

Pharmacovigilance and Clinical Environment: Utilizing OMOP-CDM and OHDSI Software Stack to Integrate EHR Data

Vlasios K. DIMITRIADIS^a, George I. GAVRIILIDIS^a and Pantelis NATSIAVAS^{a,1}
^a*Institute of Applied Biosciences, Centre for Research & Technology Hellas, Thessaloniki, Greece*

Abstract. Information Technology (IT) and specialized systems could have a prominent role towards the support of drug safety processes, both in the clinical context but also beyond that. PVClinical project aims to build an IT platform, enabling the investigation of potential Adverse Drug Reactions (ADRs). In this paper, we outline the utilization of Observational Medical Outcomes Partnership - Common Data Model (OMOP-CDM) and the openly available Observational Health Data Sciences and Informatics (OHDSI) software stack as part of PVClinical platform. OMOP-CDM offers the capacity to integrate data from Electronic Health Records (EHRs) (e.g., encounters, patients, providers, diagnoses, drugs, measurements and procedures) via an accepted data model. Furthermore, the OHDSI software stack provides valuable analytics tools which could be used to address important questions regarding drug safety quickly and efficiently, enabling the investigation of potential ADRs in the clinical environment.

Keywords. Pharmacovigilance, Drug Safety, OMOP-CDM, Electronic Health Records, Adverse Drug Reactions

1. Introduction

Adverse Drug Reactions (ADRs) represent a peril to public health. ADRs are among the most common causes of morbidity and mortality worldwide causing dismal patient outcomes in the clinical setting, severe hurdles in the development and authorization of novel therapeutics as well as significant economic burdens to public healthcare providers [1]. To this end, Pharmacovigilance (PV) is defined as the science focusing on the detection, assessment, understanding and prevention of ADRs [2].

Traditional approaches for PV “signals” (potential new or partially documented ADRs) detection, entail the investigation of Individual Case Safety Reports (ICSRs) via Spontaneous Reporting Systems (SRSs), based on “Disproportionality Analysis”

¹ Corresponding author: Pantelis Natsiavas, Institute of Applied Biosciences Centre for Research and Technology Hellas, 6th Km. Charilaou – Thessaloniki Road, P.O. BOX 60361 GR – 57001, Thessaloniki, Greece, +30 2311257608; Email: pnatsiavas@certh.gr

(DA) metrics. Beyond the investigation of ICSRs submitted by healthcare professionals and/or patients, the potential use of other data sources is investigated. In this sense, instead of "passively" waiting for ICSRs to be sent, the "Active Pharmacovigilance" (AP) model is gradually gaining a foothold by promoting the "hot pursuit" of all possible data sources [3]. To this end, modern Information Technology (IT) systems could play a prominent role by enabling heterogeneous drug safety data integration [4].

PVclinical project aims to develop a Web-based platform supporting the investigation of potential ADRs by clinicians, as well as PV experts using Knowledge Engineering (KE) as its main technical paradigm. A first design of the platform, i.e., the first "User Goals" and its information workflow was introduced in [5].

In this paper we present the use of OMOP-CDM and Observational Health Data Sciences and Informatics (OHDSI) software stack to integrate Electronic Health Record (EHR) data in the PVclinical platform. More specifically, we provide an overview interesting research initiatives based on OHDSI outcomes and KE. Then, we describe the approach used to integrate EHR data in the PVclinical platform. Finally, we discuss the potential merit and the challenges of exploiting EHR data for drug safety purposes via OHDSI based technical developments.

2. Background

While KE-based technical approaches are not yet been widely adopted in real-world products, they are actively investigated in terms of Drug Safety [6]. Indicatively, the use of ontologies is elaborated in a number of ways, e.g., to provide systematic access, discovery and explicit data interlinking capabilities, regarding safety signal report data, based on FAIR principles [7]. Other KE-related approaches include the use of Natural Language Processing (NLP) to extract potentially useful knowledge from free-text.

Regarding drug safety and real-world evidence, OHDSI is one of the most prominent initiatives shaping a wide research network for observational healthcare studies, including but not limited to data safety issues [8], including researchers and data holders. Technically, OHDSI network maintains a stack of software tools supporting visual analytics and advanced statistical processing of data stored in a reference data model, named OMOP-CDM. As a result of the above, recently, OHDSI was awarded \$10M FDA contract to support effective surveillance of vaccines [9]. Respectively, in Europe, EH DEN initiative [10] is currently forming a European health data and evidence network based on OHDSI technical developments and rationale.

