REAL-WORLD EVIDENCE ALLIANCE

info@rwealliance.org
rwealliance.org

December 14, 2023

Office of Data Science Strategy
National Institutes of Health
9000 Rockville Pike
Bethesda, MD 20892

Re: Request for Information (RFI): Inviting Comments and Suggestions on Opportunities and Challenges for the Collection, Use, and Sharing of Real-World Data (RWD) including Electronic Health Records, for NIH Supported Biomedical and Behavioral Research (Notice No. NOT-OD-23-180)

To the Office of Data Science Strategy:

The RWE Alliance appreciates the opportunity to comment on the National Institutes of Health ("NIH") Office of Data Science Strategy ("ODSS") Request for Information titled "Inviting Comments and Suggestions on Opportunities and Challenges for the Collection, Use, and Sharing of Real-World Data (RWD) including Electronic Health Records, for NIH Supported Biomedical and Behavioral Research" ("RFI on RWD").[1] We are a coalition of real-world data ("RWD") and analytics organizations with a common interest in harnessing the power of real-world evidence ("RWE") to inform regulatory decision-making to improve the lives of patients. Our members have deep knowledge and experience working with healthcare data across disease areas and patient populations, and we aim to bring these collective insights to bear in support of RWE policies.[2]

The RWE Alliance envisions a future in which data from electronic health records, administrative claims and billing records, product and disease registries, personal devices, wearables, and health applications will be used to generate evidence to support regulatory decision making related to medical product safety and effectiveness. To achieve these goals, the RWE Alliance advocates for policies that will (1) advance FDA's RWE Framework, (2) encourage the use of RWE to better understand treatment effects in underrepresented populations, (3) enhance opportunities for RWE

---

[1] https://grants.nih.gov/grants/guide/notice-files/NOT-OD-23-180.html.

[2] For information about our members, please see our website, https://rwealliance.org/who-we-are/.

organizations to consult with FDA, (4) increase communication on the generation and use of RWE, and (5) recognize the unique aspects of and opportunities for RWD/E.[3]

We appreciate that NIH has requested feedback from the RWD/E ecosystem to inform its continuing development of guidance on the use of RWD for research and its planning for mechanisms and programs for research with RWD. We believe that fit-for-purpose RWD can play an important role in medical research. RWD can bring value to research by, among other things, helping to (1) better understand treatment effects in underrepresented populations, (2) generate evidence of effectiveness for serious conditions with significant public health needs or for rare disease indications with small patient populations that cannot be studied using a randomized design, and (3) increase efficiency compared to the timeline needed to carry out randomized controlled trials.

The RWE Alliance has submitted extensive comment letters in response to FDA draft guidance documents on the topics included in NIH's RFI on RWD.[4] Where appropriate, we encourage NIH to harmonize its programs and policies on RWD with those developed by FDA and other health regulatory agencies. To that end, we suggest that NIH consider adding a position in the Office of the Director to focus on advancing use of RWD research, as well as cross-agency collaboration and harmonization.

We offer some additional comments to select questions from the RFI on RWD below.

1(2) Barriers to using RWD in research, such as bias, underrepresentation of populations in data, and technical issues of data harmonization and linkage

While barriers to using RWD in research may exist, they can be mitigated through principled approaches to study design, data selection, and study conduct.

First, advanced methodological strategies can help to mitigate potential biases. They can range from propensity score models to synthetic cohorts' formation using digital twins, and we anticipate these strategies will only advance. In addition, use of Target Trial Emulation methods can lead to a principled study design that focuses on a specific research question and results in a methodological approach that mitigates bias and uncertainty.

Second, RWD can be a strong source of data for patients historically underrepresented in clinical trials. RWD can help achieve this research aim in part because of the heterogeneity of RWD sources. There are numerous benefits of heterogeneity in patient population characteristics and clinical practices. For instance, heterogeneity in population characteristics can provide additional insight into diverse and

---

[3] Additional information about what we believe is available on our website, https://rwealliance.org/what-we-believe/.

[4] The RWE Alliance's comment letters to FDA are available on our website, https://rwealliance.org/rwe-policy-developments/.

underrepresented groups, and these insights can help advance the important public health objective of ensuring medical research provides insights for all patients and helps address current disparities in our healthcare system.

Third, we believe that data linkage can provide a more complete picture of a patient's health. When it comes to linkage, we suggest researchers (1) ensure data transfer and linking methods are appropriately tested, reliable, and accurate, (2) explain these methods clearly in their write-ups of methods, and (3) rely on established standards, such as those in the Office of the National Coordinator for Health Information Technology's U.S. Core Data for Interoperability ("USCDI"), that support linking whenever possible and appropriate. This approach provides sponsors and other stakeholders with a practical recommendation to guide the process of linking systems. A well-designed and well-executed methodology should be the most important consideration for determining an effective approach to data linkage.

2(2) How do researchers assess the validation and verification of RWD data that is used in research

The necessary degree of validation will depend on the individual study question or use case. Although complete verification of a study variable may be ideal, it is likely not feasible in all situations due to various factors, such as a very large study population or the lack of reference standard data for all study subjects. In these scenarios, other methods can be used to improve the interpretability of data. For example, researchers can utilize analogous validated algorithms, which are validated in a different data source with similar parameters, or in a similarly structured study using the same data. Additionally, researchers can conduct sensitivity analyses to improve interpretability.

3(3) Availability/utility of emerging deidentification technologies and data storage/sharing considerations; 4(1) Strategies for protecting participant privacy and autonomy

Many RWD studies involve the secondary analysis of already deidentified data.

Artificial intelligence ("AI") is one example of an emerging technology that has the potential to enable new methods of protecting patient privacy. For instance, AI-generated synthetic data based on RWD datasets may present opportunities to benefit public health and help enhance patient data privacy, provided suitable privacy metrics are included with the appropriate evaluations of privacy and utility. For example, synthetic data can produce a "digital copy" of an RWD dataset, generated from machine learning and deep learning models built from the source data, to replicate the statistical characteristics and patterns of the underlying RWD. Continued research into this area may advance the state of the art in novel ways. AI-generated synthetic data, incorporating privacy metrics such as differential privacy, can also help to address privacy considerations around data sharing by enabling privacy-protecting analyses of synthetic copies of data that otherwise might not be available for research. AI-generated synthetic data may also amplify the creation of data by taking a relatively small underlying dataset and extrapolating until there is a sufficient quantity of data for

meaningful analysis, which could help support the development of hypotheses for treatments in rare diseases and research involving treatment effects in underrepresented populations, which can then be verified with RWD and clinical trials.

* * * *

The RWE Alliance appreciates NIH's RFI on the use of RWD in research. Thank you for considering these comments, and please let us know if you have any questions. We welcome the opportunity to discuss further.

Best regards,

The RWE Alliance