Received: 15 April 2024 | Accepted: 23 June 2024

DOI: 10.1002/prp2.1239

COMMENTARY

Expanding the use of real-world data through the development of user-friendly tools to support drug development: Example of an organ impairment dashboard

Bianca Vora[1] | Erick Velasquez[1] | Rong Zhang[1] | Krystian Igras[2] | Spencer L. James[3] | Pascal Chanu[4]

[1]Clinical Pharmacology, Genentech, Inc, South San Francisco, California, USA

[2]RWD Insights Engineering Team, Contractor at FQBT, Warsaw, Poland

[3]Product Development, Genentech, Inc, South San Francisco, California, USA

[4]Clinical Pharmacology, Genentech, Inc./Roche, Lyon, France

Correspondence
Bianca Vora, Clinical Pharmacology, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA.
Email: vorab@gene.com

Real-world data (RWD), and more specifically electronic health record (EHR) data, have gained increasing interest and utilization in the pharmaceutical industry over the past few years, especially given their ability to support decision-making in drug development. From research and development to post-approval, EHR data can help inform various stages and components of the drug development process.[1] Although randomized controlled clinical trials have been the gold standard for decades, the use of RWD to complement these studies paves the way for more inclusive clinical trials, since RWD can provide a more representative composition of patients with a certain disease. Additionally, leveraging RWD when designing clinical trials allows researchers to plan for and design studies with some initial insights on the patient population of interest, enabling scientists to curate inclusion/exclusion criteria that are more directed and focused in order to get meaningful results. Recent advances in making RWD/EHR data more accessible include the OHDSI ecosystem, which aims to bring large-scale analytics and access to health data to the broader community in an open-source FAIR format.[2] In particular, there have been European efforts through the EHDEN/DARWIN projects to make EHR data more accessible.[3] We believe that by putting forward a standardized framework (particularly for a use case specific to clinical pharmacology), we might complement the efforts by the DARWIN project to showcase how EHR data can have an impact in a particular field. Overall, the use of RWD can provide qualified evidence when considering an appropriate scientific question and taking into account meaningful steps to standardize and generalize the data.[4]

However, utilization of EHR data is often limited and constrained due to various factors such as restrictions on data access (which may be out of scope at an individual scientist level) and the need for coding skills. As an example, to answer more clinically based questions, gaining insights from EHR data can be particularly challenging since it requires individuals to have both clinical expertise (to formulate the question, identify the right patient population, and interpret the results) as well as technical expertise (to understand the data, acknowledge their limitations, and wrangle the data into a more usable format). This can be addressed by developing user-friendly dashboards which enable clinical scientists to interact with complex data and standardize common analysis practices for processing and interpreting EHR data. By removing

Abbreviations: DARWIN, Data Analysis and Real World Interrogation Network; EHDEN, European Health Data Evidence Network; EHR, Electronic Health Record; FAIR, Findability, Accessibility, Interoperability, and Reuse; HTML, HyperText Markup Language; NCI, National Cancer Institute; OHDSI, Observational Health Data Sciences and Informatics; PDF, Portable Document Format; RWD, Real World Data

Bianca Vora and Erick Velasquez contributed equally to this work.

the need to be proficient in coding or wrangling big data, scientists can query RWD to help address drug development-related questions and support decision-making. Such efforts have led to the development of analytical methods to simulate and assess the impact of clinical and demographic variables on trial enrollment and outcomes, such as in the Trial Pathfinder study.[5] We present an analysis workflow that was internally developed as a dashboard; this tool was created to analyze the prevalence of hepatic and renal impairment in real-world patient cohorts as a use case example to demonstrate how to increase accessibility and utilization of RWD to support drug development. Since dedicated organ impairment studies can cost approximately 3–5 million dollars, leveraging RWD to (1) inform study design and (2) understand the need, relevance, and impact a dedicated organ impairment study may have (i.e., by providing insights on the proportion of the target population affected by these comorbidities), can be extremely beneficial.

Renal and hepatic impairment are common comorbidities for patients diagnosed with cancer. Since organ impairment can directly affect the elimination of a drug, and thus its exposure, it is important to understand the severity and extent of changes in clearance when renal and/or hepatic are the primary route(s) of elimination. However, these patients are often excluded from pivotal clinical trials, and conducting dedicated studies to evaluate the impact of organ impairment on drug exposure can be challenging due to the paucity of patients. Increasing accessibility to RWD for the evaluation of prevalence and incidence of organ impairment can greatly help in early decision-making, waiving, and/or better designing of clinical pharmacology studies for certain diseases, as seen for both ipatasertib and polatuzumab vedotin; more specifically, for both molecules, RWD was used to understand the prevalence and severity of organ impairment in their respective indications and patient populations, which in turn aided in understanding the need and feasibility to conduct dedicated organ impairment studies.[1,6,7] In addition, we also want our clinical trials to mirror the real-world population and not exclude patients from trials that could likely benefit from our medicines (advancing our diversity and inclusion goals for trials). Hence, knowing the incidence of organ impairment in the real-world setting, together with the knowledge of the drug's metabolism/elimination pathways, may also present opportunities to include a larger patient population than initially planned.

