WHITEPAPER

The Emergence of Multi-modal Real-world data:

The power of Clinico-Genomic data versus Traditional RWD Sources

PRESENTED BY:

NEO GENOMICS

ConcertAI

PUBLISHED BY:

FIERCE Pharma

Drug discovery and clinical development are highly complex processes with high precision data requirements. Historically these areas were not able to be beneficiaries of real-word data (RWD), a field that has been typified by electronic medical record, medical claims, or a combination of the two. Given these limitations, it tended to be Medical Affairs, HOER, Epidemiology Safety, and Commercial that used RWD, given their generally 'fit for purpose' coverage of approved medicines that were actively administered to patients. As the industry accelerates its shift toward precision medicine—developing therapies that are highly targeted to rare and often complex conditions—linked clinical-molecular data is poised to play an ever-more influential and critical role in internal decision-making and interactions with regulators.

To date, full coverage biomarker data has been difficult to link to traditional RWD sources. This limited its value to scientific and clinical organizations, hoping that a more complete view of the patient's disease and response to other medicines, could unlock insights guiding selection and development of highly targeted therapies. The growth of clinical networks, and RWE use guidance that was recently finalized by the FDA, have all made multiple clinical data types more accessible. This, combined with data linking and deidentification certification technologies, have allowed population scale clinic-genomic datasets to be generated supporting new insights to guide translational, first-in-human, and late phase clinical trial design strategies.

The shortfalls of traditional data sets

Today's oncology therapeutic clinical development pipeline is increasingly biomarker driven to develop highly targeted and increasingly more effective therapies. Biomarkers are markers of a disease state, potentially indicative of a specific mutation driving a cancer, reflective of a likelihood of response to a specific medicine and/or reflective or predictive of resistance.

The unfortunate aspect of most electronic medical record (EMR) systems is that they are not designed for the collection and presentation of 'machine readable' biomarker data, with the exception of routine CBC, metabolic panels for example, that are most typical of primary care requirements. This is true of most traditional sources, such as electronic health records (EHRs) and insurance claims, and emerging sources such as social determinants of health (SDoH) data and patient-reported outcomes (PROs). An electronic medical record (EMR) is a single healthcare practice's digital patient chart which replaces traditional paper-based medical records supporting the patient's history, diagnoses and treatments. On the other hand, the electronic health record (EHR) is a longitudinal view of a digital patient chart, providing a more inclusive view of the patient's complete medical history including full labs, patient registration, scheduling, and billing functions and is designed to be shared between practices. Not only do these data sets lack specificity about individual genes, proteins, or other biomarkers; they also often come in unstructured and incomplete formats. At best, researchers can only infer that a biomarker may be present based on a specific medicine being prescribed and used over a full episode of care.

Why genomic and biomarker data matter for life science

Biomarkers were defined more than two decades ago as "indicator[s] of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention."[1] In a 2018 commentary, current FDA Commissioner Dr. Robert M. Califf suggested scientific advances had "led to an avalanche of potential biomarkers for states of disease and wellness, extending beyond pure research into medical product development, clinical practice, nutrition, and environmental policy development."[2]

Biomarkers show significant promise for precision medicine, described as the approach that accounts for individual variables in determining prevention and treatment strategies at a genetic level.[3] For example, a tumor's biomarker data could be used to determine personalized radiation doses—ensuring that the right patients receive radiation while sparing those who may experience negative side effects and no therapeutic gain.[4]

Biomarkers show significant promise for precision medicine, described as the approach that accounts for individual variables in determining prevention and treatment strategies at a genetic level.[3]

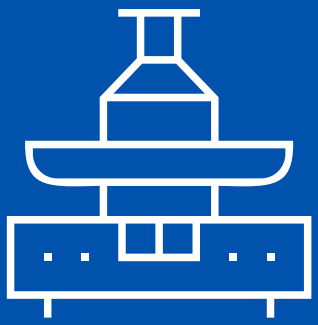
Unlock what's missing from traditional data sets

Biomarker rich RWD and clinic-genomic datasets provide utility to other data sources, such as electronic medical derived data, that would not be possible with those single sources alone. A biomarker test can determine whether an individual faces a higher-than-normal risk for certain types of cancer. If someone has been diagnosed with cancer, testing can determine what treatments the patient is most likely to respond to or become a form of surveillance of a patient's post-treatment for minimum residual disease (MRD) levels.[5]

Given that testing is done on validated instruments, using validated assays, and performed in laboratories with Clinical Laboratory Improvement Act (CLIA) certification, biomarkers are objective and precise. They are not subject to the observer bias that may be present in clinical notes,[6] such as the association of disease risk with personal behavior, or limited understanding of disease risk in ethnic populations. Of course, diagnostic testing and biomarker collection needs to be aligned to the evidence of utility and relative to the needs of certain patients given their race, ethnicity, countries of origin, and risk profile, among multiple considerations. This is critical to meeting the needs of underserved populations that are often diagnosed with cancer at later stages.

