prognos
health

Real-World Data
In Life Sciences:

Faster Launch, Increased Adoption, Higher Revenue,
and a Playbook for Commercial Success

Contents

Introduction

A 2016 analysis published in the *Journal of Health Economics* and authored by the Tufts Center for the Study of Drug Development placed the cost of bringing a drug to market, including post-approval research and development, at a staggering $2.87 billion. Meanwhile, a 2018 study from the Tufts Center noted that the timeline for new drug development ranged from 12.8 years for the average drug to 17.2 years for ultra-orphan drugs that only affect several hundred patients. This places the onus on life science organizations to find ways to get therapies in the hands of the right prescribers and patients faster – especially those who cannot wait 17 years for a potentially life-saving treatment.

Commercializing a drug requires its developer to harness various sources of real-world data to identify patient populations and refine sales and marketing strategies for those populations (among other tasks). Traditionally, this process involves purchasing large data sets from data aggregators or data platforms, if not directly from the source itself – often with little knowledge of the quality of the data.

Preparing this data for analysis is both expensive and time-consuming – so much so that it's not uncommon for organizations to outsource the process to consultants or third-party vendors. The process also tends to yield a static analysis that is difficult to modify or re-run in response to follow-up questions or potential discrepancies.

A growing number of health-data marketplaces are emerging in an effort to improve this process. Developed as open platforms that comply 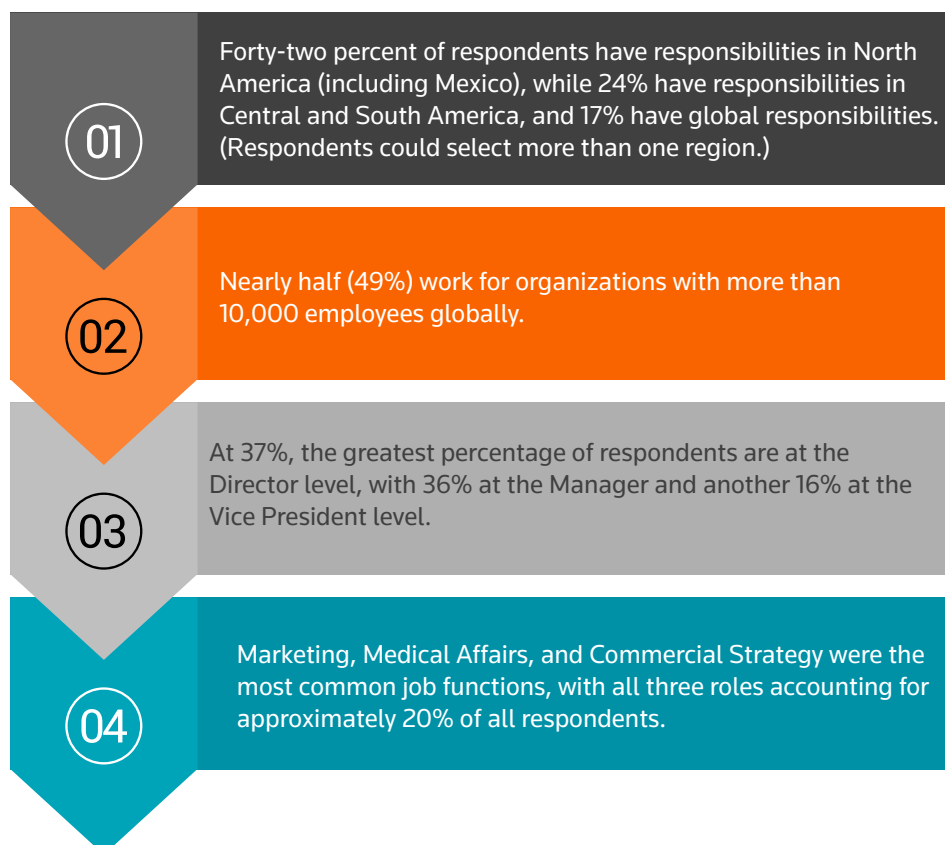with privacy regulations such as the Health Insurance Portability and Accountability Act (HIPAA), these marketplaces offer life science organizations the opportunity to glimpse a data set before buying it, to acquire only the data sets they need to complete their analysis, and in some cases to conduct an analysis within the platform itself.

Such marketplaces promise to streamline the processes of data collection, normalization, and analysis, allowing not only for faster launch but also for a nimbler approach to post-launch strategy. This can lead to higher revenue, lower costs and an increasingly data-driven approach to decision-making. It also opens the possibility for improvements to patient outcomes.

To gauge the interest of life sciences organizations in health data marketplaces, Reuters Events Pharma surveyed 215 senior pharma people around the world. Below is a snapshot of respondent demographics.

This report is based on the results of this survey, secondary research, and interviews with 10 data science professionals in the life sciences and health data platform industries about their use of data and the potential benefits of health data marketplaces.

01 Forty-two percent of respondents have responsibilities in North America (including Mexico), while 24% have responsibilities in Central and South America, and 17% have global responsibilities. (Respondents could select more than one region.)

02 Nearly half (49%) work for organizations with more than 10,000 employees globally.

03 At 37%, the greatest percentage of respondents are at the Director level, with 36% at the Manager and another 16% at the Vice President level.

04 Marketing, Medical Affairs, and Commercial Strategy were the most common job functions, with all three roles accounting for approximately 20% of all respondents.

How life science organizations leverage real-world data today

Traditionally, medical claims served as the primary external data source for life science organizations, along with internal data sources such as patient registries from previous clinical trials. In recent years, organizations have sought to augment claims with data from lab tests, prescriptions, and electronic health records (EHRs). "Life science organizations have realized that there is another avenue for them to learn from. They can gain richer insights into changes to patient behavior by looking for signals outside of claims data," says Varun Mangamoori, Head of Life Sciences Solutions, at health information technology provider Datavant.

Other sources could include but are not limited to patient-generated health data, mobile and wearable devices, and data pertaining to consumer and/or personal behavior. Multiple interview sources referred to behavioral data as a "missing piece" unavailable in traditional healthcare data sets.

