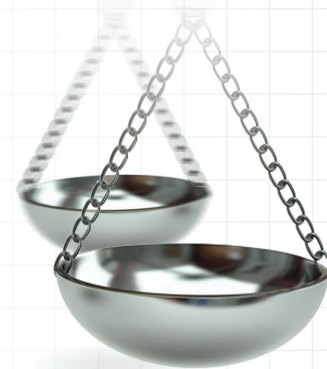




Learning from the past to advance tomorrow's real world evidence: what demonstration projects have to teach us

Journal of **Comparative Effectiveness Research**

Ashley Jaksa^{*,1}  & Nirosha Mahendraratnam¹

¹Aetion, Inc., 5 Penn Plaza, 7th Floor New York, NY 10001, USA

*Author for correspondence: Tel.: +1 810 919 0706; ashley.jaksa@aetion.com

“Untapped potential to harness RWE remains. Over the next 5 years, we believe the RWE community should focus on further developing these five areas: a consensus-driven research agenda, infrastructure, standardized process for validating RWD, and collectively assessing and adopting current best practices”

First draft submitted: 12 July 2021; Accepted for publication: 4 August 2021; Published online: 14 September 2021

Keywords: demonstration projects • guidance • health technology assessment • payers • pilots • real-world data • real-world evidence • regulators

Regulators, health technology assessment (HTA) agencies, and payers are actively exploring when, where, and how real-world evidence (RWE) can complement evidence generated from clinical trials and contribute to their decision-making. Demonstration projects bridge the gap between theory and application by evaluating untested uncertainties in the current RWE ecosystem and informing authoritative guidance recommendations [1]. Several demonstration projects have been initiated over the past few years to build credibility in RWE. Below, we summarize salient learnings from key demonstration projects where topics include understanding the underlying quality and reliability of real-world data (RWD), the role of study designs to estimate valid causal inferences, and how to incorporate RWE into healthcare decision-making. We also suggest next generation demonstration projects and actions to advance RWE adoption among decision-makers.

Key learning 1: credibility of underlying RWD is critical for widespread RWE adoption

RWD – data collected during routine clinical practice in claims, registries, electronic health records (EHR) – is the backbone of RWE. For high-quality RWE studies to inform decision-making, data reliability or the belief that data ‘adequately represent the underlying medical concept that they are intended to represent’ [2] is essential. Because RWD is often collected for purposes other than research, the data elements and the manner in which they are captured (including timing and instruments) can differ from those seen in clinical trials. For example, the Response Evaluation Criteria in Solid Tumors (RECIST) [3] is a tool often used in oncology trials to assess the response of an anticancer product on a tumor; however, it is rarely used in routine care. The lack of comparability between data collected in trials and available in RWD can lead to concerns regarding the reliability of RWD.

Demonstration projects have taken different – but complementary – approaches for building RWD credibility. Some projects focus on whether tools used in clinical trials can be retrofit into RWD sources. For example, a Flatiron project [4] found that it was challenging to retroactively apply the RECIST criteria in RWD-sources given missing imaging data in EHR and developed and tested alternative non-RECIST-based approaches to assess tumor-based outcomes. Other studies focused on validating outcomes that can be measured in RWD by comparing the correlation of different outcomes within a data source, and the consistency of outcomes [5] across real-world sources, and against trial end points to build their credibility. One such example [6] project in oncology assessed the correlation of intermediate end points such as real-world time to discontinuation with the clinical benefit, real-world overall survival. In addition, the innovative Friends of Cancer Research Real-World Evidence pilot programs aimed [7] to validate variables in RWD by bringing together multiple data partners with different curation

methods, models, and data types to assess the ability to commonly define and consistently assess variables across different datasets.

