Real-World Utility Evaluation of a Cohort Definition System Criteria2Query

Katherine Wang; Alex Butler, BA; Chunhua Weng, PhD
Department of Biomedical Informatics, Columbia University, New York City, NY, USA

Introduction
Randomized clinical trials (RCTs) are the well-regarded gold standard for generating high-quality medical evidence (Gul 2010). The success of RCTs depends on successful patient enrollment (Gul, Penberthy 2012), which remains the number one barrier to RCTs. A major bottleneck step in RCT recruitment is eligibility screening (Penberthy 2012). However, conventional methods for eligibility screening involves laborious manual review for accurate patient identification (Penberthy 2010). Much effort (Kopcke 2013) has been made to advance automated identification of eligible patients with some approaches reducing workload by up to 90% (Kopcke 2013). However, one of the most significant concerns in automatic eligibility screening lies in the criteria themselves which have been shown to be highly subjective and inconsistent in identifying patients (Penberthy 2012). Current processes rely heavily on manual review of criteria, which is not only costly, but is difficult to scale (Musen 1987).

Eligibility criteria are largely documented as unstructured free text which is not immediately amenable to computer processing, but modern developments in natural language processing (NLP) have provided a means to overcome this hurdle. A new open-source program titled Criteria2Query (C2Q, http://www.ohdsi.org/web/criteria2query/) was designed which allows for the translation of free-text eligibility criteria into executable cohort definition queries using a standard data model (Yuan 2019). C2Q works by parsing each line of the eligibility criteria, identifying medical concepts and corresponding type (e.g. condition, drug) using various technical methods, and mapping these terms to concept sets using the OMOP Common Data Model which can be formulated into logic-based queries. These queries are then submitted to a central patient data portal called ATLAS (http://www.ohdsi.org/web/atlas/) for coherent cohort identification and simplified eligibility screening. As C2Q was only recently published, it is currently undergoing other real-world evaluations, but this stands as the first user-centered study comparing it to the current state-of-the-art ATLAS cohort builder.

This study aims to further examine the real-world utility of C2Q by using eligibility criteria pulled from ClinicalTrials.gov and surveying users to more accurately assess its use in the clinical realm. We hypothesize that using C2Q will make cohort generation simpler and faster and participants will find it easier to use and more user-friendly than the ATLAS portal alone. Following these hypotheses, this study aimed to answer the following questions: (1) Does the use of C2Q save time building patient cohort queries, and if so, by how much? and (2) How do participants score the usability of C2Q for cohort definition compared to the ATLAS portal? By analyzing how C2Q impacts the cohort generation process and allowing users to provide feedback, this study will help shape the future of this platform and automated cohort generation in general.

Methods
Participant Selection
Participants were recruited from the Columbia University Department of Biomedical Informatics Department, the Summer Research Fellowship, and Columbia University Vagelos College of Physicians & Surgeons. Participants completed an online survey about their technical background and expertise (https://tinyurl.com/userexpertise) and followed a brief slideshow tutorial about the C2Q and ATLAS platforms before building patient cohorts.

Trial Selection and Cohort Pre-Processing
To allow study participants with varying technical backgrounds and ATLAS experience to participate equally, clinical trials were manually selected by the authors to have simple eligibility criteria. As such, clinical trials with fewer than 20 total eligibility criteria were selected. Within ATLAS, concept sets were manually created for all relevant concepts included in these trials to be used by participants throughout the study.

Cohort Generation
Each participant was assigned two clinical trials with distinct NCT IDs and was responsible for making one patient cohort for each trial totaling two cohorts per participant. For each participant, one query had to be generated using ATLAS alone and the other had to be generated using C2Q. The order in which these platforms were used was random for each participant. Participants also timed themselves while creating the patient cohorts. Overall time was captured for both methods but the following additional times were recorded for C2Q: time making deletions, time making corrections, and time making additions. For cohort generation using ATLAS alone, timing began when the user clicked ‘Define New Cohort.’ For cohort generation using C2Q, timing began when the user clicked ‘Check on ATLAS.’
Participant Feedback

Participants completed a questionnaire after generating both cohorts (https://tinyurl.com/c2qfeedback). The survey asks a total of three 5-Likert scale questions comparing the user satisfaction and usability of C2Q and ATLAS.

Results

Time Efficiency

Using C2Q saved an average of 47.93% in building patient cohort queries across 14 trials. 13 out of 14 trials showed a time reduction when building patient cohort queries using C2Q with a maximum reduction of 86.22% and only 1 trial took longer when using C2Q. Study participants also recorded a timing breakdown of ATLAS modifications after processing through C2Q. The most time was spent making corrections, which occupied 67.9% of the total modification time (e.g. updating incorrectly mapped concept sets, changing logic rules when multiple criteria were listed in a single line) while additions and deletions occupied 16.8% and 15.3%, respectively.

Usability

8 out of 11 participants completed a feedback form at the end of their study, which assessed ease of use of the study platforms. On a scale of 0 to 5 (with 0 being very difficult and 5 being very easy), the mean usability for the C2Q website and ATLAS website was 4.38 and 3.13, respectively. When asked how much easier was cohort generation using C2Q compared to ATLAS alone on the same scale, the average response was 4.38.

Discussion

Study participants were able to save time building patient cohorts by using C2Q. Because the study shows positive results supporting the time reduction possible with C2Q, it can be a helpful tool for advancing and expediting cohort definition, a crucial step in recruiting for clinical trials. The study assessed the current usability of C2Q by asking for ratings and feedback from participants from a variety of academic and professional experiences. Results indicate that C2Q was easy to use and showed marked improvement over cohort generation using ATLAS alone.

Limitations and Future Work

The limitations in this study were number of trials, pre-building concept sets, and capping trial complexity. The study was limited by its size in order to have a more interactive tutorial phase and aim for greater accuracy. Concept sets were pre-built in ATLAS to minimize the impact of a user’s technical expertise and previous experience in building cohort queries. The same logic was applied to limiting trial complexity to 20 or fewer criteria. In the future, C2Q improvements can increase its time efficiency. Since 67.9% of modifications in ATLAS were corrections, improving concept-mapping accuracy can greatly reduce the time it takes to make modifications. An additional future work can involve the assessment of different data standards (e.g. Fast Health Interoperability Resources [FHIR]) and their impact on query accuracy.

Conclusion

Criteria2Query saves time in creating cohort definition queries. We need to more carefully document how the time saving is related to criteria complexity and results accuracy by developing more robust measures for both.

Acknowledgements

This research is supported by National Library of Medicine grant R01LM009886 and Janssen grant JANSRD CU15-2317.

References