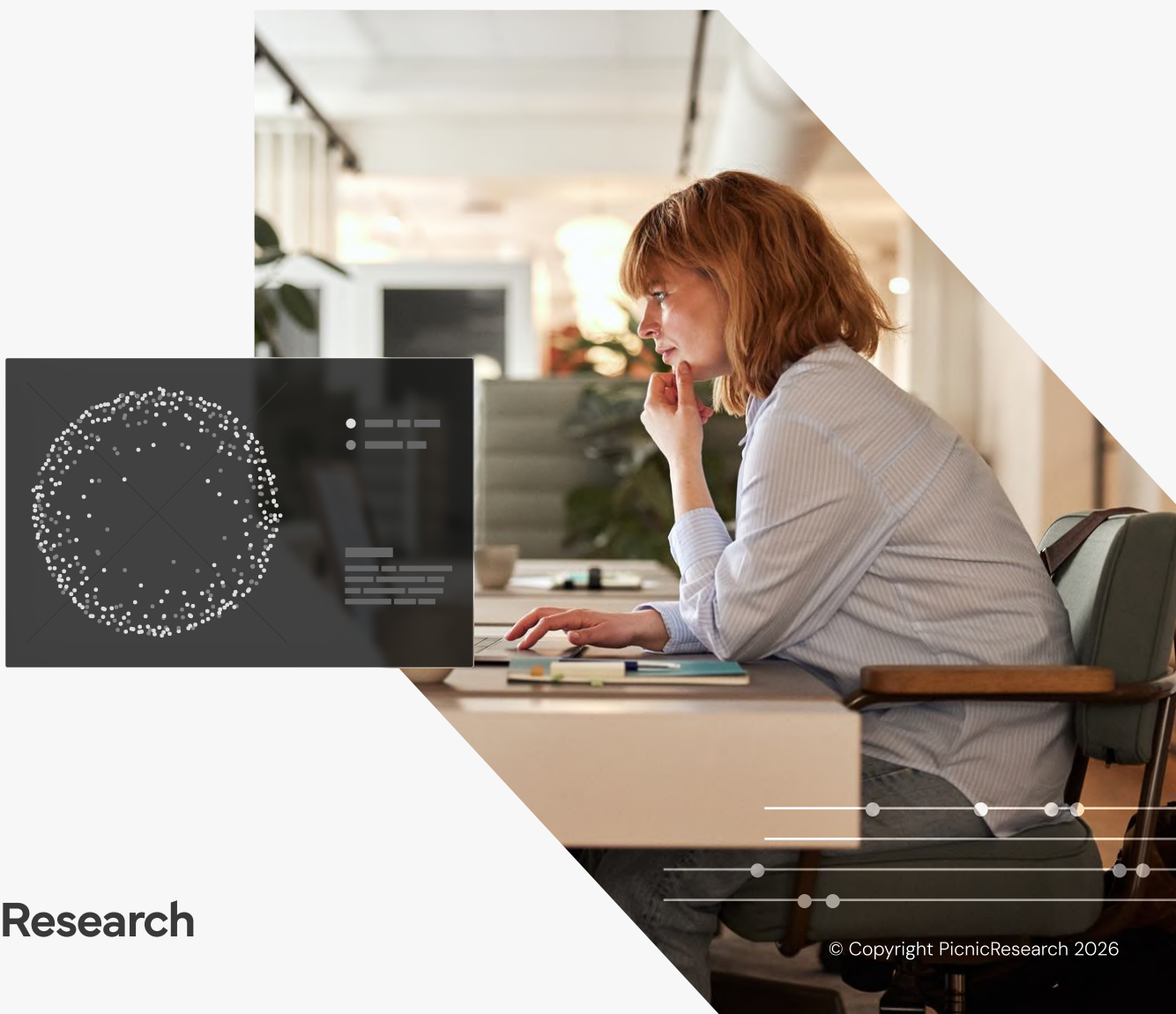


Completing the Patient Journey: Validating Multi-Source Visit Capture Against Patient- Reported Visits

AUTHORS

Shruti Menon, PhD, Epidemiology Solutions Principal, PicnicResearch

Colleen Goldberg, BSN, RN, OCN, Clinical Product Manager, PicnicResearch



Abstract

This whitepaper evaluates the visit-level completeness achieved through patient-mediated, multi-source medical record retrieval. This was based on a sample of patients over a 24 month-period (January 2022–December 2024) from two PicnicResearch registries, early breast cancer (eBC; n=10) and IgA nephropathy (IgAN; n=10). Complete medical records were retrieved under HIPAA authorization, with a study-specific focus on specialist visits. Patient-reported visits, collected through structured interviews, served as the reference standard for assessing visit completeness.

