JPMSA

JOURNAL OF THE PHARMACEUTICAL MANAGEMENT SCIENCE ASSOCIATION

SPRING 2017 | IN THIS ISSUE:

pmsa

PHARMACEUTICAL MANAGEMENT SCIENCE ASSOCIATION

Table of Contents

Official Publication of the Pharmaceutical Management Science Association (PMSA)

The mission of the Pharmaceutical Management Science Association not-for-profit organization is to efficiently meet society's pharmaceutical needs through the use of management science.

The key points in achieving this mission are:

- Raise awareness and promote use of Management Science in the pharmaceutical industry

- Foster sharing of ideas, challenges, and learning to increase overall level of knowledge and skill in this area

- Provide training opportunities to ensure continual growth within Pharmaceutical Management Science

- Encourage interaction and networking among peers in this area

Please submit correspondence to:
Pharmaceutical Management Science Association
446 E High Street, Suite 10
Lexington, KY 40507
info@pmsa.net
(877) 279-3422

Executive Director:
Stephanie Czuhajewski, CAE
sczuhajewski@pmsa.net

PMSA Journal: Spotlighting Analytics Research

Welcome to the fifth edition of the *Journal of the Pharmaceutical Management Science Association (PMSA)*, the official research publication of PMSA.

The purpose of the Journal is to promote and embody the mission of the association, by:

- Raising awareness and promoting the use of Management Science in the pharmaceutical industry

- Fostering the sharing of ideas, challenges, and learning to increase the overall level of knowledge and skill in this area.

The Journal publishes manuscripts that advance knowledge across a wide range of practical issues in the application of analytic techniques to solve Pharmaceutical Management Science problems, and that support the professional growth of PMSA members. Articles cover a wide range of peer-reviewed practice papers, research articles and professional briefings written by industry experts and academics. Articles focus on issues of key importance to pharmaceutical management science practitioners.

If you are interested in submitting content for future issues of the Journal, please send your submissions to *PMSA Headquarters* at info@pmsa.net.

Guidelines for Authors

Summary of manuscript structure: An abstract should be included, comprising approximately 150 words. Six key words are also required.

All articles and papers should be accompanied by a short description of the author(s) (approx. 100 words).

Industry submissions: For practitioners working in the pharmaceutical industry, and the consultants and other supporting professionals working with them, the Journal offers the opportunity to publish leading-edge thinking to a targeted and relevant audience.

Industry submissions should represent the work of the practical application of management science methods or techniques to solving a specific pharmaceutical marketing analytic problem. Preference will be given to papers presenting original data (qualitative or quantitative), case studies and examples. Submissions that are overtly promotional are discouraged and will not be accepted.

Industry submissions should aim for a length of 3000-5000 words and should be written in a 3rd person, objective style. They should be referenced to reflect the prior work on which the paper is based. References should be presented in Vancouver format.

Academic submissions: For academics studying the domains of management science in the pharmaceutical industry, the Journal offers an opportunity for early publication of research that is unlikely to conflict with later publication in higher-rated academic journals.

Academic submissions should represent original empirical research or critical reviews of prior work that are relevant to the pharmaceutical management science industry. Academic papers are expected to balance theoretical foundations and rigor with relevance to a non-academic readership. Submissions that are not original or that are not relevant to the industry are discouraged and will not be accepted.

Academic submissions should aim for a length of 3000-5000 words and should be written in a 3rd person, objective style. They should be referenced to reflect the prior work on which the paper is based. References should be presented in Vancouver format.

Expert Opinion Submissions: For experts working in the Pharmaceutical Management Science area, the Journal offers the opportunity to publish expert opinions to a relevant audience.

Expert opinion submissions should represent original thinking in the areas of marketing and strategic management as it relates to the pharmaceutical industry. Expert opinions could constitute a review of different methods or data sources, or a discussion of relevant advances in the industry.

Expert opinion submissions should aim for a length of 2000-3000 words and should be written in a 3rd person, objective style. While references are not essential for expert opinion submissions, they are encouraged and should be presented in Vancouver format.

Industry, academic and expert opinion authors are invited to contact the editor directly if they wish to clarify the relevance of their submission to the Journal or seek guidance regarding content before submission. In addition, academic or industry authors who wish to cooperate with other authors are welcome to contact the editor who may be able to facilitate useful introductions.

Thank you to the following reviewers for their assistance with this issue of the *PMSA Journal*:

George A. Chressanthis, Ph.D.
Principal Scientist, Axtria

Ewa J. Kleczyk, Ph.D.
Executive Director, Commercial Effectiveness Analytics, Symphony Health Solutions

Sudhakar Mandapati
Principal, Strategic Research Insights

JP Tsang, Ph.D. & MBA (INSEAD)
President, Bayser

Editor: Igor Rudychev
Global Samples Operations Lead, Pfizer

Sales Analytics and Big Data Developments Needed Now to Address Practitioner-Identified Emerging Biopharmaceutical Sales Force Strategic and Operational Issues

George A. Chressanthis, Ph.D., Principal Scientist, Axtria and Murali Mantrala, Ph.D., Sam M. Walton Distinguished Professor of Marketing, University of Missouri, Columbia

Abstract: The biopharmaceutical industry is undergoing significant changes. Most importantly, there is a shift toward greater company R&D focus on and launching of specialty medicines that use new scientific drug delivery systems, e.g., large molecules as opposed to traditional small molecule drugs, catering to smaller patient populations. While this shift solves some problems pharma has been facing, it has raised a whole new set of questions of sales and marketing, not yet addressed by academic and practitioner marketing science research. This paper focuses specifically on developments in data analytics needed now for practitioners to solve future sales force strategic and operational issues. As a starting point, the paper presents results of a survey of industry practitioners about emerging sales force science issues conducted in Spring 2015 with the help of the Pharmaceutical Management Science Association. The survey asked respondents for their perspectives on five areas of current and emerging sales force science issues: sales force strategy, sales force operations, sales analytics, big data, and environmental trend changes. The survey revealed there are big differences in the issues that are perceived as emerging and those currently preoccupying practitioners, indicating major changes in the data and analytics methods currently used are necessary to solve future sales force strategy and operational problems. A case study of this problem involves the analytics and data involved to look at the effect of detailing on drug utilization and health/economic outcomes for patients with newly diagnosed metastatic breast, lung, and colorectal cancers. Implications and challenges for the successful conduct of commercial analytics are provided based on the proposed conceptual framework.

Keywords: Biopharmaceuticals; Emerging sales force science issues; Sales force strategy, operations, analytics, big data, and environmental trends; Detailing analysis on drugs for metastatic cancer patients

1. Introduction and Industry Shift to Specialty Medicines

The growing shift to specialty medicines in the US pharma market is well documented.[1] Pricing issues are becoming more common and controversial, with questions raised about the sustainability of increasing sales revenue mainly through pricing.[2-3] One observes a rise of performance-based managed care contracts requiring pharma companies to demonstrate drug value in producing health outcomes and/or health economic benefits.[4] Providers are already adopting guidelines and treatment pathways that are driven by evidence of outcomes and value. These trends are raising new and challenging strategic and operational issues for biopharmaceutical sales forces. This paper focuses on how pharma sales data analytics need to adapt to address these issues.

We emphasize here that companies must rethink their commercial model design, supporting analytics, and data management infrastructure based on new industry dynamics. Future project/product portfolios will be

increasingly populated with expensive large molecule/biologic specialty medicines. There will be greater focus on personalized targeted therapies and significant concerns over patient access/affordability. Accordingly, the commercial model design, nature, and role of the pharma sales force must change.

2. Insights from Previous Published Research and Survey Work

Currently, the academic marketing science literature offers little insight into the kind of biopharma commercial model design now needed. Instead, extant literature focuses more on a tactical non-strategic economic model framework, emphasizing optimization of the promotion mix with the objective of increasing physician prescriptions.[5] Previously published academic works also tend to overlook the perspectives of industry practitioners as seen in this referenced review.[6]

This paper offers some insights into emerging biopharmaceutical sales force science and big data developments based on an exploratory survey of PMSA members working at biopharma firms and consulting organizations.[7] This survey is different from previous ones in the academic literature in that that it focused on biopharmaceutical sales force issues (as opposed to more general industry and broader sales & marketing perspectives), and surveys industry practitioners as opposed to academics. The survey ran from April 29, 2015 to May 22, 2015 and was sent to current members and email addresses from the Pharmaceutical Management Science Association (PMSA). The survey was also sent to members of the Pharmaceutical Marketing Research Group (PMRG). A total of 89 respondents started the survey, with 54 completing the survey.

Respondent attribute information was not broken down into finer details so as not to reveal the identities of individuals or their companies. Three tables were provided to manuscript reviewers showing: 1) profile of survey respondents, 2) respondent-company attributes (both showing broad and representative coverage of the biopharmaceutical industry, and 3) distribution of the areas of expertise across respondents from consulting firms in our sample. To conserve space, these tables are not provided here but are available upon request.

Respondents were asked to list the top 2 current and emerging issues in sales force strategy and operations per area. Table 1 notes clear differences in current and emerging issues:

1. concomitant changes in analytics and big data needed (e.g., medical claims, EMR, patient-level data, digital/social media channels) to support new solutions based on health outcomes.

2. rapidly evolving environmental trends (e.g., growing influence from IDNs, ACOs, increasing consolidation between providers and payers, increasing sales rep access restrictions to physicians).

Pharma commercial analytics are currently seen mainly as a means to support tactical execution of traditional sales and marketing channels to achieve short term financial goals, rather than as a strategic asset as a key source for competitive differentiation to sustain long term industry advantage.[5] Instead, biopharma companies need to pursue a strategic open systems based approach across the entire pharmaceutical value chain throughout the project/drug lifecycle.[5] This means pharma companies will be increasingly called upon to demonstrate value through significant

Table 1: Exploratory Survey Results on the Top 2 Identified Current vs. Emerging Issues for Each Biopharmaceutical Industry Sales Force Science Area[7]

Current Issues		Emerging Issues	
1. Sales Force Strategy			
21.4%	Targeting quality	26.2%	Institutional sales forces, especially for IDNs and ACOs
19.0%	Financial outcomes	16.7%	Outcomes and value-based messaging
2. Sales Force Operations			
21.2%	Incentive compensation	26.2%	Flexible sales force deployment
16.7%	Call planning	16.7%	Incentive compensation
3. Sales Force Analytics			
33.3%	Promotion response and ROI analytics (all channels)	16.7%	Health outcomes and cost-effectiveness analyses
11.9%	Marketing-mix optimization	14.3%	Sales analytics that drive sales force strategy and operation outcomes
4. Big Data			
21.4%	All Rx-based databases	16.7%	All Rx-based databases (physician-level and product-level)
14.3%	LRx (patient-level) data	16.7%	Activity data from social media and digital channels
		16.7%	Electronic medical records (EMR)
5. Environmental Changes			
19.0%	Increasing payer influence on physician prescribing	16.7%	Increased consolidation between provider and payer
14.3%	Increasing sales representative access restrictions	14.3%	Increasing sales representative access restrictions
		14.3%	Changes in payer influence on physician prescribing

Top 2 identified *current* (0 to 2 years out) vs. *emerging* issues (> 2 years out) for each biopharmaceutical industry sales force science area and by % of total responses.

improvements in health outcomes and reductions in treatment costs. This latter viewpoint is consistent with a newer perspective that research-based biopharma companies must think differently and apply tools beyond traditional boundaries, while engaging in interdisciplinary-type analyses to solve increasingly more complex business problems.[8]

3. Foundations Governing the Future Role and Effects of Sales and Marketing

Models must begin to connect the nature of sales rep-physician interactions on a different set of metrics ultimately tied to improvements in health outcomes, drug and total treatment costs, and cost effectiveness. Three underlying theoretical frameworks provide the foundation

for analyzing the effect of sales rep-physician interactions:

1. Pharma sales & marketing will be designed and executed as "*informative*", not "persuasive".[9]

2. The growth of more complex specialty medicines implies more weight will be placed on the quality of supporting *scientific evidence* by healthcare practitioners and payers involved in drug adoption, formulary coverage, patient access, compliance, and adherence.[10-12]

3. Variations in the diffusion of medical information create patterns of variable medical care use, which in turn, results in variations in health outcomes, expenditures (drug and treatment spending), and cost-effectiveness.[13] This fundamentally alters the commercial analytics approach. The current approach emphasizes assessing physician prescription (Rx) response to sales & marketing. The future approach must demonstrate how such channels impact health/economic outcomes. This means building new analytical capabilities based more on real world evidence (RWE) and health/economic outcomes research (HEOR) models. This is necessary to support decisions on managed care performance-based outcomes contracts on formulary coverage and patient access, and provider adoption of guidelines on treatment pathways.

4. Case Study Example in the Therapy Area of Anti-Cancer Drugs

The evolutionary role of detailing is highlighted when looking at scientifically-driven personally-targeted specialized anti-cancer drugs. Focus is on the sales force science and big data required to relate detailing and other commercial activities to variations in drug utilization and in turn health/economic outcomes for newly diagnosed metastatic patients with breast (BC), colorectal (CRC), and non-small cell lung cancer (NSCLC). There are a number of reasons for choosing these anti-cancer drugs as a case study:

1. The trend of increasing sales rep access restrictions to oncologists has been well documented[14] and was identified in the practitioner survey as an important environmental change. Oncology sales reps are also very experienced and knowledgeable, thus access restrictions to these individuals means potentially missing out on valuable information.

2. Anti-cancer drugs represent the second largest therapy class by US spending and the largest therapy area by the percentage of new drug launches.[1]

3. Keeping current with the latest information developments on anti-cancer drug R&D, clinical trials, and new novel therapies (e.g., personalized medicines, targeted cancer therapies)[15-16] by medical oncologists is challenging given the high involvement of biopharma companies.

4. A medical oncologist who falls behind on the latest anti-cancer drug developments means dire consequences and higher risks to patients given the lethality of these diseases.

5. The pricing of anti-cancer drugs and assessing the value of cancer treatment options are not only significant commercial challenges but also key public pharmaceutical policy concerns.[17-18]

A conceptual framework on how to estimate the relationship between detailing and health/economic outcomes has already been publicly presented and researched for viability.[19-20]

Figure 1: Framework How Geographic Variations in Sales Rep Detailing Affect Drug Utilization, and Health/Economic Outcomes for Newly Diagnosed BC, CRC, and NSCLC Patients[13,19]
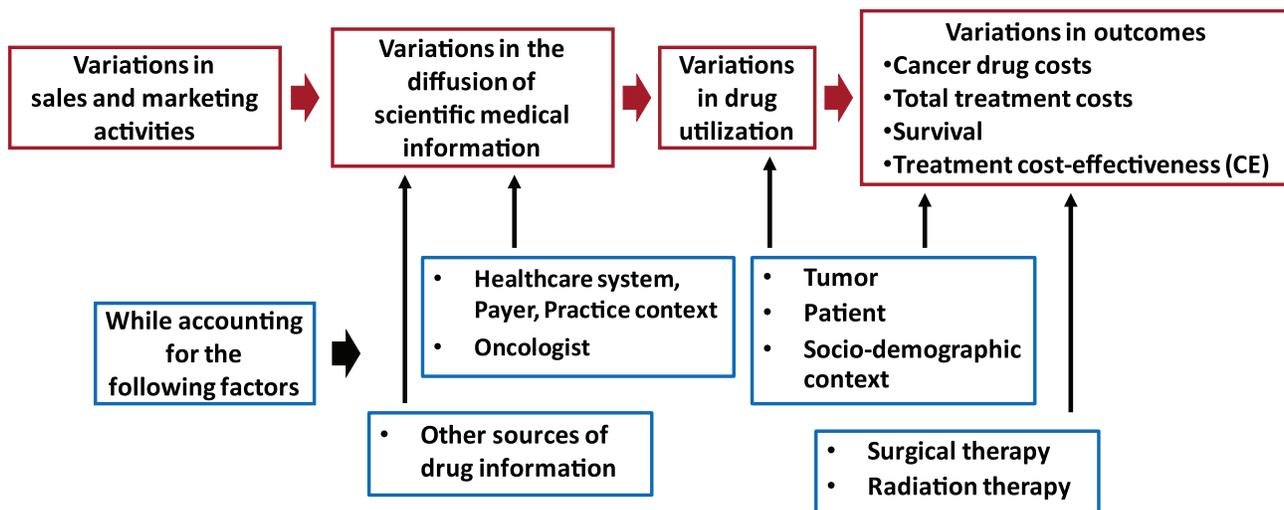
Figure 1 shows how variations in sales and marketing activity are related to outcomes along with factors impacting this chain.

Figure 2 provides a more detailed conceptual framework of key relationships and specific data elements needed to measure each link in this chain for each type of metastatic cancer patient. The data elements and physician identifier codes for data bridging needed to measure statistically key relationships are generally available to biopharmaceutical companies through an array of secondary data sources and contracts. Lastly, data from qualitative market research will be needed to determine the range of drug information sources and their weight of influence when oncologists make drug utilization decisions.

The basic modeling approaches required to test the significance and measure the effect of key relationships shown in Figure 2 is based on prior research work.[20] Important to note are the applications of research methods more common to the analysis of RWE and HEOR

models than traditional econometric methods on Rx-physician sales-response designs. The effect of detailing on drug utilization outcomes separately for each tumor type can be estimated, controlling for important factors that may confound this relationship, including those on the oncologist, health care system, payer, and practice, patient and tumor characteristics, location of treatment, time of diagnosis, and non-detailing commercial activities. Generalized Propensity Score (GPS) based weighting with bootstrap standard errors can be used to estimate the marginal effect of detailing on drug utilization, *ceteris paribus*. Full discussions of this estimation technique advantages, adjustments to account for confounding factors, sufficient covariate overlap, and misspecification concerns are well known in practice.[21-24]

The association of detailing to cancer drug costs and total cancer treatment costs per cancer site uses two methods: Kaplan-Meier Sample Average (KMSA) method[25] and an approach described by Miller & Halpern,

Figure 2: Conceptual Framework How Variations in Sales Rep Detailing Affect Drug Utilization and Health/Economic Outcomes for Newly Diagnosed BC, CRC, and NSCLC Patients[19-20]
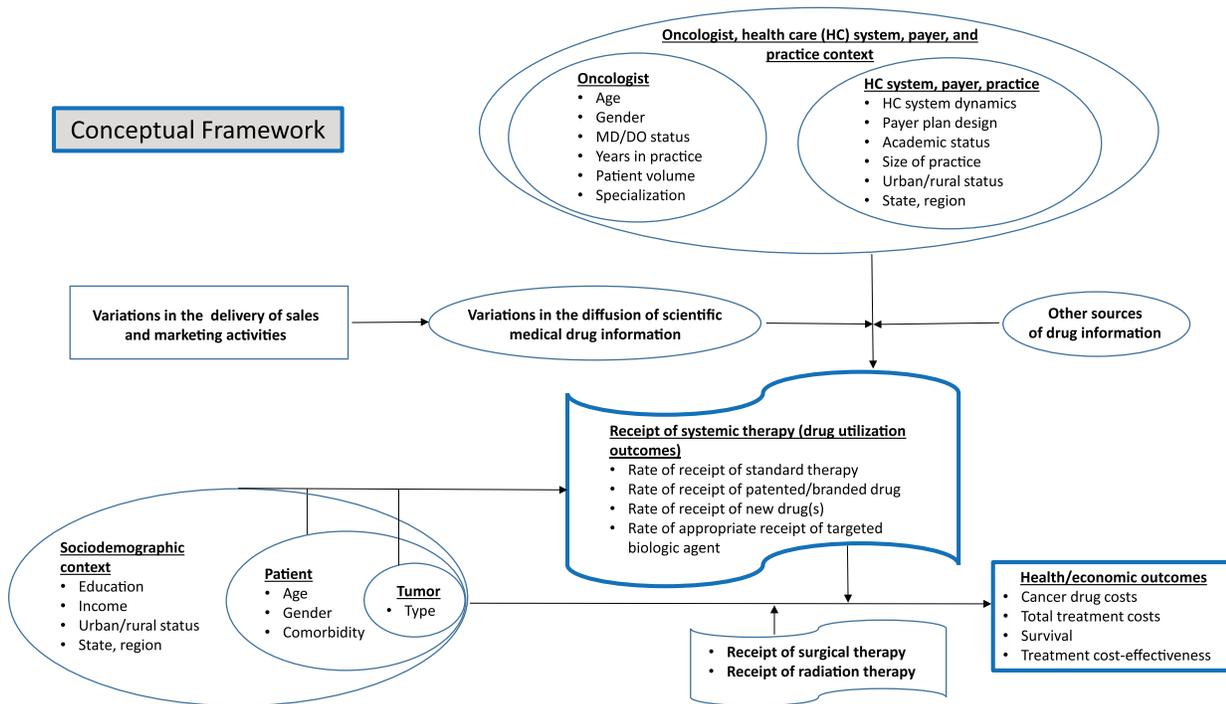

which is similar to the Cox proportional hazard model.[26] Analyses on total treatment costs and overall survival per average patient per detailing segment category can be combined per cancer site to determine incremental effects of changes in detailing on cost-effectiveness using statistical procedures previously noted.

Estimating the effects of detailing on health/economic outcomes are not without challenges, a non-exhaustive list being:

1. The effect of detailing may be tempered by sales rep access restrictions, which in turn, limits the dissemination of scientific information to physicians. Prior research work can be applied on the modeling and estimation approaches on the determination of industry sales rep access restrictions and

their effect on physician prescribing using data defined in the conceptual framework.[27-28]

2. The data needed to perform the preceding described analyses is substantially larger, potentially more expensive to acquire, and less comprehensive than traditionally-applied databases. Patient-level claims data is not as complete as existing Rx-level and even anonymized patient-level databases (APLD). Claims data is not for example IC-grade. Companies would have difficulties acquiring a national footprint of data in order to conduct various strategic and operational sales force processes, especially for diseases with smaller patient populations.

