WHITE PAPER

How to navigate European data

Life sciences industry leaders need global data—but to get it, they face the challenge of navigating the European data landscape.

By Florent Moise, Waldemar Ockert and Edith Aggarwal

Impact where it matters.®

The pursuit of European data

The life sciences industry has been dependent on data since the late 19th century, when former chemical dye companies uncovered its value to both research and manufacturing. But over the last two decades, major advances in information technology (IT) infrastructure, software and data processing have propelled data science as a driving force in the industry's evolution.

Today's life sciences companies capture, manage and use vast amounts of data. Internally sourced data comes from many arms of the organization, including manufacturing, supply chain, sales operations, research and development (R&D) and clinical trials. While these data sets offer great value to the business, many important decision-making processes hinge on data from outside company walls. As a result, life sciences leaders rely on externally sourced data to drive market access, evidence generation, field force design and optimization, product launches and operations, product performance measurement and much more.

The cost of managing all this data is substantial. While it's impossible to generate precise benchmarks, ZS's proprietary research suggests that life sciences companies generate or procure thousands of distinct data assets, spending between $60 million to $250 million each year on data licenses globally (depending on company size, therapy area focus and other factors). This benchmark doesn't include one-time data purchases or the incremental expenses incurred throughout procurement, data warehousing, analytics and insights supporting a variety of use cases.

"Life sciences companies generate or procure thousands of distinct data assets, spending between $60 million to $250 million each year on data licenses globally."

Although the direct and indirect costs of global data account for a significant portion of organizations' operational expenses, the insights, capabilities and decisions this data unlocks and the applications it enables are worth many times these costs.

The challenge of obtaining European data

Life sciences leaders around the globe recognize that data is a strategic differentiator that can help their organization secure competitive advantages. With data-driven insights, companies can make more calculated R&D decisions, reduce clinical trial failure rates, generate the evidence they need to facilitate better market access terms and target accounts strategically.

Patient-level data sets are particularly valuable because they are versatile, enabling an array of strategic applications. But to source these global data sets, organizations must navigate a complex regulatory environment that varies by region—and European data is notoriously difficult to procure.

"To obtain European data sets, life sciences companies must contend with a fragmented network of more than 2,500 distinct data sources, each with its own set of nuances and terms."

To obtain European data sets, life sciences companies must contend with a fragmented network of more than 2,500 distinct data sources, each with its own set of nuances and terms. Navigating this data landscape is not easy, but the organizations that identify the data sets best suited for their use cases and applications are poised to make better and, at times, faster strategic decisions in a range of critical business areas.

Why is European data difficult to source?

One of the major conundrums for life sciences companies is how to source and access the right data for a broad array of use cases and applications. Despite the vast amount of high-quality data being captured across the healthcare ecosystem, only a narrow spectrum of that data is accessible for commercial or even non-commercial analyses. For example, electronic medical records (EMR) collected in office-based primary care settings are available across all major European economies, whereas scaled, patient-level data sets collected in hospital care settings are not. Locating suitable patient-level data is only half the challenge. There are narrow limits on how companies can use different data sets, which come with various data governance obligations across the provider landscape. In addition, multiple legal frameworks determine how healthcare data can be accessed and used:

- **Data privacy.** Many valuable insights hinge on how patients navigate healthcare ecosystems, from initial symptomatology and diagnosis through laboratory, imaging or biomarker testing to subsequent surgical and life sciences interventions. Because insights captured by these data archetypes contain sensitive personal health information, they are governed by privacy regulations.
- **Medical confidentiality.** Access to sensitive patient data is governed by medical confidentiality laws, which vary across Europe.
- **Regional regulations**. Additionally, a myriad of pan-European, national and state-level regulatory frameworks factor into the data exchange, depending on the type of data and the care setting. These frameworks include hospital laws, social security legislation and insurance claims data usage.

The effect of these multifaceted regulatory layers—each with its own set of country-specific nuances—is a mosaic of data governance standards, limited data availability and complexities related to the permissibility of data applications. The most advanced organizations have developed capabilities and frameworks to help them navigate this environment.

The struggle to find representative data

Another challenge stemming from a complex regulatory network is adequate data representation. Across Europe, data sources have varying levels of breadth and coverage. As a result, some sources may not sufficiently represent a patient population or a set of healthcare trends, leaving gaps in data availability.

