by Robin Robinson

Big data gets all the hype, but the power of small data cannot be overlooked.

The Big Deal About **Small Data**

Small data is mighty. Small data is small enough for human comprehension, in a volume and format that are accessible, informative, and actionable. Certainly, big data has gotten a lot of hype over the past few years, but savvy marketers know the power that lies in using small data to determine trends.

Compared with big data, small data can seem to be a collection of insignificant observations, however, these observations can be the key to understanding the consumer at a human level, and help pinpoint specific information that leads to identifying an unmet consumer need.

Small data is personal, and so is the patient journey, so small data is a well-matched tool for providing pharma marketers a true understanding of patient needs and behaviors.
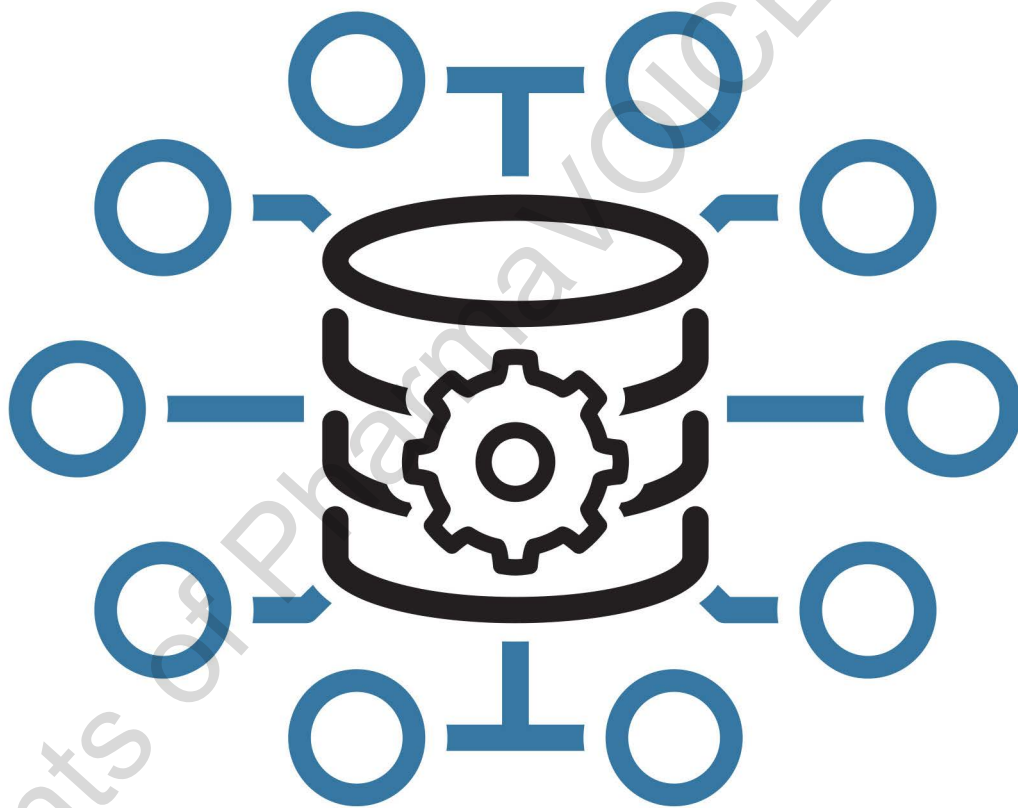
Martin Lindstrom, brand-building expert, author, and small data advocate, believes there is too much focus on big data today at the cost of losing the perspectives gained from small data. In several presentations across the nation he stresses how important small data — or clue gathering, as he calls it — is to providing accurate insights into people's lives. Small data can tell a story and create a picture of a consumer that provides guidance for solving consumer needs. In his book, Mr. Lindstrom says at least 60% to 65% of the top 100 biggest innovations in our time came from analyzing small data.

"Everything we do in life, consciously or unconsciously, reveals incredible insights of who we really are," he says. "Small data tells a story, which marketers can use to improve their efforts to better the patient experience and make the journey more pleasant, effective, and valuable for the patient."

It is the size of small data that makes it valuable in many marketing situations. It is easy to access, digest, and get to the insights needed quickly. While big data requires technology and algorithms and computing power, small data requires only a brain and an Excel spreadsheet.

"The small data set can be analyzed using spreadsheet tools such as Excel or other business intelligence tools to slice and dice the data," says Sriram Krishnan, Ph.D., executive VP of AI solutions at ProMetrics. "Small data provides a human user with a simple way to explore the data and understand the message. The human brain is incredible at identifying trends in data, and in the hands of an expert in the field, subtle insights can be gained simply by analyzing and presenting structured data in an intuitive way."

"One of small data's greatest assets is how intentional it often is," says Gaurav Kapoor, executive VP of Indegene. "It's not a find-a-needle-in-the-haystack exercise. It's often very intentional, specific, and on point. It's very purpose-built, in this way."