3. Questions may exist on the ability to merge effectively this larger array of databases, even

with available bridging components such as physician identifying codes.

4. Companies need to adjust their thinking of segmentation analysis, which currently is focused on payers and physician attitudes/behaviors to addressing patient attributes and characteristics of their disease state, and healthcare systems.

5. Current analysis of detailing is on the effect from added frequency rather than call quality and the nature of scientific evidence delivered. This means metrics on scientific evidence need to be devised and tested. Prior research work can be a guide here.[10-12]

The above challenges, while daunting, are not insurmountable and should not stop company efforts to connect sales and marketing activities to the ultimate value goal of health/economic outcomes given changes in key environmental trends as identified by survey participants.

5. Implications for Sales Force Strategy, Operations, Data Development, and Analytics

The preceding data and analytics sections chart a different course than current promotion-response econometric-modeling practices for sales force strategy, operations, data development, and analytics. Current commercial analytics is frequency-based where the end point is measuring drug utilization. Sales operations processes are geared to support this strategic approach to facilitate and incentivize detailing frequency with little regard for the effects physician interactions have on patient outcomes, drug costs, treatment costs, and cost effectiveness. Evolving commercial analytics must be structured to see drug utilization as an intermediate outcome, especially since key physician specialties that will be associated with the focus of future

new drug development are generally the most sales rep access-restricted.[14] Focus instead must leverage RWE data and the effects from delivering scientific medical information to specialty physicians on drug utilization and health/economic outcomes.[29] This means sales operations processes must structure the success of sales reps in their detailing efforts geared toward facilitating information-based physician interactions and where sales reps are focused on improvements in metrics, such as patient compliance and adherence, as leading indicators of better health/economic outcomes. A list of specific implications for sales force science issues as categorized by the survey headings noted in Table 1 are as follows:

Sales Force Strategy

1. Segmentation schemes must combine data to reflect dynamics at the following 4 levels: physicians/accounts, IDNs, payers (e.g., commercial 3rd party, Medicare, Medicaid), and patients. The goal is to be able to follow the patient through the healthcare system to ensure both patient drug adherence and health benefits are derived from continued drug utilization. This segmentation scheme will drive sales force strategy and operations process outcomes. Rx and drug adherence must be objective measures that drive sales force strategy, the latter metric being a good leading indicator of health/economic outcomes.

2. Sales force size, structure, and allocation will still be physician-based but layered differently based on the above segmentation scheme and where sales force allocation efforts are implemented to cover the patient journey.

Sales Force Operations

3. Territory alignment and call planning design will reflect the 4-level segmentation scheme.

4. Objective setting, incentive compensation, and sales force reporting/performance

management will share focus on Rxs (still needed to track company objectives for financial returns) and MBOs (management by objectives) that affect drug adherence tracked by APLD aggregated at the physician-level. These MBOs are designed to track activities related to the delivery of scientific information and efforts that improve patient adherence, thus helping outcomes on value delivered to physicians, health and economic benefits effects to patients, and to the healthcare system.[30] Examples of sales rep MBOs tied to their interactions with physicians and other healthcare personnel are as follows:

(a) qualitative assessments from physicians and office staff (nurses, office manager, etc.) on whether sales reps are adding value in their interactions.

(b) providing physicians information on and enrollment of patients in disease management programs, co-pay card programs, and coupons.

(c) connecting physicians to medical science liaisons who can provide deeper answers to medical questions.

(d) sending physicians optimal level of samples that can be helpful for physicians to try patients on new therapies when other approaches have failed to reach clinical goals.

(e) alerting physicians to new drug indications, FDA-imposed black-box warnings, and other important drug updates.

(f) measuring the proportion of call plan physicians who listen to a detail through an electronic notebook/non-paper delivery (designed to collect qualitative measurements of sales rep-physician interactions for detailing quality analysis).

(g) tracking the proportion of call plan physicians who attend and the qualitative assessment of local speaker programs organized by the local sales rep.

(h) counting the proportion of call plan physicians who seek added drug information through the company/drug website.

Sales Force Analytics
5. The preceding cancer example illustrates the need to introduce empirical techniques common to HEOR and RWE models and to combine their use with traditional statistical sales force analytics.

6. The need to demonstrate drug value through sales force activity will require the broader use of and implications from detailing quality statistical models.

Big Data
7. Figure 2 illustrates the combination of a wide array of traditional data used in sales force analytics plus newer elements utilized in HEOR/RWE models as found in claims/EMR, socio-economic, and demographic data sources. This will require efficient data management, bridge elements that allow individual data sources to be merged, and routines that allow for easy access for analytics to support these processes and insight generation.

Changes in sales force environmental trends likely mean future sales reps will need to improve their capabilities to deliver more complex scientific evidence. This will also require a team of service reps to address and focus on matters such as patient-access/affordability and payer dynamics since sales reps will have ever-decreasing time interacting with physicians.[14] Attention will have to be given to the development of technology that more

efficiently and effectively delivers scientific information to healthcare professionals while recording such data for further analysis. The introduction of complementary devices (health diagnostic and monitoring technology) will aid in improving patient compliance/adherence and demonstrating value to payers. These preceding changes mean the talents and capabilities required by people involved in commercial analytics to conduct this work will have to be upgraded to account for new modeling designs and expertise in handling different databases. Concomitant changes in the internal organizational structure of companies will be required to facilitate this type of interdisciplinary analysis among groups that currently have little to no engagement with each other (e.g., interactions between HEOR and RWE functions with sales force strategy, operations, and analytics functions). Lastly, challenges will need to be overcome to generate the necessary patient-level data (e.g., claims, EMR, etc.) size that will enable this change in sales force science approach. Patient-level databases are highly fragmented, thus a need to combine databases required to generate sufficient number of observations when analyzing diseases with small patient populations.

6. Concluding Remarks

This paper has tackled the question on the emerging future role of sales analytics and big data developments in a pharma environment increasingly focused on launching specialty medicines. The case study example can be adapted to address an array of other sales and marketing situations, other disease areas, and physician specialties. We believe fundamental commercial analytics changes are destined to happen for pharma companies in the following seven buckets consistent with comments here and trends identified by survey participants:

1. *Commercial Model Design* - the go-to-market approach and model design necessary to achieve all company strategic goals, driven by payer and patient analytics, and dependent on drug technology of the project/product portfolio that can be successfully developed and tactically executed to deliver optimal results while mitigating key risks.

2. *Payer Analytics* - focused on managed markets (e.g., private third party commercial and public drug plans), analyzing effects from changes in plan design, and their relationship to sales, marketing, and patient outcomes.

3. *Patient Analytics* - focused on analyses generated from real world evidence (RWE) and patient-level data on outcomes (e.g., drug compliance and adherence, drug costs, treatment costs, health outcomes, cost-effectiveness) resulting from drug utilization.

4. *Sales Analytics* - focused on processes and outcomes related to ensuring optimal sales force investment efficiency and effectiveness.

5. *Marketing Analytics* - focused on processes and outcomes related to ensuring optimal brand performance throughout the entire lifecycle.

6. *Commercial Analytics Innovation Center* - focused on basic research activities designed to generate new management/marketing science methods for solutions to address future commercial problems faced across the entire project/product lifecycle using experimentation, collaborations with academic researchers, and other activities to encourage innovation.

7. *Cloud Information Management* - focused on speed, agility, and scale in association with managing new data sources, elastic infrastructure, data quality & accuracy, and

actionable insight in support of activities in all of the preceding commercial analytics buckets.

These commercial analytics buckets will rapidly become interdependent activities. Health/economic outcomes from payer and patient analytics will become the principal emphasis and drivers of all future commercial decisions.[29] The construction of the right commercial model design and conduct of all remaining analytics in other areas, like sales analytics, will be done to support payer and patient outcomes. This means solving problems using commercial analytics will require greater alignment among these activities, an open system framework of thinking in solving commercial problems, data environment constructed to support all of these activities, and leadership approach and innovative analytics culture necessary to cultivate and sustain a competition advantage. A potential industry challenge is who will be leading perceived advancements in basic knowledge on key sales force science topics as outlined here? Respondents from the practitioner survey clearly noted biopharmaceutical industry consulting research companies as the leading source.[7] Yet, survey respondents from biopharmaceutical consulting companies also noted expertise focus is not in the areas identified here required for the industry to respond effectively to changing environmental trends.[7] This paper provides a commercial/sales analytics, big data management, and organizational blue print for companies on how to prepare and operate successfully in this evolving sales force pharma landscape.

About the Authors

George Chressanthis *is Principal Scientist at Axtria, a big data and analytics company, starting there in July 2016. His research focuses on pharmaceutical commercial issues and their intersection to drug industry business, public policy, clinical treatment decisions, patient outcomes, and cost of care. He spent 14.5 years from 1995-2009 working in the pharmaceutical industry after a long career in academia, with the majority of his time at AstraZeneca Pharmaceuticals LP US headquarters leading teams in support of sales force strategy, sales operations, and other commercial analytical functions. He held full-time professorships in Healthcare Management and Marketing in the Fox School of Business and a secondary professor appointment in Clinical Sciences in the School of Medicine at Temple University from 2010-2016. He also had a career as an academic economist from 1982-1995, eventually becoming a tenured full professor at Mississippi State University. He received his Ph.D. in Economics from Purdue University.*

Murali Mantrala *is the Sam M. Walton Distinguished Professor of Marketing and Chair, Department of Marketing at the University of Missouri at Columbia. He is a member of INFORMS (the Institute for Operations Research and the Management Sciences) and the American Marketing Association. His research interests include retail pricing and category management, news media pricing, pharmaceuticals marketing, marketing decision models, sales force resource allocation, and compensation strategies. His extensive publication record spans major academic journals across marketing, management and marketing science, operations research, and sales. He received his Ph.D. in Marketing from Northwestern University.*

References

1   IMS Institute for Healthcare Informatics. Medicine use and spending in the U.S.: a review of 2015 and outlook to 2020. Parsippany, NJ: April 2016.
2   Rockoff J. Pricey drugs are hurdle for new biotech CEO. *Wall Street Journal* 2016; June 7: B1-B2.
3   Walker J. Drug makers raise prices despite protests. *Wall Street Journal* 2016; July 15: B1-B2.
4   Edlin M. Performance-based pricing for pharmaceuticals. *Managed Healthcare Executive,* published online 30 September 2015, available at http://managedhealthcareexecutive.modernmedicine.com/managed-healthcare-executive/news/performance-based-pricing-pharmaceuticals?page=0,0 (accessed 22 July 2016).
5   Singh J and Jayanti R. Closing the marketing strategy-tactics gap: an institutional theory analysis of pharmaceutical value chain. In: Ding M, Eliashberg J and Stremersch S. (eds.) *Innovation and marketing in the pharmaceutical industry: emerging practices, research, and policies*. New York: Springer, 2014, pp. 710-735.
6   Stremersch S and Van Dyck W. Marketing of the life sciences: a new framework and research agenda for a nascent field. *Journal of Marketing* 2009; 73: 4-30.
7   Chressanthis G and Mantrala M. Pharmaceutical sales force science: current and emerging trends, and issues for future research. Presentation to the Pharmaceutical Management Science Association Board. Las Vegas, NV: PMSA Annual Meeting; 17 April 2016.
8   Castellani J. Keynote presentation: The changing healthcare environment. Pharmaceutical Management Science Association Annual Conference 2015. Presentation on April 20, 2015, Arlington, VA.
9   Rod M and Saunders S. The informative and persuasive components of pharmaceutical promotion. *International Journal of Advertising* 2009; 28: 313-349.
10 Azoulay P. Do pharmaceutical sales respond to scientific evidence? *Journal of Economics & Management Strategy* 2002; 11: 551-594.
11 Sood A, Kappe E and Stremersch S. The commercial contribution of clinical studies for pharmaceutical drugs. *International Journal of Research in Marketing* 2014; 13: 65-77.
12 Kappe E and Stremersch S. Drug detailing and doctors' prescription decisions: the role of information content in the face of competitive entry. *Marketing Science* 2016; 35: 915-933.
13 Phelps C. Diffusion of information in medical care. *Journal of Economic Perspectives* 1992; 6: 23-42.
14 Khedkar P and Sturgis M. AccessMonitor™ and AffinityMonitor™ 2016 executive summary: want better access to physicians? Understand what's top of mind. Evanston, IL: ZS Associates, August 2016.
15 Ingelman-Sundberg M. Personalized medicine into the next generation. *Journal of Internal Medicine* 2015; 277: 152-154.
16 Shih Y, Smieliauskas F, Geynisman D, et al. Trends in the cost and use of targeted cancer therapies for the privately insured nonelderly: 2001 to 2011. *Journal of Clinical Oncology* 2015; 33: 2190-2196.
17 Howard D, Bach P, Berndt E, et al. Pricing in the market for anticancer drugs. *Journal of Economic Perspectives* 2015; 29: 139-162.
18 Schnipper L, Davidson N, Wollins D, et al. American society of clinical oncology statement: a conceptual framework to assess the value of cancer treatment options. *Journal of Clinical Oncology* 2015; 33: 2563-2577.
19 Chressanthis G and Esnaola N. Health outcome implications from restricting the flow of FDA-regulated medical information from pharmaceutical companies to physicians. Presentation at the 2015 International Health Economics Association World Congress. Milan, Italy. Session: Marketing Drugs (14 July 2015).
20 Esnaola N and Chressanthis G. Pharmaceutical sales rep access restrictions to oncologists and cancer outcomes. NIH R01 grant proposal resubmission R01CA190551, October 2014.
21 Imbens G. The role of the propensity score in estimating dose-response functions. *Biometrika* 2000; 87: 706-710.
22 Spreeuwenberg M, Bartak A, Croon M, et al. The multiple propensity score as control for bias in the comparison of more than two treatment arms: an introduction from a case study in mental health. Medical Care 2010; 48: 166-174.
23 Austin P. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavorial Research* 2011; 46: 399-424.
24 Drake C. Effects of misspecification of the propensity score on estimators of treatment effect. *Biometrics* 1993; 49: 1231-1236.
25 Lin D, Feuer E, Etzioni R, et al. Estimating medical costs from incomplete follow-up data. *Biometrics* 1997; 53: 419-434.
26 Miller R and Halpern J. Regression with censored data. *Biometrika* 1982; 69: 521-531.
27 Chressanthis G, Sfekas A, Khedkar P, et al. Determinants of pharmaceutical sales representative access limits to physicians. *Journal of Medical Marketing* 2014; 14: 220-243.
28 Chressanthis G, Khedkar P, Jain N, et al. Can access limits on sales representatives to physicians affect clinical prescription decisions? A study of recent events with diabetes and lipid drugs. *Journal of Clinical Hypertension* 2012; 14: 435-446.
29 Looney W. Real-world evidence: from volume to value. *PharmExec.com*, published online 13 October 2016, available at http://www.pharmexec.com/real-world-evidence-volume-value?pageID=1 (accessed 7 November 2016).
30 Silverman E. Glaxo to change its compensation program for U.S. sales reps. *Pharmalot*, published online 13 April 2015, available at http://blogs.wsj.com/pharmalot/2015/04/13/glaxo-to-change-its-compensation-program-for-u-s-sales-reps/ (accessed 20 July 2016).

A New Path to Understanding the Physician's Decision Journey Using Simulated Patients

Chandra Chaterji, Consultant, and Greg Chu, Chief Operating Officer, InTask, Inc.

Abstract: Simulations built around the management of virtual patients offer healthcare market research and marketing science professionals an innovative and powerful approach to deeper understanding of the behaviors and decision-making processes of physicians. Rather than asking physicians what they do and why, as we might in a questionnaire, simulation allows us to observe physician behavior within an engaging environment in which they treat virtual patients. This data collection method using simulated patients is scalable, and can be used in both large and small sample studies. In this article, we report on early findings from pilot tests of this novel simulation software platform developed for use in market research.

Keywords: Simulation, Decision journey, Big data, Segmentation, Behavioral economics, Virtual patient

Introduction

There are two broad approaches to collecting market research data. One is observation—starting with participant observation and ethnography, and extending all the way to think-aloud methods, focus groups, IDIs and even journalism approaches. These methods have traditionally been associated with qualitative research. The other broad approach is structured questioning, as exemplified by the market research questionnaires utilized in quantitative surveys. Both approaches have strengths and weaknesses. Observational methods can provide deep insights into what respondents do and why, but do not readily scale to large samples. Structured questioning is amenable to statistical design principles, and scales easily, but does not necessarily illuminate the respondents' thought processes. Moreover, we are increasingly sensitive to problems of respondent engagement with quantitative research procedures since structured questionnaires offer little intrinsic value to the respondent. In the realm of healthcare, the questionnaire data often does not derive from ongoing medical or business processes that healthcare professionals need to engage with as part of their work. Nor does it resemble or invoke the tasks they actually undertake on a daily basis. As a result, lengthy or tedious surveys may not inspire respondents to provide their best thinking in answering survey questions.

Simulation of virtual patients who are "treated" by physicians offers an intriguing opportunity to combine the best elements of observation and questionnaire-based surveys. Within digital environments, physicians can be engaged in actual tasks, such as the diagnosis and treatment of virtual patients. Rather than *asking* physicians what they do, simulation allows researchers to *observe* what physicians do under a range of environmental situations and with different patient types tailored to address specific areas of research interest. At the same time, digital simulations can provide researchers with insights into the physician's decision-making process as reflected in the order, frequency and duration of access of patient information, as well as through the lab tests and diagnostics ordered to support treatment decisions for specific patients. Moreover, simulation allows us to observe how all this unfolds in virtual time, as patients can be programmed to return for multiple visits with evolved presentations.

Figure 1: Screenshot of Beta Version of Schizophrenia Simulation Platform

Elena
F, 28 • Private Insurance

VISIT: 1
DATE: 01/07/15
VISIT REASON: SCHIZOPHRENIA

PATIENT   TREAT

Patient Overview
Social Support
Positive Symptoms
Negative Symptoms
Patient Interaction
Insurance

POSITIVE SYMPTOMS

DATE	01/07/15		
Auditory Hallucinations	Severe		
Unusual thought content	Mild		
Paranoia	Moderate		
Delusions of persecution	Moderate		
Disorganized thinking	Mild		
Agitation/Anxiety	Moderate		
Aggression/violence	None		
Thoughts of self-harm	None		

Simulation also generates large streams of data which, relative to survey data, can reasonably be called "big"—created with no more investment of time or effort on the part of the respondent than that demanded by traditional market research questionnaires. This last point hints at additional ways in which big data can be leveraged in healthcare research. We are largely familiar with both the promise and challenge of capturing, managing and making sense of big data in healthcare. Significant value has already been generated from EMR systems, social media, and even the internet of things in supporting improved commercial decision-making. Considerable room for further progress exists, but it is interesting to note that most efforts are focused on analytic approaches rather than the sourcing and production of the data itself. Most approaches to working with big data today seem to assume that the data itself is a given—that it arises from natural processes that are separate from analytic endeavors. Accordingly, efforts to derive value from "naturally occurring" big data necessarily devote considerable effort to the cleaning and organizing of data so that it fits with the researcher's analytic goals. Given the fact that real world data can be extremely messy, this is no small task. Simulation offers the possibility of creating relatively clean, ready-to-use big data that provide insight into both what respondents do and why they do it—in essence combining some of the best traits of both observational and survey research techniques.

In this paper, we provide an overview of our experience in pilot testing a simulation platform developed for use as a market research tool. We will focus on two points: 1) The respondent interaction with the simulation platform, and 2) The types of data generated by the simulation.

Method

In order to test simulation as an alternative or complementary approach to current market research and secondary data analytics, InTask, Inc. developed an EMR-like simulation interface built around the diagnosis and treatment of virtual patients. The simulation, which is browser-based and can run on tablet, laptop or desktop devices, is structured around three tasks: the examination of the patient, treatment decisions, and scheduling of follow-up visits. The content of the simulation platform can be customized to address a wide range of therapy areas and issues. Figure 1 contains a screenshot showing the patient examination screen from a beta version utilized in a study on schizophrenia.

The simulation records every action taken by the physician when engaging with the virtual patients. While this ensures the capture of basic information such as treatment decisions made, lab tests ordered or even referrals, it also provides detailed information on what the physician chose to view in making those decisions. Thus, referring to the patient examination screen shown above, if the physician clicks on one of the information category buttons on the left side of the screen, this is recorded—as is the frequency with which he clicked on this piece of information, the order of access relative to other options and the time spent before clicking on the next information category. It is important to note that the physician is not forced to examine all groups of information; the simulation provides more potentially relevant information than the physician will actually utilize in making a treatment decision. Inasmuch as the physician pays more attention to some information items than others, the simulation helps reveal his decision-making process. In this sense, the behavioral economics concept of "bounded rationality" is "hard wired" into the simulation design.

Critical to the success of the simulation is the design of the virtual patients which populate it. The simulation can accommodate a wide range of patient parameters, including static descriptions of the patient, such as age and gender, as well as dynamic measures such as symptoms, lab values, concomitant conditions and current treatments. These parameters can be conveyed in various formats including text, image and video. Patients are designed not only to be realistic, but also representative of the pool of patients relevant to the research issue. For example, if a marketer is interested in second line therapy in type 2 diabetes, the virtual patients might be designed to fail on first line therapy with metformin. In this case, the progress of the virtual patient will be determined through a pre-determined,

branched logic largely independent of the physician's treatment decisions. In other cases, the patient progress will be determined by a Markov process, informed by outcome probabilities derived from real-world data.

