

Small But Mighty: The Use of Real-World Evidence to Inform Precision Medicine

Rebecca A. Miksad^{1,*†} , Meghna K. Samant^{1,†}, Somnath Sarkar¹ and Amy P. Abernethy^{1,2}

The intersection of information technology and biomedical breakthroughs creates new opportunities for real-world evidence (RWE). Initially used in comparative effectiveness and population health research, RWE now can bolster personalized care. Precision medicine segments diseases into small, biologically informed subgroups. For these rare populations, small-cohort RWE can inform clinical, regulatory, and development decision making as well as personalized care, but analysis and interpretation require robust and thoughtful application of quantitative and qualitative research methods.

Precision medicine has transformed clinical care areas, such as neurodegenerative diseases, rheumatology, and psychiatry but most notably oncology. Although targeted therapy has revolutionized treatment for tumors driven by aberrant but modifiable molecular features, significant unmet needs remain, particularly for patients with rare mutations.¹

Randomized controlled trials (RCTs) for precision medicine face clinical, ethical, and logistical challenges. Enrollment of rare populations can be protracted and costly. Molecular analyses requirements may delay treatment. Control arm randomization to ineffective standard therapy raises ethical concerns.

In these contexts where accrual is difficult, small-cohort RWE can supplement and complement traditional clinical trial evidence, assess real-world adverse events (AEs; e.g., postmarketing), evaluate pan-tumor effectiveness, or serve as contemporary, single-arm trial comparator. Small-cohort RWE may provide control-arm benchmark and enrollment feasibility estimates for trial design and planning. However, small cohorts impose analytical trade-offs and warrant careful quantitative and qualitative methodological considerations.

ILLUSTRATIVE USE CASE

Neurotrophic tropomyosin receptor kinase (*NTRK*) gene fusions drive oncogenesis in a small subset of solid tumors (estimated US annual cases, 3,000). In this exceedingly rare population, clinical data are limited, and traditional clinical trials are challenging. These alterations, however, are remarkably sensitive to tropomyosin receptor kinase (TRK) inhibition in many tumor types. The TRK inhibitor larotrectinib received the second-ever pan-tumor approval from the US Food and Drug Administration (FDA; 2018), based on the high overall response rate (75%) in a small clinical study (55 patients with 17 different tumor types).²

Technological innovations (e.g., high-quality electronic health record (EHR) databases) enable the identification of real-world patients with *NTRK*-mutant tumors from a geographically diverse clinicogenomic database. Even small cohorts greatly expand the evidence base for this rare population, improving generalizability of TRK inhibitor trial results (e.g., representing excluded comorbidities). Patient narratives derived from EHR data can add nuance, especially about patient outliers, such as those without a deep response (**Figure 1, left**). Real-world genomic data sources³ or prospective, genomically informed registries can deepen overall clinicobiologic understanding of patients with *NTRK* gene fusions.

In such a scenario, small-cohort RWE can be a valuable resource for the development of critically important treatment options. However, key considerations must be taken into account to ensure its validity and accuracy (see **Figure 1, right**, for a summary view).

CONSIDERATIONS FOR SMALL-COHORT RWE

Small-cohort RWE from EHRs can uniquely support quantitative assessments with the nuance and context of detailed individual-level qualitative data. Therefore, considerations relate to both quantitative and qualitative research methods.

Methodological considerations for quantitative research

1. Selection bias

The use of prespecified and objective cohort selection criteria (avoiding “cherry-picking”) is imperative. Bias occurs if baseline factors that impact outcome also influence the underlying cohort selection mechanisms (confounders). In our precision medicine example, age could influence the potentially eligible population: genomic testing practices may favor younger, fitter patients. Nonrepresentative cohorts may yield inaccurate outcome estimates and inflated type I error rates (the probability of falsely concluding a treatment is effective). This is particularly relevant in small cohorts because each data point has great leverage on outcome estimate accuracy.