Five benefits of linking biomarker and real-world data

1. **Population-scale natural history studies for rare diseases**. Biomarkers can specifically pinpoint the driver for a specific disease or cancer, care teams can prescribe targeted treatments that may be more effective for a given indication than more general-purpose therapies. Without this information, we may observe poor responders and not know they were part of a rarer sub-population that was non-responsive to current standard of care medicines.
2. **Understand the patient's journey through cancer treatment**. Routine biomarker testing can help determine whether a patient was responding or non-responding to a specific treatment approach. Coupled with PRO data on symptoms and side effects, this also gives researchers insight into how treatment affects patients' overall well-being.
3. **Utilize causal inference-based approaches in research and analysis consistent with approaches the FDA increasingly favors**. Biomarker testing is one way to obtain real-world data and reach causal inference that can be extrapolated to the larger population, and the FDA is now accepting real-world evidence in drug approval applications.[7]



Given that testing is done on validated instruments, using validated assays, and performed in laboratories with Clinical Laboratory Improvement Act (CLIA) certification, biomarkers are objective and precise.

4. **Inform translational and early clinical development strategies**. Greater depth and breadth of data sources helps life science organizations evaluate their in vivo and in vitro research data with clearer implications for a novel therapeutics' potential value for narrow cohorts of patients. Doing this is both derisking of programs, and allows for a clearer first-in-human clinical trials design. This positions organizations to better manage and prioritize each stage of the drug development process. (See sidebar below.)

The right data for each stage of drug development

Longitudinal data can support each stage of the product development life cycle. The ability to move data-driven decision-making further upstream helps organizations set priorities, allocate resources, and avoid costly Phase II and Phase III trial failures due to factors such as limited safety, efficacy, or demonstrated commercial value. [9]

Discovery and Preclinical

- **Companion diagnosis launch planning** based on observations such as treatment uptake in patients with a companion diagnosis
- **Translational discovery** assessing new precision medicine biomarker targets
- **New product planning** made possible through forecasting by biomarker sub-cohort

Clinical Trial Phases I, II, and III

- **Trial design, site selection, and trial matching** using population-level data
- **External control arm studies** leveraging biomarker-driven regulatory studies to allow for comparative analysis
- **Epidemiology and HEOR studies** that include but are not limited to prevalence, disease burden, and outcomes

Commercial Launch and Post-Market

- **Value-driven market access studies** conducted with Health Technology Assessment bodies, payers, and other stakeholders
- **Post-launch surveillance** combining clinical data with PROs and other RWD
- **Studies of trends in ordering** to assess patient outcomes, prescription patterns, and drop-off rates/adherence challenges

5. **Create a population-scale data set that truly represents disease and treatment outcomes at a national level**. When clinical, claims, and public health data sets are paired with minimally biased biomarker data, researchers gain a more representative view of disease and treatment outcomes. This can underpin their epidemiology safety, medical, and Health Economics and Outcomes Research analyses.

The power of holistic data in oncology

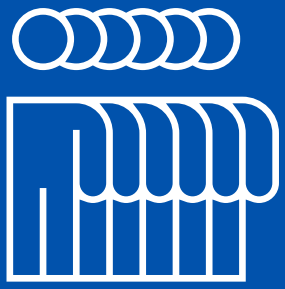
Earlier this year, NeoGenomics and ConcertAI announced a collaboration that combined ConcertAI's longitudinal clinical data and NeoGenomics' comprehensive biomarker data from hematological tests. This collaboration will support large-scale research into hematological malignancies. Cancers that affect the blood, bone marrow, and lymph nodes require far more varied treatment options than solid tumors and, as a result, more frequent clinical and diagnostic surveillance than other types of cancer.

