Drug development: Edita Karosiene

Bringing multi-omics to real-world data

The biopharmaceutical industry is generating biological and healthcare data faster than ever before. Over the past decade, real-world data (RWD) derived from a wide range of sources, has become an important part of drug development strategy. Companies widely use real-world evidence (RWE), or the analysis of this data, to plan clinical trials, support regulatory label expansions, and monitor post-marketing safety.

But even with all this clinical data, drug development remains slow, expensive and inefficient. The cost of bringing a new therapy to market remains high and can reach about $2.2 billion. Moreover, more than 90% of drug candidates fail during clinical development. Many of these failures occur in later stages, particularly in Phases 2 and 3 due to lack of efficacy or unexpected safety issues[1].

This high failure rate is related to a key limitation of traditional RWD: it shows what happens, but not necessarily why it happens. Medical and treatment records document that a patient developed a symptom, responded poorly to a drug, or experienced an adverse event, but they cannot explain the underlying biological reasons. On the other hand, preclinical discovery groups generate high-dimensional biological data in cell lines and animal models that do not fully reflect human physiology. This results in a large gap between clinical observations and what is happening at the molecular level. Therefore, it's important to integrate multi-omics data directly with real-world clinical data.

In early discovery, single-layer biology (transcriptomics, genomics, proteomics and metabolomics) is often not enough to fully understand disease mechanisms. Human genetic evidence is highly valuable and can double the probability that a target progresses successfully from Phase 1 to regulatory approval[2]. However, DNA shows who is susceptible to a disease, but it does not always reflect how a disease actively progresses or changes over time.

Human diseases are complex biological networks. Understanding what drives clinical outcomes requires looking at the full biological picture. Genomics helps identify genetic variants linked to disease risk, while transcriptomics and epigenomics show how gene regulation is altered in the context of a disease. While modern therapies increasingly target DNA or RNA directly, proteins still represent the majority of drug targets. Proteomics reveals the proteins driving biological processes and potential drug targets, while providing information on protein abundance and post-translational modifications that cannot be determined from messenger RNA alone. Finally, metabolomics measures changes in metabolism, reflecting the effects of lifestyle, environment, and the gut microbiome, all of which may have a modifying effect on a disease state[1].

A crucial part of multi-omics integration is that the different biological layers are measured from the same individual sample, allowing them to be directly linked within the same biological context. To understand disease pathophysiology and verify target engagement, researchers need multi-omic measurements extracted from the same physical biospecimen, whether it is a single tumour biopsy or a blood draw. Only then can scientists follow biological changes from DNA to proteins and better distinguish changes that drive disease from those that are simply associated with it.

Linking biology to real-world outcomes

While same-sample multi-omics provide mechanistic insights during early drug discovery, it becomes even more powerful when linked to real-world clinical data. Historically, integrating molecular profiles with health records was legally and technically challenging. Nowadays, two breakthroughs make this integration feasible at scale[3]. First, privacy-preserving record linkage can securely connect patients' multi-omics data with longitudinal health records while protecting their identity. Second, standardised data models such as OMOP, together with AI tools, make it possible to harmonise and analyse clinical data across institutions.

We can already see multi-omics being integrated with real-world clinical records in global biobank initiatives. In the US, the National Institutes of Health's 'All of Us Research Program' in mid-2026 established the world's largest integrated genomic and health database with over 535,000 whole genomes linked to 482,000 electronic health records[4]. The programme is now expanding into multi-omics, using plasma proteomics and RNA sequencing to study biological differences across large and diverse populations.

The UK Biobank has also pioneered a multimodal initiative aimed at understanding differences in drug response between individuals[5]. By combining prescription and health records with genomics, proteomics and metabolomics, researchers can better understand why the same drug works well for some patients but not for others. This approach is already delivering real value in drug discovery, and the UK Biobank Pharma Proteomics Project is a prime example.

Bringing together 10 leading biopharma companies, researchers analysed nearly 3,000 plasma proteins from 54,000 people and linked the results to their electronic health records[6]. By combining genetic variation, protein levels and medical histories, researchers can assess whether a potential drug target is actually involved in a disease. They can also look for signs that targeting a particular protein could cause unwanted effects. This provides valuable human evidence early in drug development, helping companies choose safer, more promising targets and reduce the risk of costly failures later on[6].

The value for drug development

For the pharmaceutical industry, combining multi-omics with RWD can support better decision-making across the development pipeline. In early discovery, it can help de-risk

©2026 Evernow Publishing Ltd

target safety before investing in preclinical pipelines. During clinical development, linking genomic and transcriptomic profiling with real-world outcomes allows developers to identify non-responders early. By enriching clinical trials for patients most likely to benefit based on their biological profile, companies can reduce exposure to ineffective therapies, spare patients unnecessary toxicities, and avoid unnecessary clinical trial costs[7,8].

Furthermore, while DNA is static, circulating proteins and metabolites change in real time. Combining blood-based omics with long-term prescription data can help monitor immune responses, drug interactions, and treatment effects. It can also help optimise dosing across diverse patient groups, including children and older adults who are often underrepresented in clinical trials[9, 10].

Summary

Real-world evidence is already widely used in biopharma, but on its own it provides limited information about the biology behind clinical outcomes. Combining multi-omics with real-world clinical data can help close the gap between biological mechanisms and what actually happens to patients. For biopharma, this means better-informed decisions throughout drug development, fewer costly late-stage surprises, and ultimately safer and more personalised treatments.

References

1. Li X, et al. Multi-omics driven precision medicine. iMeta. 2026; 5:e280.

2. Davitte J, et al. Integration of Real-World Data and Genetics to Support Target Identification and Validation. *Clin Pharmacol Ther*. 2021; 110(3):616–626.

3. Dagenais S, et al. Real-World Evidence to Support Internal Biopharmaceutical Decision-Making. *Clin Pharmacol Ther*. 2022; 111(1):77–89.

4. National Institutes of Health. NIH's All of Us Research Program Expands Historic Dataset with Multi-Omics and Electronic Health Records. NIH News. June 30, 2026.

5. Li Z. Systematic investigation of inter-individual variability and mechanisms of drug response using multimodal data from the UK Biobank. UK Biobank Project 1092541. 2025.

6. Sun BB, et al. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature*. 2023; 622:329–338.

7. van de Haar J, et al. Systematic identification of genomic nonresponse biomarkers to cancer therapies. ESMO Real-World Data Digit Oncol. 2026; 100721.

8. Yorio J, et al. Association of Timely Comprehensive Genomic Profiling With Precision Oncology Treatment Use and Patient Outcomes in Advanced Non–Small-Cell Lung Cancer. JCO Precis Oncol. 2024; 8:e2300292.

9. Rich P, et al. Real-world performance of blood-based proteomic profiling in first-line immunotherapy treatment in advanced stage non-small cell lung cancer. *J Immunother Cancer*. 2021; 9:e002989.

10. Brunner F, et al. Leveraging large-scale biobank EHRs to enhance pharmacogenetics of cardiometabolic disease medications. Nat Commun. 2025; 16:2913.

This article was written by Edita Karosiene, PhD, chief scientific officer at Vugene UAB in Lithuania.

Subscribing couldn't be easier...

*...with our Double Guarantee of Satisfaction**



Call +44 (0)20 7704 1015

Visit www.mednous.com >>>>>>

 ©2026 Evernow Publishing Ltd