Sarcouncil Journal of Biomedical Sciences

ISSN(Online): 2945-3666

Volume- 05| Issue- 01| 2026

Research Article

Received: 20-12-2025 | Accepted: 15-01-2026 | Published: 31-01-2026

Audit-Ready Data Governance and Quality Frameworks for Real-World Evidence In Oncology Clinical Trials

Gerald Nkogbu[1] and Isaac Nana Kissi Darko[2]

[1]*Fox Chase Cancer Center, Philadelphia, USA*

[2]*Department of Biochemistry, Cell and Molecular Biology, University of Ghana, Ghana*.

Abstract: Real-world evidence (RWE) from electronic health records, registries, claims, and other observational sources is increasingly utilized to support clinical trial design, external control arms, and regulatory decision-making. However, inconsistent governance, variable data quality, and limited provenance documentation hinder regulatory acceptance and auditability. This review synthesizes current regulatory expectations, data-quality frameworks, and oncology-specific challenges, alongside an operational "audit-ready" framework that integrates governance, technical controls, and documentation to ensure trustworthy RWE in oncology. Key recommendations include mandated provenance capture, adoption of common data models, and standardized data-quality assessment with pre-defined thresholds for audit packages submitted to regulators.

Keywords: Real-world evidence; oncology; data governance; data quality; audit-ready; regulatory.

INTRODUCTION

An "audit-ready data governance" guarantees that data is accurate, comprehensive, consistent, and traceable, enabling a company to prove compliance during an audit confidently (Weiss & Tuttle, 2006; Yap *et al*., 2022). This requires complete traceability of data collection and analysis, such as clear proof of data origin, quality checks, transparent metadata, and version management. To proactively prepare for audits and lower risk, it entails creating explicit policies, putting technologies for data visibility, and building a culture of data integrity (Siden *et al*., 2002; Weiss *et al*., 1993). To achieve this, clinical trial companies ensure access controls, automatic audit trails, and regular data quality monitoring as crucial steps to accomplish. More specifically, clinical trial companies use quality management systems to ensure robust documentation practices and provide clearly defined standard operating procedures, which are regarded as a high benchmark (Bernardo *et al*., 2024), not only to ensure audit readiness but also to guarantee overall regulatory compliance and safety during clinical trials (USFDA, 1997). Overall, these critical steps establish a thorough system of supervision to guarantee data integrity and ethical behavior necessary for audit-ready data governance, especially for oncology clinical trials (Buse *et al*., 2023).

However, In oncology clinical trials, the complexity of the disease burden and strict eligibility requirements necessitate audit readiness and high cooperation with regulatory officers, not only to ensure patient safety but also to ensure effective decision-making and ensure the availability of therapies to patients. As such, there is high reliance on Real-world Evidence data (RWE). Real-world evidence (RWE) comes from real-world data (RWD) like electronic health records (EHRs), cancer registries, claims data, and patient-generated data, among others (Agarwala *et al*., 2018). RWE data increasingly shape the development of oncology drugs, clinical guidelines, and promote regulatory decision-making. Generally, Real-world data (RWD) describes data generated or obtained outside of conventional clinical trials (Wilson & Booth, 2024). The US Food and Drug Administration (FDA) strongly encouraged using well-curated RWE for post-authorization safety studies and as additional support in clinical development. However, to ensure RWE is trusted in oncology, a field marked by complex treatment paths and different documentation practices, there is a need to ensure that data governance and quality assurance are clear and ready for regulatory oversight (Wang *et al*., 2022).

Importantly, the ever-growing use of real-world evidence in oncological research has undoubtedly shifted the paradigm in which we generate evidence in cancer care. Clinical decisions and regulatory assessments have traditionally leaned on randomized controlled trials. While these trials boast strong internal validity, they often lack external generalizability due to strict eligibility criteria and highly controlled treatment environments and differences in regulatory regimes (Chon & Alexander, 2023). As a result of

***Corresponding Author:** Isaac Nana Kissi Darko

DOI- https://doi.org/10.5281/zenodo.18444568

such stringency, the populations participating in randomized controlled trials (RCTs) often differ significantly from real-world oncology patients, who typically present with multiple comorbidities (Jiang *et al*., 2022), variable performance status, and different treatment patterns (Cho *et al*., 2023). To supplement RCTs, regulators and researchers are using an ever-increasing amount of RWE. The challenge, however, is that this data may come from several sources, including electronic health records, administrative claims, and molecular diagnostics databases (Basch *et al*., 2015; Garnett *et al*., 2011; Senderowicz, 2010). The USFDA and EMA have documented that RWE could be one source of regulatory-grade evidence for oncology, provided that the underlying data is shown to be fit for its purpose and to meet standards established for quality and transparency.

