Dermatologic Research Potential of the Observational Health Data Sciences and Informatics (OHDSI) Network

Torunn Elise Sivesind^a Taylor Runion^b Megan Branda^c Lisa M. Schilling^d
Robert P. Dellavalle^a

^aDepartment of Dermatology, University of Colorado School of Medicine, Aurora, CO, USA; ^bRocky Vista University College of Osteopathic Medicine, Parker, CO, USA; ^cDepartment of Biostatistics and Informatics, University of Colorado School of Medicine, Aurora, CO, USA; ^dDepartment of Medicine, Data Science to Patient Value Program Aurora, University of Colorado School of Medicine, Aurora, CO, USA

Keywords

Data standardization · Analytics · Networks · Observational research · Treatment guidelines · Electronic health records

Abstract

Background: The Observational Health Data Sciences and Informatics (OHDSI) network enables access to billions of de-identified, standardized health records and built-in analytics software for observational health research, with numerous potential applications to dermatology. While the use of the OHDSI has increased steadily over the past several years, review of the literature reveals few studies utilizing OHDSI in dermatology. To our knowledge, the University of Colorado School of Medicine is unique in its use of OHDSI for dermatology big data research. **Summary:** A PubMed search was conducted in August 2020, followed by a literature review, with 24 of the 72 screened articles selected for inclusion. In this review, we discuss the ways OHDSI has been used to compile and analyze data, improve prediction and estimation capabilities, and inform treatment guidelines across specialties. We also discuss the potential for OHDSI in derma-

tology – specifically, ways that it could reveal adherence to available guidelines, establish standardized protocols, and ensure health equity. **Key Messages:** OHDSI has demonstrated broad utility in medicine. Adoption of OHDSI by the field of dermatology would facilitate big data research, allow for examination of current prescribing and treatment patterns without clear best practice guidelines, improve the dermatologic knowledge base and, by extension, improve patient outcomes.

© 2021 S. Karger AG, Basel

Introduction

The Observational Health Data Sciences and Informatics (OHDSI) collaborative is a volunteer collaborative international network of researchers, with a central hub at Columbia University for organization. OHDSI is the successor of the Observational Medical Outcomes Partnership (OMOP), a public-private partnership between the FDA, pharmaceutical companies, and healthcare providers [1]. One result of OMOP is the Common Data

Fig. 1. The Common Data Model (image from Observational Health Data Sciences and Informatics, 2020). Available from: <https://www.ohdsi.org/data-standardization/the-common-data-model/>.

Model (CDM), which establishes transformation conventions for observational health data, including insurance claims, electronic health records (EHR), and hospital billing into a standardized common data model (Fig. 1). The CDM supports large-scale analytics across the many participating data partners, allowing access to billions of deidentified health records for observational health research [2]. A fundamental tool developed from OMOP is the Standard Vocabulary, which enables interoperability between systems, facilitates homogeneity and transparency of data across diverse observational databases, and supports the OHDSI network in conducting efficacious, high-quality research. The latest version of the Standard Vocabulary is available for download from the Athena vocabularies repository site (<https://athena.ohdsi.org>).

OHDSI also provides access to analytic and data quality tools for working with relevant data. Studies can be conducted on OMOP CDM data by writing custom code, or by using built-in software such as ATLAS, ACHILLES (Automated Characterization of Health Information at Large-Scale Longitudinal Evidence Systems) and its successor data quality software, the Data Quality Dashboard. For data analysis, ATLAS is a web-based component of OHDSI used to design and execute observational analyses of OMOP CDM data. Within ATLAS, specifications and codes are created to identify cohorts, outcome variables, and other analytic variables of interest, while also enabling vocabulary searching and browsing. For data quality, the Data Quality Dashboard precomputes summary statistics, describes datasets, and also creates reports allowing visualization of the previously computed summa-

ry statistics. Overall, it characterizes a database and provides data quality assessment.

The OHDSI community is open to all. Participation is as easy as joining a standing OHDSI community call, participating in a work group, joining a research network by converting data to the OMOP CDM, or leading a study across the network, and researchers do not need to have an affiliation with data partners to propose research studies (Fig. 2). Additionally, OHDSI maintains a commitment to reproducible research and making code and protocol available to the public on GIT, a code hosting platform.