Our experience in conducting multiple pilot tests and commercial projects across a range of therapeutic areas is that a good grasp of the patient presentation is critical. We always work closely with client teams—marketers and physicians—when designing the virtual patients, and have found these inputs to be indispensable. Designers of virtual patients for specific applications must know what clinical parameters are relevant, as well as the outcomes that are reasonable to include within the virtual timeframe of the simulation. They must also know what diagnostic and treatment options to include. In crafting virtual patients, input from the client's medical affairs group is often useful, as is prior market research—in particular, chart audit data. As with questionnaire development, pre-testing of the simulation is essential. Comfort with experimental design is also important, as the virtual patients may be created to tease out the impact of specific patient attributes in driving diagnostic or treatment behaviors. In our work, we have created virtual clones—pairs of patients who are essentially identical except for one or two key traits hypothesized to be determinants of behavior. This experimental approach, akin to A/B testing, has proved useful in quantifying the impact of patient parameters on prescribing decisions. But the most vital input to the design of effective simulations is a thorough understanding of the research issue. In this respect, designing a good simulation is no different from designing a good questionnaire.

A second critical design consideration is the format of the patient information presented in the simulation. Overly zealous efforts to emulate reality can cloud interpretation of

Figure 2: Combined Results From Two Pilot Studies (n=178) Conducted in 2016 in the United States with Primary Care Physicians and Psychiatrists
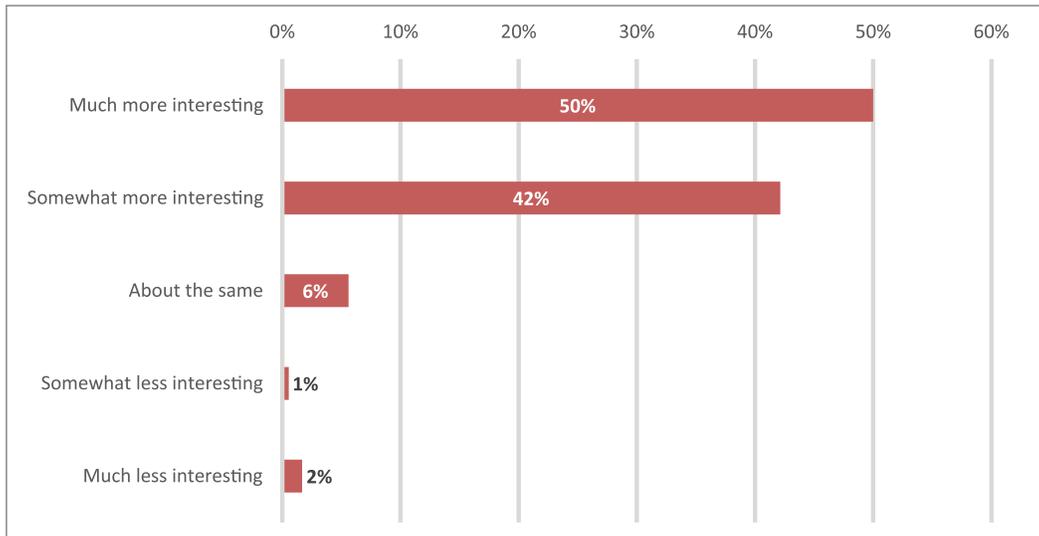
simulation results. A photo of the patient, for example, may convey information that impacts physician assessments and treatment decisions. Without validated and standardized portrait portfolios, the simulation designer may introduce unnecessary noise into the analysis of quantitative simulation data by using stock photos. Inclusion of other potentially ambiguous diagnostic input, such as an actual radiograph of a joint rather than a succinct and clear radiologist report, may similarly open the door to analytic challenges. Fortunately both these concerns can be easily addressed. Engagement with the simulation platform does not demand patient photos, and straightforward diagnostic information is usually sufficient to maintain a sense of realism. It is worthwhile noting, however, that inclusion of ambiguous stimuli can be extremely valuable in qualitative research applications of the simulation. In our experience, inclusion of these stimuli have invariably enriched discussions with physicians and generated insights that were missed in more traditional qualitative projects.

The remainder of this article focuses on actual findings and data analysis derived from two pilot

tests of the simulation platform conducted in the United States in 2016 with a total of 178 primary care physicians and psychiatrists. In each of these pilots, physicians were asked to treat eight virtual patients suffering from either schizophrenia or Alzheimer's disease over multiple visits in virtual time. Respondents were recruited from a fieldwork agency panel via an on-line screener. Their reactions to the simulation were captured using a post-simulation questionnaire.

Physician Reactions to the Simulation
The promise of simulation, as well as the theoretical efficacy of the approach, is premised in large part on the ability of simulations to better engage respondents and present data gathering tasks which more closely replicate real-world situations than do surveys. Early user experience testing with the simulation platform across a number of therapy areas suggested that simulations do in fact engage physicians to a greater extent than do typical surveys. Post-simulation surveys from our two quantitative pilots confirmed this. In both pilot tests, 9 out of 10 respondent physicians agreed that the simulation was more interesting than the online surveys they typically complete. (Figure 2)

Figure 3: Combined Results From Two Pilot Studies (n=178) Conducted in 2016 in the United States with Primary Care Physicians and Psychiatrists. Percentages Reflect Top 3 Box Agreement on a 7 Point Scale.
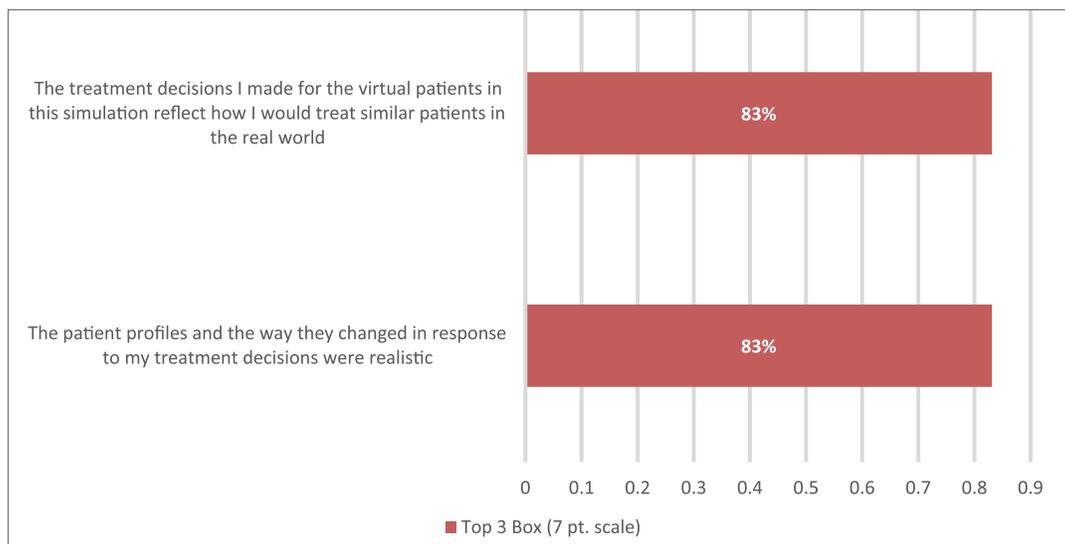
Equally important is the fact that respondents viewed the simulation exercise as reflective of reality. More than 80% of respondents in the two pilot studies agreed that the "treatment decisions they made within the simulation reflect how they would treat similar patients in the real world" and that the "patient profiles and the way they changed in response to my treatment decisions were realistic". (Figure 3)

Physician Behavior: Time Spent

If physicians claim that the simulation is more interesting and also realistic, is this reflected in how they actually interacted with the simulation? Since everything they actually do in the simulation is recorded, we have the means to provide at least a glimpse into how respondents engaged with the digital environment. One measure of interest is the amount of time physicians spent treating each virtual patient over the course of the simulation. In both pilot studies included in this analysis, physicians were asked to treat eight virtual patients over three visits each for a total of 24 virtual patient interactions. Figure 4 shows

the average time in seconds spent with each patient over each visit. As is evident from the graph, the average time spent with each patient declines quickly after the first few visits. This was anticipated and reflects a rapid learning curve as physicians became familiar with the simulation interface. Beyond this, two important points are evident in the time stamp data from which this data is derived. First, decline in time spent per virtual patient visit levels off rather than continuing to decline. This suggests that, at least within the context the patient volume presented in this simulation, respondents were not growing bored with the simulation. Second, we note that the spikes in time spent all correspond with the appearance of a new virtual patient. This is reassuring, as it reflects the real world dynamic of clinical presentations in which initial visits are longer and more extensive than follow-up visits. This analysis of time spent on virtual patient interaction supports the self-assessment of respondents that they were engaged in the research process and that the process itself reflects, at least in part, their behavior in the real world.

Figure 4: Combined Results From Two Pilot Studies (n=178) Conducted in 2016 in the United States with Primary Care Physicians and Psychiatrists. Average Virtual Patient "Visit" Time in Seconds.
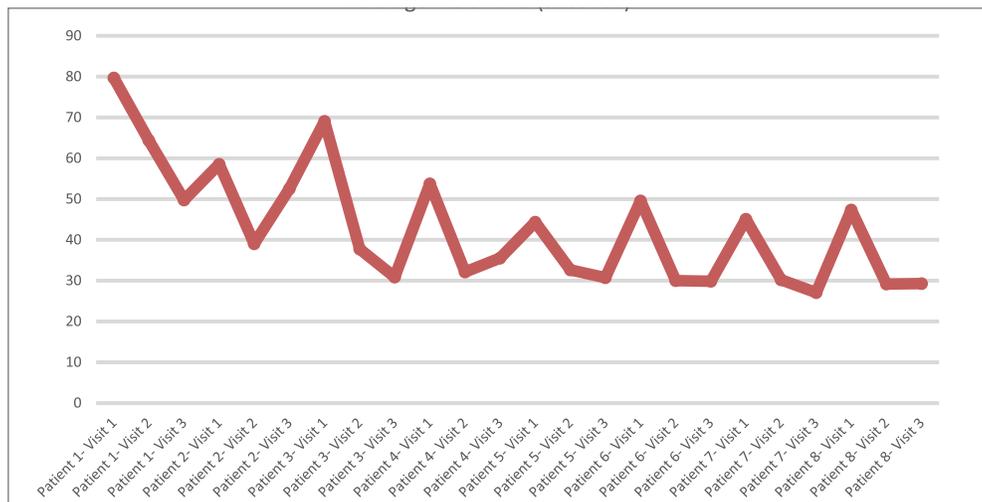
Figure 5: Data from Alzheimer's Disease Study (n=105), Conducted with Primary Care Physicians and Psychiatrists in the United States (Nov 2016)

Visit 1   Visit 2   Visit 3

■ memantine only   ■ AChEI only   ■ memantine + AChEI (including Namzaric)   ■ Neither memantine nor AChEI
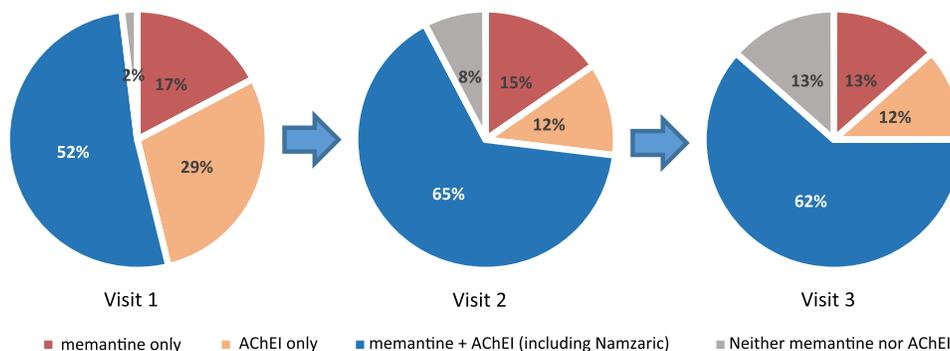
Physician Behavior: Information Seeking And Prescribing

At this point, we can move beyond physician interaction with the simulation and consider the actual data generated from it. In broad terms, the simulation produces two categories of data. The first relates to diagnostic and treatment decisions selected for specific virtual patients. When considered over multiple visits in virtual time, the resulting data is analogous to longitudinal patient data—with the exception that each individual virtual patient is seen by a sample of respondents. Hence the unit of statistical analysis is primarily the physician rather than the virtual patient. As seen in Figure 5, we can readily observe how a sample of physicians will treat a *specific* mild-moderate Alzheimer's disease patient over three visits, each separated by three months of virtual time.

While this example shows longitudinal results at the drug class level, the simulation can easily capture information that enables analysis of drug choice by molecule, strength, form and dosing. The nature of the simulation interface allows for very rapid collection of this detailed treatment information, a design characteristic of particular importance for areas like oncology,

Figure 6: Data From Schizophrenia Study (n=73), Conducted with Psychiatrists in the United States in Feb 2016 ; Average Number of Times Category Accessed per Patient, by Visit Number
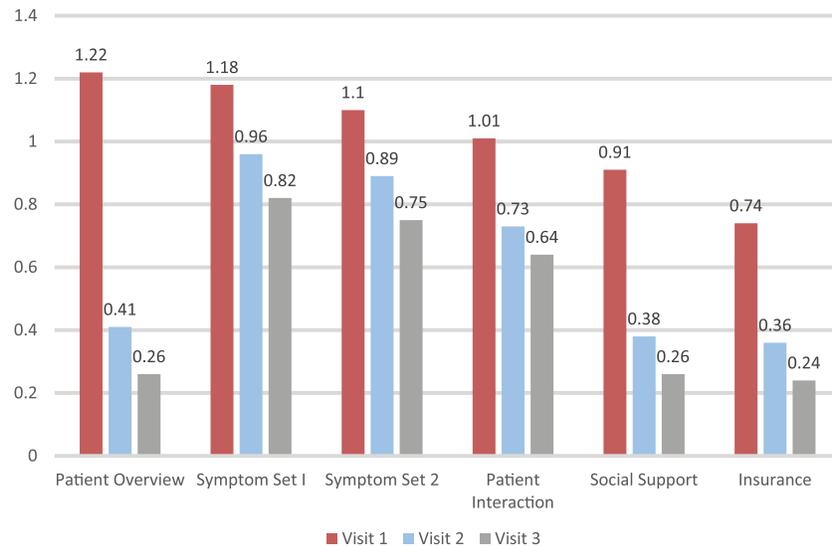
where customized cocktails of treatments may be utilized which are extremely difficult to pre-code in standard questionnaires.

The second major type of data generated by the simulation relates to the decision-making process of the physician, as reflected in the virtual patient information accessed. Unlike traditional surveys in which the survey design attempts to force respondents to view all the information provided, the virtual patient simulation as implemented here takes a radically different approach. Physicians are free to view whatever patient information they choose—multiple times or not at all, and in any particular order they wish. As in the real world, more potentially relevant information is provided than physicians will actually absorb in making their treatment decisions. What information they choose to pay attention to should correlate with information that is most likely to drive their prescribing decisions. Moreover, given that physicians are not homogeneous in their perspectives on patient treatment, information access can also provide

the means to segment physicians based on decision-making style.

Figure 6 shows the average number of times that physicians accessed different categories of information in a pilot study on schizophrenia. Each category corresponds with a category of information presented in the simulation interface and viewable with the click of a button. As is evident from the graph, on average, physicians viewed most categories of information at least once on the first visit of the virtual patient. The average frequency of access then declined on the second and third patient visits. For some categories like patient overview, which contained static information such as age and gender, it is not surprising that frequency of access declined precipitously in visits 2 and 3. Respondents quickly learned that there was no new information in these categories and therefore tended to ignore them as the simulation went on. But what about categories such as symptom set 1, symptom set 2 and patient interaction which contain dynamic information—information which

Figure 7: Results of Physician Segmentations Produced Through K-Means Procedure for Each Virtual Patient. Shaded Percentages are Significant at >95%.

	Segment Size	% Viewing Symptom Set 1	% Viewing Symptom Set 2	% Viewing Interaction
Patient 1				
Segment 1	44%	100%	100%	100%
Segment 2	48%	100%	100%	94%
Patient 2				
Segment 1	58%	100%	100%	98%
Segment 2	27%	100%	100%	100%
Patient 3				
Segment 1	47%	100%	96%	100%
Segment 2	38%	100%	96%	68%
Patient 4				
Segment 1	79%	100%	98%	100%
Segment 2	21%	73%	73%	27%
Patient 5				
Segment 1	45%	97%	100%	100%
Segment 2	44%	100%	100%	63%
Patient 6				
Segment 1	55%	100%	100%	100%
Segment 2	29%	100%	100%	76%
Patient 7				
Segment 1	42%	97%	97%	97%
Segment 2	44%	100%	97%	69%
Patient 8				
Segment 1	60%	100%	98%	100%
Segment 2	29%	91%	86%	24%

changed with each visit? Frequency of access also declined here on second and third visits, although not as dramatically as for static categories.

In order to better understand the dynamics of information access observed in the simulation, we conducted a simple segmentation of respondents (SPSS K-means procedure) based on the categories of information physicians accessed on an individual patient basis across all three visits. Included in this analysis were the three "dynamic" categories of information: symptom set 1, symptom set 2 and patient interaction, as well as two "static" categories: social support and insurance.

Our analysis produced a three cluster solution for each of the eight virtual patients included in the simulation. Across all virtual patients, the two largest of these segments accounted for at least 85% of respondents. Given the limited sample size (73 responses per patient), these two segments became the focus of our analysis. Figure 7 displays the dynamic categories of information accessed by these two segments

of respondents for visit 1 for each of the eight virtual patients. As is evident from the table, virtually all physicians, regardless of segment, viewed the two dynamic information categories related to symptom set 1 and symptom set 2. Segment One, however, extended this thoroughness to the patient interaction category, which included brief narrative descriptions of the patient or caregiver's subjective assessment of the patient's situation. In contrast, Segment Two was generally less likely to view this category of information.

As is evident in the table, both the size of segment two and the percentage of segment physicians viewing the patient interaction information varied by virtual patient. This suggests that physician decision-making style—as measured by information access—is dependent, at least in part, on the type of patient seen. In this specific case, one possible interpretation is that certain combinations of patient presentation and physician decision-making style increase the value of patient and caregiver input relative to the better defined categories of symptom sets 1 and 2.

Further work in this area needs to be conducted to develop improved analytic approaches to understanding physician decision-making within simulations. These efforts should proceed along several paths. First, information access order, frequency of access and time spent on each item should be considered alongside our current binary metric of access (viewed or did not view specific information). Second, while information access varies by virtual patient, a meta-analysis of information access across patients should be pursued to determine if there are decision-making styles which are relatively stable regardless of patient. Third, the technique gives us a unique ability to understand physician decision journeys, which is of great practical importance. Finally,

simulations such as this should be linked to real world behavioral data, such as prescribing, to determine how decision-making style is related to actual behavior.

Big Picture: What Have We Learned And What's Next?

- Our early experience demonstrates that physicians find simulation to be more engaging than questionnaires and suggests that they interact with simulation in a way that reflects their real-world practice, demonstrating face validity.

- Physician behavior during the task varies across patients and across visits with the same patient, demonstrating the validity of the procedure. As such, these simulations can provide insight into what information physicians attend to in making treatment decisions, adding to our understanding of physician decision making.

- In quantitative applications of the simulation platform, we have been able to employ simple A/B testing with virtual patient clones to quantify the impact of specific patient characteristics on prescribing.

- In qualitative applications of the simulation, we have been able to use the realism of the simulation to drive deeper insights into physician decision-making. Most notably, we have found the simulation to be extremely valuable in conducting cognitive interviews with physicians. In one cognitive interview study, the insights we developed around the treatment algorithm radically altered the client's in-going perceptions of what was originally thought to be a simple, sequential algorithm.

- In general, client teams have said that the technique provides insights and learning that they did not uncover in previous research. This has encouraged us to develop simulation as a tool for both quantitative and qualitative research applications.

- Going forward, we intend to employ more sophisticated experimental design manipulations in patient symptoms and patient interaction in conjunction with environmental parameters (insurance, competitive activity, etc.) to understand the effects of these variables on physician prescribing behavior

At this early stage, we have only scratched the surface of the analytic possibilities of such simulation data. Much work remains to be done, but initial results from these pilot tests in simulation suggest that the effort promises significant rewards.

About the Authors

Chandra Chaterji *is an independent consultant in analytics and strategic marketing. He earned his stripes in marketing science, market research and brand strategy, in companies such as Visa, J. Walter Thompson and BBDO. He has presented papers at the INFORMS Marketing Science and the AMA's Advanced Research Techniques conferences. An ex-academic, he has taught in the graduate business programs at Loyola of Chicago, Rutgers and UC Davis.*

Greg Chu *is Chief Operating Officer at InTask, Inc. where he oversees the development of simulation platforms for use in healthcare market research. His career in healthcare includes leadership positions in commercial research and analytics at Merck & Co., Forest Labs, Novartis, Synovate Healthcare and Ipsos Healthcare.*

Value Proposition of Diagnostic Test Data

Kishan Kumar, Associate Director, Axtria and Juhi Parikh, Project Lead, Axtria

Abstract: With the current focus on precision medicine, targeted therapies are already part of the standard of care for several cancers, with many more in clinical trials and on the path to commercialization. With the advancement of new age data sources, it becomes imperative for pharmaceutical manufacturers to understand the value proposition of this diagnostic test data. Building a data capability around these datasets, although not a trivial undertaking, can result in a powerful asset to be used along with other traditional or patient-level data sources. With the evolution of more personalized medicine, targeted therapies focused on specific biomarkers and a robust oncology pipeline in the near future, the value of leveraging 'data driven' insights into their commercial and clinical planning takes a higher precedence. Over the past two years, we have been engaged in exploring the value addition that this data provides to facilitate several decision making processes in commercial, clinical and market access teams. Our exploration has led to pilot implementations of several initiatives, most of which has yielded valuable insights. This paper aims to walk the readers through a journey of this exploration.