To ensure uniformity and generality of RWE in decision making and to ensure safety of clinical regulatory decisions in clinical oncology, this review brings together oncology-specific and general RWD governance frameworks. It suggests audit-ready data governance and quality model designed for oncology clinical research. The model will address some of the key gaps in tracking data sources, methods of validation, and specific use cases.

DISCUSSION

Barriers to Audit-Ready RWE in Cancer Trials

New opportunities are presented by rapidly expanding oncology of RWD resources, including decentralized clinical data sources, federated data platforms, genomic-linkage databases, and curated networks like Flatiron Health in the USA. Collectively, these opportunities enhance our knowledge of the effectiveness of treatment strategies, help accelerate post-approval safety monitoring, and inform regulatory submissions (Senderowicz, 2010; Yap *et al*., 2022).

However, the trustworthiness of RWE by stakeholders relies heavily upon the governance, origin, and quality of the data attributes, which are generally subjective and differ significantly from different clinical sites (Blonde *et al*., 2018; Corrigan-Curay *et al*., 2018; H. Yuan *et al*., 2018). Particularly in oncology clinical trials, RWE faces persistent challenges that hamper its reliability, reproducibility, and potential to be audited (Basch *et al*., 2015). Often, cancer clinical trial delays and increased trial costs are caused by the complexity of RWE from clinical trials and fragmented data across many providers and health systems, which makes treatment strategy decision-making more challenging and ineffective. Interestingly, these challenges arise from critical disease-specific variables such as stage, biomarker status, treatment sequences, and progression events that often remain embedded within the unstructured text of Natural Language Processing (NLP) or rigorous human abstraction protocols or are documented inconsistently (Chan *et al*., 2020; Eskola *et al*., 2023). Such variability increases the risk of misclassification and bias, while also contributing to missing or incomplete information about the patient journey and the broader clinical decision-making process (L. Yuan *et al*., 2024).

Furthermore, the fragmentation of clinical research to geographically dispersed sites greatly hinder participation in cancer clinical trials, even though patient access to novel therapies is becoming increasingly vital (Feinberg *et al*., 2020).

Besides the structural challenges, (Graili *et al*., 2023; Liu & Demosthenes, 2022) argued about the vast variability of RWD data models, such as Fast Healthcare Interoperability Resources, coding practices, making integration and comparison of datasets challenging, not just for different sites, but also hindering auditing and regulatory compliance. Significantly, key oncology end-point outcomes, line of therapy, real-world progression, and time to treatment discontinuation complexities turn to require multistep algorithms (Graili *et al*., 2023). Overall, this leads to non-reproducible evidence unless properly governed and validated. To continuously guarantee the use of RWD for decision making, such as label expansions, reimbursement negotiations, and updates to clinical practice guidelines, there is a need to address these challenges. Hence, the continuous calls for action by various organizations such as the USFDA, EMA, ISPOR, and leading oncology data networks through the publication of frameworks that advocate for stronger data governance and quality assessment. There is great emphasis on appropriate data review, while different quality metrics should be considered, including accuracy and timeliness, with proper documentation and explanation of ambiguities. The proposed models should improve the current governance systems, which are still uneven and poorly documented, especially in multi-institutional data contexts.

