APPLICATION NOTE

SAPIENT

Advancing Population Health insights with the addition of dynamic biomarkers

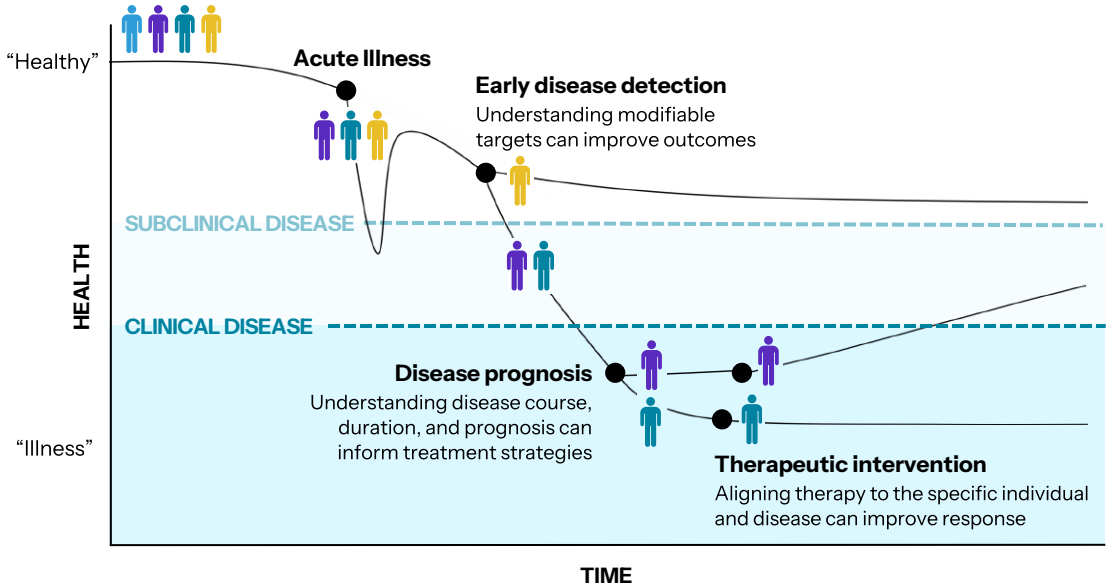
Introduction

Next-generation sequencing has transformed the rate of insight generation for population health studies, enabling rapid acquisition of genomics data across large cohorts and providing a robust foundation for understanding genetic predispositions of disease. Large-scale genomics studies have pinpointed specific genetic variations that influence both rare and common diseases, and have allowed for the development of polygenic risk scores that can be used to identify at-risk individuals that may benefit from early screening, preventative measures, or therapeutic interventions.

The challenge for population health studies is where to look next. Over a million genomes have now been sequenced worldwide, and the returns of sequencing more genomes are diminishing. Genomic information and polygenic risk scores provide a static snapshot of disease risk, but do not account for changes in health status over time. They cannot capture the full complexity of disease, which can be influenced by myriad factors such as lifestyle, environment, and other evolving genetic and non-genetic factors. In fact, across common diseases, only 15-20% is attributable to genetics.

The need for population health studies to go 'beyond the genome'

Given the dynamic nature of disease, there is a need to extend studies in population health cohorts to include the discovery of dynamic biomarkers such as metabolites, lipids, and proteins. As dynamic indicators of health, disease, and drug response, they provide insights into biological changes occurring over time. These biomarkers can allow for longitudinal tracking of disease progression to enable early disease detection and identify critical points for intervention.



The addition of dynamic biomarker discovery in population health studies can enhance and advance the impact of genomics data that has already been captured, adding a complementary layer of information to elucidate the interplay between genotype and phenotype across individuals.

The figure to the left illustrates the dynamic nature of disease, even in patients with the same diagnosis, and how dynamic biomarkers can be used to inform interventions.