When real-world data (RWD) is used as a control for evaluating treatment effectiveness in clinical trials, close alignment of the cohort selection and trial inclusion/exclusion criteria is critical. Identifying every possible patient may require creative and comprehensive efforts, such as extracting genomic information consistently and comprehensively from PDF-based reports using trained experts, unbiased machine learning, and natural language processing. Data should be collected across geographic regions and clinical sites and evaluated for representativeness using diagnostics, such

¹Flatiron Health Inc., New York, New York, USA; ²Current address: Food and Drug Administration, Silver Spring, Maryland, USA. *Correspondence: Rebecca A. Miksad (rmiksad@flatiron.com)

[†]These authors contributed equally to this work.

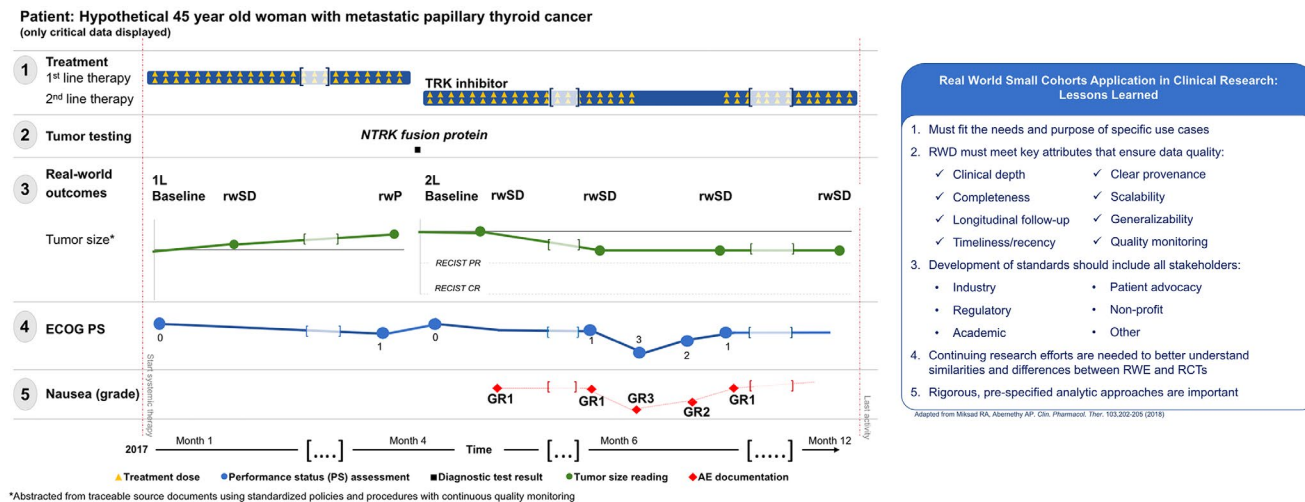


Figure 1 Summary of key aspects of real-world evidence (RWE) research. (Left) The availability of longitudinal information enables melding of quantitative and qualitative results. Visualizing the patient’s clinical course provides qualitative context for small cohort RWE quantitative results. This hypothetical case of a patient with a neurotrophic tropomyosin receptor kinase (*NTRK*) gene fusion demonstrates the relationship among critical clinical events, diagnostic results, and treatment. For example, the lack of continued decrease in tumor burden on the second scan after starting tropomyosin receptor kinase (TRK)-inhibitor treatment (row 3) may be related to missed doses during the seventh month of systemic treatment (row 2), which was associated temporally with decreased performance status (row 4) and nausea (row 5). (Right) Overview of the lessons learned for the optimal application of RWE in clinical research focused on small cohorts. RCT, randomized controlled trial; RECIST, response evaluation criteria in solid tumors; RWD, real-word data; rWP, real-world progression; rWSD, real-word stable disease.

as patient distribution across sites, comparisons of demographic and prognostic characteristics, and covariate balance plots. Where applicable, statistical methods, including inverse probability of treatment weighting, double robust regression adjustment, and matching may help address patient population differences.