Keywords: Diagnostic test data, Lab test data, Oncology, Biomarkers, Applications, Targeted therapies

Introduction

In today's world of unparalleled technological breakthroughs and scientific advancements, personalized health care has the capacity to detect the onset of disease at its earliest stages, pre-empt the progression of disease, and, at the same time, increase the efficiency of the health care system by improving quality, accessibility, and affordability. In the 12 years since the completion of the Human Genome Project (HGP), advances in genome technology have led to an exponential decrease in sequencing costs (more than 16,000-fold). Patients have benefited from major biological insights and medical advances, including the development of more than 100 drugs whose labels now include pharmacogenomics information.[1]

The opportunity that the field of personalized medicine offers is the potential for advancements in science and medicine. Pharmaceutical manufacturers can now target populations of patients into groups who have a greater susceptibility to respond to a particular treatment. Through this personalized medicine, patients will not just benefit from better treatment but also in early detection and prevention of diseases and disorders. As the advancements in personalized medicine progress, the availability of data to benefit commercial and clinical teams have also increased exponentially. In this paper, we will discuss the value proposition and potential that the data from diagnostic testing laboratories offers to commercial and clinical teams, specifically within oncology manufacturers. In this paper, the usage or references to the word lab tests or diagnostic tests refers to the biomarker testing.

Traditional and Targeted Therapies in Oncology

In order to fully realize the potential value of using biomarker data or lab data, it is important

to gather an understanding of targeted therapies. Traditional or standard chemo drugs work by killing cells in the body that grow and divide quickly. Cancer cells divide quickly, which is why these drugs often work against them. But chemo drugs can also affect other (normal) cells in the body that divide quickly, which can sometimes lead to serious side effects. The reason is that these chemo drugs are unable to differentiate between healthy cells and cancer cells. Each time chemo is given, the chemo tries to find a balance between killing the cancer cells (in order to cure or control the disease) and sparing the normal cells (to lessen side effects).

Targeted therapy, on the other hand is a newer type of cancer treatment that uses drugs or other substances to more precisely identify and attack cancer cells, usually while doing little damage to normal cells, resulting in prolonged stability of tumor. Targeted cancer therapies are drugs designed to interfere with specific molecules necessary for tumor growth and progression. Traditional cytotoxic chemotherapies usually kill rapidly dividing cells in the body by interfering with cell division. A primary goal of targeted therapies is to fight cancer cells with more precision and potentially fewer side effects.[2] Oncology has been a leader in leveraging and developing products based on these targeted therapies. To-date there are over 200 FDA-approved drugs with pharmacogenomic information in their labeling.[3]

Candidates for Targeted Therapy – Introduction to Diagnostic Testing
For certain types of cancer, most patients will have an appropriate target (a marker or mutation) for a particular targeted therapy and, thus, will be candidates to be treated with that therapy. For example, in the case of CML, most patients have the BCR-ABL fusion gene. For some other cancer types, however,

a patient's tumor tissue must be tested to determine whether or not an appropriate target is present. The use of targeted therapies may be restricted to patients whose tumor has a specific gene mutation that codes for the target; patients who do not have the mutation would not be candidates because the therapy would have nothing to target. Sometimes, a patient is a candidate for a targeted therapy only if he or she meets specific criteria (for example, their cancer did not respond to other therapies, has spread, or is inoperable). These criteria are set by the FDA when it approves a specific targeted therapy.[4]

Treatment of patients on targeted therapies starts with diagnosis and testing for cancer. Diagnostic tests are done with samples collected at the time of a first biopsy. This testing may be referred to as molecular profiling, biomarker testing or tumor testing which all implies the same meaning. The objective of these tests are to determine the mutations that may have occurred in the gene and to identify the occurrence and severity of cancer. The testing practices are intensely debated, impacting diagnostic quality and affecting pathologists, oncologists and patients. There are some slide based testing techniques such as in-situ hybridization (IHC) or Immunohistochemistry (ISH). In recent times, the most commonly used or evolved technique is the Next Generation Sequencing (NGS). NGS is increasingly used in the clinics, most commonly in the form of targeted gene panels that are custom tailored for specific diseases. The FDA has approved multiple targeted drug cancer therapies, and many more are being studied in clinical trials either alone or in combination with other treatments. Some of the commonly known, currently approved targeted therapies for solid malignancies and their molecular targets is provided in Table 1.[5]

Table 1: Approved Targeted Therapies for Solid Malignancies & Their Molecular Targets

Agent	Target(s)	FDA-approved indication(s)
Erlotinib (Tarceva)	EGFR (HER1/ ERBB1)	Non-small cell lung cancer (with EGFR exon 19 deletions or exon 21 substitution (L858R) mutations), Pancreatic cancer
Everolimus (Afinitor)	mTOR	Pancreatic, gastrointestinal, or lung origin neuroendocrine tumor, Renal cell carcinoma, Nonresectable sub ependymal giant cell astrocytoma associated with tuberous sclerosis, Breast cancer (HR+, HER2-)
Ipilimumab (Yervoy)	CTLA-4	Melanoma
Imatinib (Gleevec)	KIT, PDGFR, ABL	GI stromal tumor (KIT+), Dermatofibrosarcoma protuberans, Multiple hematologic malignancies including Philadelphia chromosome-positive ALL and CML
Sorafenib (Nexavar)	VEGFR, PDGFR, KIT, RAF	Hepatocellular carcinoma, Renal cell carcinoma, Thyroid carcinoma
Trastuzumab (Herceptin)	HER2 (ERBB2/ neu)	Breast cancer (HER2+), Gastric cancer (HER2+)
Bevacizumab (Avastin)	VEGF ligand	Cervical cancer, Colorectal cancer, Fallopian tube cancer, Glio-blastoma, Non-small cell lung cancer, Ovarian cancer, Peritoneal cancer, Renal cell carcinoma

Diagnostic Testing Landscape

Effective usage of diagnostic test data starts with a thorough understanding of the testing landscape. "Not all labs are created equal" – i.e., different genetic testing laboratories will focus on different disease states and different segments of the market. The focus may be exclusively academic or commercial testing. Despite some consolidation, the testing is still very fragmented, with many regional laboratories serving specific geographies. The two broad areas of testing are done in academic labs (that are more research focused) and commercial labs. In order for pharmaceutical manufacturers to benefit from these data sources, it is relatively easier and more accessible, in the near term, to focus on commercial laboratories. Even within the commercial lab setting, there are some labs who still are largely unaware of the potential that capturing and recording these datasets have. Within the commercial laboratories, there are the major labs, such as Quest Diagnostics™, Labcorp™, etc. Then there are the regional data providers, such as Foundation Medicine™, Genoptix™, Caris Life Sciences™, Clarient Health™, etc. Lastly, there are the private labs that limit their testing services to a select list of ZIP codes within a particular geography.

Understanding the Datasets

Given the high degree of specialization, many labs address specific, selected disease states. Even within a particular disease state, the testing process across labs vary markedly. This translates to significant variation in reporting test results. Based on the information that was gathered in our pilot experimentation with a handful list of data sets, we had a few observations:

- Presence of a unique patient identifier. However, the unique identifier was unique only to their databases (i.e. not universally unique).

Figure 1: Illustration of Positive Outcome Results From a Given Data Source

Test Panel	Outcomes That Determine Positive Result
EGFR Mutation Analysis	Presence of mutation e.g. del exon 19, L858R, ins exon 20
KRAS Mutation Analysis	Presence of mutation e.g. 12CYS, 12VAL,12ASP
ALK Rearrangement by FISH	ALK rearrangement in >= 15% cells
ROS1 by FISH	ROS1 rearrangement in >= 15% cells
RET by FISH	RET rearrangement in >= 15% cells
MET by FISH	Signals/nucleus ratio >= 5.0
EGFR Amplification Analysis Panel	% Abnormal cells > 20%
FGFR1 by FISH	Signals/nucleus ratio >= 5.0

Figure 2: Illustration of the Availability of Information Represented in Three Lab Datasets for a Two Year Period (2014 - 2016)

Laboratory Data Source	Source contents/ fields					
	Unique Patient ID	Account Name	Physician ID	ZIP	Managed Markets	Molecular information
	IDs are unique to each data source	Information on test site/ facility	Unique ID such as NPI,	ZIP of facility/ Test site	Basic info like plans, Medicare	Test Panel, Results and Clinical info
1 Source A	✓	✓	✓	✓	✓	✓
2 Source B	✓	✗	✗	✓	✓	✓
3 Source C	✓	✓	✗	✓	✓	✓

- Presence of ZIP code information corresponding to each patient. This represented the zip code of the test site or facility.

- Presence of molecular level information. Figure 1 illustrates some examples of molecular information represented in these data sets.

- Besides this information, there was other relevant information pertaining to physician ID (such as NPI, DEA, etc.)

or payer information for the patient present in these datasets.

A quick illustration of the contents in the datasets is represented in Figure 2.

The test results represented by the data sources are very granular, down to outcomes at the genomic level. As represented in Figure 3, the datasets represent mutation level data for various biomarkers. The information present shows the number of patients screened for each

26

Figure 3: Illustration of Number of Screened Patients in a Test Panel for Various Markers
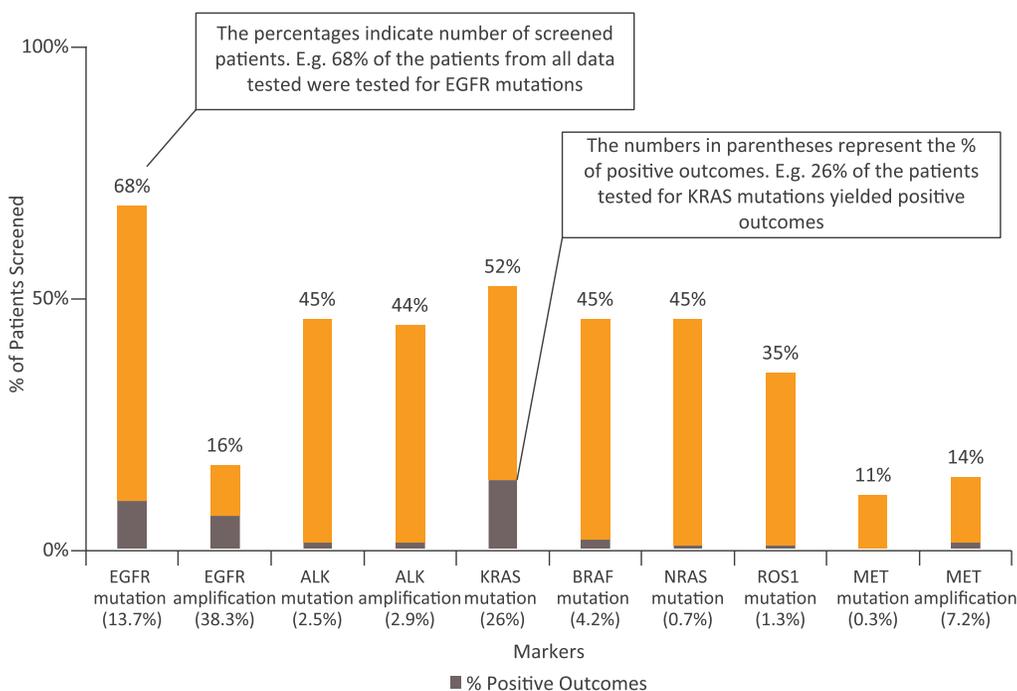

The percentages indicate number of screened patients. E.g. 68% of the patients from all data tested were tested for EGFR mutations

The numbers in parentheses represent the % of positive outcomes. E.g. 26% of the patients tested for KRAS mutations yielded positive outcomes

% of Patients Screened

100%

50%

0%

68%

16%

45%

44%

52%

45%

45%

35%

11%

14%

| EGFR mutation (13.7%) | EGFR amplification (38.3%) | ALK mutation (2.5%) | ALK amplification (2.9%) | KRAS mutation (26%) | BRAF mutation (4.2%) | NRAS mutation (0.7%) | ROS1 mutation (1.3%) | MET mutation (0.3%) | MET amplification (7.2%) |

Markers

■ % Positive Outcomes

Please note, the numbers are illustrative

Figure 4: Illustration of Information at a Granular Level From Various Data Sources

Source	% positive for EGFR mutations	Top Outcomes	Other Relevant Outcomes
Source A	12%	Del exon 19 (6%), L858R (4%), Ins exon 20 (1%)	G719X (1%), S7681(1%), T790M (0.3%),
Source B	15%	L858R (5%), E746_A750del (4%)	T790M (3%), L861Q (0.3%)
Source C	13%	Del exon 19 (5%), L858R (4%), Ins exon 20 (1%)	T790M (1%)

Please note, the numbers are illustrative

panel of testing and the number of patients that had positive outcomes.

In addition to presence of outcomes (positive or negative) for each patient, some datasets even represent point mutations.   For example, an EGFR positive outcome may be "T790M" or "L858R". This is very useful when working with targeted therapies that are appropriate for patients with very specific outcomes. Illustrated in Figure 4 are some examples from various sources.

Applications of Diagnostic Data

With the landscape of the pharmaceutical industry changing from physician-driven to patient-driven, it is imperative to tap into the potential of new age data sources such as diagnostic test data. Powerful and informed insights can be derived by combining this data with existing data sources such as patient level data, claims data, clinical trials data, etc. The advancements in medical infrastructure, not just through means of scientific breakthrough, but also the investments that are in place to track the patient records, has enabled the sales, commercial,

marketing and medical teams of pharmaceutical companies to provide customized intelligence about the physician and their patients, owing to compliancy and HIPAA regulations. More than 93% of U.S. physicians today use Electronic Health Records (EHRs)[6,7] and it continues to grow. There are frameworks in place to leverage IT investments and address critical concerns such as interoperability, data sharing and complex consent. The widespread use of EHRs creates the potential for the millions of files of data they hold to be analyzed by researchers, test developers, and regulators, to better develop, refine, and understand the underpinnings and real-world applications of personalized medicine.[8]

The presence and availability of diagnostic test data is an extension to the EHR and provides more robustness into understanding the longitudinal history of patients. Through our exploration of several datasets, we identified a few potential applications.
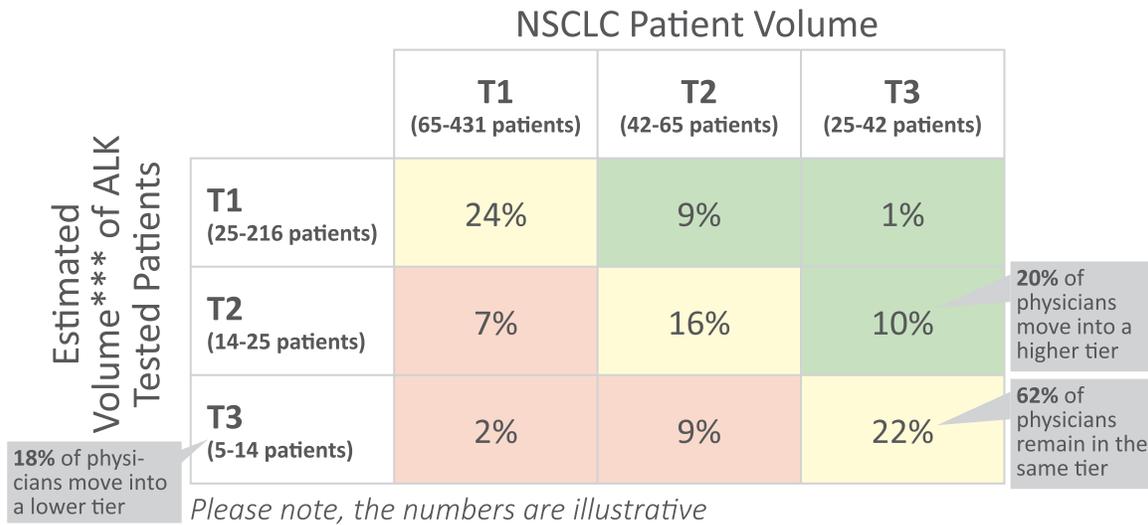
1. **Focused Physician Targeting**

The availability of patient level data enables commercial operations teams to create sophisticated targeting models. As the landscape of healthcare changed from volume driven to value driven, commercial teams leverage as many levers as possible to drive focused physician targeting. As observed in a few pharmaceutical companies, physician targeting for oncology and rare disease products differs from primary care in several ways. Because of the increased revenue per patient for an oncology product, the necessity to target the 'right' physician is exponentially increased for an oncology product compared to products in the primary care space. Information from patient level data sources and local field intelligence plays an even more crucial role here. In summary, targeting in oncology is more 'value' driven than 'volume' driven.

For oncology products, patient level data often dictates the market size depending on the type of tumor, and sales teams are structured based on the prevalence of these disease states. Using diagnostic data, teams are now able to include an additional dimension to their targeting strategy. It should be noted that diagnostic data is not likely to be used in isolation or as the primary dataset that drives targeting; it would be more useful in conjunction with other patient level and/or specialty pharmacy datasets to provide an additional layer of detail. The occurrence of datasets that allow us to see the volume of patients tested for a specific type of marker (ALK, BRAF, etc.) in combination with disease state mapping (i.e. integration with patient level data for tumor types) allows sales teams to prioritize physicians and institutions. Incorporating the diagnostic data to any existing patient level or physician level data (through unique identifiers such as physician id or ZIP codes) presents an added dimension to physician targeting.

As illustrated in Figure 5, conventional targeting approaches use information that pertains only to the physician's prescribing behavior. With the introduction of patient level data, commercial teams received some intelligence into the types of patients these physicians treat. However, this limits the group of patients to a particular broad disease state only—for example, all patients who are diagnosed with metastatic lung or breast cancer. Incorporating data from diagnostic testing allows us to identify which of these groups of patients test positive for the particular marker for which their product is being prescribed for (e.g. ALK+, BRAF+, etc.). Through extension of this, we would be able to group and differentiate physicians who treat patients for a particular type of cancer vs. a type of marker.

Figure 5: Illustration of a Segmentation Matrix That Incorporates Lab Data and Existing Patient Level Data to Group HCPs

NSCLC Patient Volume

Estimated Volume*** of ALK Tested Patients	T1 (65-431 patients)	T2 (42-65 patients)	T3 (25-42 patients)
T1 (25-216 patients)	24%	9%	1%
T2 (14-25 patients)	7%	16%	10%
T3 (5-14 patients)	2%	9%	22%

20% of physicians move into a higher tier

62% of physicians remain in the same tier

18% of physicians move into a lower tier

Please note, the numbers are illustrative

2. Uncovering Under-Tested Populations

Through exploration of the select data sources that were available, we noticed a wide range of testing variations for most markers. This pattern is similar to other existing data sources as well. This could be attributed to a number of reasons: gaps in testing (i.e. different regions have different testing rates), differentiation in prevalence of a mutation, or capturing test data. Having a national level view of a particular type of marker data, we can identify sub-national areas of opportunity through localized variations in testing. Furthermore, building a robust lab testing capability, this data could be used to identify geographical differences in testing, including trends over time.

3. Enabling Informed Forecasting

Forecasting is a critical exercise for brand teams as it feeds into and influences many other functional areas within an organization. These linkages may be unidirectional (where forecasts feed into decisions made by the other functional areas) or bidirectional (where the

forecast is used to quantify the effects of market changes envisioned by other functional areas).[9] The challenges with the correct number of patients, lines of therapies and inclusion of the appropriate biomarkers makes it increasingly difficult for forecasters to get the most accurate results. The traditional approach in forecasting for oncology manufacturers bakes in a variety of assumptions and metrics from literature as well as primary and secondary data sources to define critical inputs like testing rates and outcomes. Having insights into the real world results from diagnostic data leads to better accuracy in estimating the size of patients and positivity rates for brand teams.

4. Applications to Clinical and Medical Teams

In addition to its applications in commercial decisions, medical/clinical teams can be benefited from using these diagnostic data. One of the most efficient uses of these datasets is in identification or evaluation of clinical trial test sites. The process to select and evaluate

sites can be more streamlined and efficient, as localized testing information can provide an initial pool of candidates to choose from. They can also help in designing a trial based on the spread of population of patients. As pharmaceutical medical/clinical teams continue to educate physicians, they can influence their testing behavior to fit their clinical trial needs.

Limitations

In oncology, commercial activities like sizing the market, benchmarking versus competition, identifying the right targets, and developing the appropriate customer messaging are already complicated by factors like disease staging, line of therapy, metastatic versus adjuvant therapy, combination therapy, and off-label prescribing. In the case of targeted therapies, which are developed for very specific patient sub-populations, this becomes even harder. In order to appropriately leverage this diagnostic test data source, a deeper understanding of the limitations and gaps becomes critical. In the earlier sections of this paper, we had discussed the differentiation and 'non-standardization' amongst labs. Additionally, the fragmentation of testing leads to gaps in coverage of data availability, almost seemingly impossible to get 100% coverage, at least in the near term. As the emergence of data providers and data aggregators mature, we may get much closer to a substantial sample size. There is currently no standardized panel across labs that goes through a specific set of mutations sequentially for a particular type of tumor. The protocols for completing a panel may differ—one lab may complete a full panel for every patient, another may prioritize tests and provide varying set of results by patient. The physicians who order or request the test may or may not be able to specify or prioritize mutations to be tested. Hence, if there are manufacturers who have products that are indicated for biomarkers that may not be up the priority list in a standard test panel (e.g. KRAS or ROS1), patients may not even be tested for this biomarker. Lastly, the interpretation of results may differ by lab, depending upon criteria set by the lab and leeway given to the technician. This results in the output data vastly differing; some labs may provide results down to a mutation level while others may just point out the outcome of the results whether it be positive or negative. This creates reports that are non-standardized across various labs.

