

Growth in Health Care Data Causing an Evolution in the Pharmaceutical Industry

J. Alan Menius Jr, Matthew D. Rousculp

A health care ecosystem is evolving in which all stakeholders will need to work together, apply new technologies, and use disparate data sources to gain insights, increase efficiencies, and improve patient outcomes. The pharmaceutical industry is leveraging its experience and analytics capabilities to play an important role in this evolution.

For more than 60 years, the pharmaceutical industry has focused on research into data-driven care—a process that is fundamental in drug development and has seemingly remained unchanged [1]. Scientists within industry continue to work alongside those in academia, government, and clinical practice to study factors that play important roles in biological function, disease progression, and the mechanisms of action of potential medical interventions. This work has produced terabytes of information about genetics and molecular biology to further our understanding of diseases and to identify and validate new biological targets for the next generation of medicines [2].

Results from randomized controlled clinical trials remain the primary source of information that regulators use to decide whether and under what circumstances to make a new medicine available to patients. To support the approval of new medicines, the pharmaceutical industry has conducted thousands of clinical studies, resulting in data that can be reanalyzed to compare the outcomes of different treatments or to identify patient subgroups with varying efficacy or safety profiles. By design, clinical trials are conducted according to strict protocols, so that a medicine's safety and efficacy can be determined. However, these study designs are not sufficient to answer all of the questions raised by clinicians or population-based decision makers (eg, formulary committees). Because of this necessity, the pharmaceutical industry is now investing in disparate real-world health care data sources, such as administrative claims, electronic health records, pragmatic trials, Internet patient portals, and registries.

The future promises to provide additional sources of data that can inform both pharmaceutical research and clinical care. Initiatives that have arisen from the Patient Protection and Affordable Care Act—which involve changes in health care delivery, changes in payment models, and a focus on quality—promise to provide a continuous source of big data

relevant for understanding appropriate patient care. Over the past 3 years, the number of physicians using electronic medical records has grown substantially, with nearly three-quarters of office-based physicians using electronic records to support some form of patient care [3]. New technologies applicable to pharmaceutical research are also being tested, including wearable devices for patients, which allow patients to have less frequent study visits while providing researchers the ability to monitor patients in real time.

Efforts to Support Patient Care

While tremendous advances continue to be made in the process of drug discovery and development, the pharmaceutical industry is also focusing on using its data and analytics experience to support patient care in the changing US health care environment. For example, GlaxoSmithKline announced a collaborative analytics arrangement with Community Care of North Carolina (CCNC) in late 2013. The goal of this arrangement is to combine the strengths of these 2 organizations in order to allow health care providers to take a real-time approach to identifying potential barriers to the delivery of high-quality care. This combined initiative takes a “small data” approach, meaning that it uses health information analytics and patient-centered data to develop a predictive model that can identify patients who are at high risk of future negative outcomes, such as rehospitalization [4].

In 2013 GlaxoSmithKline also announced a data transparency program that allows greater access to deidentified patient-level data from clinical trials [5]. The goal of this program is to provide significant opportunities for additional research in the hope that giving researchers and clinicians access to patient-level clinical trial data will allow them to replicate, challenge, or combine data in ways that will facilitate scientific progress and improved patient care. Greater transparency will help ensure that the data provided by clinical research participants are used to maximum effect in the

Electronically published May 5, 2014.

Address correspondence to Mr. J. Alan Menius Jr, GlaxoSmithKline Research and Development, 5 Moore Dr, Research Triangle Park, NC 27709 (alan.j.menius@gsk.com).

N C Med J. 2014;75(3):188-190. ©2014 by the North Carolina Institute of Medicine and The Duke Endowment. All rights reserved.
0029-2559/2014/75307

creation of knowledge and understanding.

In an additional effort, GlaxoSmithKline has committed its analytics resources as part of the Observational Medical Outcomes Partnership (OMOP). OMOP is an interdisciplinary research group that is tackling the fundamental analytical task of identifying the most reliable methods for analyzing big data drawn from heterogeneous sources, such as electronic health record systems and insurance claims. OMOP was founded in 2007 by the Foundation for the National Institutes of Health as a public-private partnership and has since transitioned to the Reagan-Udall Foundation for the Food and Drug Administration [6].

