

Participatory Design of a Clinical Trial Eligibility Criteria Simplification Method

Yilu FANG^a, Jae Hyun KIM^a, Betina Ross IDNAY^{b,c}, Rebeca ARAGON GARCIA^c, Carmen E. CASTILLO^c, Yingcheng SUN^a, Hao LIU^a, Cong LIU^a, Chi YUAN^a and Chunhua WENG^{a,1}

^aDepartment of Biomedical Informatics

^bSchool of Nursing

^cDepartment of Neurology, Columbia University, New York, NY, USA

Abstract. Clinical trial eligibility criteria are important for selecting the right participants for clinical trials. However, they are often complex and not computable. This paper presents the participatory design of a human-computer collaboration method for criteria simplification that includes natural language processing followed by user-centered eligibility criteria simplification. A case study on the ARCADIA trial shows how criteria were simplified for structured database querying by clinical researchers and identifies rules for criteria simplification and concept normalization.

Keywords. named entity recognition, concept mapping, intelligence augmentation

1. Introduction

Although clinical trials are the foundation of advances in medicine, recruitment remains the No. 1 barrier to clinical trial research. We previously published a natural language processing system called Criteria2Query, which automatically transforms free-text clinical trial eligibility criteria to executable study cohort queries formatted by the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) [1]. However, despite its promising accuracy, it is not yet adopted by clinical researchers who still rely on manual chart review of electronic health records (EHR) data for cohort identification for clinical trials. Participatory Design has been widely adopted as an effective user-centered approach for developing useful and usable software by including domain experts in the iterative design process [2]. To understand the sociotechnical gap between Criteria2Query and the unmet needs of clinical research coordinators and to further improve Criteria2Query, we set out to a participatory design study that uses a clinical trial that is active at recruitment at the point of this study, the Atrial Cardiopathy and Antithrombotic Drugs in Prevention after Cryptogenic Stroke study (ARCADIA, NCT03192215 in ClinicalTrials.gov [3]), to elicit feedback and user needs of clinical research coordinators running this trial and to iteratively refine the design of Criteria2Query by incorporating their feedback. This paper reports our participatory design experience and our preliminary findings of possible rules for eligibility criteria simplification recommended by clinical researchers. Wu *et al.*

¹ Corresponding Author, Chunhua Weng, Columbia University; E-mail: chunhua@columbia.edu.

assessed the readability of trial descriptions in ClinicalTrial.gov and pointed out the need of text simplification [4]. Ross *et al.* found that clinical trial eligibility criteria text is syntactically and semantically complex [5]. Text simplification can help remove trivial, non-critical, redundant, or irrelevant texts that do not change the cohort if being removed or criteria that may introduce errors in cohort query formulation due to clinical data incompleteness or subjectiveness of criteria. This study contributes the earliest empirical knowledge for criteria simplification and promises to inspire future studies for more systematic and scalable criteria simplification.

2. Method

2.1. The Participatory Design Process

Three clinical researchers (RA, BI and CC) working on the ARCADIA trial volunteered to join our participatory design. They have 13, 10 and 7 years of experience in clinical trial recruitment, respectively. We met them weekly for a couple of months and used the “think aloud” protocol to collect their feedback and rationale on what concepts do not need to be included in the final data query and why. The recommendations and rationale were documented in detail. Then we derived rules for criteria simplification for cohort querying, discussed with the researchers until the group reached a consensus on the rules.

2.2. The Workflow for Human-Computer Collaborative Criteria Simplification

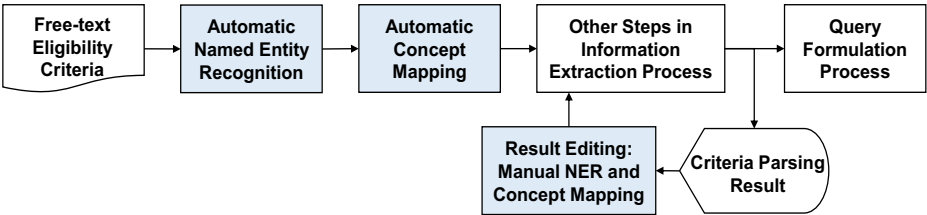


Figure 1. The workflow for human-computer collaboration for eligibility criteria simplification.

