

Identifying Core Outcome Sets in COVID-19 Clinical Trials Using ClinicalTrials.gov

Irena PARVANOVA^{a,1} and Joseph FINKELSTEIN^a

^a*Icahn School of Medicine at Mount Sinai, New York, NY, USA*

Abstract. Introduction of core outcome sets (COS) facilitates evidence synthesis, transparency in outcome reporting, and standardization in clinical research. However, development of COS may be a time consuming and expensive process. Publicly available repositories, such as ClinicalTrials.gov (CTG), provide access to a vast collection of clinical trial characteristics including primary and secondary outcomes, which can be analyzed using a comprehensive set of tools. With growing number of COVID-19 clinical trials, COS development may provide crucial means to standardize, aggregate, share, and analyze diverse research results in a harmonized way. This study was aimed at initial assessment of utility of CTG analytics for identifying COVID-19 COS. At the time of this study, January, 2021, we analyzed 120 ongoing NIH-funded COVID-19 clinical trials initiated in 2020 to inform COVID-19 COS development by evaluating and ranking clinical trial outcomes based on their structured representation in CTG. Using this approach, COS comprised of 25 major clinical outcomes has been identified with mortality, mental health status, and COVID-19 antibodies at the top of the list. We concluded that CTG analytics can be instrumental for COVID-19 COS development and that further analysis is warranted including broader number of international trials combined with more granular approach and ontology-driven pipelines for outcome extraction and curation.

Keywords. COVID-19, ClinicalTrials.gov, core outcome sets, big data

1. Introduction

Since the beginning of 2020, a novel coronavirus disease 2019 (COVID-19) has caused a worldwide pandemic. Evidence synthesis, transparency in outcome reporting, and standardization in emerging COVID-19 clinical research depends on use of a core outcome set (COS) with a minimal set of common outcomes compared across all COVID-19 trials. Identification of COVID-19 COS has been constrained to a limited number of attempts, based on opinion surveys, but lacking real-life trial characteristics. This growing body of research and voluminous data sets are integrated in data repositories, including publicly available repositories, such as ClinicalTrials.gov (CTG). CTG provides a comprehensive set of functions supporting search based on a broad list of clinical metadata. The aim of this study is to initiate an identification of COVID-19 characteristics in clinical trials to commence the establishment of evidence-based COVID-19 COS [1].

¹Irena PARVANOVA, PhD, Department of Population Health Science and Policy, Icahn School of Medicine at Mount Sinai, New York, NY, E-mail: Irena.Parvanova@mountsinai.org

2. Methods

Primary and secondary outcomes from COVID-19 trials registered in ClinicalTrials.gov were analyzed using CTG search engine. The query combined “COVID-19” in the “Other terms” field with “United States” in “Country” and “NIH” in “Funder Type” fields, generating 120 studies. We performed analysis where we examined common data elements based on the outcomes for the studies.

3. Results

We examined the outcomes for 120 COVID-19 clinical trials and identified 56 distinct outcomes. We ranked each outcome by calculating the percentage of its appearance in all 120 trials resulting in 25 major outcomes represented in more than 1% of the featured clinical trials. Overall, the clinical trials contained large quantities of diverse information that reflect COVID-19 outcomes. The most common COVID-19 clinical outcomes were: Mortality (16%), Mental health impact (12%), Anti-SARS-CoV-2 Antibodies (10%), Patients hospitalization or death (8%), SARS-CoV-2 PCR/RT-PCR test results and Severity (6%), Responsive CD8 T cells (5%), and others.

4. Discussion

ClinicalTrials.gov hosts vast amount of searchable information on clinical trials including COVID-19 trials. Standardized approaches for representation of COVID-19 research are necessary in this rapidly evolving field to allow for cross-trial comparisons. In this study, we investigated outcomes featured in 120 NIH-funded clinical trials, which underwent rigorous reviews, by panels of experts in the USA in order to identify systematic characteristics for COVID-19 research. Our findings provided an evidence-based foundation for building a comprehensive taxonomy of core outcome sets in COVID-19 clinical trials. Future plans include building an expert consensus on common data elements [2-3] in COVID-19 research and expanding our analysis beyond the NIH clinical trials, as well as implementing ontology-driven automated metadata extraction pipeline from CTG and PubMed combined with crowdsourcing functionality for engaging researchers in ranking COS candidates [4-5].

References

- [1] Elghafari A, Finkelstein J. Automated Identification of Common Disease-Specific Outcomes for Comparative Effectiveness Research Using ClinicalTrials.gov: Algorithm Development and Validation Study. *JMIR Med Inform.* 2021;9(2):e18298.
- [2] Finkelstein J, Parvanova I, Zhang, F. Informatics Approaches for Harmonized Intelligent Integration of Stem Cell Research. *Stem Cells and Cloning: Advances and Applications.* 2020 Jan 28;13:1-20.
- [3] Parvanova I, Finkelstein J. Data Integration Approaches for Representing Stem Cell Studies. *Stud Health Technol Inform.* 2020;270:1235-1236.
- [4] Elghafari A, Finkelstein J. Introducing an Ontology-Driven Pipeline for the Identification of Common Data Elements. *Stud Health Technol Inform.* 2020;272:379-382.
- [5] Borziak, K, Qi T, Evangelista, JE. Towards Intelligent Integration and Sharing of Stem Cell Research Data. *Stud Health Technol Inform.* 2020;272:334-337.