According to respondents, specialty pharmacy data ranks as the most valuable data source for commercialization efforts, followed by medical claims data. At the other end of the spectrum, imaging results rank as the least valuable data source, followed by hospital Chargemaster data. (See Figure 1.)

1.	Specialty pharmacy (Rx data)
2.	Medical claims
3.	Rx claims (non-specialty)
4.	Lab test data
5.	Medicare claims
6.	EHR
7.	Chargemaster data (Hospital)
8.	Imaging results

FIGURE 1: Most valuable and least valuable data, where 1 is most valuable.

According to survey respondents, life science organizations are more likely to first identify a use case for purchasing real-world data and then set a budget; they are less likely to first set a budget for purchasing data and then identify a use case.

Once data is in hand, there are generally four steps in using it to inform commercialization strategy: Normalize and integrate data, conduct data analysis, identify opportunities, and share insights into the value of data.

> According to respondents, specialty pharmacy data ranks as the most valuable data source for commercialization efforts, followed by medical claims data. At the other end of the spectrum, imaging results rank as the least valuable data source, followed by hospital Chargemaster data

Normalize and integrate data

Each data set comes in its own format relative to its original intended use. Much of the data in EHRs, for example, is unstructured data in the form of free-form text entered in physician notes. Imaging results contain pictures as well as text. In addition, data sets may use different terminology or abbreviations to refer to the same diagnosis or procedure.

To be useful for analysis by a machine and not a human, these disparate data sets must be normalized and then integrated into a single data set. Some life science organizations complete this step internally, while others choose to outsource to a third party. Data must also be anonymized, or stripped of types of information that qualify as protected health information (PHI) and could be used alone or in combination to identify an individual patient.

Conduct data analysis

Data analysis conducted prior to commercial launch has two key purposes. The first is to gain insight into patient populations. Here, life science organizations aim to find the patients who are most likely to realize clinical benefits from a given therapy as well as take the therapy (either as a new treatment option or by switching from an existing treatment). Organizations will explore patient populations based on factors including but not limited to demography, geography, economic status, and health status.

The second purpose is to determine which hospitals, health systems, and physicians may be positioned to target those patients. This positioning could be due to potential volume – that is, a high incidence of patients with a given condition(s) in a given area – or to the previous prescribing history of a physician or facility.

Identify opportunities

With insight in hand about patient populations and physicians to target, life science organizations can develop and refine sales and marketing strategies for reaching those physicians and patients. The overarching goal is to position a therapy on the market with written and multimedia content, take advantage of the 'influence network' that impacts prescribing decisions, and provide physicians and patients with evidence-based decision support at the point of care.

Omnichannel outreach is an important part of these strategies. Physicians may receive email correspondence, phone calls, or in-person visits from representatives of life science organizations, while patients may be targeted using materials provided at the physicians' office as well as various forms of marketing outreach.

Share insight into value of data

Data science professionals in life science organizations see value in sharing across the enterprise their in-depth knowledge of the data sets they have acquired and the best way to make the most of it.

"When I describe the role of my team, I say that we are the former chefs turned farmers sourcing key raw ingredients for the chefs in the data analyst and data scientist roles," says Jane Urban, Senior Director, Commercial Data Management and Strategy, Takeda. "We find the best and highest-quality data, and we ensure that it's connected in a way that it can be utilized. Then we offer guidance on the best practices for each data type and identify where there are trade-offs – for example, which data sets are deep and contain many patients but may be missing a lot of other points."

# Key challenges associated with the RWD status quo

The life sciences industry already has well established processes of working with real-world data, and the real-world evidence derived from its analysis, to drive evidence-based decision making and improve the positioning of a therapy on the market. This leads to better targeting of commercial activities, a more refined commercial strategy, and the ability to provide physicians with decision support at the point of care.

However, the fact that the process is well established does not mean that it comes without challenges. Life science organizations struggle with data collection, data normalization, and data analysis. Making matters worse, the industry's tepid pace of change often makes it difficult to innovate or modernize these three processes.

Clearly there is great scope for pharma companies to improve processes much further to profitable effect. Adopting more advanced real-world-data analysis could generate $300 million annually in savings or value generation for the top 20 industry players, according to consultant McKinsey.

Data collection

As noted, life science organizations most commonly acquire data by purchasing it from an aggregator or from the source itself. (Figure 2 below shows the approximate breakdown of budget allocations for common data sources.)

40%
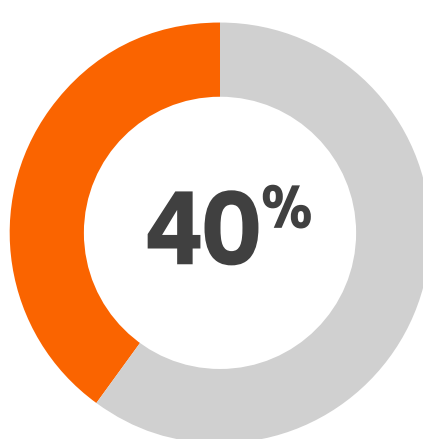
Data Source: Large Data Aggregator
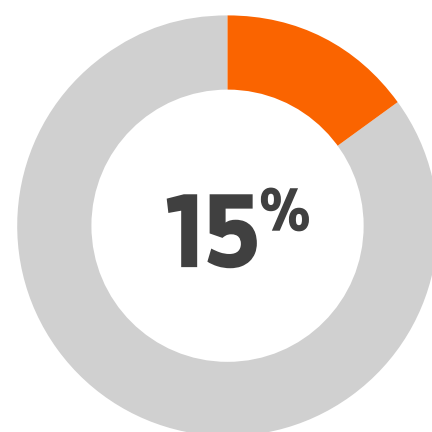**Examples:** IQVIA, Symphony Health, Definitive Healthcare

