

MULTI-SOURCE HEALTHCARE DATA

How to embrace the new research reality

By Theano Anastasopoulou and Svetlana Gogolina | April 2020

IPSOS VIEWS

GAME CHANGERS





KEY TAKEAWAYS

1. Big Data and data science are increasingly disrupting established marketing and commercial practices in pharma organisations.
2. Analysing and integrating data from a variety of sources allows pharma companies to both better understand the multiple stakeholders they interact with, and to become more customer and patient-centric.
3. Primary research has a central role in achieving a successful integration of the multiple 'real life' sources that can be leveraged to fill in any gaps and add the human stories behind the data.

The Big Data revolution, along with the emergence of multi-source data strategies, powered by mighty computing algorithms and a proliferation of data sources, is transforming the pharma industry. The speed of change is staggering, and the changes are incredibly profound. A number of key areas including drug development, clinical trial design, health service delivery and diagnostics are undergoing a transformation. Responding to the COVID-19 outbreak at the beginning of 2020 is likely to speed up this transformation as the infrastructure is rapidly evolving to allow more activities to shift online, further accelerating the growth and availability of data.

This paper will explore how marketing insight functions in pharma are also adapting to a new era of digital transformation. We can see that teams are accessing new data sources and powerful analytic tools to provide marketing insight and answer key business questions.

Although data and analytics have always been at the heart of marketing insights, adapting to the new reality of Big Data and data science comes with unprecedented challenges, while it also offers exciting opportunities.

NAVIGATING THE HEALTHCARE INFORMATION ECOSYSTEM

The ability to access and analyse a wide variety of data sources creates great opportunities, but, as each of these sources comes with their own limitations and biases, multi-source insight generation requires a good knowledge of data sources.

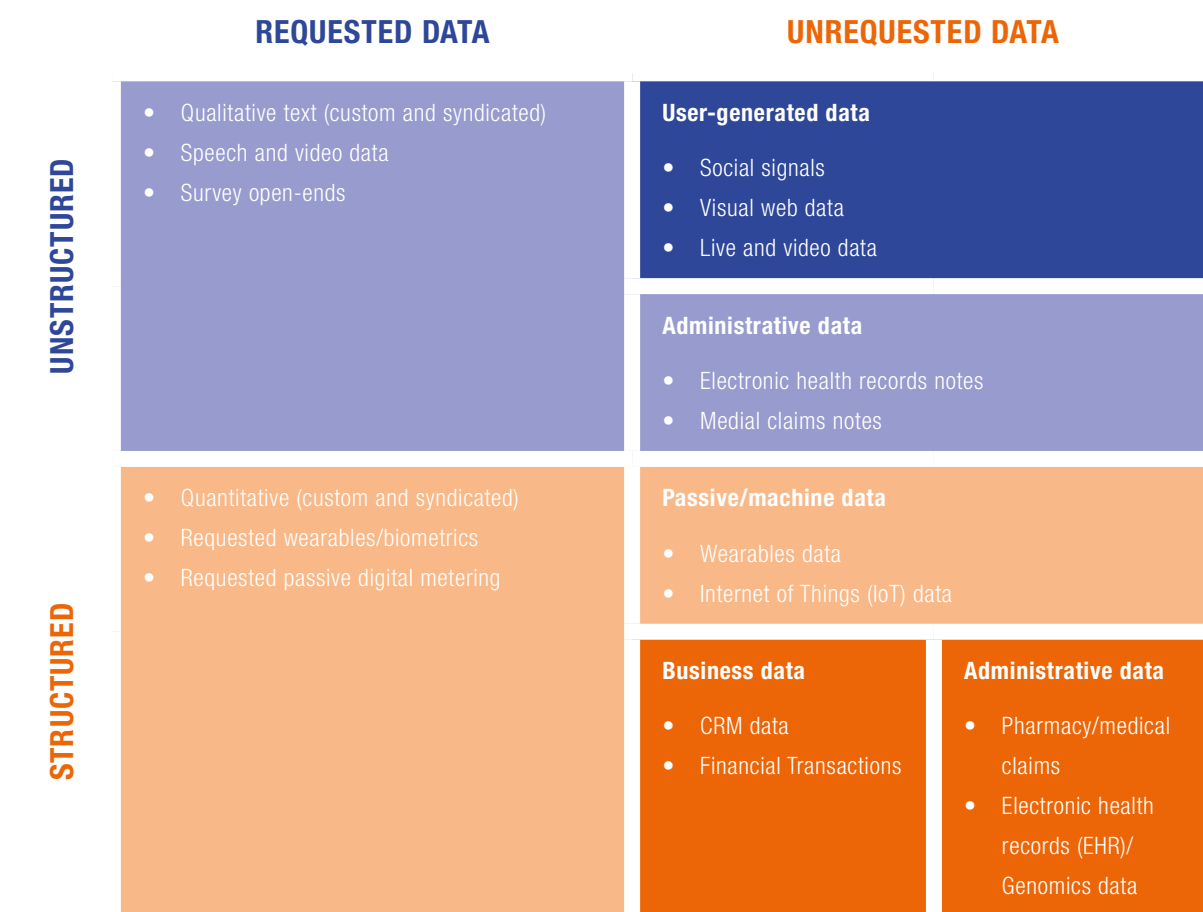
The table below summarises the healthcare information ecosystem as we see it, using two key dimensions: the original collection purpose of the data and the data type.