Along with inconsistent representation, data sources may also lack the depth and granularity required to power a broader spectrum of life sciences use cases. This lack of depth poses serious limitations on generating real-world evidence and conducting health economics and outcomes research (HEOR) studies.

When considering data availability across the EU4 (Germany, France, Italy and Spain) and U.K. regions, it's helpful to note a few patterns that hold across data sources:

- **Claims data.** Regarding administrative claims data, all EU4 countries and the U.K. have assets and providers that offer either analytics services or raw data access. These assets can be used across a broad range of use cases, particularly where the claims data captures all reimbursement healthcare system encounters and thereby provides insights across all care settings and therapy areas. Only the U.K., however, provides raw data access for other approved cases and applications.
- **EMR data.** Across the five European markets, EMR data sets are available for both primary care and secondary settings, where specialty care is provided in a physician's office-based environment. Because EMRs are typically versatile, they play a role in both commercial and non-commercial use cases such as real-world evidence studies. In fact, several leading providers have developed a network of relationships with hospitals to power evidence generation studies in the hospital care setting. More recently, these hospital provider networks have even developed proprietary solutions or capabilities fueled by EMR data. But it's important to consider the cost and time it takes to access these records, along with the strict data governance obligations for entities that permit only a narrow set of use cases.
- **Prescription data.** Long considered the bedrock of life sciences commercial analysis, prescription data is typically sourced through claims processing centers. These centers manage the product dispensation claims between pharmacies and insurers, which they complement with wholesale and retail pharmacy data. In a challenging data environment with strict data privacy obligations, prescription data provides granular and targeted geospatial insights that can help organizations track market performance within a small geographic region, but the insights aren't specific to a healthcare facility or a treating

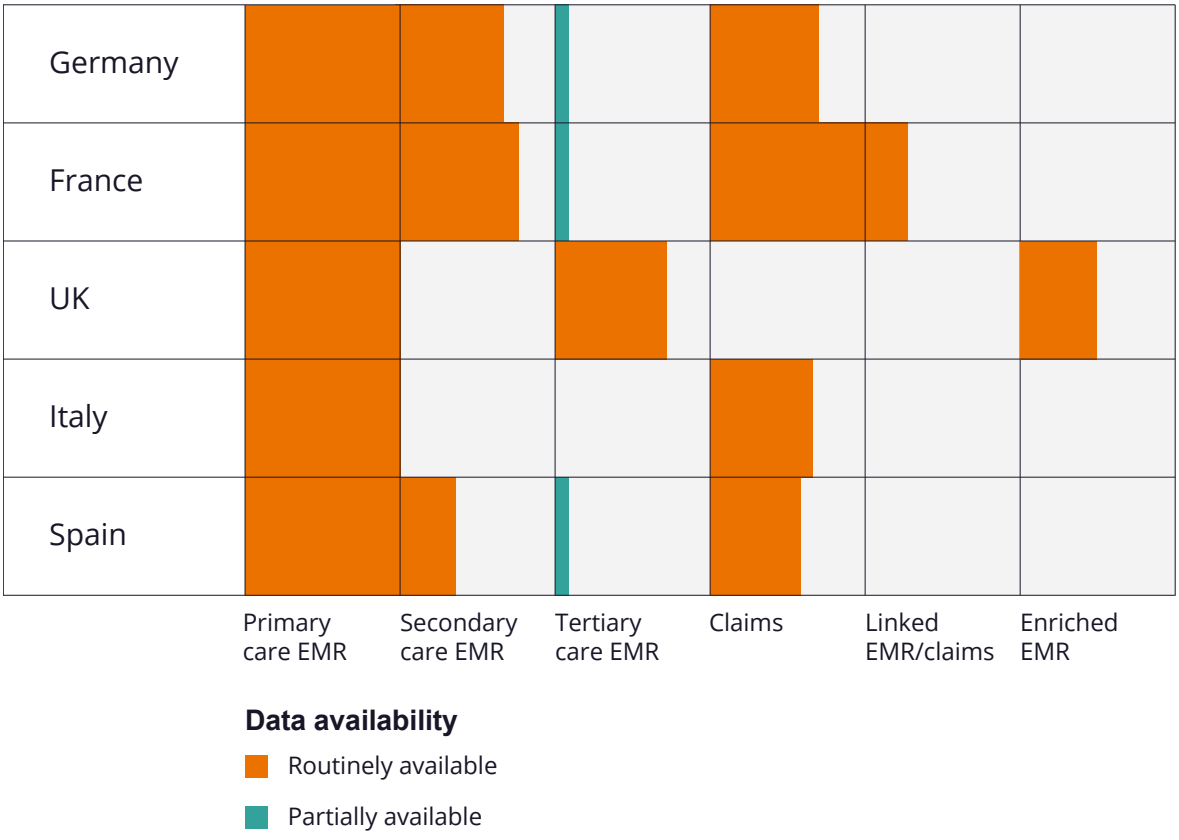
clinician. Prescription data can also help companies understand patient product switching behavior, treatment duration or persistence. But because it lacks diagnostic data and other relevant patient information, prescription data provides limited insights into the patient journey.