Conclusions

Given the vast possibilities of exploring these data sources and a realization of the current limitations, pharmaceutical and diagnostic testing companies should start thinking along the lines of developing a data asset strategy around leveraging these diagnostic data sources. The three critical areas in the near term that companies should start working towards are 1. Investments to have a repository of these data sources, 2. Building a team over time that is dedicated to investigating and refining these data sources and 3. Developing subject matter expertise within organizations and identifying means to link these data sources to their day-to-day operations for use by commercial, sales, medical and marketing teams. Soon enough, this would become a norm in a catch-up pharma world to enable better patient treatment options and more informed commercial decisions.

Acknowledgments

About the Authors

Kishan Kumar *is an Associate Director at Axtria, Inc. He has over 10 years of experience in management consulting and working across several life sciences and pharmaceutical clients with expertise in sales force effectiveness, commercial model designs and working with patient level data. Over the past few years, Kishan has worked with several oncology manufacturers to gather an understanding of their commercial models and various data assets used by their brand teams. Prior to his tenure at Axtria, Kishan worked for a few global consulting firms including Cognizant, Alexander Group and MDRx financial. He has a Master's degree in Biotechnology from University of Pennsylvania, Philadelphia.*

Juhi Parikh *is a Project Lead at Axtria, Inc., with 7 years of experience in analytics, consulting and market research. At Axtria, she has worked with several leading pharmaceutical clients on analytics related to sales, marketing and commercial operations. Juhi is experienced in analytics related to real world evidence, oncology, commercial model design, and commercial operations, with deep expertise in patient level sources such as claims data and diagnostic test data. Prior to Axtria, Juhi worked with Nielsen's Innovation Practice, conducting new product market research and forecast analytics for leading CPG and OTC pharma brands.*

References

[1] Personalized Medicine Coalition (PMC) – The case for personalized medicine – 4th Edition 2014.

[2] Felix W. Frueh, Ph.D., Shashi Amur, Ph.D., Padmaja Mummaneni, Ph.D., Robert S. Epstein, M.D.,Ronald E. Aubert, Ph.D., Teresa M. DeLuca, M.D., Robert R. Verbrugge, Ph.D.,Gilbert J. Burckart, Pharm.D., and Lawrence J. Lesko, Ph.D., Pharmacogenomic Biomarker Information in Drug Labels Approved by the United States Food and Drug Administration:Prevalence of Related Drug Use (Pharmacotherapy 2008;28(8):992–998).

[3] U.S. FDA - Table of Pharmacogenomic Biomarkers in Drug Labeling Source:http://www.fda.gov/drugs/scienceresearch/researchareas/pharmacogenetics/ucm083378.htm.

[4] National Cancer Institute - Targeted Cancer Therapies – Source: https://www.cancer.gov/about-cancer/treatment/types/targeted-therapies/targeted-therapies-fact-sheet.

[5] My Cancer Genome – Source: https://www.mycancergenome.org/content/molecular-medicine/overview-of-targeted-therapies-for-cancer/.

[6] Building Digital Trust: The role of data ethics in the digital age - Source: http://www.accenture.com/gb-en/Pages/insight-digital-doctor-is-in.aspx.

[7] Accenture newsroom - Source: https://newsroom.accenture.com/industries/health-public-service/emr-and-hie-use-increases-among-us-doctors-accenture-annual-survey-finds.htm.

[8] Degatano M, Sorokina T, Smyth C. National lab test database provides valuable marketing insights from all stages of the patient journey. *Journal of the Pharmaceutical Management Science Association.* 2015; Spring:13-21.

[9] Forecasting for the pharmaceutical industry – Arthur G. Cook. 2006.

Leveraging Predictive Analytics to Derive Patient Adherence Drivers

Ewa J. Kleczyk, Ph.D., Executive Director, Commercial Effectiveness Analytics, Symphony Health Solutions and Derek Evans, Senior Vice President, Symphony Health Solutions

Abstract: Understanding therapy adherence and its factors is an important part of managing healthcare costs and improving patients' health outcomes. Leveraging the Cox Promotional Hazard Model and health claims data can aid in identifying the 'risk factors' to staying on the treatment over time, and inform optimized and efficient resource deployment strategies. This article will review the methods often utilized for adherence evaluation, as well as introduce a case study, assessing the factors driving adherence long-term.

Keywords: Adherence, Risk factors, Cox Proportional Hazard Model, Health claims data

Introduction

Patient adherence to therapy is an essential part of improving patient health outcomes. Non-adherence to therapy can increase healthcare costs in the long term, as it can lead to increased comorbidities and concomitant therapies, as well as decreased quality of life. It is estimated that the economic cost of non-adherence to treatment is at 100 billion to 300 billion dollars.[1] This cost includes lost wages, which result from increased disease burden, in addition to the cost of the avoidable health care.

Understanding not just the current level of adherence within a population, but the factors, which impact adherence, are an important part of managing healthcare costs. This article will review the currently leveraged measures of adherence, and the 'Whys' driving patient's continuation on the prescribed treatment. It will also discuss a case study, assessing adherence and its drivers.

Practices for Measuring Adherence

Patient adherence to treatment is important to improving patient outcomes. There are a few standard measures of adherence, including Medical Possession Ratio (MPR), Proportion of Days Covered (PDC), Refill Compliance Days (RCD), Continued Measure of Medication Gaps (CMMG), Medication Refill Adherence (MRA), and Duration of Therapy (DT), to name a few.[2] The most often employed approaches to measure adherence include MPR and PDC metrics.

MPR is best used for assessing adherence for a single product by measuring the total days a product is supplied to a patient verses the total time, which elapses over the treatment period. Generally speaking, MPR is best for measuring adherence for a single product with a daily treatment schedule, such as the case with diabetes and hypertension.[3]

PDC, on the other hand, allows for the assessment of adherence when concomitant therapies should be taken into consideration. It is better suited for therapeutic areas where treatment switching and multiple-medication use is common.[3] Multi-therapy treatment often occurs in the oncological and immunological therapeutic areas.

According to the Pharmacy Quality Alliance (PQA), MPR can be biased, as it tends to overestimate adherence, as it double counts overlap periods when switching therapies.[3] Due to these limitations PDC has become the more preferred measurement of adherence; however, it may underestimate adherence rates in situations where a patient refills their medication earlier than scheduled.[3]

While these measures inform whether the level of adherence to treatment is within a given range for a given population of patients, they are static measures, which do not provide insights into the factors which are contributing to adherence within a particular patient population.

Understanding Drivers of Adherence

To enhance the adherence analysis, it is beneficial to understand the factors driving product continuation overtime. These drivers are usually referred to as 'risk factors' to adherence, and include patient's cognitive ability, attitudes towards treatment, demographics and socioeconomic variables[4], the number and types of comorbidities, and concomitant therapies. The 'risk factor' may either increase or decrease the probability that a patient will remain adherent to treatment.[5]

For example, the amount of information shared by a physician, as well as a patient's ability to retain and remember the physician's treatment recommendations, has a significant impact on the likelihood for staying on therapy long-term.[4,6] On the other hand, patient adherence tends to decrease for treatments aimed for disease prevention.[4] Demographic, socioeconomic, and health-related cost variables are also often cited as 'risk factors' in empirical research, especially as it relates to the economic variables such as out-of-pocket costs accrued by patients.[7]

Identifying significant 'risk factors' allows brand marketing teams to utilize and optimize their resources more efficiently, based on the factors' size and probability of non-adherence, and ultimately implement a marketing strategy in a way that is both effective and timely. For example, if changing managed care organizations (MCOs) decreases adherence to treatment, the brand team can design programs proactively that may offer physician support in obtaining prior authorization approvals after a plan change in order to prevent switching or dropping therapy.[8]
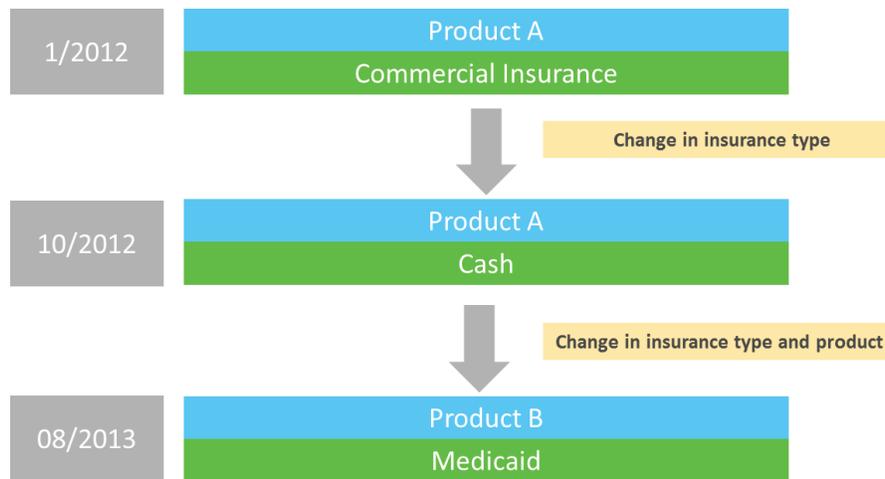
Adherence Research Approach Overview

There are many qualitative and quantitative ways to identify and measure the impact of 'risk factors' on patient adherence. For example, researchers might conduct qualitative research studies that include face-to-face interviews with patients or physicians to facilitate collection of information related to patient's adherence and perceived barriers to staying on therapy over time. These types of studies tend to inform cultural, cognitive, brand perceptions, and physician/patient relationship based factors.[9]

As qualitative research studies are often limited by the sample size of participants, patient or physician quantitative survey studies might inform additional adherence 'risk factors,' including impact of out-of-pocket cost or comorbid conditions on adherence.[9]

Working with the primary market research data may sometimes result in a limited ability to collect in-depth information on behavior-based topics of interest or track the actual behavior of the respondents. As a result, it is often recommended to leverage secondary data to help provide additional insights not captured in the survey based approaches.[10] For example, actual patient history available from health claims data can be included in analyzing adherence to medical treatment to provide more comprehensive and detailed insights into the patient treatment journey to accurately measure

Figure 1: Explaining Time Varying Covariates

1/2012	**Product A**
	Commercial Insurance

Change in insurance type

10/2012	**Product A**
	Cash

Change in insurance type and product

08/2013	**Product B**
	Medicaid

and estimate the impact of selected drivers on the length of staying on the therapy. More detailed information about health claims data will be provided in a later section of this article.

The econometric methods often leveraged to estimate the impact of the 'risk factors' on patient adherence include techniques from simple correlation analysis to multivariate cross-sectional and time-series regression analysis (i.e. logistic regressions). These types of methodologies allow to either identify correlations between variables and discontinuation of therapy, and/or in addition, identify the causality and regression estimates of each individual factor on patient adherence.[5]

In this article, the Cox Proportional Hazard Model is discussed as an approach to study adherence of medical treatments. The technique can be leveraged to identify drivers of adherence otherwise known as 'risk factors,' and the level of impact healthcare variables have on patient adherence to treatment.

Leveraging Cox Proportional Hazard Model for Deriving Adherence 'Risk Factors'
The Cox Proportional Hazard Model represents a group of survival models, often used in biostatistics, that relate the time that passes before some event occurs to one or more covariates that may be associated with that quantity of time.[11]

The original model, introduced by Cox in 1972, was adopted in economics and business areas, and several variations and changes were proposed to the base model over the years. These changes included accounting for unobserved heterogeneity and selection of entry for non-random entry models.[12]

In this article, the Cox Proportional Hazard Model is introduced for estimating the impact of 'risk factors' on drug adherence. The benefits of the model include capturing time variance, which allows modeling of not just the overall impact of an event, but the impact that the event has over time based on its timing relative to other events. This ensures that the 'risk factors' probabilities and their associated impacts account for the time dynamics on adherence. Inclusion of time allows for the estimation of the probability of adherence at a selected point in time, while taking into account the impact of multiple factors.[13] The concept of time varying covariates in the survival model is presented graphically in Figure 1.

Defining the 'Risk Factors' to Adherence
The Cox Proportional Hazard Model is often defined via the following functional form:

1. $\lambda(t|z)=\lambda_o(t)ez_1\beta_1+\cdots+z_p\beta_p = \lambda_o(t)ez_T\beta,$

where z is a p×1 vector of covariates (represented here as 'risk factors'), such as treatment indicators, prognostic factors, etc., and β is a p×1 vector of regression coefficients (Cox, 1972).[11]

Furthermore, $\lambda(t|z = 0) =\lambda_o(t)$. So, $\lambda_o(t)$ is often called the baseline hazard function, and it can be interpreted as the hazard function for the population group with z = 0. The baseline hazard function $\lambda_o(t)$ in model 1 can take any shape as a function of t. The only requirement is that $\lambda_o(t) > 0$.[11]

For any two sets of covariates (represented here by 'risk factors') z_o and z_1,

2. $\lambda(t|z_1)\lambda(t|z_o) = \lambda_o(t)ez_{T1}\beta\ \lambda_o(t)ez_{To}\beta =e\ (z_1-z_o)T\beta$, for all t ≥ 0,

which is a constant over time. Equivalently,

3. $\log\lambda(t|z_1)\lambda(t|z_o) = (\ z_1 -z_o)_T\beta$, for all t ≥ 0.

With one-unit increase in z_k while other covariate values are being held constant, the formulation is presented as:

4. $\log\lambda(t|z_{k+1})\lambda(t|z_k) = \log(\lambda(t|z_{k+1}))-\log(\lambda(t|z_k)) = \beta_k.$

Therefore, β_k is the increase in log hazard (i.e., log hazard-ratio) at any time with unit increase in the kth covariate z_k. Equivalently,

5. $\lambda(t|z_{k+1})\lambda(t|z_k) =e\beta_k$, for all t ≥ 0.

So $e\beta_k$ is the hazard ratio associated with one-unit increase in z_k. Furthermore, since

6. Probability $[t \le T<t+\Delta_t|T \ge t,z] \approx \lambda(t|z)\Delta_t,$

than

7. Probability $[t \le T<t+\Delta t|T \ge t,z_{k+1}]$ / Probability $[t \le T<t+\Delta t|T \ge t,z_k] \approx e\beta_k$, for all t ≥ 0.

$e\beta_k$ can be interpreted as the ratio of two conditional probabilities of discontinuation of prescribed therapy in the near future given that the patient is on therapy at any time t. [11]

Since

8. $\lambda(t|z_{k+1})-\lambda(t|z_k)\ \lambda(t|z_k) =e\beta_{k-1,}$

as a result, $e\beta_{k-1}$ can be interpreted as the percentage change (increase or decrease) in hazard defined as discontinuation of prescribed therapy with one-unit increase in covariates (otherwise known as 'risk factors'), z_k while adjusting for other covariates.[11]

Using Health Claims for Adherence Analysis
The healthcare industry is continuously generating large amounts of data. This is driven by record keeping, compliance & regulatory requirements, new technologies, and patient care. Historically, most data were stored in hard copy form and were very static (i.e. paper files, x-ray films, scripts). The current trend is moving toward rapid digitization of these large amounts of data, and new technologies that regularly monitor a host of variables. As a result, volume, velocity, and variety of healthcare data is rapidly changing. This rapid change is being driven by mandatory compliance and reporting requirements with the goal of improving the quality of healthcare delivery, while reducing healthcare costs. These massive quantities of data (often referred to as 'big data') hold the promise of supporting a wide range of medical and healthcare

Figure 2: Health Claims Data Overview[19]

	ICD	CPT	Labs	Medication	Clinical Notes
Availability	High	High	High	Medium	Medium
Recall	Medium	High	Medium	Inpatient:High Outpatient: Variable	Medium
Precision	Medium	High	High	Inpatient:High Outpatient: Variable	Medium High
Format	Structured	Structured	Mostly Structured	Structured and Unstructured	Unstructured
Pros	Easy to work with, a good approximation of disease status	Easy to work with, high precision	High Data Validity	High Data Validity	More details about doctor's thoughts
Cons	Disease codes often used for screening, therefore disease might not be there	Missing Data	Data Normalization and Ranges	Prescribed Not Necessary Taken	Difficult to process

functions, including among others clinical decision support, disease surveillance, and population health management.[14,15]

Using health claims data is often recommended for adherence modeling and finding 'risk factors' for staying on therapy long-term. For the purposes of this article, data for billing purposes (e.g. CPT and ICD 9/10 codes) is introduced along with medication (NDC codes) consumption in the healthcare claims bundle. While alternative data sources, such as clinical data (electronic medical or healthcare records (EMR or EHR), and medical images), lab data or genomics data and behavioral data are all readily available, the volume, variety, velocity, and veracity might not be present in alternative data sources.[16,17,18] Examples of data assets available in the healthcare industry for analysis are shown in Figure 2.

Health claims can help capture the patient's interaction with the healthcare system in a comprehensive manner over an extended duration of time. They are almost universally available, and have the benefit of being comprised of structured data, which aids in processing and analysis. Claims data are often at the patient level, and provide information on drugs dispensed by pharmacies, procedures performed, including those often used in oncological and immunological areas, as well as plan and copay information, and approvals of treatment by providers. Claims data often includes hospital transactional data that allow for linking of in- and out- patient treatment to further enhance the completeness of patient care.[20]

The ability to track patient longitudinally can often be preserved to a high degree via collecting claims data overtime based on known patient variables common between the different datasets (i.e. name, address, date of birth). Companies collecting healthcare data often have their proprietary algorithms to preserve the longitudinal data information, and merge information over time. On the other hand, changes in insurance providers might not always allow to capture the claims longitudinally if the data aggregation does not include all variables important for the process of assembling patient level datasets.[21]

Figure 3: Patient Level Health Claims Data Overview[20]

90% of Rxs dispensed in U.S. & territories

274 million ACTIVE PATIENTS

1.8 million prescribers tracked in IDV

Health Claims Data

More than 10,000 Plans Tracked

4+ YEARS Average years of patient existence in IDV

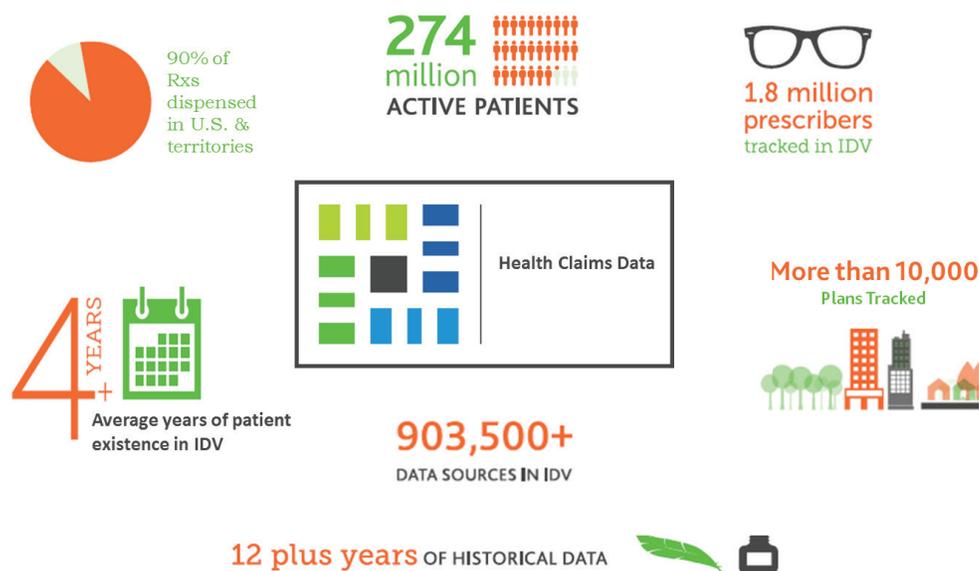903,500+ DATA SOURCES IN IDV

12 plus years OF HISTORICAL DATA

Figure 3 presents an example of a patient level health claims database, which brings together vast claims sources—medical, hospital, and prescription—to offer a consistent view across prescriber, payer, and patient dimensions. The database captures longitudinal information on more than 274 million patients over the last 12 years, which allows an understanding of patient medical history, insurance plan changes, diagnosis, comorbidities, selected treatments, and in- and out- patient care. The patient level information can be easily associated with more than 1.8 million healthcare providers, as well as 13.1 million employer groups within the healthcare environment, to allow for cost benefit analysis, as well as provider segmentation and targeting.[20]

With specific views and tools, the patient level claims data can answer key questions and facilitate critical commercial processes within sales, marketing, and managed markets. For example, it can present a comprehensive view of a given health event, allowing for its evaluation from many different angles, and development of

insight-driven strategies and programs. By enhancing claims data with information from non-retail invoices and point-of-sale data, and then applying adjustments for products abandoned at the pharmacy, the health claims data provides a detailed and nearly complete view into the brand's journey.[20]

The health claims data is often aggregated to the healthcare provider level to provide a comprehensive physician level view, allowing for in-depth analysis of treatment algorithms. The healthcare provider data can often be merged and supplemented with other data sources from other industry via using unique physician identifiers such as ME, NPI and/or DEA numbers. Most physician level data sets include at least one of the identifiers, allowing for a comprehensive analysis across disparate data sources.[20]

Although there are many benefits to using health claims data as presented above, the data might not always fully capture patients' medical history, due to the differential rates of capturing

medical claims or a lack of reporting of laboratory test results. For example, hospital claims usually have low coverage rates of rendered services and might include inconsistent reporting formats, which impact the ability to track in-patient treatments and the linking of other data sources to out-patient treatment after hospitalization.[21] The differentiated capture rates of therapies, especially for infused and injectable drugs, might also result in small sample sizes, especially for rare and orphaned disease therapies, which might cause difficulties in studying those therapeutic areas in-depth.