24%
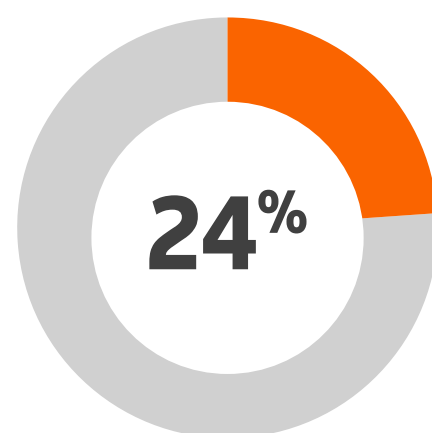
Data Source: Direct from Source
Examples: Pharmacy, lab, imaging center

18%
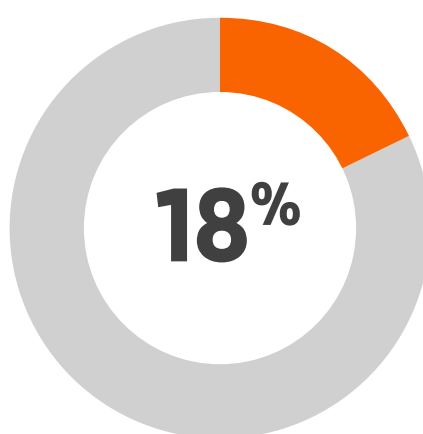
Data Source: Large Health Data Platform
Examples: AWS Data Exchange, Google, Snowflake

15%

Data Source: Health Data Marketplace
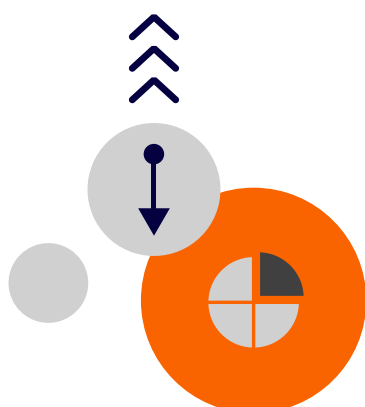**Examples:** HealthVerity, Prognos

FIGURE 2: BUDGET ALLOCATION FOR DATA SOURCES

Not only is data acquisition expensive – a single list could cost hundreds of thousands of dollars – but organizations typically purchase such data sets without knowing what information they contain. Multiple interview subjects likened this process to buying a car, whether it's paying for a car before giving it a test drive or paying for a car without knowing what color it is.

In addition, organizations run the risk of acquiring data that's similar to what they already have, further diluting its value. "If a third party tells you they have a list of 80,000 names, and you already have 50,000 names, and you know there will be some crossing of the subset, then you need to ask yourself if it's worth it to do the extra legwork," says Christopher Nugent, a digital transformation expert in the Pharma industry.

Figure 3 summarizes the value and drawbacks of several common data types, says Theresa Greco, Chief Commercial Officer, Prognos Health. "Each set of data is like a block of Swiss cheese. How big or how small are the holes?"

Data Type	Pros	Cons
Claims	Supports deep population analysis	Limited to single payer Lag time of 3+ months
EHR	Covers diagnoses, treatments, and outcomes	Mostly unstructured, requiring normalization May exclude certain care settings / specialties
Lab	Offers semantic and numeric content	Many testing outputs, even in same lab Data hard to consume in machine-ready format
Imaging	Easily mined for analysis	Same insight available in EHR, lab data
Prescriptions	Available directly from source	No insight into adherence or efficacy
Wearables	Readily available + plentiful	Limited clinical value

FIGURE 3: PROS AND CONS OF COMMON REAL-WORLD DATA TYPES

Data normalization

Because each individual data type on its own is incomplete, life science organizations must purchase data from multiple sources in order to paint a more complete picture of a patient population. At an administrative level, this poses the challenge of managing contracts (and relationships) with multiple vendors, and of aligning timelines for when data will be available for analysis.

The larger challenge is that each data set is formatted and structured differently. While hardly surprising – each data set has its own original, intended use case –it means every data set must be normalized, harmonized, and integrated into a single data set for it to be of any value to data scientists.

"Master data management can be a struggle," says Nugent. "You need to structure your internal data teams in a way that they can ingest multiple data sources that each have unique identifiers, and you need to be able to create a single data set and segment it effectively."

Our survey respondents identified the large variety of data formats and fields, and the resulting difficulty in harmonizing data sets, as the most difficult challenge associated with purchasing data from multiple vendors. (See Figure 4.) Unstructured data, especially physicians' notes in EHR data, poses a particular problem because it is difficult to tag and categorize, which in turn makes these snippets of data difficult to query.

In fact, data normalization is such a challenge for life science organizations that they allocate nearly half of their budget for the process (49%) to third-party service providers to do this work. Along with leaving fewer resources for internal stakeholders, the need to manage contracts with (and adapt to the business processes of) external consultants or vendors only lengthens the timeline for data normalization.

A final challenge associated with data normalization is the fallout that occurs after multiple data sets are integrated and duplicate records are removed. "If you're looking for something very specific, such as patients who have failed two other therapeutics for a given condition, the leakage from the process of bridging the data sets could yield an insufficient number of records for your analysis," says Greco.

This is especially true if individual data sets come from different vendors or aggregators. Each data source is likely to use different tokens in place of unique patient identifiers as defined under HIPAA regulations; the more tokens that are used, the greater the likelihood of fallout, since it is more difficult to match records.

> "Master data management can be a struggle, you need to structure your internal data teams in a way that they can ingest multiple data sources that each have unique identifiers, and you need to be able to create a single data set and segment it effectively."
>
> Christopher Nugent, digital transformation expert in the Pharma industry.