As an open community, OHDSI is essentially a collaboration of volunteers; in this regard, anyone can be an OHDSI member. Once a member of OHDSI, anyone can take the lead on a study or participate in a study, either intellectually as a contributor to the study design, analysis, or manuscript development, and/or as a data contributor. Study leads are responsible for ensuring that the data methods and distributed analytics conform to OHDSI standards. The protocol and analytic code are distributed to community members, who volunteer to run the analysis on their data and provide the aggregate results to the lead site. In this sense, there is no formal governance or prioritization of the studies conducted by partners; whatever is accomplished is determined by the follow through of the lead site, and of course, the appropriateness of the available data to the research question.

Research proposal presentations and participant recruitment are generally conducted via weekly OHDSI community meetings and by way of a Microsoft Outlook OHDSI Teams site. OHDSI partners frequently partici-

Fig. 2. OHDSI: Join the Journey (image from Observational Health Data Sciences and Informatics, 2020). Available from: <https://www.ohdsi.org/join-the-journey>.

pate in other funded research, such as from the FDA, using the community tools, but these collaborations are not governed by the community and participants are free to leverage the benefits of the OHDSI collaborative for other endeavors.

To our knowledge, the University of Colorado School of Medicine is unique in its use of OHDSI to support

large-scale, observational dermatology research for the investigation of rare conditions, therapies, and outcomes. Within the field of dermatology, there is an urgent need to establish standardized treatment guidelines for a variety of acute and chronic dermatologic conditions, to better understand adherence to presently available guidelines and further elucidate the effect of patient demo-

Fig. 3. Trends in OHDSI-related publications per year (PubMed).

graphics and population-related variables on treatment modalities and health outcomes.

To this end, we performed a review of the available literature citing OHDSI; upon review, there appears to be a relative dearth of studies utilizing OHDSI in dermatology. Here, we explore the merits of OHDSI and discuss its application as described in the literature.

Review Process

The PubMed database was used for publication review; a simple search for the term “OHDSI” was performed on August 3, 2020, returning a total of 72 articles published since 2014 (Fig. 3). These were screened by the principal author and assessed for content and relevance (Table 1). A total of 24 articles of clinical import were selected for inclusion and are discussed below.

Over the years, randomized controlled trials (RCTs) have assumed prominence in the scientific community and been widely accepted as presenting the highest-quality evidence among research methods. However, the “gold standard” status of RCTs has more recently come into question, with several publications highlighting their inherent flaws, along with the misconception that observational studies exaggerate exposure-outcome associations [3–5]. A review of meta-analyses in *The New England Journal of Medicine* that examined both RCTs and observational studies (case-control or cohort) found

the summary results of observational studies and RCTs to be extremely similar, and noted that observational studies did not overestimate the extent of exposure-outcome associations when compared to the results of RCTs on the same topics [5]. Rapid medical innovation often outpaces the slow process of clinical trials’ development and execution, and while the internal validity of RCTs is strong, external validity remains problematic [4]. The strongest evidence may be obtained with the complementary use of observational studies, which enable the assessment of both generalizability to populations of interest and the applicability of study results to individuals [3, 4]. The ease and speed of conducting observational studies have grown dramatically with the advent of large-scale databases and online research networks such as OHDSI.

The potential applications of OHDSI are many. Its large scale provides increased power and reproducibility for observational research as well as characterizing the generalizability of clinical trials to real-world populations [6–8]. The large sample sizes provided by OHDSI increase the accuracy of estimations and predictions, facilitate the study of rare exposures, diseases, and outcomes, and are ideal for investigations across diverse populations [9, 10]. The following are some illustrative studies that utilize OHDSI in a broad range of applications, including the reporting of adverse events, the estimation of heritability, guidelines for adherence to treatment, and the characterization of prescribing patterns, to name a few.