The specialist-focused design achieved high completeness for specialist visits (87% in eBC, 88% in IgAN). As the study prioritized specialist visits, retrieval of primary and acute care records was exploratory; these capture rates were correspondingly lower, consistent with the study design (primary care: 66.7–70.8%; acute care: 80–88%). Overall, patient validation indicated that approximately 80% of all patient-reported visits were successfully represented in the retrieved medical record timelines.

These results show that PicnicResearch's patient-mediated, multi-source retrieval workflow produces longitudinal, traceable datasets with high completeness for targeted visits and strong alignment with patient experience. Although patient validation is not part of the standard workflow, this one-time evaluation confirms the robustness of the methodology and its suitability for creating fit-for-purpose, regulatory-quality real-world data (RWD).

Background

Real-world evidence (RWE) now plays a central role in FDA decision-making on drug and biologic safety and effectiveness. Since the passage of the 21st Century Cures Act,¹ the FDA's RWE Program has evaluated how RWD can support new indications, label expansions, and post-approval commitments. In addition, FDA's 2023 guidance highlights that any RWD submitted under 21 CFR §505F must be "fit-for-purpose", which means it should be demonstrably complete, reliable, and relevant to the question at hand. This quality mandate dovetails with ONC's HTI-1 interoperability rule,² which tightens federal requirements for standardized, high-fidelity data exchange across certified electronic health record (EHR) systems. Together, these policies create a strong imperative for sponsors to validate the completeness of their RWD assets before relying on them for regulatory submissions.

Despite their widespread use in RWE studies, traditional EHR data sources often suffer from incomplete capture of patient visits and services. The TRUST study found that EHR data frequently omitted clinically relevant visits, lab results, or procedures, especially in fragmented care environments, where patients receive care across multiple unaffiliated health systems.³ These omissions raise concerns about the traceability and completeness of RWD used in regulatory contexts,^{3,4,5} which can compromise study endpoints, bias outcomes, and reduce the reliability of RWD in clinical research and regulatory submissions.

PicnicResearch addresses this challenge through PicnicHealth's patient-mediated, multi-source record retrieval model in which patients authorize direct access to their complete medical records across all U.S. care settings.⁶ The retrieval workflow integrates Electronic Data Network (EDN) feeds with Direct Facility Retrieval (DFR) from individual providers, enabling capture of both structured and unstructured data (e.g., clinician notes, pathology and imaging reports). Retrieved records undergo a proprietary AI-powered, human-curated abstraction process to ensure data quality, consistency, and research readiness.^{7,8} This patient-mediated approach supports comprehensive, multi-source data aggregation, producing RWD that are more complete, accurate, and traceable than traditional EHR-based sources. While the platform supports direct patient validation, this capability is distinct from the standard retrieval workflow. In this study, it was applied selectively to verify visit capture completeness, confirming the methodology's robustness in generating regulatory-grade, fit-for-purpose RWD.

Objectives

The whitepaper aims to:

- Assess the visit-level completeness of medical records obtained by PicnicResearch through PicnicHealth's patient-mediated, multi-site retrieval model across two indications: early breast cancer (eBC) and IgA nephropathy (IgAN).
- Compare record capture rates of specialist visits using patient-reported visits as the reference standard.
 - Exploratory objective: Evaluate the capture of primary care and acute care visits within this specialist-focused retrieval design. Although PicnicResearch's standard workflow retrieves records across all providers and care settings, this study intentionally prioritized specialist visits, with primary and acute care records captured only when identified through specialist documentation or automated feeds.

