Health Data Trends Part I

Top Data Types in Demand

DATAVANT

Overview

Health Data Trends Part I
Evolving Health Data

Over the last few years, connecting health data is evolving from innovation to standard practice. Pharmaceutical companies, health systems, government, payers, and analytics platforms connect data, creating a robust picture of patient health, to accelerate research, improve care, and lower healthcare costs. Examples of the types of analyses that benefit from connected health data include questions like:

1 How can I find patients with fatty liver disease for a clinical trial?

2 Is this new migraine therapy cost-effective?

3 What factors are preventing patients from coming into the clinic for follow-up care?

4 What are the risk factors contributing to this patient's high healthcare costs?

Healthcare is the Biggest Producer of Data

New health policies, technologies, and scientific discoveries have resulted in a massive increase in the volume and variety of health data. The 2009 HITECH Act accelerated the use of electronic health records ("EHRs"). The decreased cost of genomic sequencing and increase in biomarker-specific drugs has led to growth in genetic testing.[1] Smartwatches, wearables, and health apps are now commonplace.[2] The pandemic accelerated the use of sensors and remote technologies. Cloud and data science enable storage and analysis of data at scale. RBC Capital Markets estimates that healthcare generates the largest volume of data. At a 36% compound annual growth rate by 2025, healthcare data is growing faster than any other industry.[3]

2012-2025 Data – Compound Annual Growth Rate (CAGR)

Healthcare	Manufacturing	Financial Service	Media and Entertainment	Global Dataspher
36%	30%	26%	25%	27%
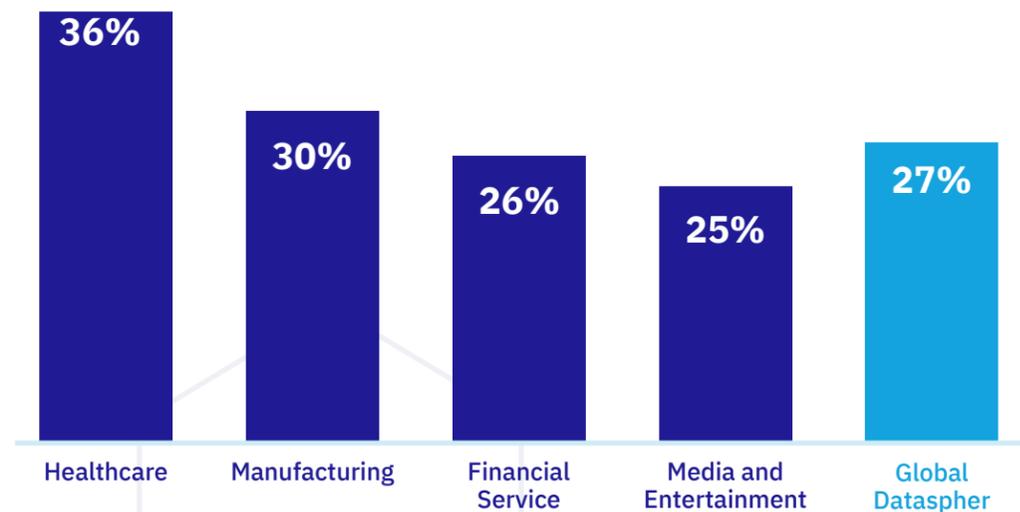
Looking to tomorrow's healthcare today: a participatory health perspective. Coughlin S, Roberts D, O'Neill K, Brooks P.Intern Med J. January 2018;48(1):92-96. doi:10.1111/imj.13661.

The Digitization of the World From Edge to Core. Reinsel, D, Grantz, Rydning, John. November 2018: IDC White Paper – #US44413318.

Source: https://www.rbccm.com/en/gib/healthcare/episode/the_healthcare_data_explosion

New Analytics Companies Are Formed Each Year

With more health data, new analytics companies have sprung up to analyze it. From patient recruitment and clinical trial technologies to population health and value-based care analytics, thousands of organizations serve biopharma, device makers, payers, patients and providers. As a result, Datavant now connects the **largest health data ecosystem** with thousands of data originators, data recipients and data platforms. We compiled top trends in the types of questions asked and the data used to answer them.

Demand Trend #1: Demographics and Social Determinants of Health (SDOH)

Many groups, such as people of color, women, and other vulnerable populations are neither adequately represented in research nor have equal access to quality care.[4] Government, life sciences, health systems and payers all want to understand the factors associated with worse health outcomes in these groups to try and close these gaps.

Demographic variables include inherent characteristics in people (e.g., age, race, gender, and marital or family status), whereas SDOH are the conditions under which people live, work, learn, etc. (i.e., education level and employment status). Datavant has numerous partners with **demographic and SDOH data**. The utility of specific SDOH attributes depends on the question a client is trying to answer.

Demand Trend #2: COVID Vaccination & Variant Data

Since the pandemic, data on COVID diagnoses, vaccination rates and variants have been in high demand and with an uptick due to the Omicron variant. Unfortunately, this data remains fragmented given the state-by-state approach to reporting COVID cases. Some of the largest data aggregator partners have vaccination status on just ~20% of all 200mm vaccinated individuals in the US, so linking data for a national view of U.S. patients is critical. Datavant lends its linking technology to support the NIH's National COVID Cohort Collaborative (N3C). Datavant connectivity technology also supports the national COVID-19 Research Database (CRDB) for academic research which was recently recognized by the Reagan Udall Foundation with an award for Innovations in Science. If you are interested in the application process to use the CRDB you can find more details here.