- Data is hard to link
- Varying formats with different fields – making harmonizing the data hard
- Data direct from sources is hard to work with
- Managing multiple vendors is time consuming and resource sapping
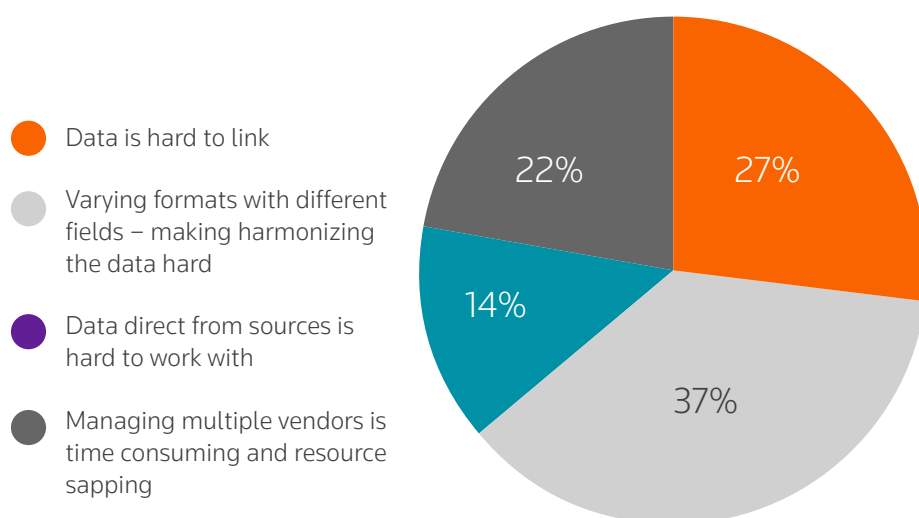
27%
37%
14%
22%

FIGURE 4: TOP CHALLENGES OF PURCHASING DATA FROM MULTIPLE SOURCES

Data analysis

As discussed, collecting and normalizing data for analysis are time-consuming processes. They are also processes that can happen several times during commercialization, particularly if an organization needs to acquire more data to get enough physician and/or patient records to run an effective sales campaign. The lack of visibility into data sets only adds to the frustration: Organizations may devote significant resources to licensing a data source, only to discover that it does not meet their needs.

Data analysis also poses a challenge due to the nature of the work itself. Querying data sets and interpreting the results of a search requires data science skills; executives who rely on this information for data-driven decision-making may have to go to their data teams instead of creating their own queries, which makes it harder to get answers quickly. In addition, effective data analysis in life sciences requires an understanding of the nuances of the patient journey – both the clinical pathways and the types of data generated at each step along that path. For many life science organizations, this knowledge and understanding is in short supply.

Finally, organizations that hope to do predictive modeling require a longitudinal view of patient history and behavior. This can be difficult to obtain from data sets with breadth but not depth.

"I primarily work on predicting the next-best action. On the patient side, this is looking at what type of future diagnosis we may be able to predict based on factors such as their underlying diseases and their drug readiness. On the provider side, we're trying to predict who will be early adopters, and how many prescriptions they may write for a given brand," says Anuvrat Chaturvedi, Associate Director, Data Science AI & Innovation, Novartis.

"For me, it's about the depth of the data. Even a data set that covers more than 300 million patients is not enough, because not every transaction is covered. With claims, lab, and EHR data, there's still a lot that's left unsaid. We need more enhanced data to make predictions."

1.	Expensive (multiple vendors, contracting, etc.)
2.	Time to value (getting answers to use case)
3.	Internal mindset shift in understanding the value of the data for decision-making purposes
4.	Managing multiple data vendors
5.	Lack of in-house technical resources or capabilities to maximize the data
6.	Inflexible (linking, integration, operationalizing)
7.	Regulatory/Compliance
8.	The insights are too 'opaque'
9.	Poor usability (clinical data)
10.	Change/Scope purchase management (nickel & diming)
11.	Double buying (across aggregators)

FIGURE 5: Rank the following challenges as relates to how to gather actionable data (1 being the most painful and 11 being the least)

> "It's about the depth of the data. Even a data set that covers more than 300 million patients is not enough, because not every transaction is covered."
>
> Anuvrat Chaturvedi, Associate Director, Data Science AI & Innovation, Novartis

Pace of change

Beyond collecting, normalizing, and analyzing real-world data, data scientists within life sciences organizations face a larger challenge in confronting the slow pace of change in their industry. In fact, survey respondents ranked 'internal mindset shift' as one of the most painful challenges in gathering actionable data, tied with the expense of contracting with and managing multiple vendor relationships. (See Figure 5.)

One reason for the slow pace of change is the structure of contracts with data aggregators. Many vendors require three-year contracts. Life science organizations that have made this investment of time and money may be less receptive to evolving towards a health data marketplace.

Another factor is the relationship that many life science organizations have with the consultancies that collect, normalize, and analyze data on their behalf. There are some benefits to these relationships, as consultants have expertise in

managing multiple vendors, performing administration and management tasks, and supporting commercial launch processes. However, consultants tend to have their own business processes, which may not complement the internal workflows of a life science organization and make it even more difficult to derive value from data.

A final obstacle is trying to realize the potential of real-world data analysis at certain stages of the commercialization process where it hasn't previously been applied. "Life science organizations are realizing that data is important, but there's still hesitation to use it during certain stages of the product development life cycle," says Paul Petraro, Global Head, Real World Evidence Analytics Center of Excellence, Boehringer Ingelheim.

"With regulatory submission and market authorization, there's hesitation because even the U.S. Food and Drug Administration and European Medicines Agency are cautious about what they will accept. On the other hand, if it's for use
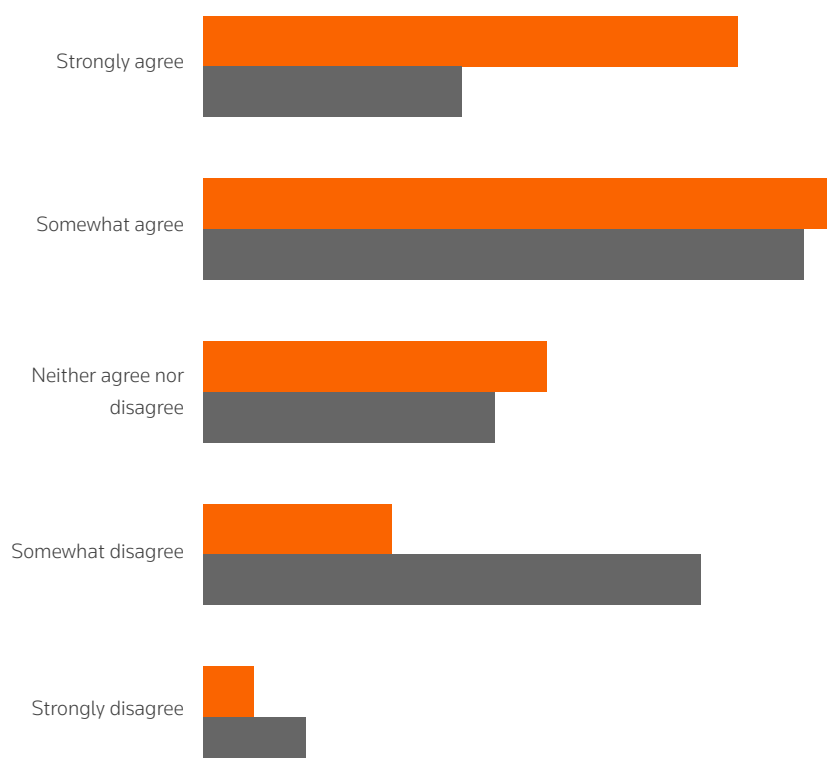
post-authorization, or for patient access and safety, that's not a hard sell, because that's where real-world data has been used for decades."