Methodology

Study design and variables

We evaluated RWD completeness using a mixed-methods approach across 20 U.S. patients, 10 with eBC and 10 with IgAN over the previous 24-month period (Jan 2022–Dec 2024). This timeframe was selected to maximize the likelihood that patients could accurately recall their care journey.

PicnicResearch identified study-specific retrieval targets during study design. Using the targets, PicnicHealth retrieved medical records under the patient's right of access using its proprietary record retrieval infrastructure, which independently identifies and sources records from all known providers and facilities. Each document was abstracted and tagged using AI and trained medical data specialists to identify visit type, provider specialty, facility, and date. Subsequently, structured one-time interviews were conducted to collect patient-reported healthcare visits. Participants were encouraged to reference personal records (e.g., appointment reminders, portals like MyChart, calendars) during recall to maximize accuracy. The retrieved medical records were then compared against patient-reported visits to assess visit completeness.

Visits were categorized into three types: (1) specialists managing the primary condition (e.g., oncology specialists for eBC; nephrology and transplant medicine for IgAN), (2) primary care providers (PCPs), and (3) acute care (e.g., emergency department visit, urgent care, inpatient visit).

Analysis

Visit-level completeness was defined as the percentage of patient-reported healthcare visits that were matched in the retrieved medical records during the same period. Descriptive analyses include:

- **Sensitivity by Visit Type:** Percent of patient-reported visits matched in the retrieved dataset (e.g., specialist, PCP, acute care).
- **Total Visit Capture:** Mean number of visits captured per patient over the 24-month period.

Key Findings

Across 20 participants, PicnicResearch captured a high proportion of patient-reported healthcare visits, particularly those central to disease management (Figure 1). Both eBC and IgAN patients generally reported high satisfaction with the quality of retrieval (Table 1), with an average satisfaction rating of 80% in eBC and 79% in IgAN, indicating strong alignment between retrieved data and patients' perceived care experiences.