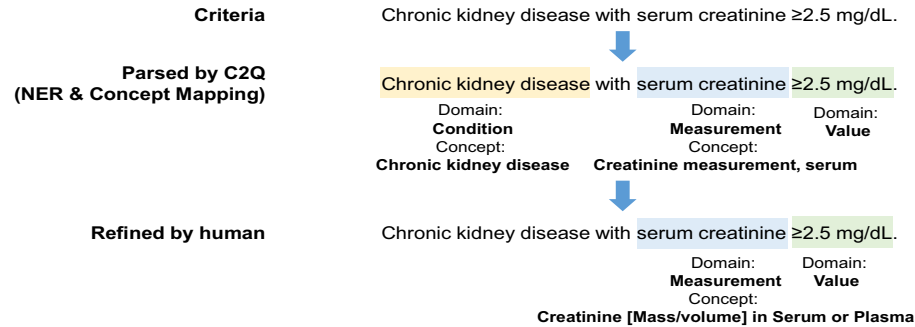


Figure 2. A simplification example: serum creatinine was preferred over chronic kidney disease.

Figure 1 shows the workflow for our human-computer collaboration for eligible criteria simplification. Eligibility terms and their semantic types (i.e., condition, drug, measurement, procedure, demographic, negation, value attribute, and temporal

constraint) are first automatically extracted and normalized by Criteria2Query. Then users can review the parsing results and edit or refine the results using a new feature (similar in Doc2HPO [6]) enhancement to Criteria2Query. This feature enables users to select criteria for inclusion in query formulation, modify its parsing results or concept mapping granularity, or change its semantic type. For concept modification, Athena supports the search and recommendation of OMOP CDM standard concepts for user selection [7]. Figure 2 illustrates an example criteria simplification result from our study. Chronic kidney disease was correctly identified as a condition concept. However, only objective pertinent laboratory measurement and its threshold were determined to be necessary for cohort querying and retained after manual review.

2.3. Iterative Abstraction of Criteria Simplification Rules

Based on the comparison and descriptive statistics of concepts in the original eligibility criteria text and those in the final Observational Health Data Sciences and Informatics (OHDSI) ATLAS (<https://atlas.ohdsi.org/>) data query, we proposed some rules for criteria simplification and concept mapping and reviewed them with the participatory designers. These rules were finalized based on the group consensus.

3. Results

3.1. Descriptive Statistics of the Simplification Results

Table 1. Descriptive statistics of the parsing result. Column A: Count of terms in 11 retained original criteria; Column B: Count of terms in the simplified cohort query, including updated and newly added terms; Column C: Count of original terms with correct NER and concept mapping; Column D: Count of original terms with correct NER (i.e., correct span detection and domain prediction) but incorrect concept mapping; Column E: Count of original terms with correct span detection but incorrect domain prediction and concept mapping.

OMOP CDM Domain	A: Original	B: Simplified	C: Correct Overall	D: Correct NER only	E: Correct Span only
Condition	27	11	7	3	0
Measurement without Value	3	0	0	0	0
Measurement with Value	4	3	0	2	0
Drug	3	1	0	1	0
Procedure	8	2	0	1	0
Observation	2	3	0	0	3
Negation_cue	2	1	1	0	0
Demographic	1	1	1	0	0
Temporal	2	3	2	0	0
Value	5	4	2	0	0
Total	57	29	13	7	3

Table 1 shows the descriptive statistics of the criteria simplification results. Only the age criterion was directly usable from the original Criteria2Query output. Due to complexity or lack of data availability, only 4/9 inclusion criteria and 7/14 exclusion

criteria were retained. 29 concepts were in the simplified cohort query, resulting in ~50% concept reduction. 23 of them are retained from the original 57 concepts, and the other 6 concepts are manually added by the coordinators. Vague measurement (without value threshold), procedure concepts, and condition were the most productive semantic types for text simplification. This result is understandable because vague measurement concepts were not suitable for database queries, procedure concepts were considered extra details, and condition concepts often require elaborative phenotyping algorithms and hence were not as easy to query as corresponding objective measurement concepts. For example, instead of using “chronic kidney disease”, the researchers preferred to use “serum creatinine ≥ 2.5 mg/L” instead.

3.2. Rules for Criteria Simplification for Cohort Querying

Based on the analyses of the criteria retention results for the ARCADIA trial and discussions with the participatory designers, we identified the following potential rules for criteria simplification at the sentence level and at the phrase level, respectively.

Rules at the sentence level:

- (1) Information about research activity authorization can be omitted from data queries, such as “*Additional cardiac imaging, such as cardiac MRI, or cardiac CT will be performed at the discretion of the local treating physician and PI*”.
- (2) Criteria that require further patient inquiry and a principal investigator’s discretion should be omitted from data queries. An example is about the pregnancy risk. Whether the female or male are sexually active or whether the female is post-menopausal may not be documented in clinical data and hence need further inquiry.
- (3) Criteria regarding informed consent or willingness for protocol compliance should be omitted. For example, the criterion, “*Inability of either participant or surrogate to provide written, informed consent for participation*” is not suitable for cohort database query formulation.