The *status quo* for collecting, normalizing, and analyzing real-world data leaves a lot to be desired for life science organizations, which are all too often simply left to hope that the data sets they are acquiring are going to meet their needs. By and large, organizations struggle to obtain the insights they need when they need them, making it difficult to develop and execute a commercialization strategy that's both comprehensive and flexible.

Two-thirds of survey respondents "strongly agree" or "somewhat agree" that the way most life science organizations buy and manage patient-level data is ripe for disruption. This compares to just 14% of respondents who strongly disagree or somewhat disagree.

The health data marketplace model is poised to provide that disruption.

● The way most pharma companies buy and manage patient-level data is ripe for disruption

● In 5 years' time, pharma companies will be buying and managing patient-level data, as easily as buying something on Amazon or watching something on Netflix

Strongly agree

Somewhat agree

Neither agree nor disagree

Somewhat disagree

Strongly disagree

Real-World Data In Life Sciences: Faster Launch, Increased Adoption, Higher Revenue, and a Playbook for Commercial Success

What is a health data marketplace?

A health data marketplace is a platform that licenses de-identified, transaction-level patient data from dozens of trusted data sources of hundreds of millions of patients. A marketplace also leverages analytics and automation to normalize data sets for faster implementation and application to a variety of use cases, including commercialization and health economics and outcomes research (HEOR).

In some cases, the marketplace can conduct certain types of data analysis that a life science organization would otherwise need to do in-house or outsource to a third party, reducing the overhead associated with technology infrastructure and staff for data analysis.

Prognos Health's Greco compares the health data marketplace to an online music streaming service. Instead of buying a physical album and playing it on a device, streaming services let users buy digital versions of songs and even listen to them before purchase, play the songs on devices of their choice, and put them in different playlists. "Music used to be a manual, analog purchase, but now it's digital and easily accessible," she says. "Real-world data is going there. Marketplaces can link data sources and provide turnkey access to answers in minutes."

The main benefits of the health data marketplace model

The health data marketplace model offers several advantages when compared to the traditional model of data collection, normalization, and analysis. One important benefit is **speed to insight**, which survey respondents identified as the top priority for improving the way they manage, store, and obtain insights from real-world data. (See Figure 7.)

"It's a much more streamlined way of accessing data. You can go to the marketplace and do it yourself, or you can work with the marketplace to find additional data sources that are available to meet your needs," says Boehringer Ingelheim's Petraro. "There's also the administrative piece. Contracting with the vendor, getting the data, working with data analysis – the marketplace expedites those steps. We get almost immediate access to the data."

The health data marketplace model also benefits from tokenization, which replaces PHI with non-sensitive data elements referred to as tokens. This makes it possible to normalize multiple data sets – and break the silos that normally stand between them – while at the same time preserving **privacy and security.**

Speed to insight with already integrated and easily linked patient-centric data

40%

Increased flexibility in compliant data selection, applications and use cases

33%

Ongoing access to unique data sets with deeper clinical specificity

31%

Cost effectiveness in data acquisition and ongoing utility

40%

FIGURE 7: How would you **rank** your key priorities for improvement in how you manage, store and gain insights from patient data?

"Without the use of tokens, there isn't a HIPAA-compliant way to de-identify and link data. To be HIPAA compliant you have to strip out HIPAA's 18 patient identifiers, or you need to have a third party certify that the risk of patient reidentification is minimal. That limits the availability and utility of the data among anyone who doesn't have authorized access to it. But with tokenization you can replace identifying information with a privacy-preserving token that is the same across the data sets," says Varun Mangamoori, Head of Life Sciences Solutions at Datavant. This makes it possible to get a more complete picture of patient data, and to put it into chronological order.

Marketplaces also support a more complete picture of data by offering **access to non-traditional data** sources. This appeals to data owners, such as small lab test companies, who may struggle to market their data directly to life science organizations alongside larger labs or data aggregators, but nonetheless have value to add.

"The marketplace is a way for niche data owners to see liquidity without standing up an entire business unit. It's a faster path to value realization," says Dan Scudder, General Manager, Data Ecosystem, Datavant. "It's like opening an art boutique. When you get a lot of small artists together, you get enough scale to make it worthwhile."

These data sources also appeal to life science organizations looking for something unique. "We're able to get access to less readily available or non-traditional datasets, and then we can link it to internal data," says Sandeep Burugupalli, Real World Evidence Strategic Partnerships & Operations Lead, Pfizer. "The biggest benefit is agility, especially when we need to move quickly. Data aggregators aren't always as quick to build this into their offerings."

> "The marketplace is a way for niche data owners to see liquidity without standing up an entire business unit. It's a faster path to value realization."
>
> Dan Scudder, General Manager, Data Ecosystem, Datavant

Eight ways health data marketplaces accelerate RWD process improvement

In the current environment, the majority of respondents (55%) say they are "somewhat confident" in their data strategy as it relates to their commercial strategy, with only minor improvements to be made. This compares to 28% of respondents who are not confident in their strategy, and see a need for many improvements, and only 11% who are very confident and see no room for improvement.