- **Chart audits.** Several providers offer chart audit data and consulting services for a range of therapy areas and indications, but they typically apply a degree of specialization. Providers maintaining readily available product solutions with a regular update cycle can offer both off-the-shelf and bespoke data and solutions. While these data sets provide quality insights, it's important to consider the size and representation of the physician and patient panels, whether the data is longitudinal and if the data relies on physician recall.

Figure 1 represents data availability in Europe by data type and region, illustrating the gaps and disparities that characterize Europe's fragmented regulatory environment.

FIGURE 1:

Data availability in Europe by data type and region



How to locate the right data sources

Because data availability in Europe varies by country, therapy area, care setting and data type, life sciences organizations often struggle to hone an effective and efficient data sourcing strategy. Yet some companies continue to demonstrate that success is possible. But to get there, a rigorous and intentional data sourcing strategy is key.

Map your use cases
Begin by listing the ways you want to use the data and anchor your search to those use cases. From there, organizations can systematically scan the available data assets by region and assess whether the breadth of available data assets fits each use case.

Reverse engineer your path
Because one data asset will not meet most of the organization's needs, a reverse engineering approach is often the most efficient way to find the best sources. Starting with the use case, follow a systematic analysis to identify the best data sources. Frameworks like these have guided dozens of life sciences companies with footprints in areas including oncology, rare disease and cardiovascular disease, allowing them to obtain data for commercial strategies, medical affairs and R&D decisions.

Plan ahead
To get the most value from your data, think about the future. Conduct an integrated use case prioritization analysis over a period of up to five years to generate a comprehensive map of data needs over the medium-term future.

Consider the therapy areas
Go a step further when mapping data sources to your use cases. Apply a therapy area-specific lens and consider how your data flows across healthcare ecosystems because this can differ significantly across various indications.

Check for strategic compatibility
Lastly, it's important to consider the fit of your potential partner's data assets with your organization. Conduct a systematic suitability analysis for existing data providers and their assets as well as potential data partners with data assets still in development. Prospective partners that do not have any data assets today may be in the process of developing the ability to capture the data sets you need. They may also have a strategic interest in co-developing data assets through a partnership.

What success looks like

By following a structured approach and developing a data sourcing strategy, ZS has helped life sciences leaders plot hundreds of potential data assets. In many cases, we also empower clients to validate the capabilities of dozens of unknown data assets and providers. In doing so, our clients can be sure that they have the best data for their use cases and research questions, acquire better data and elevate the quality of the insights the data unlocks.

CASE STUDY

A major life sciences company based in the U.S. sought ZS's help in mapping an integrated data sourcing strategy across its commercial, medical affairs and R&D functions in its five largest European markets and Canada. To find the best data quickly and efficiently, ZS helped the client work across functional boundaries to define a roadmap and a strategy. This collaboration helped minimize data redundancy, optimize the procurement processes and expenditure, and deliver better business results across the organization.

To start, ZS helped the client define its key business needs over a three-year timeframe, then applied a framework to consolidate the data requirements across all prioritized use cases. From there, the team defined an integrated data requirement map, then translated it into a structured vendor questionnaire. The questionnaire was sent to more than 200 vendors offering data or analytics services for EMR, claims, chart audit or prescription data, as well as registries, hospital provider networks and EMR software providers.

By conducting a quantitative and qualitative suitability analysis, the team could map how well each data asset enabled each of the prioritized use cases. As a result, they could easily construct an evidence-based data asset portfolio to meet the organization's diverse business and data needs.

How to navigate the European data landscape

The dynamics of the European data landscape make it a complicated and ever-changing environment. To navigate it successfully, an intentional and systematic process is critical. Here are the key factors organizations need to consider to efficiently source data.