Table 1. Summary of the included literature

First author [ref.], year	Study topic	Design	Key findings
Schumie [6], 2018	Reproducibility of observational research	Used data from OHDSI to compare all treatments for a given indication, for a large set of outcomes, via a high-throughput approach	OHDSI use can improve reproducibility. Advocate new paradigm of large-scale studies with standardized, consistent methods, yielding results for numerous hypotheses concurrently
Sen [7], 2016	Eligibility criteria for clinical trials	Propose and evaluate a multitrait metric for computing generalizability of clinical studies	OHDSI may be used to standardize a metric's algorithm, enabling portability across different EHR databases and better evaluating generalizability of clinical trials during the design phase
Bartlett [8], 2019	Feasibility of replicating clinical trials using observational health data	Feasibility study of using large-scale observational data to replace clinical trials data	Results indicate that the use of EHR and claims data complements clinical trials data and is helpful in assessing generalizability, but cannot replace clinical trials
Tong [9], 2020	Statistical analysis; heterogeneity of patient-level data across sites	Proposal and evaluation via OHDSI of a new algorithm for statistical analyses across sites that accounts for heterogeneity of data and maintains privacy	Found that performance of the new algorithm exceeds that of the existing distributed algorithm, ODAL, as well as a meta-analysis method
Gini [10], 2016	Data processing within research networks	Created a framework to evaluate local data processing, applied to 4 case studies	Evaluation of national and multinational networks including OHDSI revealed shared methods of data reorganization and centrally defined derivation of new data elements
Hripsak [2], 2015	Creation of the OHDSI collaborative	Descriptive overview	OHDSI is useful for prognosis, evaluation of comparative effectiveness, and patient-level assessment
Yu [11], 2019	Pharmacovigilance; development of OHDSI platform for adverse event reporting	Development of a platform in OHDSI for adverse event reporting	Converted spontaneous reporting data from FAERS into the OMOP-CDM to integrate spontaneous reporting and EHR data, improving pharmacovigilance
Yu [12], 2020	Pharmacovigilance; feasibility of integrating SRS and EHR data	Case study of adverse events related to immune checkpoint inhibitors, comparing OHDSI and FAERS data signal detection	Adverse event signaling from OHDSI platform allowed a specific adverse drug event to be detected 4 months earlier compared to FAERS
Li [13], 2015	Pharmacovigilance; integration of data from SRS and healthcare databases	Descriptive; propose method for combining spontaneous reporting systems with large-scale data networks such as OHDSI	Demonstrated improved detection of adverse drug reactions by combining large-scale observational data with FAERS data, compared to spontaneous reporting alone
Chandler [14], 2020	Pharmacovigilance; use of observational data from OHDSI	Combined data from OHDSI with spontaneous reports of adverse drug events to assess association of nintedanib and ischemic colitis	Found combined use of OHDSI and spontaneous reporting of adverse drug events could improve efficiency of signal management for post-authorization safety signals, improving pharmacovigilance
Jiang [15], 2019	Medical device safety	Pilot study; developed a unique device identifier vocabulary for medical devices	Proposed a framework for building a unique medical device identifier database in OHDSI CDM for research and evaluation of device safety
Duke [16], 2017	Risk of angioedema with antiepileptic agents	Retrospective cohort study using OHDSI, including 10 databases and 2 countries (USA and France)	Evaluated risk of angioedema for patients exposed to levetiracetam vs. phenytoin; found that levetiracetam carries the same or a lower risk than phenytoin for developing angioedema
Hripsak [17], 2016	Characterization of treatment pathways for chronic conditions	Used OHDSI to characterize treatment pathways for 3 chronic diseases, included 11 data sources from 4 countries	Discovered increasing consistency of treatment across diseases and locations, but variability in treatment pathways and prescribing patterns including preferences for specific drugs
Zhang [18], 2018	Characterization of treatment pathways and prescribing patterns	Feasibility study using OHDSI to characterize treatment and prescribing patterns in China for 3 chronic diseases	Proved feasibility of OHDSI for observational health studies in China
Polubriagino [19], 2018	Disease heritability estimation	Used EHR to identify familial relationships and compute heritability estimates; validated results	Heritability estimates were computed for 500 disease phenotypes, consistent between sites and with existing literature. Validated the utility of mining EHR data for genetics research
Wang [20], 2020	Prediction modeling for hemorrhagic transformation after cerebral infarction	Developed a logistic regression prediction model using OHDSI to determine risk of hemorrhagic transformation in the month after an ischemic stroke	Demonstrated that OMOP-CDM can be used for development and validation of prediction models
Choi [21], 2020	Prognosis of inflammatory bowel disease	Cohort study of early- vs. late-onset inflammatory bowel disease (IBD), using OHDSI network in Korea	Verified use of the CDM in gastroenterology research. Results demonstrated worse outcomes in patients with early-onset IBD than in those with late-onset IBD