What will help the industry boost this confidence? Looking five years into the future, nearly half of respondents (49%) say the presence of open yet compliant platforms is the most important factor in helping derive value from data for commercialization strategy. This compares to 28% of respondents who say faster ways to integrate disparate data sources is the most important factor and 23% of respondents who point to faster time to value with integrated data assets.

The health data marketplace model is designed to specifically address these wishes – open platforms, faster integration of data, and faster time to value. There are eight key ways through which this model is poised to accelerate real-world data process improvements in life sciences and shorten the time it takes to bring much-needed therapies to market.

1. Try before you buy.

Marketplaces allow customers to evaluate data sets before purchasing them. This ensures that life science organizations acquire data that will add value to the data sets that they already have without providing a large number of overlapping records. It also offers the flexibility to search for data that may meet a user's detailed search criteria or specific research needs – and to go back as needed to look for additional data as new questions arise.

2. Only pay for what you need.

Within the health data marketplace platform, users can create Boolean rules to segment the available data and buy only the subset of the data that they need. For example, if 90% of a claims data set overlaps with the EHR data that a life science organization has, the marketplace makes it possible for the organization to purchase the remaining 10% of the claims data.

3. Run no-code queries.

Marketplace platforms also leverage checkboxes and filters that let users segment data and run queries without writing SQL code. "This compresses the data analytics side of the equation. The process is simplified because all the data is there and it's accessible at Web speeds. You just need to select or unselect filters within the user interface ," says Jason Bhan, MD, Co-Founder and Chief Medical Information Officer, Prognos Health. This benefits two groups of users: Stakeholders in large organizations without a background in data science, and smaller life science organizations that may lack the resources to invest in a full data science team.

4. Do a deep dive.

Because the data on the marketplace have already been combined and standardized, users have access to data sets that are ready for in-depth analysis at the time that it's purchased. This allows data science teams to spend significantly less time managing individual data sets from different data vendors and more time to run multiple analyses on different data subsets – gaining quick access to insight that might otherwise take months to obtain. "One of the things that COVID-19 has taught us is that we need to be more self-reliant. We could have a third-party vendor do an analysis, but the more that we can bring it in house and leverage our own data team, our IT stakeholders, and our sales operations, the more we are able to create a more customized solution and figure out the best path forward," says Nugent.

5. See the full picture.

Marketplaces bring together data sets that are wide but not deep, as well as those that are deep but not wide. This helps life science organizations fill in the gaps in their normalized data sets – which is particularly valuable for researching specialized treatments or rare diseases that typically have smaller lists of targeted patients. "They can match the cohorts in the data sets that they already license against more specialized data sets such as EHR or SDOH data to get the richness they are looking for," says Datavant's Mangamoori.

49%

Presence of open yet compliant platforms is the most important factor in helping derive value from data for commercialization strategy

6. Look into labs.

As noted earlier, survey respondents ranked lab data as having a moderate amount of value – ahead of Medicare claims, imaging results, and hospital Chargemaster data but behind more traditional data sources such as prescriptions, claims, and EHRs. Because lab data is difficult to harmonize, it is traditionally unavailable from data aggregators. However, lab data can provide significant clinical insight for patient targeting that isn't found elsewhere – namely, real-time information on diagnosis, disease progression, and disease severity. As a result, health data marketplaces aim to fill this gap and make lab data available to life sciences.

7. Support your Center of Excellence.

Many global life science organizations are creating a Center of Excellence that focuses on securing and preparing real-world data for use by business units throughout the enterprise, especially those that have traditionally lacked access to data resources. The health data marketplace model complements the Center of Excellence model nicely – organizations increase their investments in data science as more data is made available, and additional data is made available because organizations see more ways to use more data. Lauren Becnel, Vice President, Real World Evidence, Pfizer, said her team focuses on offering a "menu of services" in order to best support the variety of use cases that present themselves throughout the organization. Some teams may be researching a specific vaccine, while others may be taking a broader approach.

28%

Faster ways to integrate disparate data sources is the most important factor

## 8. Link to connectivity platforms.

While health data marketplaces are a valuable resource for obtaining real-world data, it's worth noting that not every data source will be found there. Certain data sets that provide highly detailed information about a single cohort type may still be sold separately; so, too, are certain sets of claims, EHR, and specialty pharmacy data. Of course, there's also the internal data that life science organizations have acquired over time. To address this gap, some marketplaces partner with data connectivity platforms that form a link between the data available on the marketplace and the data that must be acquired elsewhere. This has the added benefit of making non-marketplace data available for analysis within the marketplace platform – allowing for faster and simpler analysis than sending data to a third party or running an analysis in business intelligence software.
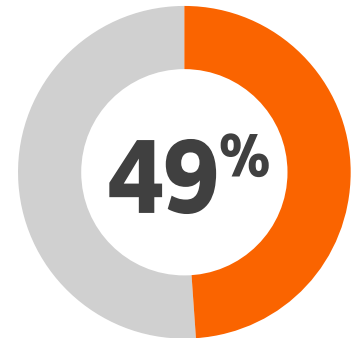
23%
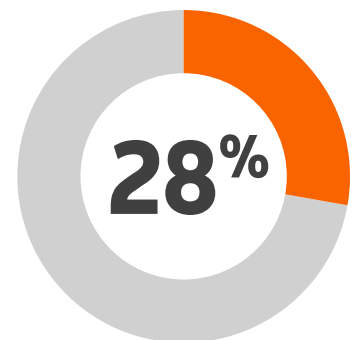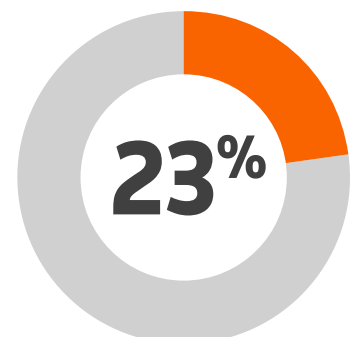
Faster time to value with integrated data assets

Impact of health data marketplaces on commercialization and post-launch activity

Life science organizations could realize the benefits of health data marketplaces across the entire product lifecycle. HEOR and patient journey teams, for example, can gain access to the full data sets they need to conduct deep-dive, patient-level analysis. There are several ways in which the health data marketplace can support a commercialization strategy as well as post-launch activity.