Publisher: SARC Publisher

QUALITY FRAMEWORKS FOR REAL-WORLD EVIDENCE IN ONCOLOGY CLINICAL TRIALS

A critical appraisal, however, continues to show entrenched limitations: first, oncology RWD has misclassified tumor descriptors, staging information and performance status, which undermines internal validity and complicates causal attribution even with advanced analytic adjustments (Agarwala *et al*., 2018; Klimek *et al*., 2022). Secondly, heterogeneous extraction involving proprietary machine-learning algorithms makes it difficult to reproduce and independently validate findings across datasets (Bernardo *et al*., 2024; Sanchini *et al*., 2025); and third, regulatory acceptance of RWE is often conditional on prespecified estimates and demonstrations that the data are fit for the stated decision context (Bright *et al*., 2023). Evolving guidance has achieved some progress, but the uptake of standard metrics and cross-database validation continues to be a challenge (Batra & Cheung, 2019). To move the field forward, stakeholders will have to adopt harmonised quality indicators, transparent pipelines, routine preregistration, and independent external validation. Only integrated, auditable frameworks that couple data engineering with epidemiologic rigor and regulatory alignment will enable RWE to augment oncology decision-making reliably (Chan *et al*., 2020).

Using data from health care databases, such as insurance claims and registry databases, is widely accepted and has been used for decades for safety evaluation, risk management, and benefit-risk assessment of medicinal products. These RWD sources offer advantages like longitudinal data, a huge population size, and easy access to data, enabling speedy study completion. However, there are disadvantages, which include limited knowledge on critical confounding variables and populations often present in cancer research.

The scope and regulatory potential of RWE derived from electronic health records, registries, and administrative claims have been articulated by (Sherman *et al*., 2016). Accordingly, robust quality frameworks for oncology RWE must specify data provenance, standardized definitions, completeness metrics, traceability, and transparent analytic procedures in establishing fitness-for-purpose (Sherman *et al*., 2016). Procedural guidance from expert emphasizes preregistration, reproducible protocols, and stakeholder engagement to boost credibility (Berger *et al*., 2017), whereas practical checklists for regulatory-grade data quality illustrate operational steps for adjudication, and audit trails (Miksad & Abernethy, 2018).

Propose Frameworks: Integrating Predictive Real-World Data Models for Regulatory Decision Support

The gap between clinical trials and RWD can be bridged by hybrid and alternative study designs that generate RWE, considering the limitations of both RCTs and observational RWD. For example, pragmatic trials offer unique opportunities to combine the potential cost savings and benefits of real-world research with the scientific rigor of RCTS. Three key characteristics must be present in pragmatic clinical trials: an aim to inform decision makers (such as clinicians, administrators, and policy makers); an aim to enroll a population relevant to the decision in practice or representative of the patients populations; and an aim to streamline procedures and data collection so that sufficient efforts can be allocated towards informing clinical and policy decisions.

The effective integration of predictive models of RWD into oncology clinical trial decision-making may provide a pragmatic pathway toward improving trial design, patient selection, and post approval safety monitoring. Predictive models such as propensity-score approaches, causal inference frameworks, and machine learning survival, can enable the approximation of counterfactual outcomes using observational data. This can be conceptualized using the concept "target trial emulation" as a structured approach for transforming RWD into causal estimates in keeping with the principles underlying randomised controlled trials (Geissler *et al*., 2023; Ninomiya & Yoshimoto, 2008). When rigorously applied, this framework diminishes bias due to treatment of heterogeneity and nonrandom initiation of treatments, challenges common in oncology datasets.

Recent evaluations demonstrate that advanced RWD models can improve endpoint estimation and external control arm construction, particularly in rare cancers where randomization is difficult. For example, (Carrigan *et al*., 2020) showed that regulatory-grade external comparator arms derived from structured datasets can approximate RCT outcomes when quality standards are enforced. However, these must address the critical challenges, including unstructured clinical variables, which often limit model accuracy,

Publisher: SARC Publisher

requiring transparent imputation strategies and rigorous sensitivity analyses. It would also address machine learning derived and algorithmic bias, particularly in racial and socioeconomically diverse cancer populations. And finally, the USFDA's current guidance stresses the need for prespecified analytic plans, auditable data provenance, and reproducibility before such models can inform regulatory decisions (Jahanshahi *et al*., 2021; Wu *et al*., 2020; Yap *et al*., 2022).