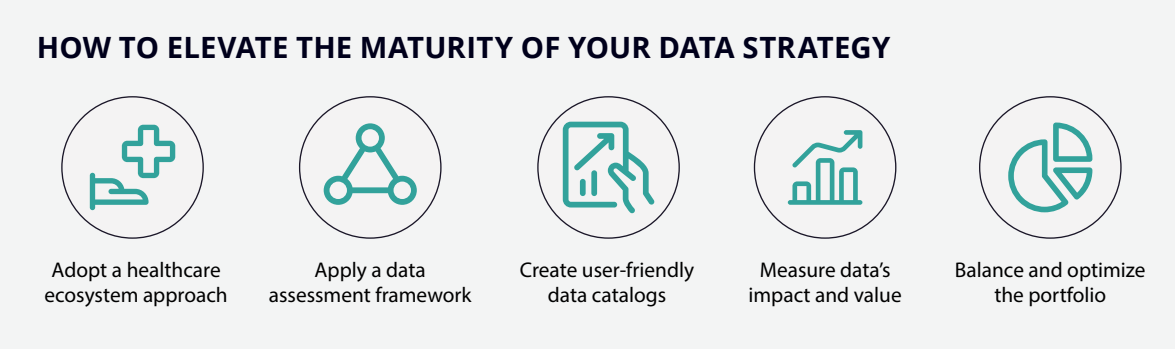
Create a network of strategic partnerships
Establishing a wide and sophisticated data partnership network can help organizations get the data sets they need faster. Networks allow companies to broaden and differentiate the capabilities they use to generate insights across commercial and medical affairs and collect effective evidence for improving market access.

Develop a cohesive data governance and operating model
The road to a successful European data strategy will look different for each organization, but those with a higher level of maturity tend to have a clear corporate-level vision and a well-defined operating model. When organizational leaders share a vision for the data strategy and have the right people and processes in place, widespread adoption and implementation comes more easily.

Fuel your data sourcing strategy
Adopting the capabilities featured in figure 2 can help the organization accelerate its data sourcing strategy. Likewise, enhancing one of more of these capabilities can significantly improve the organization's data strategy maturity level.

FIGURE 2:

Capabilities to accelerate a data sourcing



Changes on the horizon

The European data landscape is transforming. Over the last decade, innovative solutions and capabilities have shifted the market and disrupted the global data exchange—especially when it comes to patient-level data assets. Nevertheless, significant gaps remain as the industry contends with its data availability issues. These gaps, along with the influx of innovations, form a convergence of several trends that are likely to drive changes in the industry over the next few years.

Shifting governmental and regulatory initiatives

The ratification of the General Data Protection Regulation (GDPR) in 2018 and its updates in 2021 represent a monumental shift in how data privacy is managed across the European Union. With GDPR came the opportunity to balance privacy rights with legitimate use cases for healthcare data. While the legislation focuses on European individuals' rights to privacy, there are provisions that allow the processing of healthcare data for public health initiatives and even commercial use under certain conditions.

When Germany's general data protection regulation—Datenschutz-Grundverordnung (DSGVO)—came into effect a few months after GDPR, it contained many provisions allowing data use for research and analytics by both public and commercial entities under certain conditions. Just as many perceived the DSGVO's move as a nod toward an increase in healthcare data use for public and private research, the UK's proposed changes to the GDPR in 2021 may signal a shift in the market. It's likely that the UK will continue to enable a competitive regulatory environment that allows data-driven research and analysis in certain cases.

In addition to these gradual regulatory shifts, governments in several European countries have also recognized the need for improved interoperability and made significant advances in their policies. For example, the Medical Informatics Initiative (MII) in Germany sponsors the research initiatives of four university hospital consortia. The Health Data Hub in France represents many ambitious, publicly funded initiatives to improve interoperability across healthcare data systems, standardize access and accelerate the use of artificial intelligence and machine learning technologies in public health research.

An influx of consortia and public-private partnerships

By 2024, a host of alliances and big-data initiatives will change the scope and scale of European patient data. Many public-private research consortia have emerged over the last few years, such as the European Union's Innovative Medicines Initiatives (IMI) and its Big Data 4 Better Outcomes (BD4BO) program, to fill gaps and improve data circulation.