Commercialization: Targeted marketing and messaging

Market segmentation and planning. To target the right patients, life science organizations need to know as much about them as possible, from socioeconomic data to the websites they visit. In targeting physicians with marketing and educational materials, meanwhile, it's helpful to know who has submitted the most claims for a given condition, written the most prescriptions for a given therapy, or referred the most patients to a particular lab for follow-up testing. "You want as many potential touchpoints as possible, and for that you need a broad data set," says Prognos Health's Bhan.

Omnichannel marketing. Market segmentation allows for targeted marketing messaging – not just for a given area or population, but also for multiple marketing channels (print, online, social media, radio/TV, and so on). Such targeting has the added benefit of providing more detailed insight post-launch into the 'lift' in prescriptions for the targeted audience, which lets organizations shift their marketing strategy based on what is working and what is not.

Field deployment. Getting information to physicians at the point of care can positively influence their treatment decisions as well as patient outcomes. As with marketing and educational efforts, it's important for life science organizations to target the right physicians. With robust insight into factors — such as who is likely to prescribe, have a conversation with a patient about treatment options, or distribute educational material — life science organizations can better plan where to send representatives into the field. This helps ensure that in-person field team visits are saved for the physicians that an organization has prioritized.

Post-launch:
A playbook for future success

Increase sales. Leveraging data to enable more refined provider engagement, optimize sales and marketing resources, and identify the right patient cohorts helps a life science organization get the right therapy into the hands of the right patients and providers faster. This leads to increased adoption of the therapy, which increases sales and revenue. The ability to access and analyze data within the health data marketplace platform further enables organizations to identify opportunities to increase sales without going through the time- and labor-intensive process of licensing new data sets from aggregators.

Centralize data acquisition. As life science organizations create Centers of Excellence and refine their processes for enterprise-wide data use, marketplaces offer an opportunity to support their efforts to centralize data acquisition and use one data set in many ways. This is especially true when researching biomarkers that are often only found deep in clinical lab tests results. For example, if a certain biologic has been approved to treat multiple types of cancer, then the data set related to that biologic can be used to study each type of cancer. This eliminates the need to buy individual data sets for each cancer type – data sets that are likely to have significant overlap given that they all pertain to the same biologic.

Improve outcomes. More complete data sets are poised to improve patient outcomes. Life science organizations can understand how different population segments are reacting to a given treatment, which can influence current product targeting efforts as well as future product development initiatives. Looking deeper, organizations with several years' worth of harmonized data about specific patient populations also have the potential to provide insights that individual healthcare professionals may not see in a single visit. These could include hard-to-detect symptoms that may be a sign of a rare disease diagnosis, triggers of food or airborne allergies, or subtle changes to treatment patterns.

"There are patterns that the human mind can't see or find. If we can put it all together and step back, we may be able to help patients figure out what's wrong before a physician does," says Takeda's Urban. "Can we shorten the time it takes for patients to find that answer, and the time it takes to give them something to treat that problem? Where can we accelerate that process?"
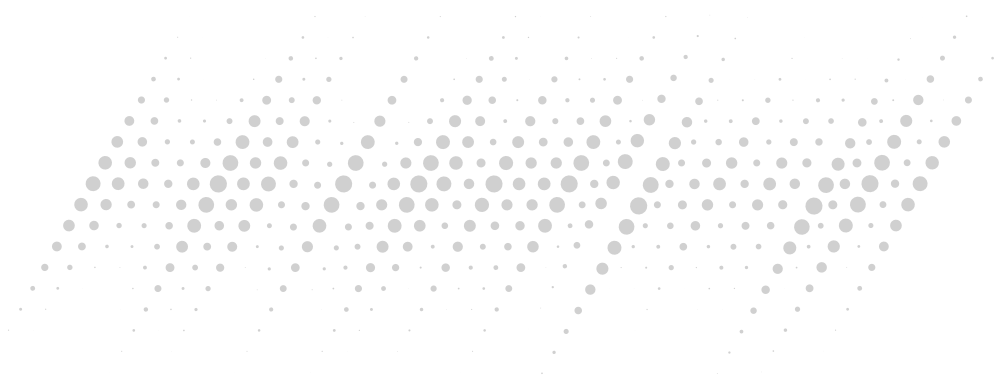
Conclusion

Traditional methods of data collection, normalization, and analysis for the commercialization process in the life sciences rely largely on acquiring data sets from aggregators for a large fee and without knowing what the data looks like. This poses a number of challenges to organizations – chief among them the length of time it takes to prepare disparate data sets to be analyzed, conduct an analysis, and bring meaningful insights to key decision-makers.

Despite the obvious flaws in this process, it remains the status quo for much of life sciences. Many organizations are locked into long-term contracts with data aggregators, leaving little else to spend on alternative data sources. The sheer complexity of data normalization

– especially for unstructured data, where the value is often derived from free text that's hard to query – also limits the appetite for incorporating more data into existing data sets.