Requested data is collected to specifically address market insight business questions. Data collected for a different purpose, but repurposed for market insight generation, is called unrequested data.

Additionally, there are two broad types of data: structured (numeric or coded) and unstructured (free text, voice, images, maps, etc).

Requested data, both structured and unstructured (i.e. quantitative and qualitative primary research respectively), has historically been the main source for marketing insight generation, but this is now rapidly changing. Unrequested data, often held by other (non-marketing insight) teams in the organisation, such as Market Access, Medical, and Digital, is now being used together with traditional primary market research data.

Figure 1 The healthcare information ecosystem



THE DATA SCIENCE OPPORTUNITY

Data science was born out of the explosion of digital data sources, the availability of powerful computers and the development of Artificial Intelligence (AI) algorithms.

Data science techniques such as Natural Language Processing are used for the analysis of free text and speech. Machine Learning techniques and Deep Learning Networks are increasingly used for identifying patterns in complex data sets and for analysing medical images for diagnostic support.

These techniques also allow marketing insight teams to utilise the wide variety of data sources and types in the healthcare data ecosystem and add new metrics and key performance indicators (KPIs) to their reporting. The vision of real-time insights via sophisticated platforms, delivery systems, real-time data collection and analytics is now possible.

Business questions along the product lifecycle can be addressed by combining and triangulating a variety of different data sources. We look at some examples in the next sections.

TRANSACTIONAL OR ADMINISTRATIVE DATA

Data science has been used for many years across industries to offer insight based on transactional data bases, such as e-commerce payments, banking transactions, insurance claims, supermarket purchases, and so on.

In the case of medical insurance, every time a patient makes a claim, details about their condition, tests, procedures, treatments or specialist visits are recorded, evaluated, approved and finally processed electronically. All this information is analysed by insurance companies to optimise treatment pathways and improve outcomes, while also minimising the number of future claims. It is also increasingly used by pharma companies to obtain evidence on treatment safety and real-world outcomes for their products.

The world of pharma business insight is likely to change as it embraces a multi-data source reality.

Where possible, the medical claims data can be further enriched by integrating medical claims with pharmacy claims (Rx) or patient treatment data from electronic health records or primary patient record forms (PRFs) – also described as patient charts. This integration can provide valuable KPIs for brand tracking and also information on the position of a product within a treatment pathway and real-life patient treatment journeys.

Unrequested administrative sources inevitably come with challenges of limited and patchy coverage. The data is sometimes of suboptimal accuracy, with gaps and missing information that may limit its value. In certain cases, identified gaps can be filled by data science techniques: for example, machine learning algorithms can be called upon to complete missing information. Integrating or cross-referencing multiple sources could also be used to fill gaps. This inevitably increases the processing time required and the cost of accurate and defensible insight from real-world sources. In some markets and some complex therapy areas, compatibility and consistency of data sources is still a big challenge.

While lacking the scale and breadth of the unrequested data sources, the well-established primary sources based on PRFs could offer the quality, consistency and depth when it is missing from the transactional sources - and in a more cost-efficient way.