Why dynamic biomarker discovery is ideally suited for population health studies

The insight value of metabolomics, lipidomics, and proteomics is clear, and with thousands of molecules to explore, there is tremendous discovery potential in these approaches. The metabolome in particular is comprised of a vast array of small molecules with varying chemical properties, and that can change in levels over time – even within the course of a day – in response to environmental exposures, physiological conditions, and genetic variations. While in small studies, it can be harder to gain the statistical power necessary to confirm the biological significance of these changes over a time course, population health cohorts are ideal for running analyses to profile dynamic biomarkers.

Their large sample sizes facilitate identification of small molecule and protein biomarker patterns that are reliably reproduced across a large number of individuals. Their biological significance can be statistically validated within the population, bringing confidence that the findings will also replicate across independent cohorts. Longitudinal samples from these cohorts can be assayed to correlate biomarker changes over time with specific subpopulations or health outcomes, providing guidance for tailored prevention and treatment strategies.

How Sapient enables population-level metabolomics and multi-omics analyses

One reason that genomics has been the leading omics approach for population health studies is that next-generation sequencing has unlocked the ability to broadly and deeply measure the genome at speed. There was not a like technology for metabolomics, which could accommodate the simultaneous capture of diverse chemistries in an efficient manner, until recent years. With the development of Sapient's rapid liquid chromatography-mass spectrometry (rLC-MS) systems, population health studies now have an approach to support dynamic biomarker discovery within their large number of samples.

Our rLC-MS technology captures >15,000 small molecule biomarkers per biosample – inclusive of metabolites and lipids – in a single run. With an analytical cycle time of less than one minute, we have the capacity to process more than 5,000 samples per day, providing the scale and efficiency required for population-level analysis. Our nontargeted approach supports discovery of the most biologically relevant biomarkers, including those that have yet to be characterized, to greatly amplify the potential of identifying significant signals across the population.

These mass spectrometry systems can also be used to assay thousands of proteins per sample, obtaining peptide-level information for high specificity of measurements. Together this additional data adds new layers of information to extend beyond a genetic-centric to a multi-omic understanding of disease. As part of Sapient's discovery pipeline, we have built an expert data science team that leverages diverse biocomputational approaches to integrate these complex, multi-dimensional multi-omics datasets. Through their analyses they can elucidate disease mechanisms, map identified biomarkers to disease and/or drug response, and derive insights into biomarker-genotype and biomarker-phenotype associations across populations.

>15,000
small molecule biomarkers captured in every biosample

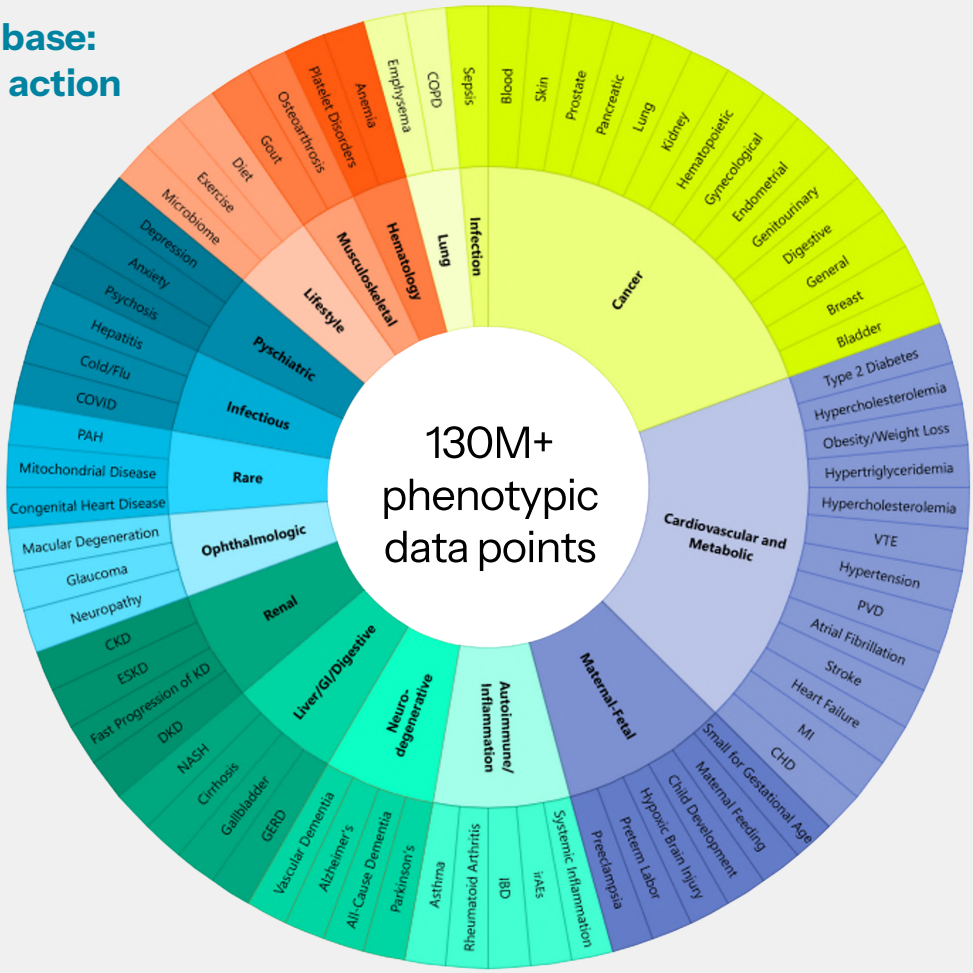
>5,000
biosamples can be processed per day