A number of research works focusing on drug safety have been produced as an outcome of the OHDSI initiative. LAERTES [11] integrated multiple data sources of evidence, potentially useful for safety PV signal detection and evaluation. LAERTES focused on providing an open and scalable architecture, using standardized terminology to associate drugs with health outcomes of interest (HOI) via OMOP-CDM. The Common Evidence Model (CEM) [12] is the next step after LAERTES, emphasizing on the identification of negative controls for the calibration of statistical methods aiming to detect PV signals from various heterogeneous databases [13]. Finally, ADEpedia-on-OHDSI [14] is an informatics platform, aiming to support PV signal detection via the integration of SRS data from FDA's Adverse Event Reporting System (FAERS), into the standardized OMOP-CDM format and its further combination with EHR data. The relevant study conducted for the evaluation of ADEpedia-on-OHDSI, demonstrated the capacity to analyze both SRS and EHR data for PV purposes.

3. Methods

Based on the user requirements and design process outcomes, namely the User Goals (UGs), and the corresponding Business Processes (BPs) as defined in [15], the PVClinical web platform provides one workspace for each data source type, i.e. SRS data provided via FAERS, literature data provided via PubMed, social media data provided via Twitter and EHR data.

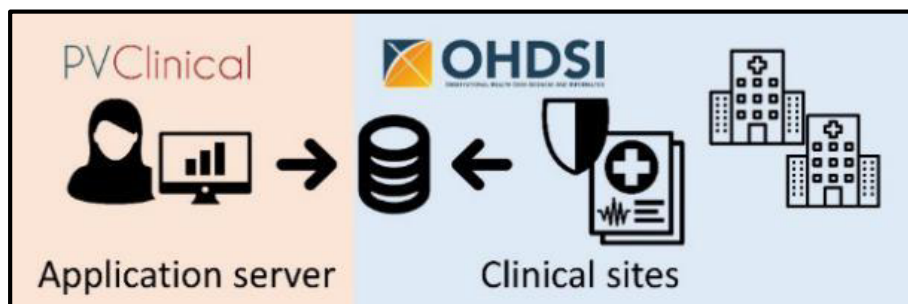


Figure 1. Integration of EHR data via OMOP-CDM and OHDSI software stack

The OHDSI workspace provides a gateway to OMOP-CDM data, via OHDSI software stack and analytics [16] modified to support PV use cases. OMOP-CDM is used as the intermediate model to integrate EHR data provided by the PVClinical project clinical partners which are stored locally in the respective partner's premises to prevent privacy issues. For development purposes, we use SynPUFs data [17] in CDM Version 5.0 [18].

4. Results

After the end-user defines the investigation scenario in the PVClinical platform, i.e., the drug and the ADR of interest, the OHDSI workspace provides population-level analytics on the respective EHR data, based on OHDSI software stack analytics. The provided views can be summarized as follows:

- The *Incidence Rates* view depicts the rates of incidence of the respective outcomes within the target/specific population (patient cohorts) based on age, gender and study time window criteria.
- The *Characterizations* view, enables the generation of cohort level descriptive summary statistics from person level covarying data. Selection criteria include categorization features such as Charlson Index to facilitate comorbidity investigations, age group, gender, and drug eras, i.e. successive periods of drug exposures.
- The *Cohort Pathways* view, depicts the sequence of events occurring for the target patients cohort based on various criteria in the form of a sunburst plot.
- Finally, the *Drug Exposure* and *Condition Occurrence* views, depict an overview of the specific EHR data source providing insights in terms of data quality as they depict information regarding the numbers of patients exposed to drugs or developing various medical conditions, respectively.

These views were modified compared to the original views provided by OHDSI software stack to simplify not only the end-user's visual interactions (while retaining

both useful forms of representation for analytics; tabular and graphical), but also aiming to automate/simplify processes like patient cohort formulation. For example, target and outcome patient cohorts are formed transparently, based on the drug and the ADR of interest selected while defining the investigation scenario in the PVclinical platform. These user interface modifications were defined via end-users focus groups as part of the overall project's iterative and user oriented design approach.

5. Discussion

In this paper, we provided an overview of how OMOP-CDM and OHDSI software stack are used in the PVclinical web platform to support the integration of EHR data. It should be noted that during the development of the PVclinical OHDSI workspace, several challenges were faced. These challenges can be summarized as following:

- Setup of OHDSI software stack and integration of synthetic data to be used for development purposes was proven to be far from trivial
- Notwithstanding the published methodologies on harnessing OMOP-CDM and OHDSI software stack for drug safety scenarios, the integration of OHDSI statistical tools for PV purposes proved rather daunting as disproportionality or PV related analysis measures are not directly implemented in the OHDSI analytics tools.