THE IMPORTANCE OF THE HUMAN PERSPECTIVE

Real-world evidence (RWE) data sets do not provide the understanding of the personal element: the 'human stories' behind the numbers. A wide range of tried and tested primary qualitative and quantitative research techniques remain the best way of understanding the perceptions, needs, motivations and emotions of the key stakeholders. These techniques are the essential complement of the insights generated from unrequested big data as they offer an understanding of the 'why' in addition to the understanding of the 'what'.

USING SOCIAL INTELLIGENCE ANALYTICS

Another source of unrequested unstructured data, allowing us to draw valuable insights, are the rich patient conversations on social media platforms, including some public online patient communities and forums.

Historically, we have accessed publicly available sources of patient conversations and distilled the main insights via techniques traditionally used in qualitative primary market research. This, of course, can only be done on a small subset

of social data. Social Intelligence Analytics (SIA), based on data science and powerful AI algorithms, now allows the analysis of increasing volumes of social media discussions quantitatively, by automatically coding the data and analysing patterns, themes and topics within online conversations.

SIA allows access to increasing volumes of data on patient experiences, challenges and everyday needs. Although these conversations naturally reflect the views of the patients that are active online, there is no interviewer and no questionnaire design bias: the experiences are described in the patients' own words. SIA enables us to uncover unexpected insights; the 'unknown unknowns', as opposed to 'known unknowns', which well-designed primary research traditionally addressed.

Like all other unrequested data sources, social listening data is likely to provide wide and rich coverage but could also be biased and patchy when looking at specific therapy areas or specific patient touch points. When there are clear gaps in specific areas of interest, primary market research that is specifically designed to fill in the gaps of the social listening outputs is invaluable. We may need limited scale and focused qualitative research, but also representative quantitative research depending on the business objectives.



THE 'DESCRIPTIVE, PREDICTIVE AND PRESCRIPTIVE' ANALYTICS PATHWAY

Data science offers exciting opportunities to extract insights from big and otherwise unmanageable data sources. It also helps businesses to optimise their performance with the use of predictive and prescriptive algorithms.

A HISTORY OF PREDICTIVE INTELLIGENCE

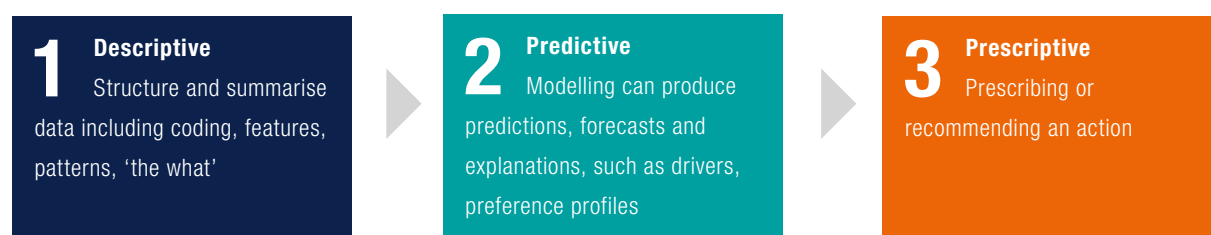
Outside pharma, extensive usage of predictive intelligence for targeting optimisation according to customer needs started some years ago, first being adopted by large e-commerce companies. Organisations like Amazon started to analyse individual purchase journeys, using all available data from online purchases, advertising exposure, product reviews and past purchase behaviours.

To start with, companies were able to see which were the most clicked-on advertisements, most viewed products and most successful product placements in a recommendation list. We call this a 'descriptive analytics' stage; the first step that is used to identify performance measures that may be useful in predicting future customer behaviour.

Predicting the future behaviour of each individual consumer based on their past behaviour, clicks and views is described as 'predictive analytics', and is at the heart of an important data science application.

Data science algorithms also identify the right actions and interventions at different touchpoints. This is often referred to as 'prescriptive analytics'. It allows companies to identify in advance the most likely outcome for each individual customer based on their past transactions and online behaviour, enabling them to create a targeted intervention at each touchpoint, such as suggested offers and advertising – both personalised and targeted. Such techniques, although sometimes seen as intrusive by consumers, are proven to deliver commercial gains.