One of the limitations of claims data is its lack of specific confirmatory information, which results in medium recall and medium precision for characterizing patients.[22] For example, health claims data often does not include results of lab tests, which disallow tracking precisely disease progression and understanding of the physician decision making process in selecting treatments. This might be especially important in oncology and immunology, where laboratory test results impact the treatment pathways chosen for each patient.[23]

Some of the gaps in the healthcare claims data can be supplemented with the Electronic Medical Records (EMR) or laboratory claims data to provide missing information in the health claims data. For example, EMR data can often provide an in-depth and comprehensive view of patient's history over time. The EMR data can often be merged in with the health claims data; however, due to an often limited sample size for each EMR vendor, only a limited sample of patients can have fully supplemented medical history. In addition, other data assets, such as imaging, genomics, biosensor readings, and consumer and promotional events datasets, can be merged with the health claims data to provide a more comprehensive view of patient

disease progression, treatment pathways, and exposures to promotional events during the treatment decision making process.[24]

Case Study: Estimating the Adherence Drivers via Survival Analysis

Introduction
Consider a chronic therapeutic area that impacts about 10% of Americans, of whom nearly half of the affected individuals are aware or properly diagnosed. The often recommended therapies for diagnosed patients include a variety of treatments from branded products and generics, to over the counter medications. Generic and branded products account for more than 95% of the prescribed and recommended treatments.[25]

Product A is a leading branded treatment in the therapeutic area. There is also a generic version of the product, accounting for more than 80% of the market.[25] Marketing leaders would like to understand the positive and negative drivers of adherence to ensure efficient and optimized allocation of resources to drive the brand's maximum performance in-field.

The objectives of the case study, therefore, are:
1. To measure patient adherence for a key branded drug, called Product A.

2. To understand the 'risk factors' impacting patients staying on the treatment over time.

Methodology
The case study leverages patient level health claims data, as well as the Cox Proportional Hazard Model to identify 'risk factors,' and the associated level of impact of healthcare variables on patient adherence for Product A.

Analysis Assumptions:
- Timeframe: 2011 - 2015

- A stable panel of patients who used more than one treatment (including Product A), and were new to the therapeutic area at the time of study

- Right censoring: a proper adjustment based on month of entry into the market

- Testing for the statistical significance at the minimum of 90% CI of various patient groups for each of the following covariates:
  - Demographics
  - Comorbidities
  - MCO Plan Changes
  - Procedures
  - Prices
  - Transaction Types (Mail Order vs. Retail)

- Cox Proportional Hazard Model is leveraged to estimate the impact of covariates on drug adherence, defined here as the 'risk factors' to discontinuing Product A.

- The model outputs are presented as:
  - Relative Risk Ratios that represent the strength of the association with the covariate:
    - Less than 1, implies that the covariate decreases the risk for discontinuing Product A.
    - Greater than 1, implies that the covariate increases the risk for discontinuing Product A.
  - Risk Probabilities that are provided in absolute values, and their direction is noted by the accompanying Relative Risk Ratios. The probability values represent the percentage increase or decrease in risk for therapy discontinuation.

Case Study Results

Within the case study therapeutic area, overall most patients maintained their treatments long term, including Product A, and were highly loyal to their product of choice with more than 70% of patients continuing treatment after 3 years. These results are consistent with other empirical studies, in which patients treated for chronic diseases were more likely to stay on therapy long-term.[4,26] The results were also consistent between MPR and PDC measures of adherence. The patients in the stable panel had a variety of comorbid conditions, including fatigue and hormonal issues.[25]

The predictive analytics results, leveraging the Cox Proportional Hazard Model, suggest that the following 'risk factors' are important at the minimum of 90% CI in driving adherence or as defined in this study the risk of discontinuing Product A over time. As shown in Figure 4, there are several 'risk factors' impacting adherence for Product A (or increasing the risk for discontinuing Product A): plan changes, gender (F), medication burden, out-of-pocket costs, as well as selected comorbidities, including mental health diagnosis.

These results are consistent with previous empirical research, studying the impact of covariates on staying on therapy long-term. For example, in her article, Fullman found that medication adherence levels increased with third party co-insurance payments, while decreased with a higher cash burden accrued by patients.[7] In addition, mental health and specifically, depression diagnosis, was also found as one of the strongest predictors of patient non-adherence to medical treatments.[6] In this study, mental health diagnosis increases the risk for Product A discontinuation by 29%.

As cited above, gender is also an important predictor in negatively driving adherence, and

Figure 4: Risk Factors Driving Patient Adherence / Risk for Discontinuation of Product A at 90% CI

Risk Factors	Relative Risk Ratios	Risk Probabilities
Plan Changes	3.92	292%
Mail Order Delivery	0.47	53%
Mental Health	1.29	29%
Vitamin Deficiency	0.74	26%
Gender (Females)	1.15	15%
High Medication Burden	1.09	9%
Out-of-Pocket Costs	1.01	1%

increasing the risk for discontinuing Product A by 15%. This result, although perhaps counter intuitive, has been mentioned in previous empirical research studies. There are many causes of this phenomenon, including often cited income disparities between women and men, medications' side effects and efficacy impact, as well as the role women play in caring for their families in society. Since women are more often the primary caregivers, they tend to ensure their loved ones are cared for first before attending to their own needs, and are also more likely to consider how the prescribed medication will affect them and impact their day to day activities.[27]

In this case study, the highest impact of 'risk factor' driving non-adherence are changes in insurance plans. As health claims data spans across multiple insurance plans, it often allows us to investigate the price elasticity of prescribed treatments, and the impact of local managed care organizations on patient treatment choices compared to national-footprint plans. Results of such studies can aid revisions in contracting with MCOs to improve coverage, and therefore patient adherence long-term.

In this case study, insurance plan changes cause the highest rate of switching away from Product A (nearly 300% increase in risk), most often

progressing to generics. The changes usually occur January through March each year, causing some patients to drop from the market altogether, if they are not able to remain on Product A as their preferred choice. Interestingly, many patients might be on the drug in the following year, due to favorable changes to insurance coverage or help with product pre-authorization. This finding was confirmed in a previous empirical study authored by Fendrick and Chernew, in which allowing health plans the flexibility to cover more services outside of the deductible, enhanced consumer choice and increased adherence.[28]

On the other hand, mail order deliveries of Product A decrease the risk of patients discontinuing the therapy long-term by 53%. Mail order deliveries provide a more convenient way of refilling the medication, and therefore drive adherence over time. These results are in agreement with an empirical article conducted by All Scripts Holding Company, which analyzed the association between drug delivery channels and adherence. The study offered strong evidence that drug home delivery was associated with greater odds of being adherent with prescribed treatments for Medicare patients.[29]

Other covariates of interest were also included and analyzed in this case study, but they were not found statistically significant at the 90% CI. For example, patient education and income levels were not statically significant in driving Product A's adherence. Previous empirical studies present mixed results on the impact of socioeconomics variables on non-adherence, with some citing significant impact while others implying a limited correlation. The results usually are dependent on the homogeneity of the studied population, and the therapeutic area of interest.[30,31] Furthermore, blood tests, often performed to monitor progression of the condition, did not impact staying on therapy long-term either. Other empirical studies cite, however, that frequent blood testing is a good measure of chronic disease progression monitoring and optimizing patient treatment, while increasing drug adherence.[32] The insignificant results might suggest the need for tracking the test outcomes to inform changes in the treatment from disease progression, which may in turn lead to increased patient adherence.

Recommendations

Leveraging the survival analysis model provides the opportunity for realizing additional insights into not only the adherence level for Product A, but the marketing and sales strategies, as well as programs that could help lower the probability for non-adherence.

Understanding the key drivers and their associated impacts on non-adherence can aid in driving brand success in an efficient manner over time, while greatly impacting patient outcomes via improved compliance and adherence. For example, knowing that mail order deliveries of medications increase patient adherence may lead to extension of the programs, while ensuring patient convenience in refilling their medications.

Research Limitations and Future Direction

The primary limitation of the study is leveraging only health claims data and not accounting for other factors related to the patient's treatment journey, including laboratory results. Understanding in-depth the progression or sub-type of the disease might add significantly to the evaluation of drivers of adherence.

In addition, including more information related to patient cognitive abilities in comprehending the prescribed treatment dosing and schedule recommendations, as well as understanding patients' perceptions of the prescribed treatment, disease progression, and seriousness of the condition, may also provide an additional explanation for the adherence levels, as well as its drivers and barriers.

As mentioned in the earlier section of the article, linking other datasets, including primary market research information, EMR, and genomics, might further define and help evaluate the drivers and 'risk' factors of patient adherence via presenting a more comprehensive view of patients and their experiences on the prescribed therapy.

Furthermore, the data may allow investigating the health and economic outcomes of therapy, as well as evaluating how the current sales and marketing strategies and tactics, as well as other programs implemented to increase patient education and adherence, drive the desired outcomes. These results could inform an optimized and efficient allocation of resources to maximize desired health outcomes.

Finally, future developments and direction in this research area may lead adherence assessment into leveraging other analytics approaches, including exploratory machine

learning algorithms, in order to increase predictive accuracy of the survival models. Through the use of social media data and natural language processing, additional insights might also be gained into the more 'personal' effects currently not captured in health claims data.

Acknowledgment

About the Authors

Ewa J. Kleczyk, *PhD is an Executive Director of Commercial Effectiveness Analytics with Symphony Health Solutions, with experience in primary and secondary market research and marketing analytics for pharmaceutical clients.*

Derek Evans *is a Senior Vice President leading the Commercial Effectiveness Practice at Symphony Health Solutions, with experience in primary and secondary market research, marketing analytics and strategic consulting for pharmaceutical clients.*

References

1   Iuga AO, McGuire MJ. Adherence and Health Care Costs. Risk Management Healthcare Policy. 2014; 7: p. 35-44.

2   Sudeep K, Cleves MA, Helm M, Hudson TJ, West DS, Martin BC. Empirical Basis for Standardizing Adherence Measures Derived from Administrative Claims Data among Diabetic Patients. Medical Care. 2008 November; 46(11): p. 1125-1133.

3   Nau D. Proportion of Days Covered (PDC) as a Preferred Method of Measuring Medication Adherence. Article. Springfield, VA.

4   Beena J, Jimmy J. Patient Medication Adherence: Measures in Daily Practice. 2011 May; 26(3): p. 155-159.

5   Mann DM, Howard D, Ponieman H, Leventhal EA. Predictors of Adherence to Diabetes Medications: The Role of Disease and Medication Beliefs. *Journal of Behavioral Medicine*. 2009; 32: p. 278-284.

6   Martin LR, Williams SL, Haskard KB, DiMatteo MR. The Challenge of Patient Adherence. *Therapy Clinical Risk Management*. 2005; 1(3): p. 189-199.

7   Fullman P. The Drivers that Impact the Patient Adherence. PM360. 2012 October 1.

8   McKethan A, Benner J, Brookhart A. Health Affair Blog. [Online].; 2012 [cited 2017 January 10. Available from: http://healthaffairs.org/blog/2012/08/28/seizing-the-opportunity-to-improve-medication-adherence/.

9   Muller S, Kohlmann T, Wilke T. Validation of the Adherence Barriers Questionnaire - an Instrument for Identifying Potential Risk Factors Associated with Medication Related Non-Adherence. BMC Health Services Research. 2015; 15: p. 153.

10 What Researchers Mean by Primary and Secondary Data. *A Quarterly Publication of the Institute for Work & Health: At Work*. 2015 Fall; 82: p. 2.

11 Cox DR. Regression Models and Life-Tables. 1972; 34(2): p. 187-220.

12 Huynh KP, Voia MC. Mixed Proportional Hazard Models with Continuous Finite Mixture Unobserved Heterogeneity. [Online].; 2016 [cited 2017 January [Syracuse University]. Available from: https://www.maxwell.syr.edu/uploadedFiles/cpr/events/cpr/papers2016/voia.pdf.

13 Therneau T, Crowson C, Atkinson E. The Comprehensive R Archive Network. [Online].; 2016 [cited 2017 January 10. Available from: https://cran.r-project.org/web/packages/survival/vignettes/timedep.pdf..

14 Raghupathi W. Healthcare Informatics: Improving Efficiency and Productivity. In S K, editor. Data Mining in Health Care.; 2010. p. 211-223.

15 Burghard C. Big Data and Analytics Key to Accountable Care Success; 2012.

16 Dembosky A. Data Prescription for Better Healthcare. *Financial Times*. 2012 December 12; 19.

17 Feldman B, Martin EM, Skotnes T. Big Data in Healthcare Hype and Hope. 2012 October.

18 Fernandes L, O'Connor M, Weaver V. Big Data, Bigger Outcomes. *Journal of AHIMA*. 2012;: p. 38-42.

19 Denny JC. Chapter 13: Mining Electronic Health Records in the Genomics Era. PLoS Computational Biology. 2012; 8(12): p. e1002823.

20 Symphony Health Solutions. [Online]. [cited 2017 January 10. Available from: www.symphonyhealth.com.

21 The Future of Health Insurance: A Road Map Through Change. Report. ; 2015. Report No.: CSG No. 1506-1546714 Northeast.

22 Big Data Analytics for Healthcare. In Tutorial Presentation at the SIAM International Conference on Data Mining; 2013; Austin, TX.

23 Emons M. Integrated Patient Data for Optimal PAtient Management: The Value of Laboratory Data in Quality Improvement. In Proceedings of the 24th Arnold O. Beckman Conference.; 2001: Clinical Chemistry. p. 1516-1520.

24 Wilson J, Bock A. The Benefit of Using Both Claims Data and Electronic Medical Record Data in Health Care Analysis. White Paper. Eden Prairie, MN:; 2012. Report No.: 12-27912.

25 Maughn K, Kleczyk E. Assessing Adherence and Usage Behavior to Identify Product Usage among Chronically Ill Patients with Idiosyncratic Treatment Plans. In PBIRG Annual General Meeting; 2016; Bocca Raton, FL.

26 Ho PM, Bryson CL, Rumsfeld JS. Medication Adherence. Its Importance in Cardiovascular Outcomes. *Circulation*. 2009 June; 119(23): p. 3028-35.

27 Pearson C. Medication Adherence: Are Women Worse then Men at Taking their Meds? The Huffington Post. 2013 May 5.

28 Fendrick M, Chernew ME. Precision Benefit Design - Using "Smarter" Deductibles to Better Engage Consumers and Mitigate Cost-Related Nonadherence. *JAMA Internal Medicine*. 2017 January 7.

29 Iyengar RN, Lefrancois AL, Frazee SG. Association of Pharmacy Dispensing Channel with Medication Adherence Among Medicare Part D Patients Having Comorbid Conditions. 2014. Poster Presentation.

30 Billings ME, Auckley D, Benca R, Foldvary-Schaefer N, Iber C, Redline S, et al. Race and Residential Socioeconomics as Predictors of CPAP Adherence. Sleep. 2011; 34(12): p. 1653-1658.

31 Falagas ME, Zarkadoulia EA, Pliatsika PA, Panos G. Socioeconomic Status (SES) as a Determinant of Adherence to Treatment in HIV Infected Patients: a Systematic Review of the Literature. *Retrovirology*. 2008 February 1; 5(13).

32 White T. Stanford Medicine. [Online].; 2015 [cited 2017 February 15. Available from: http://med.stanford.edu/news/all-news/2015/04/adherence-to-blood-thinner-best-with-pharmacist-management.html.

Anticipating the Market Access Outlook for Drugs in R&D

Walter Brooks, Vice President, Equinox Group; David Godolphin, Vice President, Equinox Group; and Allan Miller, Vice President, Equinox Group

Abstract: This article describes an analytical framework to evaluate the market access outlook for drugs in R&D. Based on the anticipated clinical attributes and planned price for a new drug, this method provides an objective basis for quantifying a drug's clinical benefits and its impact on direct costs of treatment (both drug price and cost off-sets). The resulting ratio of benefit to cost is then compared to the corresponding ratios for drugs launched in recent years. This historical dataset shows there is a strong correlation between clinical benefit vs. cost and the ability to achieve market access. Comparing a new drug to those in the historical dataset reveals where the new drug's benefit vs. cost ratio falls along the spectrum of good to poor market access. Finally, this modeling technique easily accommodates sensitivity analysis – useful when there is high uncertainty about the drug's clinical attributes or planned price.

The market access new drugs ultimately achieve is affected by decisions made at all stages of R&D, including those undertaken in early development. Examples of such decisions include responses to questions such as:

- In what patient segments should the drug be developed – only the most severely ill, or more broadly for all patients with the disease; patients with prognostic markers; patients at specific lines of therapy?
- With which agents, if any, should the drug be combined?
- Against what comparators should the drug be tested in clinical trials?

Most biopharmaceutical companies recognize the need to apply analytics to address these and related questions early and at each decision point in the R&D cycle. Traditional market research tools provide essential information from clinical experts. But those tools also have inherent weaknesses, which will be familiar to anyone who

has watched their newly launched drug vastly underperform its commercial expectations. As valuable as primary research is, its weaknesses should be accounted for by adding other analytical methods to reduce the risk of mistaken commercial assessments, especially those driving important development decisions.

To ensure that a new drug is positioned for favorable market access, development decisions should be made with a clear notion of how they might affect the magnitude of clinical benefit in the target patient population. Will the clinical benefits warrant the desired price? Drugs that miss that goal are unlikely to achieve favorable market access, adequate patient share, or a revenue stream that pays back the R&D investment. This paper presents historical data showing that the balance between a drug's clinical benefit and its price helps explain its ability to achieve market access and patient share. Having an historical perspective provides guidance about what will be required of drugs

in R&D. The anticipated benefit/cost balance of a developmental drug can be compared to drugs in the historical data set to determine if its balance aligns better with recent historical successes or with failures. This paper suggests a practical solution to improve the analytics needed to inform these decisions in a time- and cost-effective way.

Primary Market Research is Crucial, But Its Inherent Weaknesses Must Be Addressed

Primary research results frequently contain "framing problems"–the introduction of bias arising from the manner in which information is provided and questions are posed. Framing problems are hard to avoid. For instance, if a series of questions puts great emphasis on dosing, the results are likely to artificially raise the importance of dosing in relation to efficacy and safety/tolerability. Moreover, questions posed in market research are typically affected by individuals with a stake in the outcome of the research (product team members), so the tendency is to elicit feedback that is favorable to the product in question.

Another weakness of primary research is in predicting patient share, specifically, the conversion of preference share to patient share, which is notoriously imprecise and usually at the heart of forecasts gone bad.

Finally, primary research focuses on the clinical attributes of drugs with little direct assessment of disease burden (mortality, morbidity, and costs). But disease burden is a crucial factor affecting payer policies. This limitation should be addressed directly.

We argue that the weaknesses described here explain why many new drugs miss their prelaunch commercial expectations. To address these weaknesses, we propose using an objective and comprehensive framework to compare a new drug to relevant competitors and measure its benefits and costs compared to those alternatives.

An Additional Analytical Perspective

There is a powerful source of information about the clinical benefits of a new drug–the hard clinical data upon which drug approvals are based. Such data has been collected under the supervision of regulatory bodies that insist on valid measurement of clinical advantages and disadvantages. Clinical trial data is an ideal source of information to measure the strengths and weaknesses of any drug. Other researchers have also shown that hard clinical data directly affects prescribing.[1,2,3]

To assess the magnitude of clinical improvement offered by a developmental drug, its anticipated clinical characteristics can be compared, attribute by attribute, to those of the standard of care (SOC). By exploiting clinical trial data we gain deeper and more precise insights into the strengths and weaknesses of the new drug relative to the SOC. This approach can also measure changes in disease burden–mortality, morbidity, and costs–precisely.

Measuring Clinical Improvement–Benchmarking Against the SOC

The first step in this method is to measure a drug's benefits and costs in an objective and consistent way. We illustrate this concept using the metrics Equinox Group has developed for this purpose, but other holistic metrics can serve the same goal (the sidebar on page 55 provides more details and observations about the design of a consistent system for measuring clinical benefit).

The clinical benefit of the new drug is its marginal improvement relative to the SOC. The starting point should be to determine how much unmet medical need remains after a

Table 1: Domains of Medical Need

Efficacy	cure, prophylactic success, symptom relief, slowing of progression, damage reversal, pharmacokinetics
Safety/Tolerability	frequency and severity of each side effect, net of placebo, warnings and monitoring requirements
Convenience	mode and frequency of dosing
Mortality	age-adjusted excess risk of mortality
Morbidity	pain, disability, hospitalization, quality of life, complications
Costs	drug price, non-drug costs, lost work time

typical patient has been treated with the SOC. The domains of medical need that should be included are shown in Table 1–note the inclusion of both drug attributes and elements of disease burden.

Unmet need in each domain should be quantified using data from clinical trials and peer-reviewed literature. For instance, it is possible to quantify mortality in virtually any disease. For the SOC's efficacy, we use endpoint data collected in pivotal trials to measure factors such as impact on symptoms and disease progression.

These clinical data are mapped onto scales for each domain to reflect the extent to which remaining need is low or high. The scores for each domain are weighted (see sidebar, page 55) to reflect the inherent importance of each domain. The sum of the weighted scores reflects overall unmet medical need for the SOC in the target population.

The process can then be repeated with data or assumptions for the candidate drug, using the same scales, to calculate its total unmet need score. Improvements in efficacy reduce disease burden. For instance, symptom relief might reduce pain scores or improve quality of life. Reduction in cardiovascular events should reduce mortality.

This approach provides objective and consistently derived scores for both the SOC and the new drug, so they can be validly compared. The new drug's reduction in medical need relative to the standard of care is a measure of its clinical improvement.

