Innovation in Applied Nursing Informatics
G. Strudwick et al. (Eds.)

doi:10.3233/SHTI240292

Building Bridges for Federated Learning in Healthcare: Review on Approaches for Common Data Model Development

Hanna VON GERICH[a,1], Taridzo CHOMUTARE[b] and Laura-Maria PELTONEN[a,c]
[a] *Department of Nursing Science, University of Turku, Finland*
[b] *Norwegian Centre for E-health Research, Norway*
[c] *Turku University Hospital, Finland*
ORCiD ID: Hanna von Gerich https://orcid.org/0000-0003-1036-2163
Taridzo Chomutare https://orcid.org/0000-0003-3603-175X
Laura-Maria Peltonen https://orcid.org/0000-0001-5740-6480

Abstract. Common data models provide a standardized way to represent data used in federated learning tasks. The aim of this review was to explore the development and use of common data models to harmonize electronic health record data in health research. The data search yielded 724 records, of which 19 were included for this study. None of the research focused on nursing specific topics. All studies either utilized the Observational Medical Outcomes Partnership (OMOP) common data model, or developed a model partly based on the OMOP. A roadmap to guide research for the development of common data models for federated learning are warranted.

Keywords. Common data model, Federated learning, Healthcare, Scoping review

1. Introduction

Artificial intelligence -based models trained with patient data are bound to revolutionize nursing by enhancing identification of health problems, care provision, and patient outcomes. Vast amounts of electronic health record (EHR) data from different sources are needed to ensure the accuracy and generalizability of these models [1]. Federated learning is an approach to train the models using EHR data sets from different institutions and beyond the national borders without the need to share the data itself [2]. Common data models (CDM) provide a standardized way to represent the data used in federated learning tasks [3]. The aim of this review was to explore the development and use of CDMs to harmonize EHR data in health research. The results can be used to support federated learning tasks in nursing research.

[1] Corresponding Author: Hanna von Gerich, Department of Nursing Science, University of Turku, Finland. Email: hanna.m.vongerich@utu.fi

2. Methods

For this review, two databases (PubMed, CINAHL) were searched for articles describing the development and use of CDMs to harmonize EHR data in health research. The search included no limitations regarding the setting, patient population or methodology. Studies that did not describe the transformation of clinical EHR data for a CDM were excluded.

3. Results

A total of 724 records were retrieved from the databases. After the removal of 195 duplicates, 529 records were screened through title and abstract screening. Full text review included 239 articles, of which 19 were included for this study. None of the research focused on nursing specific topics**.** All studies either used the Observational Medical Outcomes Partnership (OMOP) Common Data Model (n=16) or developed a CDM partly based on the OMOP (n=3). The development of CDMs were not described in the articles, nor did the articles describe the use or development of CDMs for federated learning purposes. Nine of the articles described the harmonization of different data sources (e.g. electronic health record data, health claims data), eight described the harmonization of one EHR dataset to align with the commonly used OMOP CDM, and two described the harmonization of EHR datasets from different nations for an international collaboration.

4. Conclusions

CDMs have gained traction in health research, yet their utilization in nursing research is still lacking, and comprehensive descriptions regarding their development or utilization were not identified in this review. A roadmap to guide research and development of CDM for federated learning are warranted, considering the varying contexts and environments, as well as different legal and ethical aspects related to the use of highly confidential data in transnational federated learning tasks.

Acknowledgements

This work is supported by Nordic Innovation.

References

[1] Rieke N, Hancox J, Li W, Milletarì F, Roth HR, Albarqouni S, et al. The future of digital health with federated learning. Npj Digit Med. 2020 Sep 14;3(1):119.
[2] Xu Y, Liu Y, Shu T, Yang W, Liang M. Variations in the quality of care at large public hospitals in Beijing, China: A condition-based outcome approach. PLoS ONE. 2015;10(10).
[3] Park K, Cho M, Song M, Yoo S, Baek H, Kim S, et al. Exploring the potential of OMOP common data model for process mining in healthcare. PLOS ONE. 2023 Jan 3;18(1):e0279641.