Table 1 (continued)

First author [ref.], year	Study topic	Design	Key findings
Meysire [22], 2019	Eligibility criteria for clinical trials	Extracted data with natural language processing application to determine clinical trials eligibility	Use of OMOP-CDM datasets and automatic extraction enabled detection of patients' eligibility for clinical trials with good accuracy
Maier [23], 2018	Data integration; use of OHDSI for a consortium of hospitals	Pilot study implementing OHDSI/OMOP in a consortium of 8 hospitals in Germany	Code mapping was 98.8% successful, but procedure codes could not be mapped to standard terminologies. Concluded OMOP/OHDSI will be a viable method of data integration for the consortium once terminology problems are resolved
Yoon [24], 2016	Conversion of EHR data to the OHDSI CDM	Data conversion of EHR from a single medical center in Korea to OMOP-CDM	The transformed dataset was validated and quality was improved by using the open-source data characterization platform ACHILLES
You [25], 2017	Conversion of EHR data to the OHDSI CDM	Converted data from a nationwide cohort in Korea into the OHDSI OMOP-CDM	Were able to convert data from 1.13 million subjects into the OMOP-CDM. Found OHDSI promotes high-quality evidence across nations
Lima [26], 2019	Conversion of EHR data to the OHDSI CDM; ETL process for data conversion	Descriptive; describes data standardization experience	Successfully transformed 20 years of data from a single institution into the OMOP-CDM standard; found the new database consistent with the original data
Alnofal [27], 2020	Conversion of EHR data to the OHDSI CDM	Pilot study of data conversion	Converted EHR from a single institution in Saudi Arabia to OMOP-CDM, successfully creating a pilot database that could be expanded across the nation
Boyce [28], 2016	Conversion of EHR data to the OHDSI CDM for patient-level prediction of falls	Case study of data conversion from 5 skilled nursing facilities	Successfully converted EHR data to OMOP-CDM to create a longitudinal dataset that supports observational research. Data characterization with OHDSI's ACHILLES tool improved the quality of the data

Yu et al. [11] converted spontaneous reporting system (SRS) data from the FDA's Adverse Event Reporting System (FAERS) into the OMOP CDM, creating and assessing a platform "ADEpedia-on-OHDSI," by combining SRS and EHR data to improve pharmacovigilance. Their follow-up to this study was conducted in 2020, in which the detection of adverse event signals was compared between FAERS and EHR data for a specific adverse drug event (ipilimumab-induced hypopituitarism) [12]. It found that detection of EHR signaling occurred 4 months earlier than with FAERS and concluded that an integrated CDM-based approach to the detection of adverse events would enhance drug safety monitoring. Li et al. [13] likewise advocate integrating SRS data with large-scale, standardized healthcare databases, noting that this can speed up the detection of adverse events.

Chandler [14] reached a similar conclusion in a 2020 study on the association of nintedanib with ischemic colitis, in which safety signaling from SRS and observational health data was assessed. This research suggested that the use of real-world data from the OHDSI complements the spontaneous reporting of adverse drug reactions and has the potential to increase the efficiency of signal management [14].

A 2019 study by Jiang et al. [15] represents one possible extension of the OHDSI CDM to medical device safety, in the form of a unique device identifier (UDI) research database. This pilot study developed a UDI vocabulary for medical devices and proposed a framework for building a UDI research database in the CDM format.

Hripcsak et al. [2] discussed the efficacy of OHDSI in yielding prognostic information useful for comparing treatment options and gaining patient-level assessment of these options. This concept was adopted by Duke et al. [16] in a comparative study of antiepileptic agents spanning 10 databases and 2 countries, in which the risk of experiencing angioedema in new users exposed to levetiracetam was compared to that of phenytoin. Notably, they found the risk associated with levetiracetam to be equal to or lower than that associated with phenytoin (a drug that, unlike levetiracetam, does not carry a warning label for angioedema).

In a 2016 study by Hripcsak et al. [17], 11 data sources from 4 countries (encompassing 250 million patients) were mapped to the OHDSI CDM standard to investigate treatment pathways for type II diabetes mellitus, hypertension, and depression. This study was able to reveal a trend toward more consistent treatment across diseases and locations, although the authors did note some heterogeneity in treatment pathways and prescribing pat-

terns. Zhang et al. [18] conducted a similar study in 2018, converting national data from China to the OMOP CDM to: examine treatment pathways for hypertension, type II diabetes mellitus and depression, compare these to recommended guidelines, uncover medication differences in clinical practice and, ultimately, urge the standardization of such guidelines.