If they are to have maximum impact, predictive models using RWD must be developed within harmonised ontologies, validated across multiple real-world datasets, and evaluated against predefined specifications. Only then will they reliably improve trial efficiency and support evidence generation in the development of U.S. oncology drugs.

Policy and Implications

The application of audit-ready data governance and quality frameworks directly impacts oncology research, clinical practice, and regulatory decision-making. Researchers and sponsors will find that the adoption of harmonized data standards, transparent curation pipelines, and reproducible analytic workflows offers a base for improving the reliability of RWE and reducing the challenges in conducting multi-institutional studies. Well-governed RWE will allow clinicians and health systems to inform treatment decisions based on evidence, optimize care pathways, and continuously monitor real-world safety and effectiveness, especially in patient populations that are typically underrepresented in randomized controlled trials. Additionally, regulators and stakeholders can apply robustly curated RWE to inform regulatory approvals, label expansions, and coverage decisions to ensure the maintenance of public trust in oncology evidence generation. Embedding governance principles throughout the data life cycle facilitates the creation of evidence that is not only scientifically rigorous but also auditable, transparent, and ethically sound, enabling data-driven decisions that are equitable and reproducible across diverse oncological settings.

CONCLUSION

For RWE generated in cancer clinical trials to be reliable and accepted by regulatory authorities, data governance and quality frameworks are essential. This should include mandatory provenance capture, adoption of standard data models, and standardized data-quality assessment with predefined thresholds for audits. The need for strict procedures that ensure data provenance and traceability has grown due to increasing reliance on multiple data sources, including genomic platforms. While standardized data models, metadata management, and regulatory guidelines have progressed, combining datasets from different sources. However, verifying complex endpoints, and maintaining transparency in multi-institutional settings remain challenging. This review highlights that audit-ready RWE requires more than just technical infrastructure; it demands an integrated approach combining governance policies, documented standard operating procedures, harmonized quality metrics, and reproducible analytic workflows. Moreover, oncology-specific issues such as unstructured clinical variables, evolving biomarker definitions, and inconsistent reporting of disease progression and treatment lines necessitate customized quality frameworks. Predictive real-world data models and target trial emulation methods can improve decision-making and ensure methodological rigor. Ultimately, data custodians share the responsibility for developing audit-ready RWE in oncology.

REFERENCES

1. Food and Drug Administration. "Electronic records; electronic signatures." *Federal register* 62.54 (1997): 13430-13466.
2. Agarwala, V., Khozin, S., Singal, G., O'Connell, C., Kuk, D., Li, G., & Abernethy, A. P. "Real-world evidence in support of precision medicine: clinico-genomic cancer data as a case study." *Health affairs* 37.5 (2018): 765-772.
3. Basch, E., Geoghegan, C., Coons, S. J., Gnanasakthy, A., Slagle, A. F., Papadopoulos, E. J., & Kluetz, P. G. "Patient-reported outcomes in cancer drug development and US regulatory review: perspectives from industry, the Food and Drug Administration, and the patient." *JAMA oncology* 1.3 (2015): 375-379.
4. Batra, A., & Cheung, W. Y. "Role of real-world evidence in informing cancer care: lessons from colorectal cancer." *Current Oncology* 26.Suppl 1 (2019): S53.
5. Berger, M. L., Sox, H., Willke, R. J., Brixner, D. L., Eichler, H. G., Goettsch, W., & Mullins, C. D. "Good practices for real-world data studies of treatment and/or comparative effectiveness: recommendations from the joint ISPOR-ISPE Special Task Force on real-world evidence in health care decision