Time period selection is also important. Temporal outcome drifts (due to supportive care improvements, introduction of new interventions, or other factors) can significantly bias results, leading to erroneous conclusions about treatment effectiveness (type I and II errors).⁴ The extent and impact of such temporal drifts need to be scrutinized by aggregating RWE over time, or with sensitivity analyses restricted to contemporaneous time periods with relatively stable clinical practice for the disease of interest. Finally, long-term follow-up is important to assess result robustness over time; continuous RWD aggregation makes this possible, unlike the typically limited RCT follow-up.

2. Variability and precision

Small cohorts tend toward variability and imprecision in outcome estimation, because precision correlates with sample size (Figure 2, left), often resulting in low statistical power to detect small but clinically meaningful differences. The latter, however, may be mitigated by the anticipated large effect sizes of precision-medicine therapeutics compared with standard of care (Figure 2, right).

3. Data quality and completeness

Given the high impact of individual data points on the accuracy and precision of small-cohort outcome estimates, data quality and completeness is paramount. Rigorous systematic and continuous

monitoring, including automated checks and clinical review, can help maximize data quality by evaluating the extent and impact of missing data, scrubbing implausible values, and pressure-testing edge cases. Maximal completeness on key prognostic, clinical, and genomic information is critical to describe the patient population and its underlying heterogeneity. In small cohorts, complementing structured data with detailed data abstracted from unstructured documents can be a crucial bridge to qualitative research.

4. End-point selection

Careful selection of relevant real-world end points to address the scientific objectives of a given study is important. RWE may necessitate end point definitions distinct from highly regimented clinical trial criteria, such as the response evaluation criteria in solid tumors (RECIST). End points (e.g., response, progression, and mortality) need to be reliable and validated.

5. Appropriate statistical methodology and clinicogenomic result contextualization

Finally, good scientific and statistical principles, such as prespecification of objectives, patient population, analytical methods, and assumptions, via the estimand framework or otherwise, ensure high result integrity and credibility. Bayesian approaches that naturally allow incorporation of external information to improve estimation may also be considered. To promote efficiency in trial designs, combination of RWD and clinical trial data for hybrid study arms warrants further exploration. Critically, result interpretation must be informed by a deep understanding of clinical and genomic details and analytic limitations.