As mentioned in the work of Hripcsak et al. [2], discussed previously, OHDSI has demonstrated usefulness in obtaining prognostic information. Polubriaginof et al. [19] employed next-of-kin information from EHR standardized to OHDSI's CDM to identify familial relationships, computing heritability estimates for 500 disease phenotypes and yielding knowledge useful for estimating disease risk, identifying risk factors, and tailoring the care and treatment of patients. Wang et al. [20] used OHDSI to develop a logistic regression prediction model able to assess a patient's risk of hemorrhagic transformation within 30 days of an ischemic stroke, thereby showing OMOP CDM to be a valuable tool for the development and validation of prognostic prediction models.

Choi et al. [21] utilized data converted to the OHDSI CDM to analyze how age at diagnosis of inflammatory bowel disease (IBD) impacted the prognosis in Korea. This study verified the use of CDM in gastrointestinal research and found that early-onset IBD portends a worse prognosis than late-onset IBD does. The authors defined "worsened disease prognosis" as the initiation of a biologic agent (infliximab or adalimumab), reporting odds ratios of experiencing a biologic agent to be 2.3 and 5.4 in patients with early- and late-onset ulcerative colitis and Crohn's disease, respectively.

OHDSI has proven to be a complement to tools used in clinical trials research. Meystre et al. [22] used the OHDSI's open-source data analytics platform (ATLAS) and natural language processing (NLP) to find patients with breast cancer who were eligible for specific clinical trials. NLP matched patients with clinical trials with an area under the curve (AUC) of up to 89.8%, illustrating its potential to decrease the demands on human workers who screen for clinical trials.

Bartlett et al. [8] looked at the use of OHDSI as it relates to clinical trials, performing a study aimed at determining whether clinical trials could be replicated by using only observational data obtained from EHR or insurance claims. The study found that while only 15% of US-based clinical trials published in 2017 could be entirely replicated via observational data alone, many additional clinical trial end points, interventions, indications, and inclusion/exclusion criteria could be identified from observa-

tional data, ultimately suggesting that the use of real-world data is complementary to clinical trials, particularly for comparing the generalizability of trial populations to real-world populations of interest. In their 2016 study, Sen et al. [7] also found OHDSI to be a useful tool for assessing the generalizability of clinical trials.

The advantages of data standardization and the development of large-scale national and international databases are increasingly being recognized and addressed at diverse sites around the globe. The literature includes articles from a range of locales, including Germany, Korea, Brazil, and Saudi Arabia, describing the experience of standardizing existing data to the OMOP CDM to facilitate the interinstitutional sharing of information [23–27]. Boyce et al. [28] recognized the benefit of using OHDSI's data characterization tools for longitudinal research on topics important to skilled nursing facilities (SNFs), e.g., patient-level falls prediction, and therefore converted EHR from 5 SNFs to the OMOP CDM.

Conclusion

For clinicians and researchers alike, the OHDSI network presents a means to enhance knowledge and improve patient care. For dermatology, specifically, OHDSI has the potential to assess prescribing patterns and treatment pathways for diseases that presently lack clear best practice guidelines. The rapid expansion of available biologic agents for the management of dermatologic conditions necessitates deliberate efforts to support pharmacovigilance, explore variations in the prescribing of novel and costly pharmaceuticals according to demographic and socioeconomic factors and, ultimately, to ensure health equity and inform clinical practice guidelines. Predictive studies utilizing biologic prescription data, such as that of Choi et al. [21], could be adapted to immune-mediated skin diseases and enable more individualized patient care. While this paper provides only a brief introduction to the OHDSI and its diverse applications, it is clear that, in this increasingly globalized and digitized world, the OHDSI's practicality and value will likewise continue to grow.

Key Message

The OHDSI network has demonstrable merit; its use in dermatology has the potential to advance observational research.

Appendix

Table of acronyms and abbreviations used in this article.