Several demonstration projects are also underway to understand and validate the use of novel measurement tools such as sensors and other digital measurement tools. For example, Apple, Evidation and Eli Lilly are developing cognitive health measures based on multiple sensor streams [8] including iPhone applications, Apple Watch and Beddit sleep monitor versus traditional tools that are used to measure cognitive impairment that rely on physician or patient-based questionnaires (e.g., the Mini-Mental State Examination). While these novel measurement tools have been widely adopted in the general population, the use of these tools in healthcare decision-making is limited, making further validation testing an essential step toward their use.

While each RWD demonstration project provides narrow validation insights in the targeted therapeutic area(s), data source(s), and research question(s), these projects demonstrate the utility of RWD and the circumstances in which RWD are reliable and that proxy or alternative measures can be used when the specific measure of interest is not captured in the real-world.

Key learning 2: credibility of causal inference is driven by study design

The lack of randomization in RWE studies has led to skepticism in the ability to study causal questions in RWD [9], which is an essential proof point for broad RWE adoption in regulatory, HTA and payer decision-making. A number of projects have focused on validating real-world findings based on causal conclusions by replicating randomized controlled trials (RCTs) in RWD. The goal of these replication studies is to compare and calibrate the RWE results against the ‘gold-standard’ RCT.

Several projects have focused on emulating RCTs in administrative claims or EHR data [10–13]. These projects attempt to use the same inclusion/exclusion criteria, exposure, and outcomes, as the RCT to estimate the same treatment effect and thus make the same regulatory conclusion. For example, the RCT Duplicate Initiative [12] is emulating 30 RCTs and an additional seven ongoing RCTs in claims data in Diabetes and Cardiovascular disease states. These replication efforts demonstrate the ability of RWE to answer causal questions when principled epidemiologic methods are employed. Furthermore, study design outweighs analytical methods [14] as the most critical component of making valid inference – analytical methods cannot fix study design flaws or poor quality data.

While emulating RCTs is not the end goal of RWE, numerous successful emulations are creating a repository of cases that can increase the predictability of future RWE studies, identify challenging areas for RWE, and increase confidence in common RWE methodological approaches. Already we see RWE can lead to valid inference [14] when there is a large effect size, objective end point, active comparator and evidence that residual confounding is unlikely.

Key learning 3: role of RWE in decision-making on a global scale

Stakeholders have released several RWD/RWE recommendations on conducting RWE [1]. To operationalize these recommendations, decision-makers are developing policies, recommendations and guidance on RWE use cases both pre- and postlaunch. For example, the US FDA Center for Devices and Radiological Health published a review [15] of 90 RWE submissions between 2012 and 2019 that detailed when RWE was used through the medical device total product lifecycle and how it was used. The US-based Institute for Clinical and Economic Review is piloting re-assessments [16] of accelerated approval products 24-months after its initial assessment. The goal is to determine how RWE can be used to reduce uncertainties in the first assessment and further refine estimates of the drug’s cost–effectiveness. Similarly, the Dental and Pharmaceutical Benefits Agency (TLV), the HTA agency in Sweden, is conducting numerous studies [17] to evaluate how RWD can be used to continuously perform follow-up studies on utilization and treatment effect in clinical practice to inform gaps where RCT data cannot be used. The National Institute for Health and Care Excellence in collaboration with Flatiron [18] are focused on identifying the most appropriate methods to evaluate real-world survival in oncology patients after treatment launch and how results compare to survival estimated in the RCTs. These RWE use cases highlight the gaps in knowledge between RCTs and clinical practice and where RWE can complement RCTs.

Next generation of demonstration projects to propel RWE use

With the help of demonstration projects, the RWE community has made substantial headway in accepting RWE for decision-making. Regulatory, HTA, and payer decision-makers are focused on integrating RWE into their processes; however, full integration has not occurred. Untapped potential to harness RWE remains. Over the next

5 years, we believe the RWE community should focus on further developing these five areas: a consensus-driven research agenda, infrastructure, standardized process for validating RWD, and collectively assessing and adopting current best practices.