The emerging health data marketplace model couples de-identified, normalized transaction-level patient data with easy-to-use filter features and analytics capabilities. With the difficult work of preparing data already done, users can focus on running detailed analyses on in-depth data sets that take days instead of months. This can drive significant process improvements for the use of real-world data, both during commercialization and after launch – enabling organizations to better target the right therapies to the right prescribers and patients, boost sales, and improve patient outcomes.
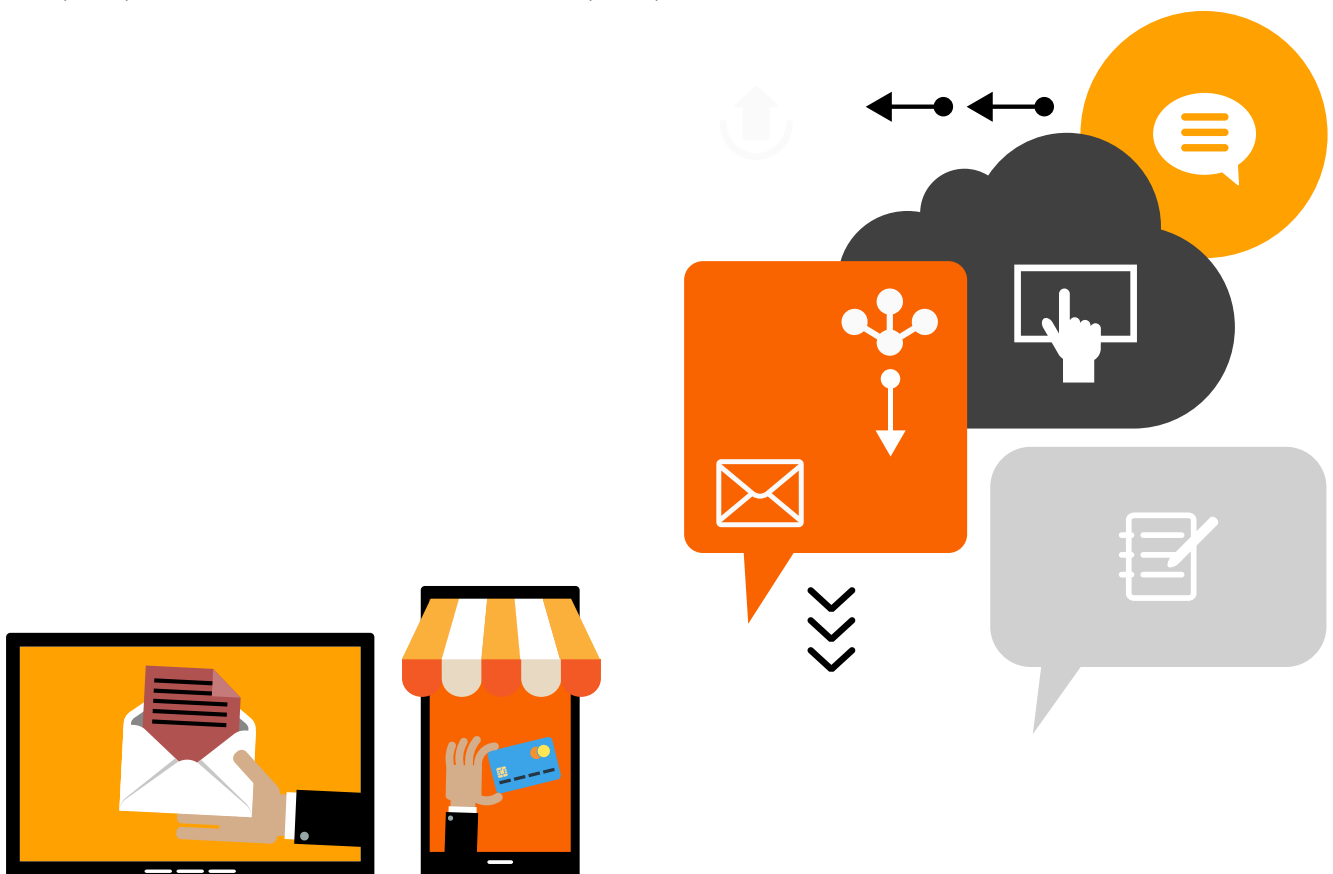
Contributors

Reuters Events would like to thank the following contributors for sharing their time and insights for this report:

Lauren Becnel, Vice President, Real World Evidence, Pfizer

Jason Bhan, MD, Co-Founder and Chief Medical Information Officer, Prognos Health

Sandeep Burugupalli, Real World Evidence Strategic Partnerships & Operations Lead, Pfizer

Anuvrat Chaturvedi, Associate Director, Data Science AI & Innovation, Novartis

Theresa Greco, Chief Commercial Officer, Prognos Health

Varun Mangamoori, Head of Life Sciences Solutions, Datavant

Christopher Nugent, a Digital transformation expert in the Pharma industry

Paul Petraro, Global Head, Real World Evidence Analytics Center of Excellence, Boehringer Ingelheim

Dan Scudder, General Manager, Data Ecosystem, Datavant

Jane Urban, Senior Director, Commercial Data Management and Strategy, Takeda

About Prognos

Prognos® Health is a patient-centric data and analytics company that harnesses the value of real-world data (RWD) to accelerate the path to clinical truth. Our HIPAA-compliant software solutions powered by Prognos Factor®, start with access to the largest RWD Marketplace of integrated record-level sources that include 325M de-identified patients, multi-lab test data results, open and closed medical claims, as well as Rx data. Prognos enables healthcare organizations to better and more accurately target specific populations, optimize commercial and HEOR efforts, and assess member underwriting risk. For more information visit prognoshealth.com

About Reuters Events

The pharmaceutical division at Reuters Events strives to make Pharma more open and valued. More open so that the strongest ideas and insights are brought to the fore in a transparent, trustworthy manner. More valued by taking an authentic approach to building products and services that matter to patients.

To do this, Reuters Events provides a hub for senior-level Pharma executives, patient groups and other health stakeholders to exchange ideas and observe shifting trends and practices. We actively respond to the aims and interests of our audience, so please get in touch.

Accelerate Your Path to Clinical Truth.

RWD insights throughout the patient care journey

Prognos Marketplace is a patient-centric healthcare data analytics solution that increases speed to data value.

100B+
lab results, medical, and Rx claims

325M+
de-identified patients

2M+
treating providers

largest collection of closed medical claims

ANSWERS IN MINUTES, NOT MONTHS

prognos health

Prognos® Health is a patient-centric data and analytics company that harnesses the value of real-world data (RWD) to accelerate the path to clinical truth. Access the largest RWD marketplace of integrated record-level sources powered by our HIPAA-compliant software, Prognos Factor®, enabling healthcare organizations to more accurately target specific populations, optimize commercial and HEOR efforts, and assess member underwriting risk. **Visit prognoshealth.com to learn more.**