Publisher: SARC Publisher

making." *Value in Health* 20.8 (2017): 1003-1008.
6. Bernardo, B. M. V., São Mamede, H., Barroso, J. M. P., & dos Santos, V. M. P. D. "Data governance & quality management—Innovation and breakthroughs across different fields." *Journal of Innovation & Knowledge* 9.4 (2024): 100598.
7. Blonde, L., Khunti, K., Harris, S. B., Meizinger, C., & Skolnik, N. S. "Interpretation and impact of real-world clinical data for the practicing clinician." *Advances in therapy* 35.11 (2018): 1763-1774.
8. Bright, K., Mills, A., Bradford, J. P., & Stewart, D. J. "RAPID framework for improved access to precision oncology for lethal disease: Results from a modified multi-round delphi study." *Frontiers in Health Services* 3 (2023): 1015621.
9. Buse, J. B., Austin, C. P., Johnston, S. C., Lewis-Hall, F., March, A. N., Shore, C. K., & Rutter, J. L. "A framework for assessing clinical trial site readiness." *Journal of clinical and translational science* 7.1 (2023): e151.
10. Carrigan, G., Whipple, S., Capra, W. B., Taylor, M. D., Brown, J. S., Lu, M., & Rothman, K. J. "Using electronic health records to derive control arms for early phase single-arm lung cancer trials: proof-of-concept in randomized controlled trials." *Clinical Pharmacology & Therapeutics* 107.2 (2020): 369-377.
11. Chan, K., Nam, S., Evans, B., de Oliveira, C., Chambers, A., Gavura, S., & Isaranuwatchai, W. "Developing a framework to incorporate real-world evidence in cancer drug funding decisions: the Canadian Real-world Evidence for Value of Cancer Drugs (CanREValue) collaboration." *BMJ open* 10.1 (2020): e032884.
12. Cho, Y., Shang, S., & Zhou, W. "Comorbidities were associated with cancer clinical trial discussion and participation: findings from the Health Information National Trends Survey–Surveillance, Epidemiology, and End Results Program (2021)." *Journal of Clinical Epidemiology* 163 (2023): 62-69.
13. Chon, Z., & Alexander, D. "Data Governance Frameworks for Ensuring Data Integrity in Clinical Trials Informatics." (2023).
14. Corrigan-Curay, J., Sacks, L., & Woodcock, J. "Real-world evidence and real-world data for evaluating drug safety and effectiveness." *Jama* 320.9 (2018): 867-868.
15. Eskola, S. M., Leufkens, H. G., Bate, A., De Bruin, M. L., & Gardarsdottir, H. "The role of real-world data and evidence in oncology medicines approved in EU in 2018–2019." *Journal of Cancer Policy* 36 (2023): 100424.
16. Feinberg, B. A., Gajra, A., Zettler, M. E., Phillips, T. D., Phillips Jr, E. G., & Kish, J. K. "Use of real-world evidence to support FDA approval of oncology drugs." *Value in Health* 23.10 (2020): 1358-1365.
17. Garnett, C. E., Lee, J. Y., & Gobburu, J. V. "Contribution of modeling and simulation in the regulatory review and decision-making: US FDA perspective." *Clinical Trial Simulations: Applications and Trends*. New York, NY: Springer New York, 2010. 37-57.
18. Geissler, J., Makaroff, L. E., Söhlke, B., & Bokemeyer, C. "Precision oncology medicines and the need for real world evidence acceptance in health technology assessment: Importance of patient involvement in sustainable healthcare." *European Journal of Cancer* 193 (2023): 113323.
19. Graili, P., Guertin, J. R., Chan, K. K., & Tadrous, M. "Integration of real-world evidence from different data sources in health technology assessment." *Journal of Pharmacy & Pharmaceutical Sciences* 26 (2023): 11460.
20. Jahanshahi, M., Gregg, K., Davis, G., Ndu, A., Miller, V., Vockley, J., ... & Sakai, S. "The use of external controls in FDA regulatory decision making." *Therapeutic Innovation & Regulatory Science* 55.5 (2021): 1019-1035.
21. Jiang, C., Deng, L., Karr, M. A., Wen, Y., Wang, Q., Perimbeti, S., & Han, X. "Chronic comorbid conditions among adult cancer survivors in the United States: results from the National Health Interview Survey, 2002-2018." *Cancer* 128.4 (2022): 828-838.
22. Klimek, P., Baltic, D., Brunner, M., Degelsegger-Marquez, A., Garhöfer, G., Gouya-Lechner, G., & Pleiner-Duxneuner, J. "Quality criteria for real-world data in pharmaceutical research and health care decision-making: Austrian expert consensus." *JMIR Medical Informatics* 10.6 (2022): e34204.
23. Liu, F., & Panagiotakos, D. "Real-world data: a brief review of the methods, applications, challenges and opportunities." *BMC Medical Research Methodology* 22.1 (2022): 287.
24. Miksad, R. A., & Abernethy, A. P. "Harnessing the power of real-world evidence (RWE): a checklist to ensure