High throughput capacity for
>1.5M
biosamples per year

Sapient Human Biology Database: Population health analysis in action

Given the speed of Sapient's technologies, we have already generated rLC-MS data in more than 100,000 human biosamples collected from diverse individuals around the world. These measures are paired with phenotypic and other -omics data, including demographic features, lifestyle factors, genetics information, microbiome data, laboratory test values, and adjudicated clinical outcomes. The database includes patient follow-up data spanning 10-30 years across individuals, facilitating deep phenotyping and longitudinal analysis of factors influencing health and disease over time.

We leverage this multi-dimensional data to bring more biological context to biomarkers, from performing genome-wide association studies (GWAS) to integrate genomics and small molecule biomarkers, to evaluating biomarker stability and dynamics across geography, time of day, treatment, clinical outcomes, and other dimensions.



Sapient's proprietary database includes disease-centric data for more than 60 diseases and disorders, as well as data from samples from healthy individuals across populations worldwide.

Applications for dynamic biomarkers in population health studies

The inclusion of dynamic biomarkers in population health studies, derived from small molecule biomarker analyses and related multi-omics approaches, can enrich understanding of disease biology and enhance drug development efforts in several key ways, including to:

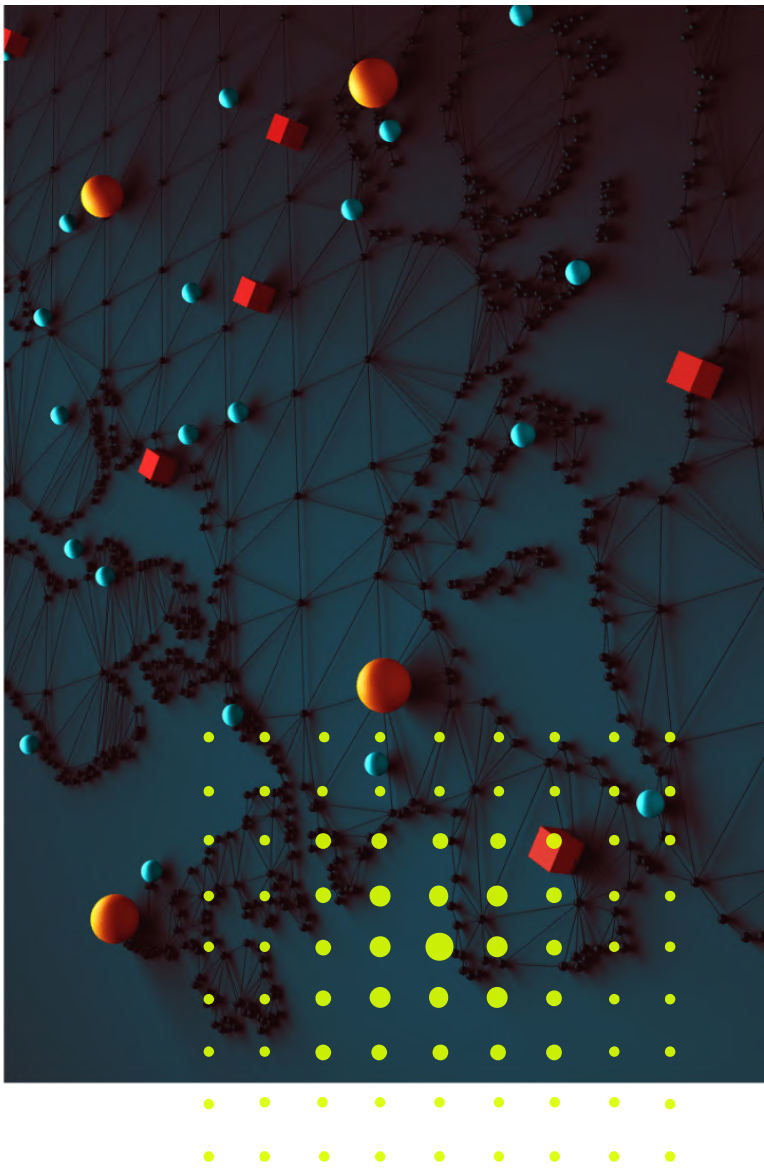
Elucidate Genotype-Phenotype Correlations

The integration of small molecule biomarkers with genomics allows for a more nuanced understanding of dynamic human biological processes, and how host-disease and host-environment interactions influence disease risk, disease progression, and drug responsiveness.

For example, GWAS can be used to map the relation of small molecules to specific genetic variants or protein targets, offering a deeper understanding of how genetic predispositions manifest at the molecular level.

Predict Disease Risk and Progression