Using the nationwide Flatiron Health EHR-derived de-identified database, a longitudinal database comprising of de-identified patient-level structured (i.e., analysis dataset similar to clinical trial data) and unstructured (e.g., physician notes) data curated via technology-enabled abstraction, we can evaluate the real-world incidence of renal and/or hepatic impairment in oncology patient populations of interest, that is, a target population for a particular drug.[8,9]

We developed an easy-to-use dashboard to assess the prevalence of organ impairment. Users are able to create virtual cohorts that most accurately represent their clinical trial population of interest, following which the dashboard categorizes patients based on a predefined renal (Cockcroft & Gault) or hepatic (NCI organ dysfunction working group) impairment criteria[10,11] (Table 1). Outputs include attrition tables and distribution figures which can be downloaded as PDF or HTML files (Data S1). The example presented in Data S1 demonstrates a case example where organ impairment prevalence and incidence were evaluated for patients diagnosed with advanced non-small cell lung cancer. In this example, the "Reference Date" chosen for both renal and hepatic impairment was *Diagnosis date* and the "Time Range (months)" from which the dashboard should select laboratory values was defined as *3 months prior to diagnosis to 3 months following diagnosis*. Additionally, in this example, it was defined, in the "Lab Selection Criteria," that if there is more than one laboratory

TABLE 1 Prevalence of renal and hepatic impairment across selected indications.

Organ function			Advanced melanoma	Advanced non-small cell lung cancer	Hepatocellular carcinoma	Metastatic breast cancer
Renal function	Total number of patients (*N*)		7540	64 537	6906	15 324
	Normal		3910 (51.86)	23 807 (36.89)	3143 (45.51)	7917 (51.66)
	Mild impairment		2345 (31.1)	24 131 (37.39)	2171 (31.44)	4709 (30.73)
	Moderate impairment		1176 (15.6)	15 093 (23.39)	1346 (19.49)	2399 (15.66)
	Severe impairment		87 (1.15)	1280 (1.98)	195 (2.82)	240 (1.57)
	Kidney failure		22 (0.29)	226 (0.35)	51 (0.74)	59 (0.39)
Hepatic function	Total number of patients (*N*)		6393	59 446	6341	13 935
	Normal		5701 (89.18)	53 639 (90.23)	1684 (26.56)	11 882 (85.27)
	Mild impairment	Group 1	384 (6.01)	4303 (7.24)	2307 (36.38)	1661 (11.92)
		Group 2	188 (2.94)	932 (1.57)	897 (14.15)	227 (1.63)
	Moderate impairment		75 (1.17)	384 (0.65)	909 (14.34)	90 (0.65)
	Severe impairment		45 (0.70)	188 (0.32)	544 (8.58)	75 (0.54)

Note: Data are represented as *N* (%) for all impairment categories. Patients were categorized based on a predefined renal (Cockcroft & Gault) or hepatic (NCI organ dysfunction working group) impairment criteria. Data cut: August 2023. The laboratory result closest to the patient's initial diagnosis (but within 3 months before or after their initial diagnosis) was chosen for the categorization of each patient.

result present in the specified time range, the dashboard should use the laboratory value closest to the reference date. Additional details on each of these parameters are provided in Section 1.3 of both the renal and hepatic impairment workflows in Data S1. With this given input, the dashboard was able to select the appropriate laboratory values (in addition to other demographic variables needed for renal impairment determination) and categorize patients based on the predefined renal and hepatic criteria. The results of these analyses are presented in Sections 4 and 5 of the report. Code can also be generated and downloaded to allow replication of the analysis. This dashboard empowers study teams to easily get insights into the real-world prevalence and incidence of mild, moderate, and severe renal/hepatic impairment.

By incorporating the analysis pipeline into a modularized, easy-to-use application, it becomes more accessible to scientists without coding skills. More specifically, we first focused on creating relevant toggles in the graphical user interface that would aid the user in building a cohort of interest while incorporating all available data. Secondly, we modularized functions that would then appropriately handle the data in the backend. Taking advantage of "shinydashboard" (an R package that makes it easy to build R shiny applications; https://rstudio.github.io/shinydashboard/), it was easy to implement the backend functions with input toggles that could easily be manipulated by a user in an outward facing interface. Lastly, we incorporated intuitive visuals using "ggplot" (https://ggplot2.tidyverse.org/) and interactive plots using "plotly" (https://plotly.com/) so that the user would easily be able to convey results to their team and interact with the data in real time. We believe that standardizing the analysis workflow through an easy-to-use application reduces the risk of misinterpretation while creating more reproducible results.

Overall, this organ impairment dashboard puts RWD and analyses at the fingertips of scientists without requiring any coding skills, allowing for widespread use and adoption of RWD as well as assessing the feasibility and optimizing clinical trial design, informing drug development decisions, and supporting health authority interactions. We believe that developing additional dashboards with multiple RWD sources (i.e., electronic health records and claims data) to answer various drug development questions can be informative and helpful for the scientific community to not only complement randomized clinical trials but also to promote more representative, diverse studies and reduce patient burden. The use of standardized dashboards may also increase the regulatory acceptance of the data generated. We hope this dashboard can serve as a template and be a springboard for other institutions to create their own dashboards, expanding the use of RWD among the scientific community to help supplement drug development.