Individually these endeavors may seem unconnected, but they share a common thread—the need to adapt to the changing US health care environment. Because the pharmaceutical industry has expertise in medical analytics, a focus on protecting patient-level data, and a need to better understand the true effectiveness of treatments, the industry is working in many areas to ensure that data are accessible and that appropriate analytics are available to realize the full potential of these data.

Best Practices in Data-Driven Care

The experience of conducting hundreds of studies using real-world health care data has presented challenges, but it has also provided us with several best practices. Although the following list is not exhaustive, we believe these are some important areas of concern and best practices for those interested in data-driven care:

Involve analytics experts and guard against misinformation. Because data have varying degrees of robustness, analytics experts should be involved to prevent the production and communication of misinformation. In the absence of analysts with appropriate quantitative expertise, researchers who analyze real-world data may be fooled by false but seemingly plausible findings. Even when trained analytics experts are engaged in this research, there is a danger that findings may not be reproducible. Thus the production of reproducible results will be an extremely important way of guarding against the dissemination of misinformation [7].

Tune the analysis system. It is very important to empirically test assumptions regarding databases, study design, and analysis methods together—as components of an analysis system—in order to fully understand the inherent biases that may impact results when studies use large observational databases. Initiatives such as OMOP and efforts funded by the Patient Centered Outcomes Research Institute (PCORI) and the Office of the National Coordinator of Health System Technology (ONC) demonstrate that further research is required to improve the accuracy of studies that rely on observational data [8].

Decide who owns health care data. Our society is still trying to answer questions regarding how we should use health information and who should have access to it. This cultural debate is readily apparent as patients, physicians, and gov-

ernment agencies consider who owns a patient's health care history and what ethics may come into play with regard to use of that data. The recent decision by the US Food and Drug Administration that required the genetic testing company 23andMe to temporarily halt the sales of its genetics kits to patients [9] is an example of this struggle.

Ensure that patients and key stakeholders will benefit. To ensure that high-quality data will be captured, systems should ensure that whoever is providing the data will benefit from making it available. For example, patients who provide data describing their condition should receive information that helps them to better understand or manage their health. Similarly, health information systems should reward health care professionals for inputting high-quality data by instantly giving them information that will help guide their decision making, making their job easier or more efficient. Companies such as PatientsLikeMe have pioneered this approach to collecting data, and they are now using those data to achieve new insights about diseases such as amyotrophic lateral sclerosis [10].

Make health care data available. While randomized clinical trials will continue to be an important tool for conducting research, if we are to reach our goal of optimizing real-world outcomes for all patients, then data that capture the health care journey of those patients must be made available for research. Only by analyzing those combined data will we then be able to accurately make patient-specific predictions. In order to ensure the relevance of our work, we must create an environment of data transparency and data access, an expectation of free communication of study results, and a commitment to increasing patients' involvement in research. To overcome these challenges and reach our goals, we must foster a health care ecosystem in which all stakeholders work together to find ways of practically and responsibly using technologies and information to drive new efficiencies, to improve patients' understanding of their current health status, and ultimately to improve health outcomes.

The concept of getting all stakeholders involved is not new. In fact, PCORI's primary research agenda is geared toward creating patient-centered systems. Patient-centered research networks—collaborations among digitally enabled health care systems, academics, and patient groups—are being fostered by federal investments that are intended to facilitate more timely and robust patient-centered outcomes research. Examples of such networks include PCORI's clinical data research networks; the National Institutes of Health (NIH) Health Care Systems Research Collaboratory; and the Electronic Data Methods Forum for Comparative Effectiveness Research, which is an initiative of the Agency for Healthcare Research and Quality. In the not-too-distant future, we may see "machine learning systems that will allow clinicians to match a patient's electronic health record directly to RCT [randomized controlled trial] and observational study data sets for better, individualized therapeutic

decisions" [11]. Through this lens, we seem tantalizingly close to the goal of the Institute of Medicine of the National Academies, which is to create a learning health care system in which "clinical decisions will be supported by accurate, timely, and up-to-date clinical information and will reflect the best available evidence" [12].