With funding from both public and private sources, these programs aim to improve the capture and curation of data for scientific research that is often hampered by inadequate standards and limited resources. By co-creating a network of pan-European partners, they are building ambitious data platforms and research networks that will be completed between 2022 and 2024. Flagship initiatives include:

- The European Health Data & Evidence Network (EHDEN) which is creating a federated research and data network for 100 million patients across Europe.
- PIONEER, which focuses on pancreatic cancer research in the UK.
- The Healthcare Alliance for Resourceful Medicine Offensive against Neoplasms in Hematology (HARMONY), which explores hematological malignancies.
- The ROADMAP project for Alzheimer's disease.
- The IMI's BigData@Heart project, which is creating a translational research platform for cardiovascular disease.

The European Medicines Agency (EMA) has also formed its own consortium—the Data Analysis and Real World Interrogation Network. DARWIN EU aims to establish a network of observational data sources that can be used to collect real-world evidence on the safety and efficacy of medicines to support regulatory approvals. Likewise, many academic and registry networks have already established their own consortia or they are in the process of doing so within the next decade.

As alliances take shape across Europe, it's important to note some of the challenges they face and the limitations of their data. For example, the data governance frameworks adopted by consortia-led initiatives are often designed to enable a narrow set of scientific or clinical research projects. As a result, the data they produce is probably best suited for real-world evidence studies and may not fit the modalities of a broader spectrum of use cases.

Hospital provider networks and EMR software providers break with tradition

Despite the risk to public perception, several sizable hospital provider networks have established either independent capabilities to conduct or facilitate real-world evidence studies or have entered partnerships with established data providers. Others have developed a hybrid approach where, in some cases, they have chosen not to release the details to the public.

While these efforts are already making a significant impact on the European data landscape, they are restricted by a narrowly defined data governance framework. For example, data use may only be permitted with explicit patient consent.

Often, these networks operate across legacy systems with business technology (BT) and EMR software provider strategies that are not unified. As a result, they will only cover some use cases, studies take a long time to set up and they are typically more expensive. These factors may limit the use of the data sets to advanced applications such as published real-world evidence, HEOR studies, clinical trial design or feasibility assessments. In the future, they may apply to synthetic or hybrid clinical trial execution as well.

Commercial analytics and data provider investments

In recent years, commercial vendors and providers have made significant advances in their data capabilities. Notable examples include TriNetX, Clinerion, Sensyne Health, and Graticule—all of which have established extensive networks of hospital partners to enable and accelerate a spectrum of use cases and capabilities. Some of these organizations are focused on conducting clinical trial feasibility assessments to pave the way to data use for publishable real-word evidence studies.

Elevated scale and volume

The European data landscape is changing with each innovation, regulation and initiative—and recent efforts across public and commercial entities are picking up steam. As the scale of these initiatives increases, the trajectory of Europe's data capabilities looks increasingly positive. The trend suggests that major improvements in the depth and granularity of new data assets are on the horizon. It is, however, unlikely that existing or novel data assets will offer improved scale and data volume—characteristics that are primarily a function of the underlying source of data and the regulatory environment.

As hospitals, EMR software providers and networks enter the data exchange with elevated capabilities, scientific and clinical research communities can hope for increased access to data for real-world evidence, HEOR studies and even real-world data-driven clinical trial execution. While this is a positive development, it remains to be seen whether the access and permissible use cases will extend to commercial and early-stage R&D applications. Without material changes in the range of applications, organizations may still struggle to generate scientific and commercial insights for key decision-making processes.

Hope for the future

For many years, the European data landscape has been characterized by limited, nuanced and complex data availability. But many organizations including data providers, hospital provider networks, EMR software providers, lab testing providers, data integrators, technology firms and others are making bold and significant investments to counteract these challenges.

As a result, a growing number of innovative solutions and capabilities are in the pipeline, any of which could propel the evolution of the industry's data strategies. Over the next few years, significant technological advances and a somewhat clearer regulatory environment are likely to accelerate the pace of progress. And as compelling data capabilities continue to emerge, the nuance and complexity of the European data landscape will likely increase.