Consensus-based research agenda

Recent research has shown that many stakeholders have issued policies, recommendations or guidance on similar, often overlapping RWE topics [1]. These recommendations have high-levels of agreement but are not completely aligned; collaboration and alignment between decision-makers on recommendations may speed development of guidance and increase validity of RWE studies. We suggest a community-wide research agenda to help prioritize future research questions and infrastructure projects all competing for funding by the public and private sectors. This research agenda would not only detail the next generation of demonstration projects (e.g., advanced analytics, tools such as master RWE protocols), but also prevent unnecessary duplication of efforts. Third-party, independent conveners such as Innovative Medicines Institute, National Academies of Sciences, Engineering and Medicine or the Duke-Margolis Center for Health Policy can help develop and execute on this agenda with major stakeholders (e.g., sponsors, ISPE, ISPOR, academics and decision-makers) and government input.

RWD/E infrastructure

A variety of infrastructure enhancements encompassing data systems, evaluation tools and research guidance are needed to buttress current standards.

From a data perspective, essential data elements required to answer research questions such as race and ethnicity data are not consistently collected or accessible [19]. For example, administrative claims sources rarely capture race/ethnicity data. Furthermore, race and ethnicity data collection in EHR sources are variable and may not be available due to privacy restrictions. Several efforts are underway to develop a set of minimum required data elements with the goal of enhancing research data sets and interoperability [20,21]. The practicality and feasibility of implementing requisite data collection in routine care and their usability in research have yet to be tested. Once these minimally required data are collected, standards on how to evaluate whether they are sufficiently reliable to inform decision-making are also necessary. While there has been progress in developing tools to evaluate RWD quality (e.g., REQuEST [22]), these tools are often limited to a specific type of RWD (e.g. registries) and lack concrete criteria to determine if quality metrics are met and thresholds for what is ‘good’ quality [1].

From a study design perspective, we must continue to advance methods and research tools to make valid causal inference. Hand-in-hand, we must accelerate access to this knowledge by developing tools for researchers, promoting transparency and creating programs to democratize the RWE landscape assessment to enable the generation of high-quality research from different perspectives. While tools like the START RWE Template [23] enable researchers with the appropriate capabilities and expertise (i.e., epidemiology, clinical, biostatistical and data science) to execute principled epidemiological studies, not all research teams have access to such expertise or know how to implement these tools. Publicly publishing protocols not only enables transparency, but also allows researchers to leverage state-of-the-art protocols to guide their studies in other data sets. More efforts to democratize RWE study conduct and facilitate RWE learnings can empower a new generation of researchers to continue to build robust evidence on a research topic. For example, the COVID-19 Evidence Accelerator [24] convenes government, clinical, academic, data, analytics, technology and payer stakeholders to answer critical COVID-19 questions using a common protocol and set of data definitions in real-time. The COVID-19 Evidence Accelerator also hosts weekly research meetings to foster an open-forum to discuss in-progress work and share collective learning and expertise for the greater research community’s benefit.

Develop standardized process for validating RWD

One of the most compelling benefits of RWD is the ability to readily access data to evaluate outcomes that are important to patients but are often absent from clinical trials. However, it is impractical to develop demonstration projects for every measurement necessary across every disease state to demonstrate its reliability. Instead, a standardized and harmonized process for validating real-world measures based on pharmacoepidemiology’s long-history for developing and validating algorithms in claims data, innovative curation tools such as machine learning, and potential frameworks such as the Duke Margolis Center for Health Policy’s Developing Real-World End points for Regulatory Use Roadmap [25] of the Digital Medicine Society’s Playbook for Digital Measures [26] can provide structure for a repeatable process that instills confidence in their credibility.

Adopt current best practices for critical appraisal of RWE

While the RWE research community has best practices to follow, it is not often clear if decision-makers embrace or adopt them [1]. Where appropriate, future demonstration projects should focus on validating these tools and checklists so that decision-makers can officially adopt them in their RWE guidelines. For example, the SPACE framework [27] provides researchers with a step-by-step process for designing RWE studies from articulating the research question through providing decision-makers with justification for design choices. It is unclear if regulators and HTAs will accept these templates in submissions. We recommend dedicating demonstration studies to evaluate the utility of these standard templates for RWE-related regulatory and HTA submissions to demonstrate the template's value and identify potential shortcomings preventing decision-makers from endorsing their wide-spread use.