Patterns of metabolites, lipids, and/or proteins associated with increased or decreased disease risk can be identified, allowing for the development of predictive models that help in early disease detection and prevention. Such patterns can be followed over time to understand disease progression and identify points of intervention to potentially slow or halt the disease.



Stratify Patient Populations

The integration of dynamic biomarkers in population health studies can further refine patient stratification based on distinct metabolic profiles, which is crucial to stratify heterogeneous populations and to understand variations in health outcomes. Targeted therapies can then be developed to specific metabolic or genetic subtypes.

Advance Precision Medicine

The combined analysis of dynamic biomarkers and genomics data can accelerate the realization of precision medicine at scale. Tailored interventions can be developed based on both the genetic and non-genetic factors that contribute to an individual's disease for more effective, personalized treatment strategies.

Conclusion

The addition of dynamic biomarkers in population health studies will enable pivotal advancements to deepen our understanding of the non-genetic factors underlying health and disease across heterogeneous human cohorts. Now there is finally an efficient, scalable approach to 'go beyond the genome' and enable large-scale discovery of metabolite, lipid, and protein biomarkers that read out these influences. Sapient's next-generation rLC-MS systems and biocomputational framework were built from the ground up for population-scale data generation and insight delivery, making our services ideal for population health cohorts.

Our Human Biology Database, comprised of data from >100,000 human biosamples run on our platform, proves the speed and scale of measurement we can achieve and the critical discovery insight that can be derived through integrative analysis of large-scale, multi-omics datasets. From population stratification to disease risk prediction, dynamic biomarkers can enrich existing genomics-based population health research and pave the way for more personalized and effective healthcare interventions.

SAPIENT

Ready to discover more? Schedule a time to discuss your programs.

Visit: sapient.bio | Email: discover@sapient.bio | Call: 858.290.7010