The collaboration represents the first population-scale hematological data set available in the U.S., covering more than 1 million patients across over 1,000 oncology clinics. The combined data set provides coverage of key biomarkers throughout the patient journey, across multiple lines of therapy and can support each stage of therapeutic product development.

Further, the application of generative artificial intelligence tools to the data set will enable researchers to better define patterns associated with biomarkers, treatments, and outcomes. This is positioned to provide valuable context that is otherwise inaccessible when viewing siloed data sources in isolation, and it increases the likelihood that data analysis provides actionable insights.

In patients with multiple myeloma, it is common for a treatment to keep a patient stable for many years but suddenly stop working if a biologic affects the immune system in an unexpected way, for example. Looking solely at an individual's medical record, a care team may have little choice to continue to the next recommended line of therapy without knowing how a patient will respond.

On the other hand, the ability to analyze data at a population level makes it possible to predict which therapy is the next-best option, including those currently being studied in a clinical trials. Going a step further, generative AI models could assess a patient's outcomes during treatment and potentially predict whether a patient will stop responding. This would enable care teams to discontinue the treatment before it stops working and recommend a new treatment without disrupting continuity of care or causing a patient to experience side effects or diminished quality of life.



The collaboration represents the first population-scale hematological data set available in the U.S., covering more than 1 million patient lives across over 1,000 oncology clinics.

Tokenization's role in linking data and supporting product development

Tokenization is the process of replacing pieces of personally identifiable information with a token, or unique string of characters. The token applies to an individual patient – and, critically, each of the variables previously scattered among disparate clinical, claims, and public data sources.

Tokenization reduces false positives and other errors as it lets organizations avoid the effort of matching, integrating, and processing data sources, and it allows for the integration of genomic and sociographic data sources. Ultimately, tokenization provides a more complete picture of both individual patients and population segments impacted by a particular diagnosis.

This complete picture benefits multiple stages of the drug development lifecycle.

- Gathering PROs supplied by wearable devices through early-stage clinical trials could eliminate the need for frequent in-person site visits to enter PROs into clinical applications and allow for more decentralized trials.
- Compiling RWD and RWE and successfully linking it to patients already using a given therapy could significantly shorten the timeline for Phase IV and post-market studies.[8]
- Tokenization's ability to verify that all data points apply to the same patient – and only that patient – ensures clinical research meets the scrutiny of peer review and regulatory approval.

Conclusion

Holistic data sets that combine clinical, claims, sociographic, and biomarker data sources can provide unprecedented insight into which subsets of patient populations are poised to benefit from targeted therapies for their rare conditions. These data sets have traditionally been costly and time-consuming for life science organizations to assemble. Leveraging tokenization makes it possible to securely match data on individual patients across data sets—and unlock insights that drive innovation in drug development and serve patients with previously limited treatment options.

To learn more about the collaboration between NeoGenomics Laboratories and ConcertAI, please visit NeoGenomics.com/Informatics and https://www.concertai.com/content or contact: Informatics@NeoGenomics.com to discuss your next project.

Holistic data sets that combine clinical, claims, sociographic, and biomarker data sources can provide unprecedented insight into which subsets of patient populations are poised to benefit from targeted therapies for their rare conditions.

Endnotes

1 Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clinical Pharmacology & Therapeutics*. March 2001.

2 Biomarker definitions and their applications. *Experimental Biology and Medicine*. February 2018.

3 Biomarker Testing for Cancer Treatment. National Cancer Institute. December 2021.

4 Precision Medicine and the Role of Biomarkers of Radiotherapy Response in Breast Cancer. *Frontiers in Oncology*. April 2020.

5 Biomarker Tests and Cancer Treatment. American Cancer Society. September 2022.

6 Adding value to real-world data: the role of biomarkers. *Rheumatology*. January 2020.

7 FDA's Califf: Expect to see more RWE-based regulatory decisions. *Regulatory Focus: A Regulatory Affairs Professionals Society Publication*. September 2023.

8 The Case For Tokenizing Data On Clinical Trial Participants. *Clinical Research News*. March 2022.

9 Clinical Study Failures Vary Widely by Phase of Study and Therapeutic Class. Tufts Center for the Study of Drug Development. September 2013.

NeoGenomics, Inc. specializes in cancer genetics testing and information services, providing one of the most comprehensive oncology-focused testing menus in the world for physicians to help them diagnose and treat cancer. The Company's Advanced Diagnostics Division serves pharmaceutical clients in clinical trials and drug development.