Figure 1: Visit completeness stratified by provider type and indications

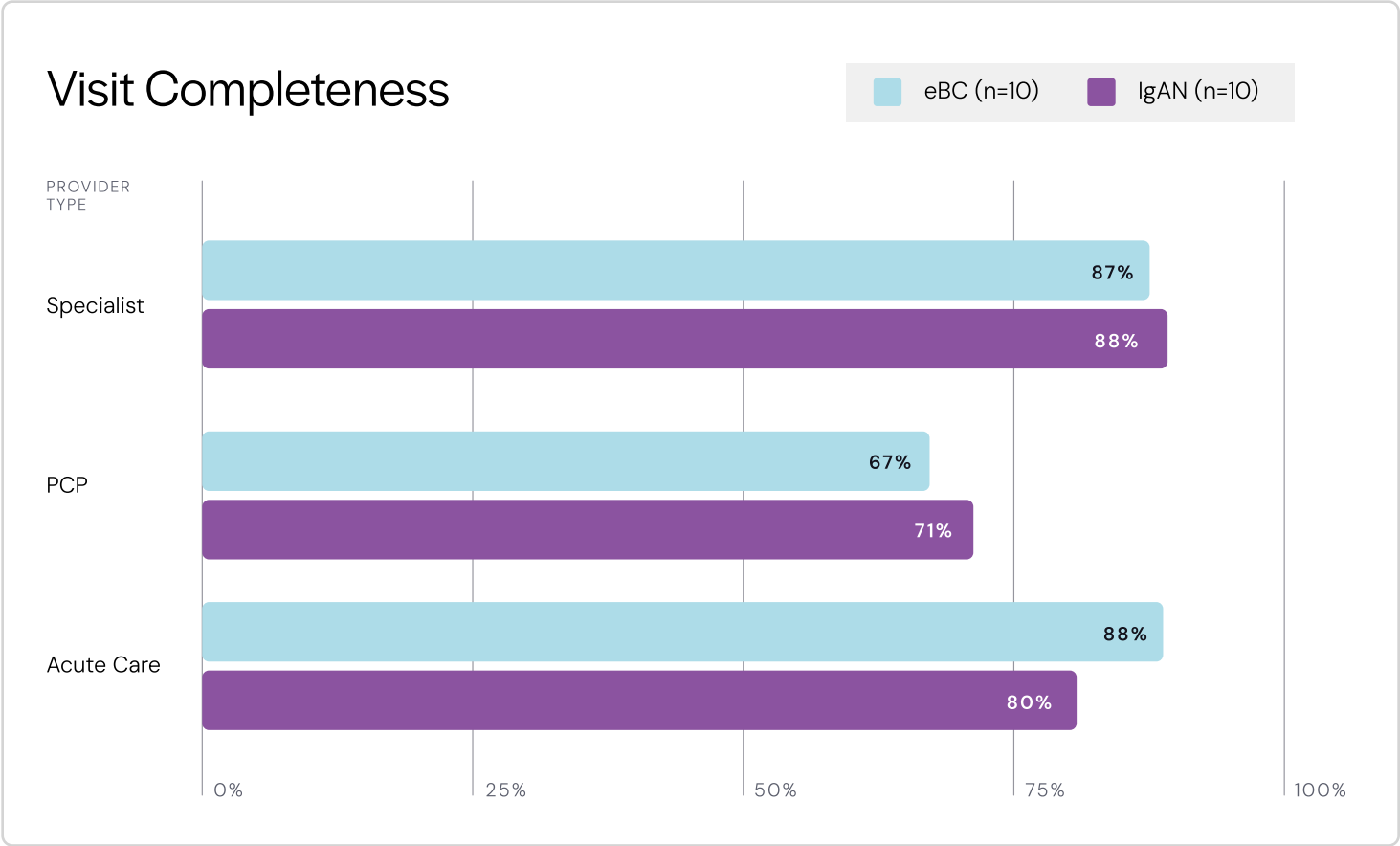


Table 1: Patient perspective on quality of data retrieval by indication

Registry	Mean retrieval quality score*
eBC (n = 10)	80%
IgAN (n = 10)	79%

*Patients were asked to rate on a scale from 1–10 how representative is your current visit timeline of your interaction with the providers you see regularly; a score of 1 indicated that none of the visits were captured, and a score of 10 indicated that all visits were captured. Scores were normalized to a percentage (e.g., a mean score of 8.0/10 is reported as 80%), and a score of 100% indicates the patient believes all visits were captured.

High completeness for specialists

The analyses demonstrated that PicnicResearch achieved high completeness (≥87%) for specialist visits, consistent across both eBC and IgAN registries.

In the eBC registry, records across a diverse range of specialties were retrieved, including oncology, surgical oncology, radiation oncology, and plastic surgery. These records were sourced through EDN and DFR, validating the robustness of both retrieval pathways. Completeness for specialist visits in this group ranged from 85% to 100%, with an average of 14.3 visits per patient (maximum: 24).

For IgAN patients, key specialties included nephrology and transplant surgery. Visit capture completeness in this registry ranged from 60% to 100%, with an average of 8.3 visits per patient (maximum: 15).

Despite different care delivery models in oncology and nephrology, the consistently high specialist capture rates (≥87%) support the generalizability of the retrieval approach, particularly for disease areas characterized by complex, recurring specialist visits and multi-site care coordination, such as autoimmune or neurologic conditions.