Acronym or abbreviation	Definition
ACHILLES	Automated Characterization of Health Information at Large-Scale Longitudinal Evidence Systems
AUC	area under the curve
CDM	Common Data Model
EHR	electronic health records
FAERS	FDA's Adverse Event Reporting System
FDA	Federal Drug Administration
IBD	inflammatory bowel disease
NLP	natural language processing
OHDSI	Observational Health Data Sciences and Informatics
OMOP	Observational Medical Outcomes Partnership
RCT	randomized controlled trial
SNF	skilled nursing facility
SRS	spontaneous reporting system
UDI	Unique Device Identifier

Acknowledgements

The authors wish to acknowledge John Barbieri, David Margolis, and Jonathan Silverberg for their ongoing collaboration and review of this manuscript.

References

- [Internet] Welcome to OMOP – OHDSI [cited 12 Nov 2020]. Available from: <https://ohdsi.org/omop/>
- Hripcsak G, Duke JD, Shah NH, Reich CG, Huser V, Schuemie MJ, et al. Observational Health Data Sciences and Informatics (OHDSI): Opportunities for Observational Researchers. *Stud Health Technol Inform*. 2015; 216:574–8.
- Frieden TR. Evidence for Health Decision Making – Beyond Randomized, Controlled Trials. *N Engl J Med*. 2017 Aug 3;377(5):465–75.
- Bothwell LE, Greene JA, Podolsky SH, Jones DS. Assessing the Gold Standard – Lessons from the History of RCTs. *N Engl J Med*. 2016 Jun 2;374(22):2175–81.
- Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med*. 2000 Jun;342(25):1887–92.
- Schuemie MJ, Ryan PB, Hripcsak G, Madigan D, Suchard MA. Improving reproducibility by using high-throughput observational studies with empirical calibration. *Philos Trans A Math Phys Eng Sci*. 2018;376(2128):20170356.
- Sen A, Chakrabarti S, Goldstein A, Wang S, Ryan PB, Weng C. GIST 2.0: A scalable multi-trait metric for quantifying population representativeness of individual clinical studies. *J Biomed Inform*. 2016 Oct;63:325–36.
- Bartlett VL, Dhruva SS, Shah ND, Ryan P, Ross JS. Feasibility of Using Real-World Data to Replicate Clinical Trial Evidence. *JAMA Netw Open*. 2019 Oct;2(10):e1912869.
- Tong J, Duan R, Li R, Schuemie MJ, Moore JH, Chen Y. Robust-ODAL: learning from heterogeneous health systems without sharing patient-level data. *Pac Symp Biocomput*. 2020;25:695–706.
- Gini R, Schuemie M, Brown J, Ryan P, Vacchi E, Coppola M, et al. Data Extraction and Management in Networks of Observational Health Care Databases for Scientific Research: A Comparison of EU-ADR, OMOP, Mini-Sentinel and MATRICE Strategies. *EGEMS (Wash DC)*. 2016 Feb;4(1):1189.
- Yu Y, Ruddy KJ, Hong N, Tsuji S, Wen A, Shah ND, et al. ADEpedia-on-OHDSI: A next generation pharmacovigilance signal detection platform using the OHDSI common data model. *J Biomed Inform*. 2019 Mar;91:103119.
- Yu Y, Ruddy KJ, Wen A, Zong N, Tsuji S, Chen J, et al. Integrating Electronic Health Record Data into the ADEpedia-on-OHDSI Platform for Improved Signal Detection: A Case Study of Immune-related Adverse Events. *AMIA Jt Summits Transl Sci Proc*. 2020 May;2020:710–9.
- Li Y, Ryan PB, Wei Y, Friedman C. A Method to Combine Signals from Spontaneous Reporting Systems and Observational Health-care Data to Detect Adverse Drug Reactions. *Drug Saf*. 2015 Oct;38(10):895–908.
- Chandler RE. Nintedanib and ischemic colitis: signal assessment with the integrated use of two types of real-world evidence, spontaneous reports of suspected adverse drug reactions, and observational data from large health-care databases. *Pharmacoevidenciol Drug Saf*. 2020 Aug;29(8):951–7.
- Jiang G, Yu Y, Kingsbury PR, Shah N. Augmenting Medical Device Evaluation Using a Reusable Unique Device Identifier Interoperability Solution Based on the OHDSI Common Data Model. *Stud Health Technol Inform*. 2019 Aug;264:1502–3.