Demand Trend #3: Oncology Patient Journey

Linking data to understand the full oncology patient journey is the most frequent type of linkage we see. The oncology patient journey is complex. Patients may present at a primary care office and be referred to an oncologist who orders labs such as a biopsy, diagnostic imaging or genetic testing. They may visit radiologists, undergo surgery, and take drug regimens that include infused and oral drugs. While survival is a key outcome in oncology studies, mortality data is not always captured in EHR datasets.[6]

Understanding the oncology patient journey often requires linking claims data including medical, retail pharmacy and specialty pharmacy claims, to EHR data from ambulatory care, integrated delivery networks, and community oncology clinics. In particular, clients struggle to find oncology EHR data from academic medical centers, which are hospital systems that are affiliated with a research university. Academic medical centers (AMCs) are very important care settings that conduct research, provide education, and deliver clinical care. They have access to cutting-edge technologies and can serve patients with rare cancers by offering advanced procedures like bone marrow transplants and novel drug trials. Clients seek pathology, radiology, labs, imaging and genetic data which necessitates linking data from multiple partners specializing in cancer. In addition to **mortality data that covers more than 85% of weekly death events**, the Datavant ecosystem includes many oncology-specific data partners:

1   5 community-focused oncology RWD providers

2   3 AMC-focused oncology RWD providers

3   8 genomic data providers (3 of whom also have germline testing data)

4   9 providers of ambulatory EHR data

5   6 of the top 10 AMCs in cancer

Demand Trend #4: Specialty Drug Data

Specialty drugs dominate pharmaceutical company pipelines.[6] These drugs are distributed by specialty pharmacies (SPs). Many SPs only provide usage data on that drug back to the drug manufacturer, known as Limited Distribution Drugs ("LDDs"). In such cases, the drug and often the entire drug class is not available to aggregators and represents a gap in commercial databases. To compensate, life science companies connect their 1st party SP data feeds to 3rd party datasets for a more complete picture of patients' adherence to therapy and to understand their drug's utilization versus the competition.

Demand Trend #5: Rare Disease Patient Data

Rare diseases affect fewer than 200,000 people in the U.S.[7] Rare disease patients often endure a diagnostic odyssey that runs 5-7 years. They may see more than 7 specialists, before being properly diagnosed which creates a highly fragmented data journey.[8] Lastly, drugs for these patients are primarily specialty pharmacy-distributed, so data is often unavailable within commercial datasets. Clients seek our help linking data across many partners to find patients for clinical trials and observational studies.

Demand Trend #6: Linking First-Party Data

Recently, biopharma clients have begun linking internal 1st party data to external 3rd party data assets they license. First-party data includes clinical trial data, specialty pharmacy data, hub services data, patient registries, sponsored testing data, digital engagement data, and other assets. Linking first-party and third-party data creates a comprehensive view of patients' health. For example, connecting data from hub services to Rx claims can shed light on the time between a patient's attempt to get on therapy to script fulfillment and ongoing adherence. Connecting clinical trial data to EHR and claims datasets extends data collected for long-term outcomes analyses. There has been significant demand in linking these pharma-proprietary data assets to create more value and insights from enterprise data.

Demand Trend #7: Enterprise Data Linking

One of the most exciting trends we are seeing with biopharma clients is an enterprise-level data strategy that includes **tokenizing every trial and health economics and outcomes study** (tokenizing refers to de-identification and assignment of an ID we call a token). Centralized evidence generation teams are using the approach to maximize external partnerships as a data partner may have data that matches patients in several studies. Tokenizing all studies enables rapid partner identification and negotiation of data for multiple studies at one time.

Summary: Implications for Data Users, Data Originators and Platforms

This year we expect to see growth in use cases that require granular clinical, genetic, pathologic, imaging, and biomarker features. Those seeking data need to define clear criteria for the population of interest, the questions they're trying to answer, and the use case. Preparing these in advance will accelerate finding a partner with the best data to meet those needs.

For data originators, aggregators and analytics providers, understanding demand-side trends can help them maximize the relevance of their data to answer key questions. Making data available to run **on-demand overlap analysis** helps those seeking data identify shared patients faster. Once partners with overlapping populations have been identified, coming to the initial partner meeting with information about data content, quality and curation processes speeds partner diligence.

At Datavant, we believe that every decision in healthcare should be powered by complete patient data, and the more we can help partners find each other, the more we enable everyone's effort to improve patient outcomes.

References

The Cost of Sequencing a Human Genome. The National Human Genome Research Institute. Published November 1, 2021. Accessed January 5, 2022.

How Fitbit, Whoop and Other Gadgets Are Measuring Brain Activity, Glucose and Sleep. Betsy Morris, The Wall Street Journal. Published April 12, 2021. Accessed January 5, 2022.

The Healthcare Data Explosion. RBC Capital Markets. Accessed January 5, 2022.

Representation in Clinical Trials: A Review on Reaching Underrepresented Populations in Research. Yates I, Byrne J., Clinical Researcher. 2020;34(7). Accessed January 6, 2022.

Validation of a Mortality Composite Score in the Real-World Setting: Overcoming Source-Specific Disparities and Biases. Michelle H. Lerman, Benjamin Holmes, Daniel St Hilaire, Mary Tran, Matthew Rioth, Vinod Subramanian, Alissa M. Winzeler, and Thomas Brown. JCO Clinical Cancer Informatics 2021:5, 401-413.

A look at Specialty Pharmacy Dynamics. Shah P, CVS Health. Published July 13, 2020. Accessed January 5, 2022.

How Many Rare Diseases Are There? Haendel M, Vasilevsky N, Bologa C, Harris N., Nature Reviews Drug Discovery; 2019.

The Diagnostic Journey For Rare Disease Patients. Avalere. Published June 2021. Accessed January 5, 2022.