Completeness for primary care visits and acute care

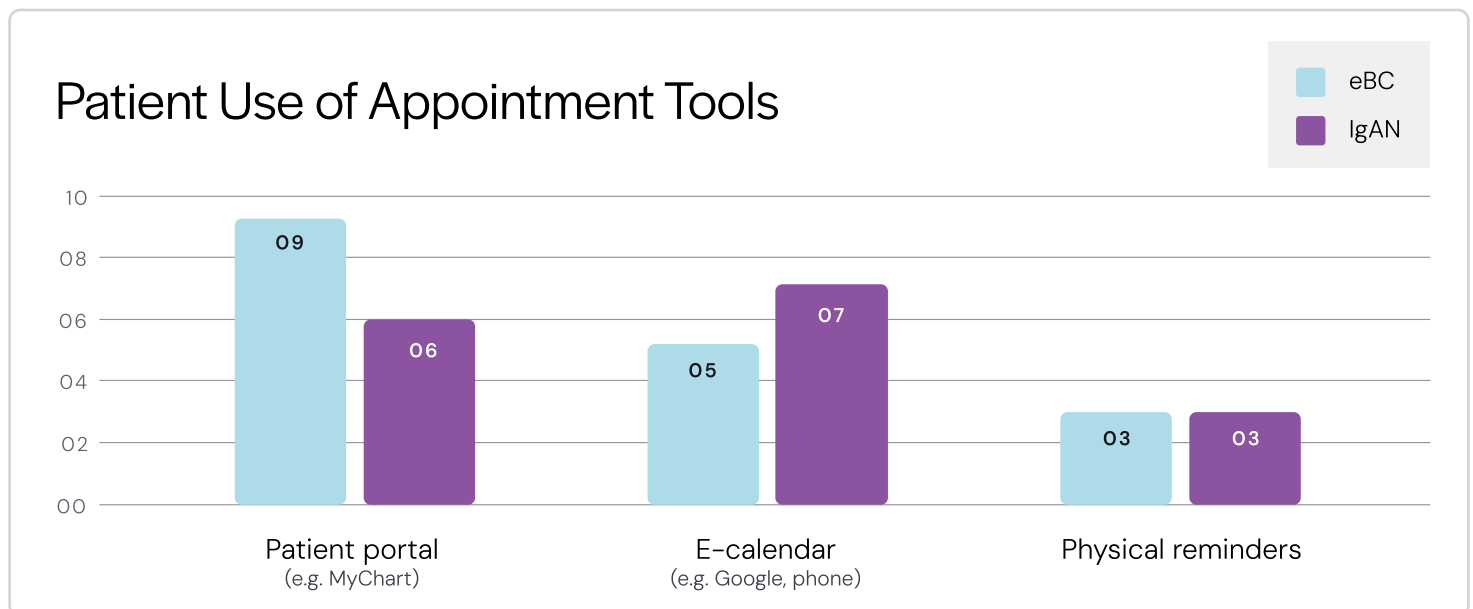
PCP visits and acute care visits showed lower capture rates relative to specialist visits. In the eBC registry, 66.7% of patient-reported PCP visits were retrieved. In IgAN, this rate was slightly higher at 70.8%.

For this validation exercise, retrieval efforts were intentionally focused on oncology and specialist visits. PCP and acute care visits were included only when they appeared in specialist notes or were available through automated EDN feeds. As expected, routine visits not referenced in specialist records were less likely to be captured, reflecting the scope of this specialist-focused validation rather than limitations of PicnicHealth's standard retrieval workflow.

Patient tools for tracking visits

Patients used a range of tools to recall and track healthcare visits. Both eBC and IgAN patients commonly utilize patient portals (e.g., MyChart), electronic calendars (e.g., Google, phone apps), and physical reminders (e.g., paper calendars, notebooks) to track their healthcare visits. Notably, eBC patients predominantly relied on patient portals, while IgAN patients frequently preferred electronic calendars (Figure 2).

Figure 2: Patient-reported methods for managing healthcare appointments across eBC and IgAN registries



Limitations

Our study acknowledges several potential limitations inherent in its design and data collection methodology. First, the small sample size ($N = 20$) limits the statistical precision of our findings and may not capture rare event patterns or less common visit types. Second, participant selection may introduce bias, as those willing to engage in structured interviews may differ from the broader patient population in terms of engagement, health literacy, or technology use. Accurate reporting depended on patients' meticulous tracking of their visits over the 24-month period. Despite these limitations, the structured methodology and direct cross-validation with full medical records help mitigate potential inaccuracies and provide a meaningful assessment of data completeness.

Implications for RWD Research and Regulatory Use

The findings of this validation study demonstrate that a patient-mediated, multi-source retrieval model reliably captures the specialist visits most critical to generating fit-for-purpose real-world datasets. High completeness for disease-relevant specialist visits ($\geq 87\%$), strong alignment with patient-reported visits, and positive patient assessments collectively support the robustness of PicnicResearch's approach for research and regulatory applications.