We have conducted these analyses for hundreds of drugs and have data that show what constitutes low, medium, and high clinical improvement, based on historical observations of peak patient share achieved. We have validated this framework through a statistical model that predicts peak-year patient share as a function of clinical improvement and other causative variables. The correlation is strong (R-squared of 85%).*

---

* The causative variables explaining peak-year patient share are: 1) level of clinical innovation offered by the new drug, 2) the number of competitors, 3) the price of the new drug, 4) the size of the population (the last two affecting payer budgets). To test robustness we performed "Leave-One-Out Cross Validation".
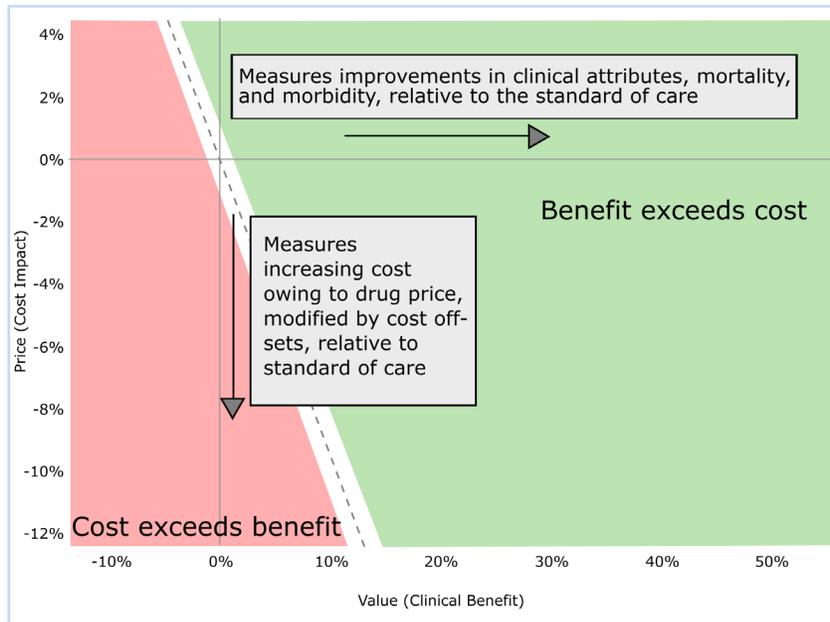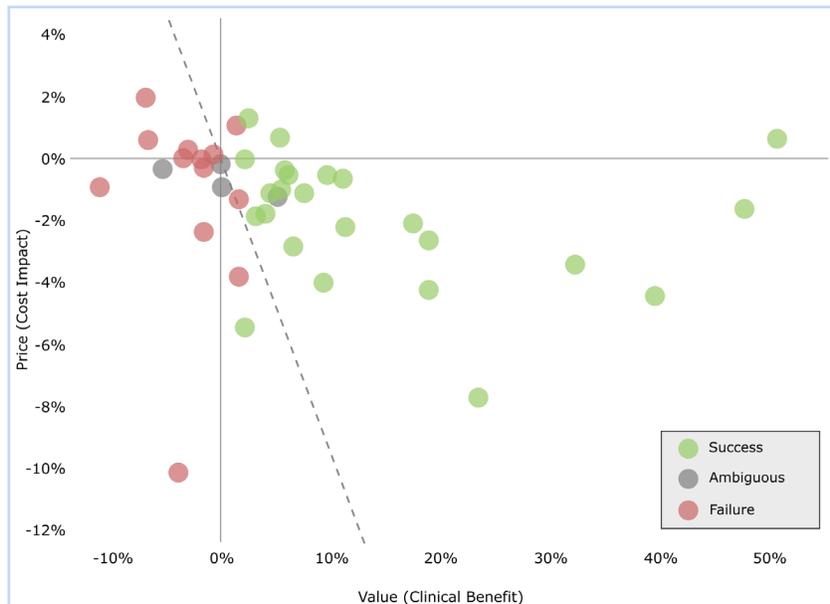
Figure 1

Measures improvements in clinical attributes, mortality, and morbidity, relative to the standard of care

Benefit exceeds cost

Measures increasing cost owing to drug price, modified by cost offsets, relative to standard of care
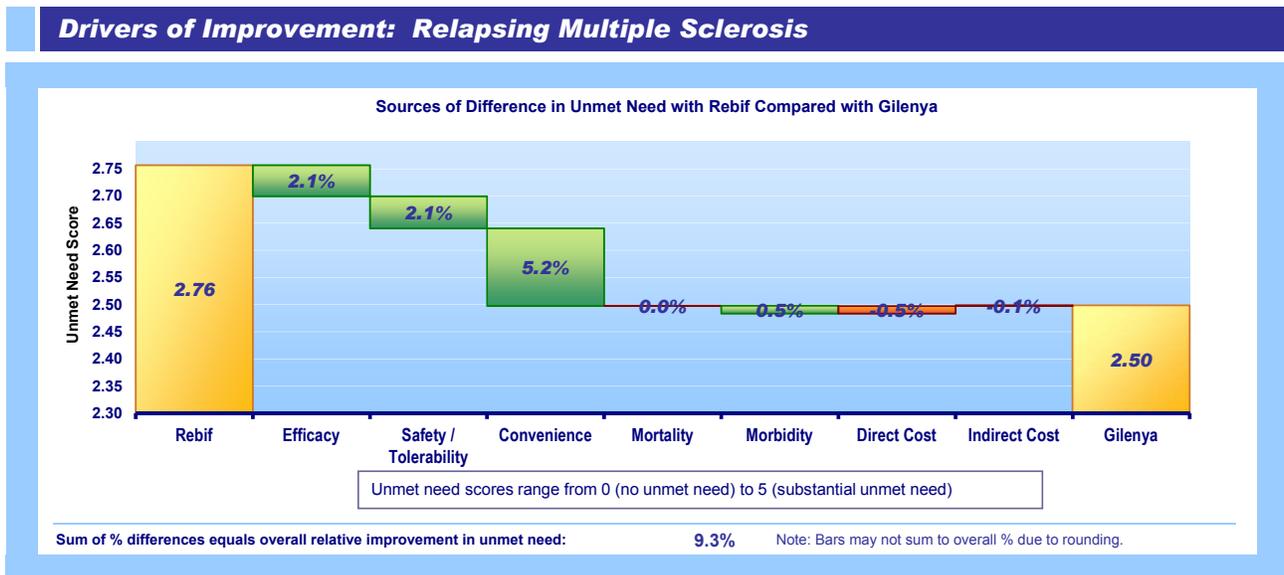
Cost exceeds benefit

Figure 2

Defining the Balance Between a New Drug's Clinical Benefit and Its Cost (Price)

Figure 1 shows a simple framework to compare the balance between clinical benefit (X-axis) and cost impact (Y-axis) for a prescription drug. The greater a drug's clinical benefit, the further to the right it is plotted. The higher the drug's net cost impact, the lower it is plotted. The dotted line running through the origin represents "value equal to cost" or, put another way, the price at which all of the drug's clinical value is clawed back by the developer. Drugs to the right of the dotted line offer more benefit than cost and drugs to the left cost more than the clinical benefit they deliver. Our analysis shows that there is a strong correlation between successful market access and being on the right side of the "value equal to cost" line (Figure 2).

Figure 3: Comparing Gilenya to Rebif in RR-MS

Drivers of Improvement: Relapsing Multiple Sclerosis

Sources of Difference in Unmet Need with Rebif Compared with Gilenya

Unmet need scores range from 0 (no unmet need) to 5 (substantial unmet need)

Sum of % differences equals overall relative improvement in unmet need: 9.3%   Note: Bars may not sum to overall % due to rounding.

Government regulators, academia, and medical societies have grappled with initiatives to measure the value of prescription drugs. Peter Neumann and Joshua Cohen at the Institute for Clinical Research and Health Policy Studies at Tufts Medical Center have written about the strengths and weaknesses of many of these approaches.[4] Most of these initiatives focus on disease burden–the drug's impact on mortality, morbidity, and cost. For instance, The National Institute of Clinical and Health and Care Excellence (NICE) and The Institute of Clinical and Economic Review (ICER) emphasize disease burden in their analyses. The quality-adjusted-life-year (QALY) is a well-established metric for this purpose, although it is not clear how it can be used to predict commercial performance and market access. Clearly, developers need to understand how their developmental drugs affect disease burden.

Focusing exclusively on disease burden, as most of these initiatives do, can miss other benefits arising from new therapies; these factors affect the patient's experience and therefore influence market success. For instance, the once-a-day

oral dosing of Gilenya (fingolimod) offers an important advantage over the subcutaneous injections of beta-interferons in the treatment of relapsing multiple sclerosis. That dosing advantage contributed substantially to Gilenya's rapid adoption. The safety/tolerability advantages of Stelara (ustekinumab) over the TNF-alpha inhibitors in the treatment of psoriasis does not affect disease burden, but it is a significant benefit for patients and was the main contributor to Stelara's commercial success. QALYs do not fully capture these benefits. Any method to accurately measure the "value" of a new drug must take into account both product attributes (efficacy, safety/tolerability, dosing) and disease burden.

Measuring a New Drug's Reduction in Unmet Need

A model containing all of the domains included in Table 1 reflects both disease burden and other clinical benefits (e.g., safety/tolerability and dosing) that QALYs miss. To illustrate the concept, Figure 3 compares Gilenya to Rebif in relapsing multiple sclerosis. Gilenya reduces medical need by 9.3%, and the waterfall

Figure 4: Mapping Gilenya's Clinical Benefit and Cost Into the Pricing and Market Access Graph
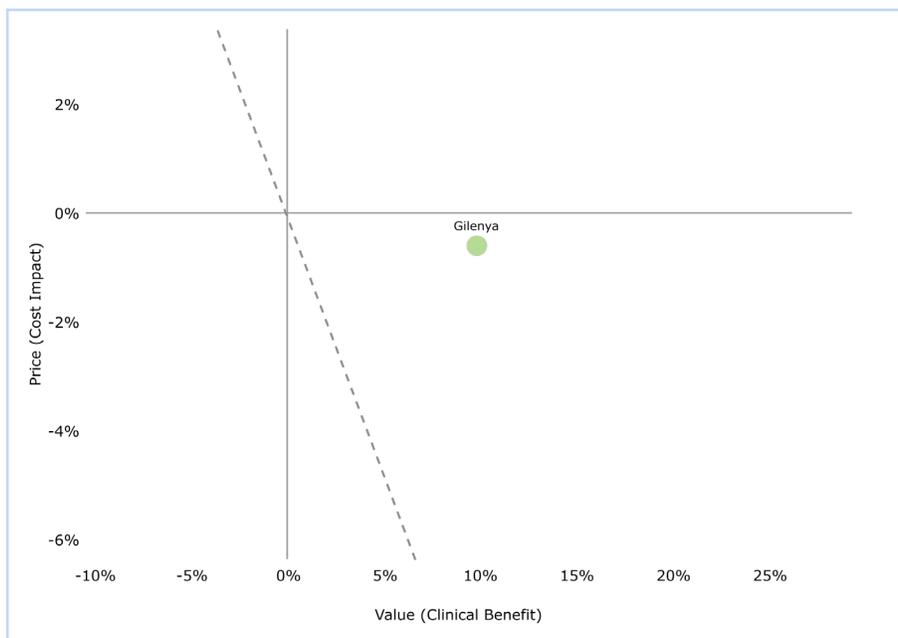

chart shows the sources of its advantages and disadvantages of Rebif for each domain of medical need.

Using a 0-5 scale that reflects how high unmet need is given treatment with each drug (see side-bar, page 55), the total unmet need score is 2.76 for Rebif and 2.50 for Gilenya. Gilenya reduces medical need; that is, it offers clinical improvement over Rebif. The overall improvement of Gilenya is 9.3%, and the sources and magnitudes of Gilenya's advantages appear in the graphic – modest improvements in efficacy and safety/tolerability, with a more substantial benefit in dosing (convenience). The slightly higher price for Gilenya is a disadvantage, reflected in the direct cost domain (where drug costs and cost offsets are captured).
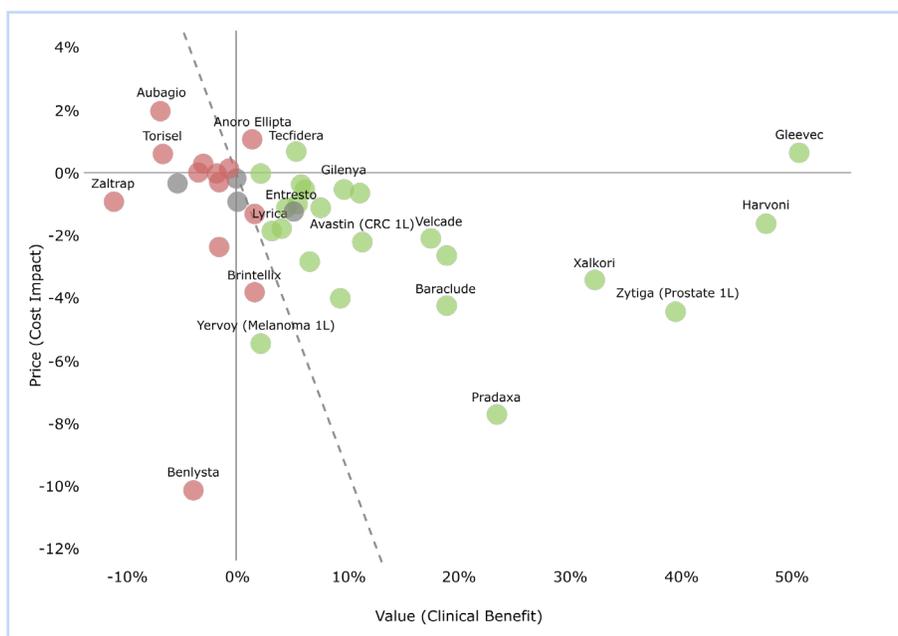
Because this analysis measures how much of the overall improvement comes from each domain of medical need, we can separate out the impact of cost (driven mainly by drug price) from other drivers. "Clinical Benefit" is defined as the net impact of all non-direct-cost factors combined – i.e., improvements in mortality, morbidity, efficacy, safety/tolerability, dosing, and indirect cost. Gilenya's "clinical benefit" (compared to Rebif) is 9.8%:

- Efficacy 2.1%
- Safety/tolerability 2.1%
- Convenience (dosing) 5.2%
- Mortality 0%
- Morbidity 0.5%
- Indirect costs -0.1%

} 9.8%

The "cost" value for Gilenya (driven by its higher price) is -0.5%, as shown in Figure 3. These two values are used to map Gilenya onto the benefit vs. cost graph, shown here in Figure 4. It shows that Gilenya's coordinates place it well to the right of "value equal to cost" line, in the favorable zone. Gilenya had rapid market penetration and has achieved annual sales of

Figure 5: Clinical Benefit vs. Cost for 40 Recent Launches

Adding Observations to the "Benefit vs. Cost" Graph

Figure 5 shows the results for 40 drugs launched in recent years analyzed through this method. In addition to mapping benefit vs. cost, we have colored the points representing each drug to indicate whether it has achieved high patient share (green), low patient share (red), or patient share somewhere in between or not yet clear (grey). Naturally, the ability to achieve strong patient share requires that market access be favorable. Drugs to the right of the "benefit equals cost" line have tended to achieve strong patient share, while those to the left have not.

Here we comment on results for several agents shown above:

- **Harvoni** – compared to the Incivek regimen in the treatment of HCV genotype 1. Despite the controversy surrounding Harvoni's high price, this

more than $3 billion, indicating that market access was favorable.

analysis shows that its clinical benefits are well worth that price. Much of the benefit is attributable to the elimination of pegylated interferon and ribavirin from the regimen, resulting in large improvements in safety/tolerability and dosing. The improvement in efficacy also confers gains in mortality and morbidity. The net "clinical benefit" measured here is over 48% – very high by historical standards. The net drug price impact (an increase of $12,000 for a course of treatment) is small relative to the benefits. FDA approved in October 2014; Harvoni generated revenue of $13.9 billion in 2015, indicating that market access was not problematic.

- **Xalkori** – compared to carboplatin + paclitaxel in first line ALK+ non-small-cell lung cancer. Xalkori is an example of a branded agent competing successfully against inexpensive generics. Xalkori's revenue is not high by old "blockbuster"

standards due to the small target patient population. But because its clinical benefit is so high (substantial improvements in efficacy and mortality, with added improvements in safety/tolerability and dosing), Xalkori quickly went on to achieve overwhelming patient share in this small patient population.

- **Pradaxa** – compared to warfarin in stroke prevention in atrial fibrillation (SPAF). As the first of the novel oral anticoagulants indicated for SPAF, Pradaxa offered strong clinical benefits at a price premium to generic warfarin, mainly from much improved efficacy. Pradaxa achieved rapid revenue growth despite its high price before competition from other novel oral agents slowed its growth.

- **Brintellix** – compared to escitalopram in major depressive disorders. Brintellix offers modest improvements in efficacy and safety/tolerability, but the magnitude of these advantages are low, earning a clinical benefit value of only 1.7%. But the branded price against generic escitalopram results in a cost disadvantage of -3.7%. Brintellix has struggled to achieve patient share.

- **Anoro Ellipta** – compared to Spiriva in chronic obstructive pulmonary disorder. Anoro Ellipta offers improvements in efficacy and safety/tolerability over Spiriva, but the magnitude of those improvements is modest at 1.5%. It is priced at a significant discount to Spiriva, providing a cost advantage at 1.1% (note its placement above the X axis). Despite being on the favorable side of the Benefit-Cost line, since its approval in December of 2013, Anoro Ellipta has had modest sales. In recent years, drugs with overall clinical improvement at levels

this modest tend to struggle to achieve differentiation.

- **Entresto** regimen – compared to a traditional heart failure regimen. Entresto (replacing the traditional ACEi or ARB) offers a moderate improvement in efficacy, which confers benefits in mortality and morbidity, and with modest disadvantages in safety/tolerability and dosing. The net clinical benefit is 5.3%. The price disadvantage is -1.1%. This places Entresto on the favorable side of the benefit-cost line. Approved in July 2015, Entresto has underperformed Wall Street predictions. This is another example of a new agent with relatively modest clinical improvement entering a market dominated by many effective and safe generic alternatives. Our analysis suggests that Entresto's sales will improve, but will fall well short of the $5 billion predicted by Wall Street analysts.

Applying This Method to Assess a New Drug Program

To analyze a new drug's market access outlook through this lens requires characterizing the extent to which the current SOC meets the medical need in the target population (see the domains of need listed in Table 1). The analysis of the SOC creates a baseline against which the new drug is compared. For the SOC, information on product attributes (efficacy, safety/tolerability, and dosing) can be taken directly from its package insert or clinical trial data published in peer-reviewed journals.

Data describing the SOC's efficacy endpoint values are translated into a numerical score that reflects how efficacious the SOC is. The new drug's efficacy score is based on endpoint values it is expected to achieve and calculated using the same scales and weights that were used for the SOC. The same process is applied to safety/tolerability and dosing.

Table 2: Clinical Attributes of MS Drug in Two Profiles vs. Tecfidera

Relapsing Multiple Sclerosis	Current SOC: Tecfidera (Dimethyl fumarate)	Low Case	High Case
Relative Reduction in Annual Relapse Rates (net of placebo)	49%	55%	60%
Relative Reduction in % Progressing at 2 years (net of placebo)	38%	38%	50%
Damage Reversal (net of placebo)	0%	0%	20%
Price	$54,750/year	$49,274/year	$60,225/year
Safety/Tolerability **Dosing**	Oral BID	Safety/tolerability and dosing equal to Tecfidera	

To illustrate the concepts, imagine a team is evaluating a drug targeted for relapsing multiple sclerosis. There is high uncertainty about its efficacy. Management wants to know how the market access outlook is affected by assumed efficacy improvements. The uncertain range in efficacy is bounded by the Low Case and High Case profiles (Table 2).

The team decides that Tecfidera (dimethyl fumarate) is the appropriate SOC to benchmark against their new drug's profiles. Table 2 shows key clinical attributes for Tecfidera and the assumed efficacy for the Low Case and High Case profiles. Note that the low profile is expected to offer only a modest improvement in relapse rates. The high profile is expected to have greater reduction in relapses, improved reduction in progression, and damage reversal. Furthermore, the team thinks that the High Case profile will command a premium price relative to the SOC, and the Low Case profile will require a lower price.

In addition to differences in clinical attributes, the model measures how the new drug's improved efficacy will reduce disease burden; for instance, how symptom reduction diminishes pain or improves quality of life. Objective data to measure disease burden (mortality, morbidity, and costs) is also obtained from peer reviewed literature.

Each of the subdomains of morbidity listed in Table 1 is analyzed and quantified. For instance, reduction in the progression of MS will yield improvements in disability, an important benefit. Table 3 shows the disability scores for Tecfidera, and the Low and High cases. The High Case reduces disability because it offers slower disease progression and reverses damage compared to Tecfidera. We assume that these efficacy improvements shift the distribution of disability from the most severe level (ADL-impaired) to lower severities (activities limited, minor disability, no disability). The lower total chronic disability score for the High Case reflects improvement (reduced medical need), and is included as part of the total clinical benefit attributable to that profile.