AUTHOR CONTRIBUTIONS

All authors participated in research design; Bianca Vora, Erick Velasquez, Rong Zhang, and Krystian Igras contributed to or performed the data analysis; all authors wrote or contributed to the writing of the manuscript.

ACKNOWLEDGMENTS

The authors acknowledge Dale Miles, Benjamin Wu, Rui Zhu, Rucha Sane, Scott Pivirotto, Chunze Li, and Amita Joshi for their helpful feedback and discussions during the development of this dashboard and writing of this manuscript.

FUNDING INFORMATION

No funding was received for this work.

CONFLICT OF INTEREST STATEMENT

All authors are employees and stockholders of Roche/Genentech, Inc.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study were originated by and are the property of Flatiron Health, Inc., which has restrictions prohibiting the authors from making the data set publicly available. Requests for data sharing by license or by permission for the specific purpose of replicating results in this manuscript can be submitted to PublicationsDataAccess@flatiron.com.

ETHICS STATEMENT

None.

ORCID

Bianca Vora https://orcid.org/0000-0002-0501-1804
Erick Velasquez https://orcid.org/0000-0002-0874-4317
Rong Zhang https://orcid.org/0000-0001-5107-1974
Krystian Igras https://orcid.org/0009-0002-7722-9228
Spencer L. James https://orcid.org/0000-0003-4653-2507
Pascal Chanu https://orcid.org/0000-0001-9011-5431

REFERENCES

1. Zhu R, Vora B, Menon S, et al. Clinical pharmacology applications of real-world data and real-world evidence in drug development and approval-an industry perspective. *Clin Pharmacol Ther*. 2023;114(4):751-767. doi:10.1002/cpt.2988
2. Hripcsak G, Duke JD, Shah NH, et al. Observational health data sciences and informatics (OHDSI): opportunities for observational researchers. *Stud Health Technol Inform*. 2015;216:574-578.
3. EMA. Data analysis and real world interrogation network (DARWIN EU), European Medicines Agency. 2021. Accessed June 14, 2024. https://www.ema.europa.eu/en/about-us/how-we-work/big-data/data-analysis-real-world-interrogation-network-darwin-eu
4. Breckenridge AM, Breckenridge RA, Peck CC. Report on the current status of the use of real-world data (RWD) and real-world evidence (RWE) in drug development and regulation. *Br J Clin Pharmacol*. 2019;85(9):1874-1877. doi:10.1111/bcp.14026
5. Liu R, Rizzo S, Whipple S, et al. Evaluating eligibility criteria of oncology trials using real-world data and AI. *Nature*. 2021;592(7855):629-633. doi:10.1038/s41586-021-03430-5
6. Sane R, Malhi V, Sutaria DS, et al. Pharmacokinetics of Ipatasertib in subjects with hepatic impairment using 2 methods of classification of hepatic function. *J Clin Pharmacol*. 2022;62(2):171-181. doi:10.1002/jcph.1941
7. Lu D, Lu T, Shi R, et al. Application of a two-analyte integrated population pharmacokinetic model to evaluate the impact of intrinsic and extrinsic factors on the pharmacokinetics of Polatuzumab

Vedotin in patients with non-Hodgkin lymphoma. *Pharm Res*. 2020;37(12):252. doi:10.1007/s11095-020-02933-6
8. Ma X, Long L, Moon S, Adamson BJS, Baxi SS. Comparison of population characteristics in real-world clinical oncology databases in the US: flatiron health, SEER, and NPCR *medRxiv*. 2023. doi:10.1101/2020.03.16.20037143
9. Birnbaum B, Nussbaum N, Seidl-Rathkopf K, et al. Model-assisted cohort selection with bias analysis for generating large-scale cohorts from the EHR for oncology research. *arXiv preprint arXiv:200109765*. 2020.
10. Lichtman SM, Harvey RD, Damiette Smit MA, et al. Modernizing clinical trial eligibility criteria: recommendations of the American Society of Clinical Oncology-friends of cancer research organ dysfunction, prior or concurrent malignancy, and comorbidities working group. *J Clin Oncol*. 2017;35(33):3753-3759. doi:10.1200/JCO.2017.74.4102
11. U.S. Department of Health and Human Services, Food and Drug Administration (FDA), Oncology Center of Excellence Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER). Cancer Clinical Trial Eligibility Criteria: Patients with Organ Dysfunction or Prior or Concurrent Malignancies. Guidance for Industry. https://www.fda.gov/media/123745/download

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Vora B, Velasquez E, Zhang R, Igras K, James SL, Chanu P. Expanding the use of real-world data through the development of user-friendly tools to support drug development: Example of an organ impairment dashboard. *Pharmacol Res Perspect*. 2024;12:e1239. doi:10.1002/prp2.1239