This patient-mediated medical record retrieval methodology offers a distinct advantage over traditional site-based retrieval methods. Existing literature, including a 2022 whitepaper published in collaboration with CERIS and the American Health Information Management Association,⁹ highlights the logistical and administrative barriers that frequently prevent complete record capture across fragmented provider networks. Similarly, automated pipelines often underperform in identifying complete patient histories due to gaps in visit capture and inconsistent coding practices.¹⁰ By securing direct authorization from patients and retrieving records at the source, many of these institutional bottlenecks are bypassed and the data "missingness" that commonly affects traditional RWD sources is reduced.

Unlike aggregators that depend primarily on claims feeds or automated EHR integrations, PicnicResearch leverages a direct, active relationship with the patient. This model not only enables consent for record access, but also the ability to perform active validation when necessary. Validating retrieved data with patient-reported visits mitigates metadata errors and helps ensure a more complete and representative reconstruction of each patient's care journey. The resulting dataset is longitudinal, multi-system, and demonstrably fit-for-purpose for regulatory-grade research, distinguishing it from the partial or fragmented datasets produced through indirect retrieval approaches.

Finally, the analysis also underscores the importance of integrating multiple data sources to mitigate missingness and enhance data quality. A 2018 PLOS ONE study, for instance, found that patient recall is influenced by factors such as the salience of the visit and the time elapsed since the visit.⁷ However, when systematically collected, patient recall can effectively supplement and validate administrative data. The patient-centered validation approach adopted in this study served as a proof-of-concept, demonstrating the reliability of the retrieval methodology. While not part of PicnicResearch's standard workflow, this validation provides additional assurance of completeness.

Optimizing Data Completeness

To further enhance the completeness, accuracy, and usability of the data retrieval process, the PicnicHealth app serves as a key platform for improving patient engagement. Automated, user-friendly post-visit notifications and personalized reminders, help patients more easily confirm visits and maintain up-to-date care timelines, thereby reducing missed visits and supporting more complete data capture across care settings. Leveraging these patient-driven interactions can significantly reduce missed visits, particularly routine PCP and out-of-network specialist visits. Additionally, future validation efforts will include expanding the patient-centered methodology to larger and more diverse patient populations beyond oncology and nephrology. Extending the approach to other therapeutic areas and broader registries will help demonstrate greater generalizability, reinforcing the robustness and regulatory value of PicnicResearch's solutions.

Learn more about how PicnicResearch can transform your evidence generation



Conclusion

This analysis demonstrates the effectiveness of PicnicResearch's patient-mediated retrieval model, demonstrated quantitatively through robust capture rates for specialist visits ($\geq 87\%$) and qualitatively through patient validation and satisfaction ratings (80% eBC, 79% IgAN). While PCP and acute care visits exhibited lower completeness rates due to study's intentional specialist-focused design, the methodology proved highly successful in capturing data essential to the core study objectives. By integrating data from multiple sources and applying rigorous abstraction, PicnicResearch routinely generates longitudinal, multi-system datasets with documented completeness and provenance.

A key strength of PicnicResearch's approach is the trusted relationship built with patients. The ability to selectively engage patients for validation, though not part of the standard process, provides an additional layer of assurance when completeness must be explicitly demonstrated for research or regulatory needs. Together, these capabilities enhance the credibility and reliability of RWE, supporting more robust insights and greater confidence in data used to inform clinical and regulatory decision-making.

Key Takeaways

- 01 **Robust retrieval across specialists:** PicnicResearch's retrieval model effectively captures a high proportion of patient-reported encounters, particularly for specialists (e.g., 87% for eBC and 88% for IgAN cohorts).
- 02 **Unique and differentiated retrieval methodology:** PicnicResearch combines multi-source document retrieval with human-led abstraction. This offers a distinct advantage over relying on fragmented EHR data or claims feeds.
- 03 **Patient validation boosts reliability:** Direct patient validation and satisfaction ratings (e.g., 80% for eBC, 79% for IgAN) confirm the records' reliability and suitability for research. This approach also captures encounters that traditional EHR systems often miss, such as out-of-network referrals or urgent care visits.
- 04 **Regulatory-grade RWD:** The completeness, traceability, and patient linkage of the data meet the FDA's "fit-for-purpose" standards under §505F and align with ONC HTI-1 interoperability goals. This supports use in label expansions, post-marketing surveillance, and regulatory submissions.