However, we argue that using OMOP-CDM as a reference data model to integrate EHR data for PV purposes entails significant benefits. First, given that OMOP-CDM is a widely accepted data model, it provides a potential gateway to a big data sharing ecosystem. Furthermore, OMOP-CDM could be used to integrate other kinds of data (e.g., claims databases). Finally, the statistical and analytics tools developed by OHDSI are supported by an ecosystem of widely recognized experts, with a proven record of highly valuable results on observational studies. Although these tools are not explicitly developed for PV purposes, there is significant merit in repurposing them towards PV due their analytical prowess. This repurposing approach could be an important next step, providing significant value for both the drug safety and OHDSI ecosystem.

Acknowledgements

This research has been co - financed by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH – CREATE - INNOVATE (project code: T1EDK-03789). The authors would also like to highlight the contribution of the recently deceased Vassilis Koutkias, initial Coordinator/Principal Investigator of PVclinical project.

References

- [1] Formica D, Sultana J, Cutroneo PM, Lucchesi S, Angelica R, Crisafulli S, Ingrassiotta Y, Salvo F, Spina E, Trifirò G. The economic burden of preventable adverse drug reactions: a systematic review of observational studies. *Expert Opin Drug Saf.* 2018 Jul;17(7):681-695.
- [2] W. C. C. for I. D. M. World Health Organization, The importance of pharmacovigilance. World Health Organization, 2002.
- [3] Wiktorowicz M, Lexchin J, Moscou K. Pharmacovigilance in Europe and North America: divergent approaches. *Soc Sci Med.* 2012 Jul;75(1):165-70.

- [4] Harpaz R, DuMochel W, Shah NH, Big data and adverse drug reaction detection. *Clin. Pharmacol. Ther.* 2016;99(3):268–270.
- [5] Natsiavas P, Jaulent MC, Koutkias V. A Knowledge-Based Platform for Assessing Potential Adverse Drug Reactions at the Point of Care: User Requirements and Design. *Stud Health Technol Inform.* 2019 Aug 21;264:1007-1011.
- [6] Natsiavas P, Malousi A, Bousquet C, Jaulent MC, Koutkias V. Computational Advances in Drug Safety: Systematic and Mapping Review of Knowledge Engineering Based Approaches. *Front Pharmacol.* 2019 May 17;10:415.
- [7] Natsiavas P, Boyce RD, Jaulent MC, Koutkias V. OpenPVSignal: Advancing Information Search, Sharing and Reuse on Pharmacovigilance Signals via FAIR Principles and Semantic Web Technologies. *Front Pharmacol.* 2018 Jun 26;9:609.
- [8] Observational Health Data Sciences and Informatics, Available at <https://www.ohdsi.org/>.
- [9] OHDSI Awarded \$10 Million FDA Contract to Support Safety/Effectiveness Surveillance of Vaccines, Other Biological Products, Available at <https://www.ohdsi.org/ohdsi-news-updates/ohdsi-fda-120120/>.
- [10] EHDEN, Available at: <https://www.ehden.eu/>.
- [11] Knowledge Base workgroup of the Observational Health Data Sciences and Informatics (OHDSI) collaborative. Large-scale adverse effects related to treatment evidence standardization (LAERTES): an open scalable system for linking pharmacovigilance evidence sources with clinical data. *J Biomed Semantics.* 2017 Mar 7;8(1):11.
- [12] OHDSI/CommonEvidenceModel, Available at <https://github.com/OHDSI/CommonEvidenceModel/wiki>.
- [13] Voss EA, Boyce RD, Ryan PB, van der Lei J, Rijnbeek PR, Schuemie MJ. Accuracy of an automated knowledge base for identifying drug adverse reactions. *J Biomed Inform.* 2017 Feb;66:72-81.
- [14] Yu Y, Ruddy KJ, Hong N, Tsuji S, Wen A, Shah ND, Jiang G. ADEpedia-on-OHDSI: A next generation pharmacovigilance signal detection platform using the OHDSI common data model. *J Biomed Inform.* 2019 Mar;91:103119.
- [15] Natsiavas P, Rasmussen J, Voss-Knude M, Votis K, Coppolino L, Campegnani P, Cano I, Mari D, Faiella G, Clemente F, Nalin M, Grivas E, Stan O, Gelenbe E, Dumortier J, Petersen J, Tzovaras D, Romano L, Komnios I, Koutkias V. Comprehensive user requirements engineering methodology for secure and interoperable health data exchange. *BMC Med Inform Decis Mak.* 2018 Oct 16;18(1):85.
- [16] OHDSI/Atlas, Available at <https://github.com/OHDSI/Atlas>.
- [17] CMS.gov Centers for Medicare & Medicaid Services, Available at <https://www.cms.gov/research-statistics-data-and-systems/downloadable-public-use-files/synpufs/index>.
- [18] mustafaascha/ubuntu-synpuf-cdm, Available at <https://github.com/mustafaascha/ubuntu-synpuf-cdm>.