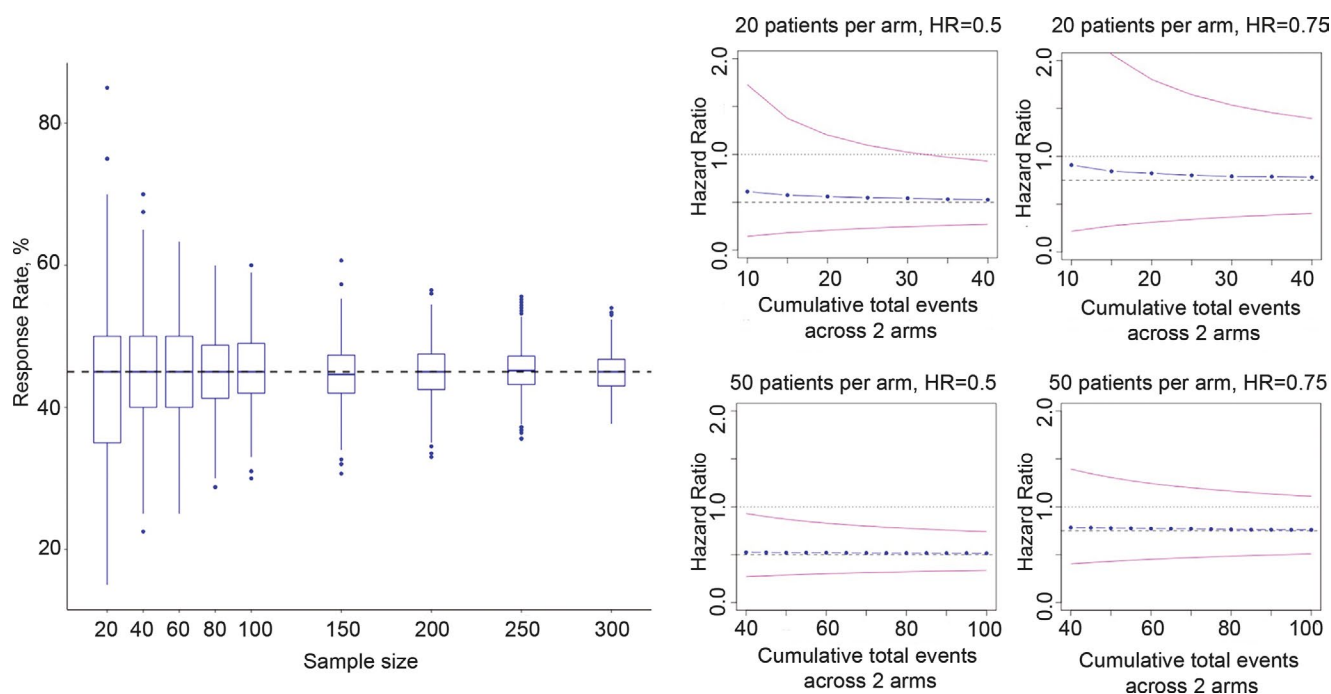


Figure 2 Variability and precision in small cohort analyses. (Left) Box plot of estimated response rates based on 1,000 simulated samples using true response rate of 45% (dashed horizontal black line). Blue dots represent outliers. This simulation evaluates the relationship between sample size and precision for response rate estimation (precision being the inverse of the variability represented by the box plot span). (Right) Simulation results evaluating the relationship among the number of events, effect size, and precision of hazard ratio (HR) estimate. Note that when the true HR between an experimental regimen and a real-world comparator is 0.5, HR estimates have high precision and meaningful confidence intervals even with low absolute event counts (blue dots, median estimated HR over 1,000 simulations; pink lines, 95% confidence intervals).

Considerations to capitalize on qualitative research

1. Longitudinality

Individual clinical narratives (collated in detail from clinical notes, radiology, or pathology reports, etc.) can add qualitative nuance and depth to quantitative results. Longitudinal visual displays can strengthen/confirm and contextualize overall findings (Figure 1, left, shows a hypothetical example of a real-world patient in an *NTRK* fusion-positive cancer small cohort). In our hypothetical case, a patient receiving an TRK inhibitor experiences tumor burden reduction that is durable but short of response by RECIST, possibly due to treatment hold (Figure 1, left). Quantitatively, this is “stable disease”; however, the qualitative narratives of individual patients in the cohort may highlight near-ubiquitous “tumor shrinkage.”

2. Narrative depth

Narrative review can be highly specific and sensitive for adverse event detection. For example, cardiac toxicity identification in patients with underlying heart disease (typically excluded from clinical trials) may be enhanced by review of hospital admission records for “shortness of breath,” to determine whether: (i) hospitalization temporally followed drug administration, (ii) an echocardiogram documented decreasing ejection fraction, and/or (iii) congestive heart failure treatment was initiated. Advanced technology solutions enable a fuller qualitative understanding of the complexity of care delivered throughout the clinical course.

3. Traceability to source

Qualitative source documents (e.g., medical notes, radiology reports, etc.) can clarify uncertain quantitative findings (e.g., do abnormal liver function results reflect viral hepatitis or drug toxicity?). Contemporary technology allows rapid RWD verification for accuracy, relevance, and context, as well as annotation for uses, such as machine learning, balancing rich data extraction, and patient privacy concerns.