References

1. U.S. Food and Drug Administration. "Considerations for the Use of Real-World Data and Real-World Evidence to Support Regulatory Decision-Making for Drug and Biological Products." August 30, 2023. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-use-real-world-data-and-real-world-evidence-support-regulatory-decision-making-drug>.
2. The Office of the National Coordinator for Health Information Technology (ONC). "Health Data, Technology, and Interoperability: Certification Program Updates, Algorithm Transparency, and Information Sharing (HTI-1) Final Rule." HealthIT.gov. Accessed January 9, 2026. <https://www.healthit.gov/topic/laws-regulation-and-policy/health-data-technology-and-interoperability-certification-program>.
3. Riskin, Dan J., Keri L. Monda, Joshua J. Gagne, Robert Reynolds, A. Reshad Garan, Nancy Dreyer, et al. "Implementing Accuracy, Completeness, and Traceability for Data Reliability." JAMA Network Open 8 (2025): e250128.
4. Savitz, Samuel T., Lucy A. Savitz, Neil S. Fleming, Nilay D. Shah, and Alan S. Go. "How Much Can We Trust Electronic Health Record Data?" Healthcare 8, no. 3 (2020): 100444.
5. Pavlik, Valory, Anthony E. Brown, Stephen Nash, and J. Thomas Gossey. "Association of Patient Recall, Satisfaction, and Adherence to Content of an Electronic Health Record (EHR)-Generated After Visit Summary: A Randomized Clinical Trial." Journal of the American Board of Family Medicine 27, no. 2 (2014): 209–18.
6. PicnicHealth. "Why PicnicHealth Data?" Accessed November 11, 2025. https://assets-global.website-files.com/6074d822c3c38b07ef63efe5/608abcf69bafeb1b2a49b16e_Why%20PicnicHealth%20Data.pdf.
7. Laws, Michael Barton, Yael Lee, Tali Taubin, William H. Rogers, and Ira B. Wilson. "Factors Associated with Patient Recall of Key Information in Ambulatory Specialty Care Visits: Results of an Innovative Methodology." PLOS ONE 13, no. 2 (2018): e0191940.
8. PicnicResearch. "Multi-Modal Data Retrieval for Enhanced Patient Journey Capture." PicnicResearch Blog. Accessed November 11, 2025. <https://picnicresearch.com/resources/multi-modal-data-retrieval-for-enhanced-patient-journey-capture-electronic-data-network-edn-direct-facility-retrieval-dfr>.

References

9. CERiS and American Health Information Management Association (AHIMA). "Medical Record Retrieval and Procurement: Challenges and Solutions." White paper. Accessed January 9, 2026. https://www.ceris.com/static/052b020b2f01406497f2bb45280ddedf/CERiS_Medical-Record-Retrieval-Whitepaper_AHIMA_Final.
10. Thai, Duy N., Vahan Ardulov, Jose Juan Mena, Sarthak Tiwari, Gleb Erofeev, Ramez Eskander, et al. "ACR: A Benchmark for Automatic Cohort Retrieval." arXiv, June 20, 2024. <http://arxiv.org/abs/2406.14780>.

An abstract graphic consisting of several horizontal lines of varying lengths, each ending in a small circular dot. The lines and dots are light gray and are positioned in the upper half of the image, creating a sense of motion or data points.

PicnicAI builds intelligence to accelerate human health. The company harnesses the power of AI to address the complexity of healthcare delivery and clinical research through two product lines: PicnicHealth and PicnicResearch. PicnicHealth empowers patients with personalized guidance, actionable insights, and clinical support throughout their healthcare journey. PicnicResearch accelerates the pace of clinical research by making studies faster, more efficient, and more reliable, helping bring new treatments to patients sooner. Together, these efforts reflect PicnicAI's mission to use AI to transform human health and expand access to life-changing therapies. Learn more at picnic.ai.