Rules at the phrase level:

- (1) Phrases related to demographics (e.g., age and gender) should always be retained.
- (2) Phrases that refer to imaging results for ruling out a certain condition can be discarded, and most of the imaging results are not queryable in a structured EHR database. For example, in the criterion “*clinical diagnosis of ischemic stroke and brain imaging to rule out hemorrhagic stroke*”, since it contains “*hemorrhagic stroke*” condition, the imaging procedure is not critically needed.
- (3) Phrases that refer to broad and unspecified concepts can be discarded. For example, in the criterion “*No major-risk cardioembolic source of embolism, including intracardiac thrombus, mechanical prosthetic cardiac valve, atrial myxoma or other cardiac tumors, mitral stenosis, myocardial infarction within the last 4 weeks, left ventricular ejection fraction < 30 percent, valvular vegetations, or infective endocarditis*”, “*major-risk cardioembolic source of embolism*” term can be ignored, but we should still include their examples.
- (4) Phrases that refer to general disease conditions followed by more specific measurements to define them can be discarded. For example, “*chronic kidney disease with serum creatinine ≥ 2.5 mg/dL*” contains “*chronic kidney disease*”

condition and its relevant measurement “*serum creatinine*”. More objective measurements are preferred to general condition concepts for data querying.

- (5) Phrases that refer to measurements without their explicit or implicit value thresholds (numeric value or categorical value) can be discarded. The implicit value threshold refers to the negation term around the measurement.

4. Discussion and Conclusion

In the ARCADIA trial, more than half of the information was not needed for cohort querying, six terms and their concept sets need to be added, and ten terms need to be refined manually. This result is new and significant because it shows a large gap between the automated Criteria2query output and the user's need. Moreover, this result also shows the feasibility of criteria simplification. Complex eligibility criteria tend to introduce errors during natural language processing given the limitation of the state-of-the-art solutions. Wrong entity extraction, incorrect concept mapping, and redundant concepts can lead to the failure of receiving expected cohort results. Simplification can remove complex and yet not critical information and minimize the errors in criteria parsing by Criteria2Query. Our participatory design users are open to this solution and agree that not all criteria are appropriate for inclusion in the data query. Although methods and decisions for the simplification or prioritization of eligibility criteria may vary among people, we can still find some patterns in criteria simplification. Simplification can start with removing the not queryable sentences or paragraphs whose information is not stored in the database and in which the extracted terms and concepts may interfere with the result before the automatic named entity recognition. Then, it can turn to delete the unnecessary terms and concepts in the remaining criteria according to the potential rules we proposed. Future work can be developed to automate criteria simplification following these rules.

Acknowledgments. This study was sponsored by the National Library of Medicine grant 5R01LM009886-11 and the National Center for Advancing Clinical and Translational Science grants UL1TR001873 and 1OT2TR003434-01.

References

- [1] Yuan C, Ryan PB, Ta C, Guo Y, Li Z, Hardin J, Makadia R, Jin P, Shang N, Kang T, Weng C. Criteria2Query: a natural language interface to clinical databases for cohort definition. *J Am Med Inform Assoc.* 2019 Apr 1;26(4):294-305.
- [2] Schuler D, Namioka A. Participatory design: Principles and practices: CRC Press; 1993.
- [3] AtRial Cardiopathy and Antithrombotic Drugs In Prevention After Cryptogenic Stroke (ARCADIA) Available at: <https://clinicaltrials.gov/ct2/show/NCT03192215>. Accessed Jan 2021.
- [4] Wu DT, Hanauer DA, Mei Q, Clark PM, An LC, Proulx J, Zeng QT, Vydiswaran VG, Collins-Thompson K, Zheng K. Assessing the readability of ClinicalTrials.gov. *J Am Med Inform Assoc.* 2016 Mar;23(2):269-75.
- [5] Ross J, Tu S, Carini S, Sim I. Analysis of eligibility criteria complexity in clinical trials. *Summit Transl Bioinform.* 2010 Mar 1;2010:46-50.
- [6] Liu C, Peres Kury FS, Li Z, Ta C, Wang K, Weng C. Doc2Hpo: a web application for efficient and accurate HPO concept curation. *Nucleic Acids Res.* 2019 Jul 2;47(W1):W566-W570.
- [7] Athena. Available at: <https://athena.ohdsi.org/>. Accessed Jan 2021.