Make health care data relevant. In addition, as the United States and other countries move to value-based systems for delivering health care, the pharmaceutical industry is adapting by seeking to better understand the kind of research that patients, clinicians, payers, and policymakers will need to make more informed decisions. The collaboration between GlaxoSmithKline and CCNC leverages CCNC's considerable experience in providing analytically based community care. CCNC's strengths are in developing real-time interventions that guide patient engagement and can be accessed by a range of health care providers in different settings; such interventions are particularly useful at key transition points in care, such as when a patient is discharged from the hospital. GlaxoSmithKline is contributing its expertise in comprehensive medication management and its knowledge of data analytics and information technology systems. This project will provide the appropriate health care providers with an analytic tool and suggested interventions that are likely to facilitate better patient engagement and improved patient outcomes. The goal of the project is to assist these health care professionals in recognizing and responding to medication-related problems that often lead to poor adherence with the prescribed treatment regimen.

The desire to better understand how to provide appropriate health care to the right patient at the right time drives innovative research design. GlaxoSmithKline's initiatives and similar work in academia, government, and industry are part of the broader effort to create a learning health care environment. When making high-quality data available for use, it is important to protect patient privacy and to ensure ethical behavior. We must therefore work together to find ways to use these data while also guarding against the promotion of misinformation. Ultimately the key to getting better data will be to get patients involved and to make sure that they are the final beneficiaries of these efforts. *NCMJ*

J. Alan Menius Jr, MHS senior director, Quantitative Sciences, GlaxoSmithKline Research and Development, Research Triangle Park, North Carolina.

Matthew D. Rousculp, PhD, MPH senior director, Comparative Effectiveness Research and Health, GlaxoSmithKline Research and Development, Research Triangle Park, North Carolina.

Acknowledgments

Potential conflicts of interest. J.A.M. and M.D.R. are employees of GlaxoSmithKline.

References

1. Kaitin KI. Deconstructing the drug development process: the new face of innovation. *Clin Pharmacol Ther*. 2010;87(3):356-361.
2. Jørgensen JT. A challenging drug development process in the era of personalized medicine. *Drug Discov Today*. 2011;16(19-20):891-897.
3. King J, Patel V, Furukawa MF. Physician adoption of electronic health record technology to meet meaningful use objectives: 2009-2012. *ONC Data Brief no. 7*. Washington, DC: Office of the National Coordinator for Health Information Technology (ONC); 2012. <http://www.healthit.gov/sites/default/files/nc-onc-data-brief-7-december-2012.pdf>. Accessed March 2, 2014.
4. GlaxoSmithKline, Community Care of North Carolina harness "small data" to improve patient care and reduce health care costs [press release]. Community Care of North Carolina Web site. <https://www.communitycarenc.org/media/files/harnessing-small-data-improve-patient-care-and-reduce-costs.pdf>. September 18, 2013. Accessed February 4, 2014.
5. Nisen P, Rockhold F. Access to patient-level data from GlaxoSmithKline clinical trials. *N Engl J Med*. 2013;369(5):475-478.
6. Reagan-Udall Foundation for the Food and Drug Administration. Observational Medical Outcomes Partnership (OMOP). OMOP Web site. <http://omop.org/>. Accessed February 4, 2014.
7. Ioannidis JP. Why most published research findings are false. *PLoS Med*. 2005;2(8):e124. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1182327/>. Published August 30, 2005. Accessed March 3, 2014.
8. Madigan D, Ryan PB, Schuemie M, et al. Evaluating the impact of database heterogeneity on observational study results. *Am J Epidemiol*. 2013;178(4):645-651.
9. Annas GJ, Elias S. 23andMe and the FDA. *N Engl J Med*. [published online ahead of print February 12, 2014]. doi:10.1056/NEJMp1316367. <http://www.nejm.org/doi/full/10.1056/NEJMp1316367>. Accessed March 3, 2014.
10. Wicks P, Vaughan TE, Massagli MP, Heywood J. Accelerated clinical discovery using self-reported patient data collected online and a patient-matching algorithm. *Nature Biotechnol*. 2011;29(5):411-414.
11. Eichler HG, Abadie E, Breckenridge A, Leufkens H, Rasi G. Open clinical trial data for all? A view from regulators. *PLoS Med*. 2012;9(4):e1001202. doi:10.1371/journal.pmed.1001201. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3323505/>. Published April 10, 2012. Accessed March 3, 2014.
12. Olsen L, Grossman C, McGinnis JM. Learning What Works: Infrastructure Required for Comparative Effectiveness Research: Workshop Summary. The Learning Healthcare System Series. Washington, DC: Institute of Medicine; 2011.