.* Mapping a fixed duration such as *6 months* to an implied interval *the past 6 months* can be problematic because in some criteria, the fixed duration is intended while in others the implied past interval is intended.

Fourth, a criterion contains an English word that is the name of an Allen temporal relation, such as *before* or *after,* whose contextual meaning does not correspond to its Allen meaning. Disambiguating when to use the Allen temporal relation to represent the English word can cause difficulties. For example, the criterion *subject has agreed to continue adequate contraception during the entire treatment period and for 2 months after completing of the vaccination series* contains two TLEs, both of which use the Allen temporal relation *during*. The first TLE captures the meaning that contraception be used *during* the entire treatment period and the second TLE captures the meaning that contraception be used *during* a relative time interval with begin_point = *end of the treatment period* and end_point = *2 months after end of treatment period*. In this context, *after* indicates a time lag and not an Allen temporal relation.

5. Conclusion

We developed a simple but comprehensive temporal knowledge representation for eligibility criteria, called EliXR-TIME, based on the selective reuse of existing temporal knowledge representations designed for clinical narratives or clinical trial study calendars. We used the small number of classes and attributes in this model to successfully annotate 96% of sentence constituents in a test set of eligibility criteria and to demonstrate its suitability to facilitate manual or automatic semantic annotation of temporal expressions in clinical research eligibility criteria.

6. Acknowledgements

References

1. Thadani S, Weng C, Bigger J, Ennever J, Wajngurt D. Electronic Screening Improves Efficiency in Clinical Trial Recruitment. *Journal of American Medical Informatics Association.* 2009;16(6):869-873.
2. Tu SW, Peleg M, Carini S, et al. A practical method for transforming free-text eligibility criteria into computable criteria. *Journal of Biomedical Informatics.* 2011;44(2):239-250.
3. Weng C, Tu SW, Sim I, Richesson R. Formal representation of eligibility criteria: A literature review. *Journal of Biomedical Informatics.* 2010;43(3):451-467.
4. Shankar R, Martins S, O'Connor M, Das A. An ontological approach to representing and reasoning with temporal constraints in clinical trial protocols. *Proceedings HEALTHINF – International Conference on Health Information* 2008:87-93.
5. Shankar R, Martins S, Parrish D, O'Connor M, Das A. Representing and reasoning with temporal constraints in clinical trials using semantic technologies. *Biomedical Engineering Systems and Technologies, Communications in Computer and Information Science, Springer.* 2009;25(4):520-530.
6. Shankar RD, Martins SB, O'Connor MJ, Parrish DB, Das AK. Epoch: an ontological framework to support clinical trials management. *Proceedings of the international workshop on Healthcare information and knowledge management*. Arlington, Virginia, USA: ACM; 2006:25-32.
7. Tao C, Solbrig H, Sharma D, Wei W, Savova G, Chute C. Time-oriented question answering from clinical narratives using semantic-web techniques. *ISWC 2010* 2010;6497:241-256.
8. Tao C, Wei W. A semantic web ontology for temporal relation inferencing in clinical narratives. *Proc. AMIA Symp.* 2010:787-791.
9. Weng C, Kahn M, Gennari J. Temporal knowledge representation for scheduling tasks in clinical trial protocols. . *Proc. AMIA Symp.* 2002:879-883.
10. Zhou L, Hripcsak G. Temporal reasoning with medical data - A review with emphasis on medical natural language processing. *Journal of Biomedical Informatics.* 9 August 2006 2006;40:183-202.
11. Zhou L, Melton GB, Parsons S, Hripcsak G. A temporal constraint structure for extracting temporal information from clinical narrative. *Journal of Biomedical Informatics.* 2006;39:424-439.
12. Shahar Y. A framework for knowledge-based temporal abstraction. *Artificial Intelligence.* 1997;90(1-2):79-133.
13. Hripcsak G, Zhou L, Parsons S, Das A, Johnson S. Modeling Electronic Discharge Summaries as a Simple Temporal Constraint Satisfaction Problem. *Journal of American Medical Informatics Association.* Jan/Feb 2005 2005;12(1):55-63.
14. O'Connor M, Shankar, R, Das, A. An Ontology-Driven Mediator for Querying Time-Oriented Biomedical Data. *Proceedings of the 19th IEEE International Symposium on Computer-Based Medical Systems.* 2006:264-269.
15. Protege 3.4.6. http://protege.stanford.edu/download/protege/3.4/installanywhere/Web_Installers/. Accessed July 28, 2011.
16. McCray A. Better access to information about clinical trials. *Annals of Internal. Medicine.* 2000;133(8):609-614.
17. Luo Z, Johnson S, Lai A, Weng C. Extracting Temporal Constraints from Clinical Research Eligibility Criteria Using Conditional Random Fields. *Proc of AMIA 2011 Fall Symposium.* 2011.
18. Pustejovsky P, Castaño J, Ingria R, et al. TimeML: Robust Specification of Event and Temporal Expressions in Text. *AAI Spring Symposium on New Directions in Question-Answering (Working Papers).* 2003:28-34.
19. Sauri R, Goldberg L, Verhagen M, Pustejovsky J. Annotating Events in English. *Brandeis University.* 2009;Version TempEval-2010.
20. Weng C, Wu X, Luo Z, Boland MR, Theodoratos D, Johnson SB. EliXR: An Approach to Eligibility Criteria Extraction and Representation. *JAMIA* 2011;Accepted.
21. Allen JF. Maintaining knowledge about temporal intervals. *Commun. ACM.* 1983;26(11):832-843.
22. Allen JF. Time and time again: The many ways to represent time. *Int'l. Jr. of Intelligent Systems.* 1991;6(4):341-356.