Figure 2 Analytics: three-stage pathway



This requires an integrated set of data that encompasses all sources, leading to a better understanding on which to base commercial decisions.

BUILDING CUSTOMER RELATIONSHIPS IN PHARMA

Predictive and prescriptive analytics could be successfully applied to Customer Relationship Management (CRM) systems data and can be used to improve the effectiveness of digital communication campaigns.

Pharma companies, such as Novartis¹ and MSD², have been sharing with the industry how they started applying predictive analytics to their CRM system data to achieve personalised targeting and greater return of investment for their communication campaigns.

The first step of descriptive analytics measures communication campaign performance using digital behavioural metrics. The next steps of developing predictive and prescriptive models will require a deep understanding of customer needs, motivations and preferences. Without this understanding companies cannot deliver 'right time, right place' communications that attract attention and influence behaviours.

Achieving this level of data integration and customer understanding in pharma will require a major shift in customer relationship management and digital data capture and integration.



The next steps of developing predictive and prescriptive models will require a deep understanding of customer needs, motivations and preferences. Without this understanding companies cannot deliver 'right time, right place' communications.

PUTTING THE CUSTOMER AT THE HEART OF COMMERCIAL DECISIONS

Ideally, to develop targeted interventions, companies should have every piece of information, including both historical and current data on behaviours, attitudes, communication preferences and touchpoints for each customer in a CRM system. In reality, most pharma companies have limited information on each individual customer, so it becomes even more important to make the best use of all the information that is available.

Primary research and the right analytics of all the available data sources can help companies move closer to the predictive and prescriptive steps within the real-life constraints of healthcare. A healthcare professionals' (HCP) customer segmentation based on primary research can measure behaviours, attitudes, product perceptions, media preferences and initial reactions to alternative

communication material. This can offer clear direction on the customer segments or 'personas' to focus on and their unique communication needs.

A customer segmentation can be developed from survey data of HCPs recruited from primary market research panels. Integration techniques, such as data fusion, can then be used to get an estimate of the most likely communication preferences of similar physicians from pharma companies' CRM systems. This is done by using common characteristics between the HCPs from the primary research sample and those on the CRM system to predict the likely match.

Follow-up qualitative research can then be the ideal way to enrich the segments and develop targeted communications.



LOOKING TO THE FUTURE

Big Data and data science will continue to disrupt marketing insights. The recent need to quickly understand and adapt to the COVID-19 pandemic has highlighted the value of immediate access to reliable data and powerful analytics. A variety of sources including real-time healthcare data, HCP and patient behavioural data, social media conversations, video ethnographies, short regular survey pulses and in-depth interviews are integrated to provide a complete understanding of the impact of a major crisis as it unfolds. This in turn has exposed more marketing teams to new data sources and analytics, and further highlighted that as powerful and as helpful as data science and Big Data is, it is unable to address all pharma industry information needs. We believe there remains a firm need for primary research, as the optimal way to identify and understand the real person behind the data.

The world of pharma business insight is likely to change fast as it embraces a multi-data source reality. Primary research projects are likely to become shorter, more agile and focused, and will be used in combination with other sources more often.

In addition, the unconventional combination of data science and qualitative research will be key for marketing insight professionals of the future, with qualitative research likely to become the best friend of a data scientist, and vice versa.

For more information on Ipsos solutions on data science and data integration advisory please contact **Theano.Anastasopoulou@Ipsos.com**



REFERENCES

1. Novartis puts AI on the job to help reps say the right things to the right doctors, fiercepharma, January 9, 2019
<https://www.fiercepharma.com/pharma/novartis-equips-sales-team-ai-bid-to-boost-productivity-exec>
2. Step Aside, Manual Marketing. Automation Is Coming, eyeforpharma https://eloqua.eyeforpharma.com/LP=22867?utm_campaign=4913-06DEC18-WK14-TA%20%28Digital%29&utm_medium=email&utm_

MULTI-SOURCE HEALTHCARE DATA

Theano Anastasopoulou Director, Data Advisory, Ipsos, UK

Svetlana Gogolina Head of Data Science and Advanced Analytics, Ipsos, UK

The **Ipsos Views** papers
are produced by the
Ipsos Knowledge Centre.

www.ipsos.com
@Ipsos