Expanding on appropriate use of RWD & RWE in decision-making

Demonstration projects have already begun to identify when and how RWE can supplement RCT data. However, as RWE science evolves, additional demonstration projects and use cases can accelerate the continued expansion of RWE use and pinpoint novel circumstances where RWE can be used. For example, one of the most common RWD use cases is supporting regulatory decisions for serious and life-threatening rare diseases [28], where it is often infeasible or unethical to conduct RCTs. In these circumstances, high-quality RWD can be used to contextualize the safety and effectiveness results of single-arm studies. Could RWE expand into highly crowded disease areas where there are limited patients for trials? For example, could a single high-quality RWD control arm be created and implemented through a precompetitive collaboration that implements a platform trial? Lessons learned from such use cases should be collected, centralized, and shared with the broader community to continue to advance the field.

Conclusion

Demonstration projects are an essential bridge to move to wider and appropriate use of RWE in healthcare decision-making. While progress has been made, we believe it is important to reflect on what we have learned thus far and develop a consensus on the next generation of demonstration projects. Continuing to invest in projects that strengthen infrastructure, adopt current best practice and explore expanded use cases of RWE are the high priority areas that will move us toward maximizing the utility of RWE.

Financial & competing interests disclosure

A Jaksa and N Mahendraratnam are both employees at Aetion Inc., and A Jaksa owns stock options at Aetion Inc. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Acknowledgments

We would like to thank Patra Mattox for her editing and valuable feedback and Nicolle Gatto and Mark Stewart for sharing their deep expertise.

Open access

This work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>

References

1. Jaksa A, Wu J, Jónsson P, Eichler H-G, Vittoe S, Gatto NM. Organized structure of real-world evidence best practices: moving from fragmented recommendations to comprehensive guidance. *J. Comp. Eff. Res.* 10(9), 711–731 (2021).
2. FDA. Framework for FDA's Real-World Evidence Program. US Food and Drug Administration (2018). www.fda.gov/media/120060/download
3. Eisenhauer EA, Therasse P, Bogaerts J *et al.* New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *Eur. J. Cancer* 45(2), 228–247 (2009).
4. Griffith SD, Tucker M, Bowser B *et al.* Generating real-world tumor burden endpoints from electronic health record data: comparison of RECIST, radiology-anchored, and clinician-anchored approaches for abstracting real-world progression in non-small cell lung cancer. *Adv. Ther.* 36(8), 2122–2136 (2019).