Statement of Ethics

This study complies with internationally accepted standards for research practice. This study did not involve human subjects so no institutional Ethics Committee approval was required.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Funding Sources

T.E.S. receives funding from the 58858477 Pfizer Global Medical Grant – Dermatology Fellowship 2020 (principal investigator: R.D.). The authors have no other funding or financial interests to declare.

Author Contributions

T.E.S. contributed to study design, data acquisition and interpretation, manuscript drafting and revising, and design and formulation of tables. T.R. contributed to data acquisition, manuscript drafting, and designing of the figures. L.M.S. contributed to review and revision of the manuscript and OHDSI expertise. M.B. contributed to review of the manuscript and OHDSI expertise. R.D. contributed to study design, manuscript revision and review, and approval of the final version.

- 16 Duke JD, Ryan PB, Suchard MA, Hripscak G, Jin P, Reich C, et al. Risk of angioedema associated with levetiracetam compared with phenytoin: findings of the observational health data sciences and informatics research network. *Epilepsia*. 2017 Aug;58(8):e101–6.
- 17 Hripscak G, Ryan PB, Duke JD, Shah NH, Park RW, Huser V, et al. Characterizing treatment pathways at scale using the OHDSI network. *Proc Natl Acad Sci USA*. 2016 Jul; 113(27):7329–36.
- 18 Zhang X, Wang L, Miao S, Xu H, Yin Y, Zhu Y, et al. Analysis of treatment pathways for three chronic diseases using OMOP CDM. *J Med Syst*. 2018 Nov;42(12):260.
- 19 Polubriaginof FC, Vanguri R, Quinnes K, Belbin GM, Yahi A, Salmasian H, et al. Disease Heritability Inferred from Familial Relationships Reported in Medical Records. *Cell*. 2018 Jun;173(7):1692–1704.e11.
- 20 Wang Q, Reips JM, Kostka KF, Ryan PB, Zou Y, Voss EA, et al. Development and validation of a prognostic model predicting symptomatic hemorrhagic transformation in acute ischemic stroke at scale in the OHDSI network. *PLoS One*. 2020 Jan;15(1):e0226718.
- 21 Choi YI, Kim YJ, Chung JW, Kim KO, Kim H, Park RW, et al. Effect of Age on the Initiation of Biologic Agent Therapy in Patients with Inflammatory Bowel Disease: Korean Common Data Model Cohort Study. *JMIR Med Inform*. 2020 Apr;8(4):e15124.
- 22 Meystre SM, Heider PM, Kim Y, Aruch DB, Britten CD. Automatic trial eligibility surveillance based on unstructured clinical data. *Int J Med Inform*. 2019 Sep;129:13–9.
- 23 Maier C, Lang L, Storf H, Vormstein P, Bieber R, Bernarding J, et al. Towards Implementation of OMOP in a German University Hospital Consortium. *Appl Clin Inform*. 2018 Jan;9(1):54–61.
- 24 Yoon D, Ahn EK, Park MY, Cho SY, Ryan P, Schuemie MJ, et al. Conversion and Data Quality Assessment of Electronic Health Record Data at a Korean Tertiary Teaching Hospital to a Common Data Model for Distributed Network Research. *Healthc Inform Res*. 2016 Jan;22(1):54–8.
- 25 You SC, Lee S, Cho SZ, Park H, Jung S, Cho J, et al. Conversion of National Health Insurance Service-National Sample Cohort (NHIS-NSC) Database into Observational Medical Outcomes Partnership- Common Data Model (OMOP-CDM). *Stud Health Technol Inform*. 2017;245:467–70.
- 26 Lima DM, Rodrigues-Jr JF, Traina AJ, Pires FA, Gutierrez MA. Transforming Two Decades of ePR Data to OMOP CDM for Clinical Research. *Stud Health Technol Inform*. 2019 Aug;264:233–7.
- 27 Alnofal FA, Alrwisan AA, Alshammari TM. Real-world data in Saudi Arabia: current situation and challenges for regulatory decision-making. *Pharmacoepidemiol Drug Saf*. 2020 Oct;29(10):1303–6.
- 28 Boyce RD, Handler SM, Karp JF, Perera S, Reynolds CF 3rd. Preparing Nursing Home Data from Multiple Sites for Clinical Research - A Case Study Using Observational Health Data Sciences and Informatics. *EGEMS (Wash DC)*. 2016 Oct;4(1):1252.