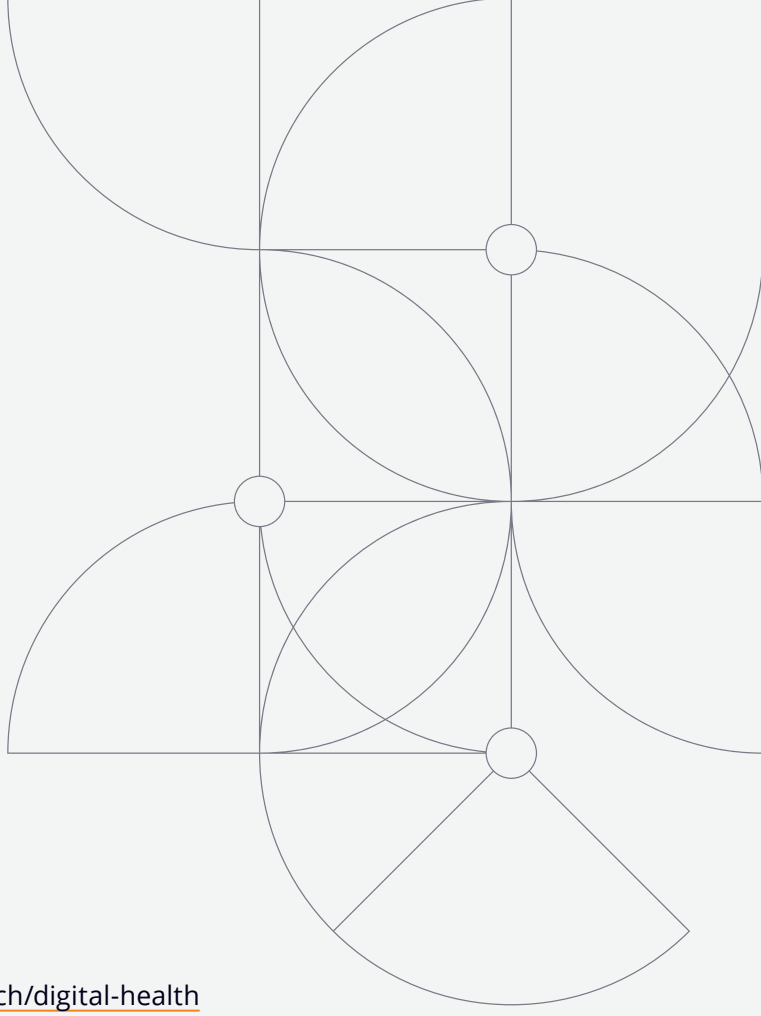
For life sciences companies, the challenge lies in keeping track of existing data assets, pushing for continued improvement and seeking novel capabilities. There are no quick fixes, but long-term solutions are taking shape. The plans and data sourcing strategies that are built tactically and intentionally today will continue to drive changes in an industry eager to transform.

About the authors

Florent Moise is the managing principal leading ZS's data strategy and partnerships team. With more than 19 years of life sciences consulting experience, Florent has helped many clients in key areas such as data strategy, governance, operating models, data evaluation, value creation and data management. Florent also serves as ZS's internal chief data officer (CDO), developing partnerships with data providers to enhance access and broaden data usage. Florent holds an MBA degree from the Wharton School of the University of Pennsylvania and a Master of Science degree in Aerospace Engineering from the Institut Supérieur de l'Aéronautique et de l'Espace in Toulouse, France.

Waldemar Ockert leads ZS's European data strategy and partnership team, where he helps clients navigate a broad range of data strategy issues across commercial areas, real-world evidence and R&D. He has more than 10 years of experience in analytics, data product development and data partnerships. Before joining ZS, Waldemar served as director for real-world data analytics at a global pharmaceutical company and developed the real-world evidence and analytics practice in Europe at Decision Resources Group (DRG). Waldemar holds a Ph.D. in neuroscience from the University of Manchester and a Bachelor of Science degree in Pharmacology from the University of Southampton.

Edith Aggarwal is a manager within ZS's data strategy and partnerships practice based in London. Wielding deep knowledge in the European and Latin American healthcare data ecosystems, she has helped clients identify new and innovative data solutions to overcome existing data challenges. Edith also assists clients with a range of data strategy issues, including data value measurement, portfolio optimization and operating models. Before joining ZS, Edith worked for a medical device company. She holds a Master of Arts degree from the Ludwig Maximilian University of Munich (LMU) and an MBA from Manchester Business School.



About ZS

ZS is a management consulting and technology firm focused on transforming global healthcare and beyond. We leverage our leading-edge analytics, plus the power of data, science and products, to help our clients make more intelligent decisions, deliver innovative solutions and improve outcomes for all. Founded in 1983, ZS has more than 12,000 employees in 35 offices worldwide.

Learn more: www.zs.com/industries/pharmaceuticals-biotech/digital-health