5. Khozin S, Miksad RA, Adami J *et al.* Real-world progression, treatment, and survival outcomes during rapid adoption of immunotherapy for advanced non-small cell lung cancer. *Cancer* 125(22), 4019–4032 (2019).
6. Stewart M, Norden AD, Dreyer N *et al.* An exploratory analysis of real-world end points for assessing outcomes among immunotherapy-treated patients with advanced non-small-cell lung cancer. *JCO Clin. Cancer Inform.* 3, doi: 10.1200/CCI.18.00155 (2019) (Epub ahead of print).
7. Friends of Cancer Research. Considerations for use of real-world evidence in oncology (2020). https://friendsofcancerresearch.org/sites/default/files/2020-10/Use_of_Real-World_Evidence_in_Oncology_0.pdf
8. Chen R, Jankovic F, Marinsek N *et al.* Developing measures of cognitive impairment in the real world from consumer-grade multimodal sensor streams. In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM, AK, USA, 2145–2155 (2019). <https://dl.acm.org/doi/10.1145/3292500.3330690>
9. Collins R, Bowman L, Landray M, Peto R. The magic of randomization versus the myth of real-world evidence. *N. Engl. J. Med.* 382(7), 674–678 (2020).
10. Using Veradigm's EHR Dataset to Replicate ROCKET-AF Clinical Trial (2021). <https://veradigm.com/veradigm-news/ehr-dataset-to-replicate-rocket-af-clinical-trial/>
11. Optum Labs. Using real-world data in regulatory decision-making (2021). www.optumlabs.com/work/data-regulatory-decision.html
12. Franklin JM, Pawar A, Martin D, *et al.* Nonrandomized real-world evidence to support regulatory decision making: process for a randomized trial replication project. *Clin. Pharmacol. Ther.* 107(4), 817–826 (2020).
13. US Food and Drug Administration. Leveraging real-world data and shared clinical trial data to inform regulatory decision-making, Yale University-Mayo Clinic CERSI. U.S. Food and Drug Administration (2019). www.fda.gov/science-research/advancing-regulatory-science/leveraging-real-world-data-and-shared-clinical-trial-data-inform-regulatory-decision-making-yale
14. Franklin JM, Patorno E, Desai RJ *et al.* Emulating randomized clinical trials with nonrandomized real-world evidence studies: first results from the RCT DUPLICATE Initiative. *Circulation* 143(10), 1002–1013 (2021).
15. US Food and Drug Administration. Examples of real-world evidence (RWE) used in medical device regulatory decisions (2021). www.fda.gov/media/146258/download
16. Institute for Quality and Efficiency in Health Care. Hereditary angioedema. ICER (2018). <https://icer.org/assessment/hereditary-angioedema-2018/>
17. Gustafsson S, Johansson P, Ponten J, Stromgren A, Viber A. Follow-up of drug utilisation and treatment effects in clinical practice (2018). www.tv.se/in-english/reports/arkiv/2020-12-09-follow-up-of-drug-utilisation-and-treatment-effects-in-clinical-practice.html
18. Flatiron Health. NICE Partners with Flatiron Health to develop real-world evidence research methodologies (2020). <https://flatiron.com/press/press-release/nice-partnership-2020/>
19. Tarver ME. Race and ethnicity in real-world data sources: considerations for medical device regulatory efforts. *J. Prim. Care Community Health* 12, <https://doi.org/10.1177/2150132721994040> (2021) (Epub ahead of print).
20. Kush RD, Warzel D, Kush MA *et al.* FAIR data sharing: the roles of common data elements and harmonization. *J. Biomed. Inform.* 107, 103421 (2020).
21. HealthIT.gov. United States Core Data for Interoperability (USCDI) | Interoperability Standards Advisory (ISA) (2021). www.healthit.gov/isa/united-states-core-data-interoperability-uscdi
22. EUnetHTA JA3. Registry Evaluation and Quality Standards Tool (REQueST®). EUnetHTA (2019). <https://eunetha.eu/request-tool-and-its-vision-paper/>
23. Wang SV, Pinheiro S, Hua W *et al.* STaRT-RWE: structured template for planning and reporting on the implementation of real world evidence studies. *BMJ* 372, m4856 (2021).
24. Evidence Accelerator (2021). <https://evidenceaccelerator.org/>
25. Mercon K, Mahendratnam N, Eckert J *et al.* A Roadmap for Developing Study Endpoints in Real-World Settings. *Margolis Center for Health Policy* (2020). <https://healthpolicy.duke.edu/publications/roadmap-developing-study-endpoints-real-world-settings>
26. The Digital Medicine Society. The Playbook – Digital Clinical Measures (2021). *The Playbook* <https://playbook.dimesociety.org/>
27. Gatto NM, Reynolds RF, Campbell UB. A structured preapproval and postapproval comparative study design framework to generate valid and transparent real-world evidence for regulatory decisions. *Clin. Pharmacol. Ther.* 106(1), 103–115 (2019).
28. Mahendratnam N, Mercon K, Gill M, Benzing L, McClellan MB. Understanding use of real-world data and real-world evidence to support regulatory decisions on medical product effectiveness. *Clin. Pharmacol. Ther.* doi: 10.1002/cpt.2272 (2021) (Epub ahead of print).