A small data set lends itself to clearly defined data, variables, hypothesis, understandable analysis, and action, and these traits can translate into speed, efficiency, and impact, says Iyiola Obayomi, marketing analytics, practice lead, Ogilvy Health. "Speed is increased because it takes less time to access and collate the information needed for analysis; efficiency grows because smaller datasets are usually cheaper to acquire and take fewer resources to store, manage, and analyze; and lastly, small data has value because the outputs and inputs are easier to comprehend, show, explain, share, and adopt, which potentially translates into better business impact," Mr. Obayomi explains.

"Small data allows for greater focus," says Diane Hayes, Ph.D., president and co-founder, InCrowd. "With small data, researchers can zero in on a specific learning, identify questions they may still have, and go back and ask these questions to refine the insight in an iterative manner. Small data is critical in creating segments that allow for focused messaging, data collection, and subsequent insights. It's a much more targeted effort that enables an intervention for that specific group being analyzed. Big data doesn't lend itself so well to focus, iteration, and informed intervention."

Despite all the upsides, experts say pharma marketers are not using small data to its fullest potential. Some may be distracted by the next new thing — big data — or entrenched in yesterday's traditional marketing strategies. Though traditional methods may appear to have cost efficiencies and are long proven,

> A small data set can be analyzed using spreadsheet tools such as Excel or other business intelligence tools to slice and dice the data.
>
> **DR. SRIRAM KRISHNAN**
> ProMetrics

lack of sufficient exploration in, and adoption with, newer methods limits pharma market researchers' ability to gain many and as deep as possible insights.

"At this point in time, small data is not being used to its full potential in pharma marketing — nor is big data, really," Dr. Hayes says. "Traditional market research methods in pharma marketing, which comprise the majority of commercialization research still rule the day."

Dr. Hayes says for the best success, it is necessary to combine the results from both big and small data for a complete view of consumer insights. Traditional methods are neither computationally analyzable on a large scale, nor accelerated and focused on a small scale. Though many pharma marketers apply big and small data as peripheral sources of data, few embrace a truly blended approach to capturing the strongest and most meaningful insights for their brands.

Mr. Obayomi says he can understand why pharma marketers might overlook the benefits of small data. "There is a lot of noise from the data community and marketers fear missing out on the next promising competitive advantage," he says. "Some marketers are still grappling with the basics of performance analytics and modeling based on small data or with identifying the right KPIs for the programs and getting comfortable with forecasts and projections. Yet these marketers want to talk about AI and ML and how these are the next frontier. The low-hanging fruits of small data often fall to the ground and are even trampled on in the search for the next titillating fruits."

While new avenues need to be explored and the potential of big data can't be ignored, the focus on small data should continue to be pragmatically prioritized as it addresses, informs, and provides practical actionable decisions for marketing effectiveness in most situations, Mr. Obayomi says.

According to Dr. Krishnan, big and small data should be used to complement each other. The decision is rarely to use one over the other. Sometimes, the questions can be answered simply through someone's knowledge and

insight, but the question of why is one that may require more data. To determine a viable answer to why, it may take deeper and bigger data. "In this sense, the insights gained from small data may be enough to drive a business decision, or it may drive a larger project leveraging big data, with more data sources to mine. For example, to understand a geographic gap, CRM data may be combined with other purchased or publicly available data to try and identify root causes in geographic gaps to better direct solutions, if simply identifying the gaps is not sufficient."

Tim Hare, VP, head of data science, 81qd, calls this a hybrid approach. "Where appropriate, human experts can form testable hypotheses and glean insights from small data directly, while coupling this with different types of machine learning," he says. "Such hybrid human-machine approaches blend rule-based machine learning, clinical expert review of candidate rules, hypotheses formulation, small sample statistical testing, and literature research on the disease of interest."

Mr. Kapoor says he also believes the industry is not using either data approach to their full potential. "There is always additional potential for how we use data, however, small data has been demonstrated to be of huge value in many cases," he says.

For example, simple A/B testing of site content has yielded enormous insights into how brand managers should message their audiences. "There are plenty of tools to do this, even for small audiences coming to some pharma sites," he says.

> Small data has value because the outputs and inputs are easier to comprehend, show, explain, share, and adopt, which potentially translates into better business impact.
>
> **IYIOLA OBAYOMI**
> Ogilvy Health

Mr. Kapoor has also seen success in creating cost savings by using small data to track the days and dollars consumed along the content development lifecycle. "By managing the internal review processes (MLR) and de-coupling collateralization of content from agencies, we are able to drive significant time and cost savings," he says.

Big vs. Small

While both big and small data are necessary to determine appropriate and effective marketing decisions, there are situations where small data can provide the answers more quickly.

"Small data works better than big data in situations where the desired questions or hypothesis are clearly defined, the insight needed pertain to a large audience, and data are available only for a small group of the population,"

> Small data is a much more targeted effort that enables an intervention for the specific group being analyzed. Big data doesn't lend itself so well to focus, iteration, and informed intervention.
>
> **DR. DIANE HAYES**
> InCrowd

Mr. Obayomi says. "An example is where a marketer seeks to understand the preference of intended audience for a new offer."