Other qualitative research methods, such as triangulation (using multiple sources to produce rich, robust, comprehensive, and well-developed understanding), consideration of the impact of reliability and validity on result interpretation, and saturation assessment (i.e., do additional data lead to any new emergent themes?⁵) will likely contribute to the evolution and maturation of small-cohort RWE science.

Finally, RWE limitations are well documented elsewhere. RWE is not suitable for all use cases. In the context of precision medicine, RWE may be limited by practical factors, including lags in uptake of testing/treatment, diagnostic assay variability, and differences in outcome assessments.

CONCLUSIONS

By supplementing and complementing the clinical trial data for rare patient populations, small-cohort RWE can contribute mightily to the decision-making process in precision medicine, from understanding prevalence to assessing effectiveness and

safety of real-world practice patterns. Careful RWD generation and curation, thoughtful planning and rigorous quantitative and qualitative research execution, and close attention to the clinicogenomic context maximizes the value of small-cohort RWE.

ACKNOWLEDGMENT

The authors wish to acknowledge Julia Saiz (Flatiron Health, Inc.) for editorial support.

FUNDING

This paper does not report results from a funded study. All authors are or have been employees of Flatiron Health.

CONFLICT OF INTEREST

All authors are or have been employees of Flatiron Health, Inc., an independent subsidiary from the Roche Group, and own or have owned equity in Flatiron Health, and stock in Roche. In addition, R.A.M. is an advisor for the De Luca Foundation. M.S. owns stocks of Hoffmann-La Roche, and has served on the executive committee of EPPIC Global. At the time of this work, A.P.A. was chief medical, chief scientific officer, and senior vice president of oncology at Flatiron Health, Inc., which is an independent subsidiary of the Roche Group, and had stock ownership in Roche. At that time, she also declared the following: serving on the board of directors and stock ownership of Athenahealth and CareDx; owner of Orange Leaf Associates, LLC; senior advisor of Highlander Partners; advisor of SignalPath Research, RobinCare, and KelaHealth, Inc.; special advisor of The One Health Company; receiving honoraria from Roche/Genentech (<US \$10,000 per year); and having a patent pending for a technology that facilitates the extraction of unstructured information from medical records. All of these relationships ended on or before January 31, 2019, predating her joining the US Food and Drug Administration (FDA), except for Orange Leaf Associates, LLC. Since joining the FDA, A.P.A. has complied with applicable ethics laws.

DISCLAIMER

Dr Abernethy participated in this work prior to joining the FDA. This work and related conclusions reflect the independent work of study authors

and does not necessarily represent the views of the FDA or the United States.

© 2019 Flatiron Health. *Clinical Pharmacology & Therapeutics* published by Wiley Periodicals, Inc. on behalf of American Society for Clinical Pharmacology and Therapeutics.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

[The copyright line for this article was changed on 04 June 2019 after original online publication].

1. Goldberg, K.B., Blumenthal, G.M., McKee, A.E. & Pazdur, R. The FDA oncology center of excellence and precision medicine. *Exp. Biol. Med. (Maywood)* **243**, 308–312 (2018).
2. US Food and Drug Administration (FDA). FDA approves an oncology drug that targets a key genetic driver of cancer, rather than a specific type of tumor [News Release] <<https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm626710.htm>> (2018, November 26). Accessed December 20, 2018.
3. Agarwala, V. et al. Real-world evidence in support of precision medicine: clinico-genomic cancer data as a case study. *Health. Aff. (Millwood)* **37**, 765–772 (2018).
4. Tang, H., Foster, N.R., Grothey, A., Ansell, S.M., Goldberg, R.M. & Sargent, D.J. Comparison of error rates in single-arm versus randomized phase II cancer clinical trials. *J. Clin. Oncol.* **28**, 1936–1941 (2010).
5. Saunders, B. et al. Saturation in qualitative research: exploring its conceptualization and operationalization. *Qual. Quant.* **52**, 1893–1907 (2018).