Publisher: SARC Publisher

regulatory-grade data quality." *Clinical Pharmacology & Therapeutics* 103.2 (2018): 202-205.
25. Ninomiya, Y., & Yoshimoto, A. "Statistical method for detecting structural change in the growth process." *Biometrics* 64.1 (2008): 46-53.
26. Sanchini, V., Marelli, L., Monturano, M., Bonizzi, G., Peruzzotti, G., Orecchia, R., & Pravettoni, G. "A comprehensive ethics and data governance framework for data-intensive health research: Lessons from an Italian cancer research institute." *Accountability in Research* 32.1 (2025): 59-76.
27. Senderowicz, A. M. "Information needed to conduct first-in-human oncology trials in the United States: a view from a former FDA medical reviewer." *Clinical cancer research* 16.6 (2010): 1719-1725.
28. Sherman, R. E., Anderson, S. A., Dal Pan, G. J., Gray, G. W., Gross, T., Hunter, N. L., & Califf, R. M. "Real-world evidence—what is it and what can it tell us." *N Engl J Med* 375.23 (2016): 2293-2297.
29. Siden, R., Tankanow, R. M., & Tamer, H. R. "Understanding and preparing for clinical drug trial audits." *American journal of health-system pharmacy* 59.23 (2002): 2301-2301.
30. Wang, S. V., Pottegård, A., Crown, W., Arlett, P., Ashcroft, D. M., Benchimol, E. I., & Williams, R. J. "HARmonized Protocol Template to Enhance Reproducibility of hypothesis evaluating real-world evidence studies on treatment effects: a good practices report of a joint ISPE/ISPOR task force." *Value in Health* 25.10 (2022): 1663-1672.
31. Weiss, R. B., & Tuttle, S. S. "Preparing for clinical trial data audits." *Journal of Oncology Practice* 2.4 (2006): 157.
32. Weiss, R. B., Vogelzang, N. J., Peterson, B. A., Panasci, L. C., Carpenter, J. T., Gavigan, M., ... & McIntyre, O. R. "A successful system of scientific data audits for clinical trials: a report from the Cancer and Leukemia Group B." *JAMA* 270.4 (1993): 459-464.
33. Wilson, B. E., & Booth, C. M. "Real-world data: bridging the gap between clinical trials and practice." *EClinicalMedicine* 78 (2024).
34. Wu, J., Wang, C., Toh, S., Pisa, F. E., & Bauer, L. "Use of real-world evidence in regulatory decisions for rare diseases in the United States—current status and future directions." *Pharmacoepidemiology and drug safety* 29.10 (2020): 1213-1218.
35. Yap, T. A., Jacobs, I., Baumfeld Andre, E., Lee, L. J., Beaupre, D., & Azoulay, L. "Application of real-world data to external control groups in oncology clinical trial drug development." *Frontiers in Oncology* 11 (2022): 695936.
36. Yuan, H., Ali, M. S., Brouwer, E. S., Girman, C. J., Guo, J. J., Lund, J. L., & ISPE Comparative Effectiveness Research Special Interest Group. "Real-world evidence: what it is and what it can tell us according to the International Society for Pharmacoepidemiology (ISPE) comparative effectiveness research (CER) special interest group (SIG)." *Clinical Pharmacology & Therapeutics* 104.2 (2018): 239-241.
37. Yuan, L., Rahman, M., & Concato, J. "Comparison of two assessments of real-world data and real-world evidence for regulatory decision-making." *Clinical and Translational Science* 17.1 (2024): e13702.

Source of support: Nil; **Conflict of interest:** Nil.

Cite this article as:
Nkogbu, G. & Darko, I. N. K. “Audit-Ready Data Governance and Quality Frameworks for Real-World Evidence In Oncology Clinical Trials." *Sarcouncil Journal of Biomedical Sciences* 5.1 (2026): pp 10-15.

Publisher: SARC Publisher