With the clear hypothesis or specific insights, marketers can design the structure of the analysis, identify the data and variables that are needed, proceed to analyze and explore the hypothesis, and understand and explain the results to the audience in intuitively understandable terms.

Small data is preferred in such situations because more is not always necessarily better, Mr. Obayomi says. Statistical inference from a representative sample of the population can provide the needed insights efficiently. The variables used in the analysis are understandable and the findings can be explained and generalized.

Dr. Hayes notes that she defines big data as large data sets that may include both primary and secondary data and are intended for computational analysis designed to hypothesize themes regarding respondent choice, behavior, and trends. On the other hand, InCrowd defines small data as small, timely, tech-enabled primary data sets. "Though they may have

large sample sizes, small-data efforts focus on discrete problems, subjects, or discovery," she says. "As such, we regard big data as good for understanding general behavior themes for specific topics."

Dr. Hayes adds that small data is needed to get to the meaning of big data.

"Being able to get to the why of behavior and understand unique segments within a big data set requires small data investigation," Dr. Hayes says. "A focused, iterative, small-data approach to unique behavioral categories identified through big data is an optimal way to leverage the benefits of each type of data set."

A situation where this becomes relevant is tracking how consumers change their behavior when a new product comes on the market. Marketers will want to know if physicians will use the same product or if they switch to a newer product, and what motivated them to make a change, for example.

Other parameters to be uncovered may include whether safety or efficacy are the only considerations in the physician's choice, or do cost, delivery mechanism, frequency of use, or other variables influence a physician's selection of treatment for a particular disorder.

"Small data provides great insight into these decision points as physicians seek to best manage each patient," Dr. Hayes says.

Challenges of Small Data

As useful as small data is, there are several challenges in using it to generate actionable insights. The first is, because of its size, that data quality

must absolutely be preserved. Unlike big data, where the size of the data sets may enable insights even with slightly "dirty" or incomplete data, analysis of small data is not as resilient to errors, Dr. Krishnan says.

While data quality is always a consideration in any kind of analysis, the acceptance rate of errors using small data is smaller to ensure proper analysis.

The second challenge of using small data is that analysis is often dependent not just on the data, but the presentation of the data. Often, humans are exploring and identifying results, and so proper tools must be in place not just to slice and dice the data, but also visualize the data in a meaningful way.

Finally, while small data can help identify trends, the lack of the variety of data often prevents answering any questions of why something is happening, Dr. Krishnan advises. "Trends and alerts can be identified, but understanding underlying root causes becomes much more difficult, since the data available is often not sufficient."

According to Mr. Obayomi, small data can address 70% to 80% of marketing needs, however, there are three limitations to this rule.

Marketers could run into data limitations when seeking deeper understanding of subgroups when using only small data. A well-defined analysis scope will ensure key questions are addressed, but it may be difficult to address questions that emerge during the analysis due to sample size limitations. Second, small data does not lend itself easily to extensive computational exploration to identify potentially new insights or perspectives.

And third, small data is not useful for training datasets. Seeking a training dataset for the increasingly appealing exploration of machine learning or AI requires many datasets thrown at the systems for self-exploration, self-learning, and new algorithm development, Mr. Obayomi says.

Big data analysis is better to use for broader behavioral themes and market trends around a subject, because even the same result can be achieved by small data efforts it's not ideal, Dr. Hayes says. "The nimbleness of small data is less critical for big data mass analysis," she says. "Small data helps dig in to an issue or specific actions of a cohort within the big data set, one that requires additional investigation, consideration of the respondents, and nuance."

Mr. Kapoor adds two more challenges to the list. First, detailed follow-on questions are often not able to be answered by small data, and second, multiple small data sets can often be tied together only by a time dimension. "That is, they are often not granular enough for detailed, integrated analytics," he says. "Therefore, trade-offs rule the day." PV

> Human experts can form testable hypotheses and glean insights from small data directly, while coupling this with different types of machine learning.
>
> **TIM HARE**
> 81qd
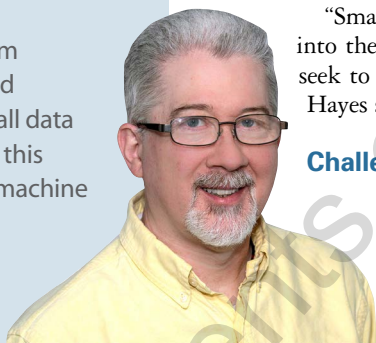
> One of small data's greatest assets is how intentional it is; it's not a find-a-needle-in-the-haystack exercise.
>
> **GAURAV KAPOOR**
> Indegene

Be brilliant.

Be here.

PharmaVOICE

THE FORUM FOR THE INDUSTRY EXECUTIVE