Similar analyses are conducted for each element of disease burden to calculate scores that reflect the total unmet medical need for Tecfidera and the Low and High Case profiles. Both profiles are compared to Tecfidera (as in the "Drivers"

Table 3: Disability Scores for the SOC, Low and High Cases

	Tecfidera	**Low Case**	**High Case**
Chronic Disability Score	**2.17**	**2.17**	**1.93**
Minor (% of patients)	46.5%	46.5%	41.1%
Activities Limited (% of patients)	16.1%	16.1%	13.2%
ADL-Impaired (% of patients)	1.4%	1.4%	1.1%

Figure 6: Mapping Low and High Case Profiles in Market Access Grid
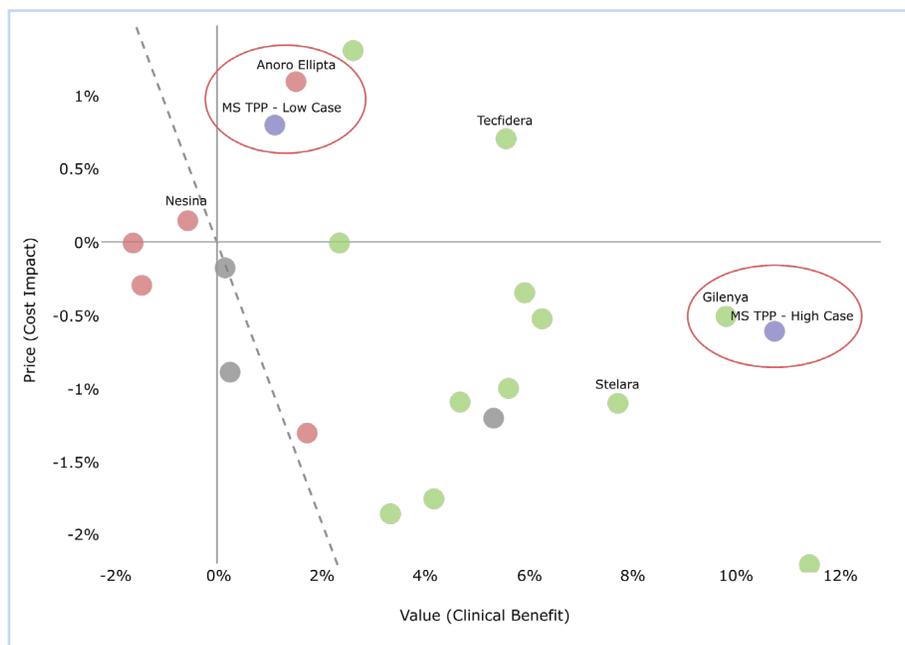
analysis shown in Figure 3) to calculate total clinical benefit and the cost impact of each profile, relative to Tecfidera. The outputs of that analysis are plotted on the clinical benefit vs. cost graph to identify launched drugs that have similar benefit-to-cost ratios; drugs with similar ratios represent plausible market access analogs.

In this instance, the Low Case profile has a benefit to cost ratio that is similar to Anoro Ellipta, a drug that has struggled commercially (Figure 6). Alternatively, should the new agent's

efficacy reflect the values of the High Case, its balance of benefit to cost more closely matches Gilenya, a drug that has achieved relatively favorable market access. Knowing where a drug is likely to fall in this space provides the team with information to anticipate how challenging market access is likely to be.

Limitations of This Approach
The agents we have evaluated through this method treat a wide range of disorders, including those treated by primary care doctors and by

Design of a Consistent System for Measuring Clinical Benefit

To design a system that measures unmet need and clinical benefit using objective clinical data, a few specific characteristics are essential. We describe here how we have implemented those characteristics in the model Equinox Group has developed and applied for more than 20 years. In that model, the contribution of each domain to total need depends on 1) the objectively measured level of need for a patient treated with the standard of care, and 2) the intrinsic importance of the domain. We multiply these two values; the product drives the final contribution to medical need of each domain in a given disease. The total unmet need score for the disease equals the sum of those weighted scores for all domains.

The Level of Unmet Need in a Domain

To measure the level of need requires clinical data. We translate that data into a score on a 0-5 scale that indicates a spectrum of low (0) to high (5) unmet need. For example, in relapsing multiple sclerosis (MS), Gilenya reduced relapse rates by 55%, translating into an unmet need score of 2.25 (55% of the distance from 5 to 0) in symptom relief efficacy. The use of objective data on consistently applied scales minimizes subjectivity and inter-rater variability in the scoring method.

The Importance of Each Domain of Unmet Need

The intrinsic importance of each domain is captured in the weights. For instance, the weights reflect the observation that efficacy has higher intrinsic importance than side effects, and side effects have higher importance than dosing. The score for each domain, when multiplied by its corresponding weight, reflects the contribution of that domain to unmet need in that disease. It is the contribution to unmet need that matches intuitive notions of the importance of a domain in a particular disease.

The weights do not differ between diseases. This may seem counter intuitive, but consider the following: myocardial infarction inflicts high mortality, and psoriasis inflicts no mortality. In both cases, the respective mortality scores are multiplied by the fixed mortality weight (20%) to arrive at the contribution to unmet need. Due to the high mortality score in MI (4.5 on the $0 - 5$ score), it has a high contribution to need, whereas in psoriasis, the mortality score is 0, and therefore mortality's contribution to medical need is 0 x 20% = 0.

specialists (e.g., oncology and immunology). While we have analyzed agents targeted to rare and ultra-rare disorders, to date we have too few observations to say anything definitive. Neither does this approach take into account how promotion affects patient share.

Conclusion: Anticipating the Market Access Outlook for Drugs in R&D

Crucial decisions made at all stages of R&D affect the ultimate pricing and market access outlook for a new drug. Analytics to assess the effect of alternative development strategies on market access should be applied early and refreshed at each decision point in the R&D cycle.

Traditional primary market research provides valuable insights to inform these decisions, but it has serious weaknesses, most typically leading to inflated forecasts. Primary research

collects expert opinion; as valuable as that opinion is, it is nonetheless opinion and therefore a subjective basis of measurement. Second, primary research focuses on the clinical attributes of drugs, with little direct assessment of disease burden. But disease burden is a crucial factor affecting payer policies. Despite the important insights arising from primary research, its weaknesses explain why many new drugs miss prelaunch commercial expectations.

To gain a more accurate view of the market access potential and commercial outlook for developmental drugs, biopharmaceutical companies should supplement traditional tools with alternative analytics that address traditional tools' weaknesses. We propose adding a method that explicitly measures improvement in both clinical attributes and disease burden to provide a robust comparison of a drug's clinical benefit against its cost as seen by payers. Such an approach is grounded in objective clinical data, providing a valid basis for evaluating that critical trade-off. Arrived at independently from primary research techniques, this can confirm or challenge the findings from traditional methods. The redundancy of two independent approaches strengthens crucial developmental decisions.

This approach can be applied to new drug programs as soon as there is a hypothesis for clinical attributes and target price. Uncertain clinical attributes can be modeled in multiple scenarios to determine thresholds for commercial success (e.g., what level of efficacy will be required to achieve an adequate patient share, given anticipated safety issues?), and can inform decisions even at early phases of R&D, when clinical uncertainty is high.

Before important development decisions are taken, such as which patient segments to pursue, which comparator to select for clinical trials, which drugs to be combined with, overall go/no-go and related decisions, the alternatives should be modeled and their effects assessed in a framework like the one described here. Applying this analytical lens to each developmental program can maximize the value of each asset and that of the overall portfolio.

References

[1] Azoulay P. (2002). Do pharmaceutical sales respond to scientific evidence? *Journal of Economics and Management Strategy*, 11(4), 551-594.

[2] Ashish S, Kappe E, Stremersch S. (2014). The commercial contribution of clinical studies for pharmaceutical drugs. *International Journal of Research in Marketing*, 31, 65-77.

[3] Kappe E, Stremersch S. (2016). Drug detailing and doctors' prescription decisions: the role of information content in the face of competitive entry. *Marketing Science*, 35(6), 915-933.

[4] Neumann P, Cohen J. (2015) Measuring the value of prescription drugs. *New England Journal of Medicine*, 373;27, 2595-2597.

About the Authors

Walter Brooks *is a founder and vice president of Equinox Group. He has specialized in pharmaceutical new product development strategy and R&D resource allocation for more than 25 years. Using elements from decision analysis, microeconomics, statistics and probability theory, and market research, Mr. Brooks has contributed to the creation of analytical methods to estimate the range of market potential for a new drug based on alternative assumptions about its clinical performance. Mr. Brooks holds a B.A. in economics and an M.S. in policy analysis from the State University of New York at Stony Brook.*

David Godolphin, *a vice president of Equinox Group, helped found the company in 1995. At Equinox, he has led more than 400 studies for U.S. and European clients, ranging from commercial outlooks for individual products to reviews of therapeutic-area portfolios and strategies. He has a special interest in forecasting techniques that reflect*

factors such as unmet need performance, compliance, reimbursement, and the strengths of emerging competitors. Mr. Godolphin received his A.B. from Harvard College.

Allan Miller, *a Vice President at Equinox Group, joined the company in 2007 as a Consultant after nearly 20 years of senior management experience in both commercial and scientific roles at several biotechnology and pharmaceutical firms. At Equinox Group he has focused on oncology and on the development of sophisticated simulation-based forecast models. Dr. Miller holds a M.A. in Natural Sciences (Biochemistry), from the University of Cambridge and a Ph.D. in Molecular Biology from the Laboratory of Molecular Biology at Cambridge. He conducted post-doctoral research and taught at Harvard University.*

Leveraging CMS Open Payments Data to Identify Channel Preferences and Gather Competitive Intelligence, Thereby Improving HCP Targeting

Rahul Anand, Engagement Manager, Mu Sigma; Duggan Collier, Deputy Director, Commercial Analytics, Bayer; Yan Jiang, Director, Business Insights, Rare Diseases, Bayer; Janardhan Vellore, Director, Commercial Analytics, Bayer

Abstract: Organizations continuously explore means to gain competitive intelligence to understand market potential or improve physician targeting. Some of the predominant ways to gain this advantage has been to leverage traditional third-party data sources such as IMS, Symphony Health, and market research among others. Since 2014, the Centers for Medicare and Medicaid Services has published transparency data which captures transfer of value to physicians from pharmaceutical companies through different interaction channels. This paper discusses examples of how a pharmaceutical company has leveraged this data to identify new targets and further improve targeting by identifying channel preferences for existing and new targets.

Keywords: CMS open payments, Physician targeting, Competitive intelligence, Physician interactions, Channel preferences

Introduction

Today, while competition is getting tougher with new product launches and pharmaceutical organizations are struggling to meet their goals, such organizations are exploring different ways to gather competitive intelligence. The objective for gaining this intelligence is to better understand market potential, improve HCP targeting, and potentially determine competitive loyalists. Some of the predominant ways in which organizations try to achieve these objectives are:

    a)   Leveraging traditional third-party data such as IMS, HMS and others

    b)   Measuring the share of voice of their own products and their competitor's products through market research

    c)   Identifying office accessibility or channel preferences of HCPs through licensed data sources such as AccessMonitor™ and AffinityMonitor™

Organizations leverage this information to either refine/expand their target universe or make targeting more personalized, which aids sales and marketing teams. However, there are challenges that the approaches mentioned above pose—content participation limits and associated cost. (Figure 1)

This paper will review examples of how sales and marketing teams can use publicly-available CMS data to identify new targets and gather competitive intelligence. It also showcases how this data can be leveraged to improve HCP targeting by identifying channel preferences for existing and potential new targets.

Overview of CMS Open Payments Data

As a result of the Physician Payments Sunshine Act and collection of relevant data associated with the act, additional data resources became available for healthcare analytics to inform targeting decisions. The Act was passed in 2010

Figure 1: Currently Available Data Sources for Targeting

Targeting Data Sources	Pros	Cons
Commercial Claims Data (i.e. QuintilesIMS[1])	• Rich patient and provider transaction data • Ideal for determining treatment pathways and provider behaviors • Visibility to competitive products	• Incomplete geographic coverage • Not projected. Complete only for reporting services • Associated cost
Projected Prescription Data (i.e. QuintilesIMS[1], Symphony Health[2])	• Rx projected to national coverage • Provides most complete prescription coverage • Visibility to competitive products	• Projection limitations • Effectively limited to Retail/Mail Order channels • Specialty pharmacy distribution generally not available • Associated cost
CMS Claims Data (i.e. LDS files[3,4])	• Complete Medicare A, B, D inpatient and outpatient claims regardless of therapy class • Low cost	• Data Latency – 9-12 months after close of period • Cumbersome process for acquisition
CMS NPI Registry[5]	• Complete national listing of practitioner and facility healthcare providers	• No utilization metric • Providers limited to those receiving/ requesting reimbursement

to increase transparency of financial relationships between health care providers and pharmaceutical manufacturers. The Centers for Medicare and Medicaid Services (CMS) is managing the compliance data collection activity with the first data published in 2014 for H2'13. Since then, CMS has published *transparency data[6]* for FY 2014 and 2015 which is publicly available. It captures transfer of value to physicians through various *Nature of Payments[7]* (Consulting fees, Honoraria, Gifts, Entertainment, Food and beverage, etc.). CMS has mandated reporting of ToV greater than $10 per activity or greater than $100 per year. Hence capture rate for low-expense channels can sometimes be as low as 60%. The data is reported at Company – Brand – HCP – *Nature of Payment[1]* level.

Approach
The first stage of the analysis is to create a database which links *CMS Open Payments data[6]* with internal sales/call activity data. This helps in identifying overlap with the existing target list. Certain keys (combination of First Name, Last Name and Zip) can be used to pull *NPI#[8]* on the *CMS data[6]* which is then mapped to internal sales/call activity data through the data mastering process. Once the database is created, the market and competitors are identified. The next step is to establish business rules to map "*Nature of Payments[6]*" to "Contact Channels" which defines the channel to which the transfer of value is made (Figure 2). A combination of *Nature of Payments[7]* across a certain time period is defined as the Contact channel.

At this stage, validation of the business rules becomes critical to get buy-in from sales/marketing teams. Scenarios with varying *Nature of Payments[7]* and time period definition are generated to arrive at an acceptable capture rate (Figure 3).

Figure 2: Defining Contact Channel

Time Period	Nature of Payment(s)	Contact Channel
One Date	Consulting Fee	Consulting/Ad-Board
One Date	Honararia	R&D—Product Development/Improvements
One Date	Research	R&D—Clinical Trial
One Date	F&B (<$25)	In-Service Activities/Others*
Date+-1	F&B (>$25)	Speaker Program—Attendee
Date+-1	F&B + T&L —Compensation for services	Speaker Program—Speaker
Date+-1	T&L	Speaker Program Training
Date+-1	F&B + T&L	Speaker Program Training
One Date	Education	Textbook/Educational Materials

Figure 3: Validation Table

Metric	Actual	Analytical Data*	% Capture Rate
# Speaker Programs	358	269	75%
# Unique Speakers	86	68	80%
Calls	48,761	4,019 (includes ToV for In-Service Activities/Others)	8% (Low capture rate for calls is attributed to >$10 reporting cut-off of CMS data)

* *Based on business rules*

Limitations of the Approach

The data mastering process might not accurately link physicians in *CMS Open Payments data[2]* to internal sales/call activity data due to lack of, or too constraining, mastering rules. This will lead to low capture rates during validation or capturing false positive physicians.

Additionally, the combination of the nature of the payments to define Contact Channels might vary across organizations. It should be analyzed for different scenarios to arrive at an acceptable capture rate in partnership with the sales teams.

Case Study 1: Gathering Competitive Intelligence

For sales and marketing teams, having a deep understanding of competitors' level of interactions with HCPs across different channels of influence gives a huge competitive advantage. In this case study, the analysis identified a physician pool and measured level of contact through different channels (Figure 4).

Distribution across states was also studied to understand anomalies in resource allocation (Figure 5). It helped the teams optimize its resources across channels by providing visibility to the targeting strategy deployed by competitors.

## Case Study 2: Identifying New Targets for Call Plan

For 2016, the sales team of the cardiopulmonary business unit of the company wanted to expand its sales force target list. The existing call plan was created using latest claims and internal

Figure 4: HCPs Contacted and Touchpoints by Product and Contact Channels
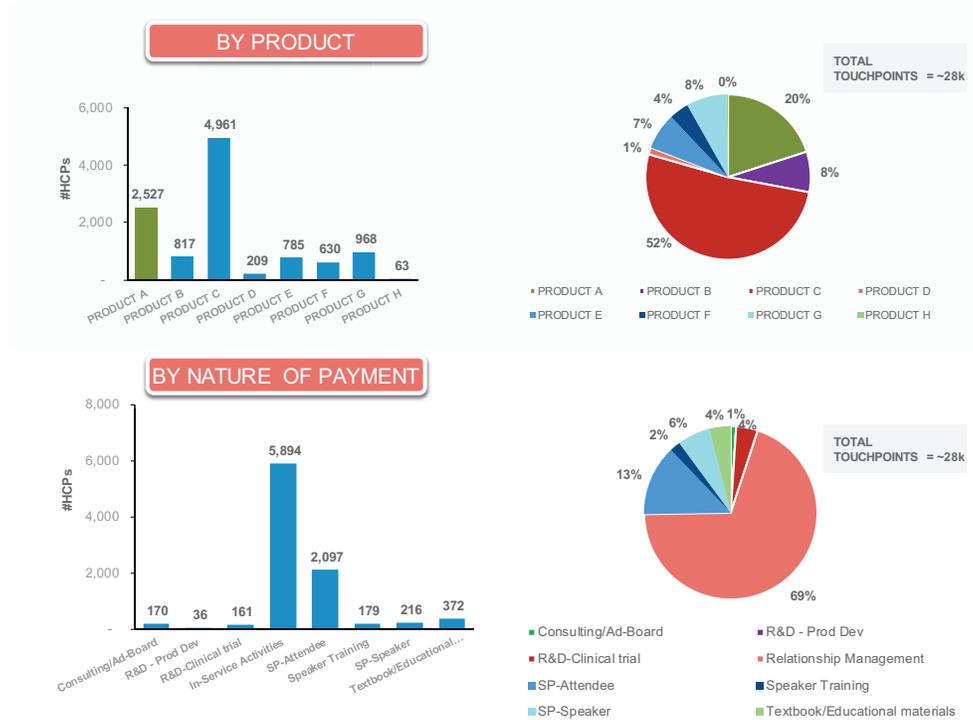
Comparing ratio of #touchpoint to #HCPs helped identify opportunity to boost targeting effort. For example, Product A has lower ration than Product C, which means that there is an opportunity to increase targeting effort.

Figure 5: Distribution of Touchpoints (Including All Contact Channels) by States
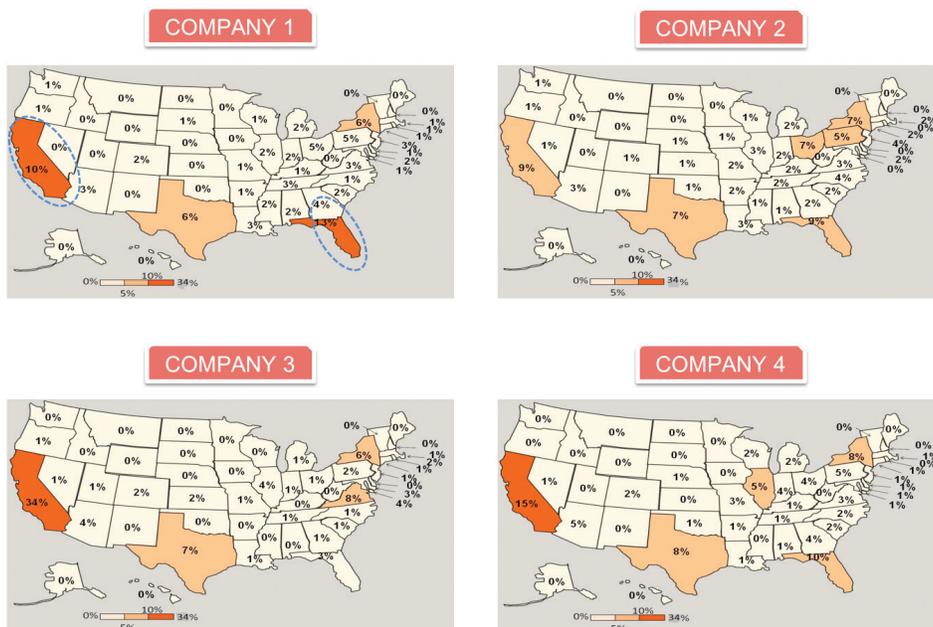
Comparison of Company 1's touchpoints against other companies' touchpoints helped identify anomaly in targeting and potential to improve and optimize resource allocation. For example, Company 1 has the potential to reallocate a high concentration of touchpoint in Florida to California, where other companies have higher concentration.
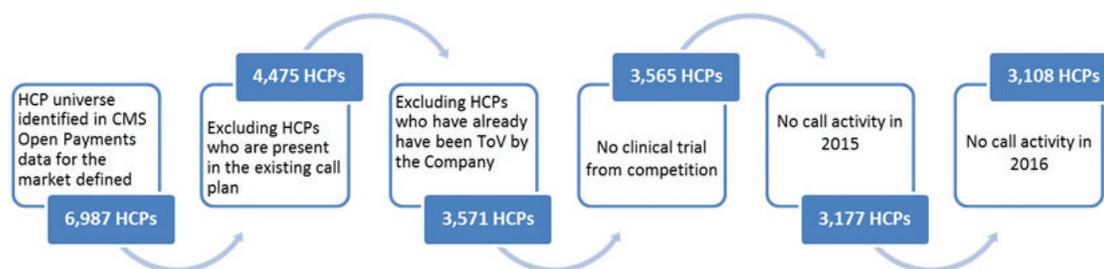
Figure 6: Flow to Identify Potential Target List


Figure 7: Targets in the Current Target List

Call Class	# HCPs	# Touchpoints	Touchpoint per HCP	Claims per HCP	Patients per HCP
Tier 1	842	6,496	7.7	13	1.5
Tier 2	925	6,122	6.6	6	0.8
Tier 3	315	1,716	5.4	1	0.3
TOTAL	2,082	14,334	6.9	80	10

Decile (Touch Points)	# HCPs	# Touchpoints	Touchpoint per HCP	Claims per HCP	Patients per HCP	
10	50	618	12	2.6	0.48	
9	91	621	7	2.2	0.36	*Recommended additional 450 targets*
8	132	628	5	2.4	0.42	
7	176	621	4	1.9	0.41	
6	208	624	3	1.6	0.33	
5	293	622	2	1.2	0.26	
4	311	622	2	1.3	0.26	
3	602	623	1	0.8	0.18	
2	622	622	1	0.9	0.19	
1	623	623	1	1.8	0.28	
Total	3,108	6,224	2	1.4	0.26	

Only commercial channels (Speaker Program – Speaker, Speaker Program – Attendee/Trainings, Speaker Training and In-Service Activities/Others) are included for deciling

Figure 8: Exploring the Data Shows the Potential to Solve a Variety of Interconnected Problems
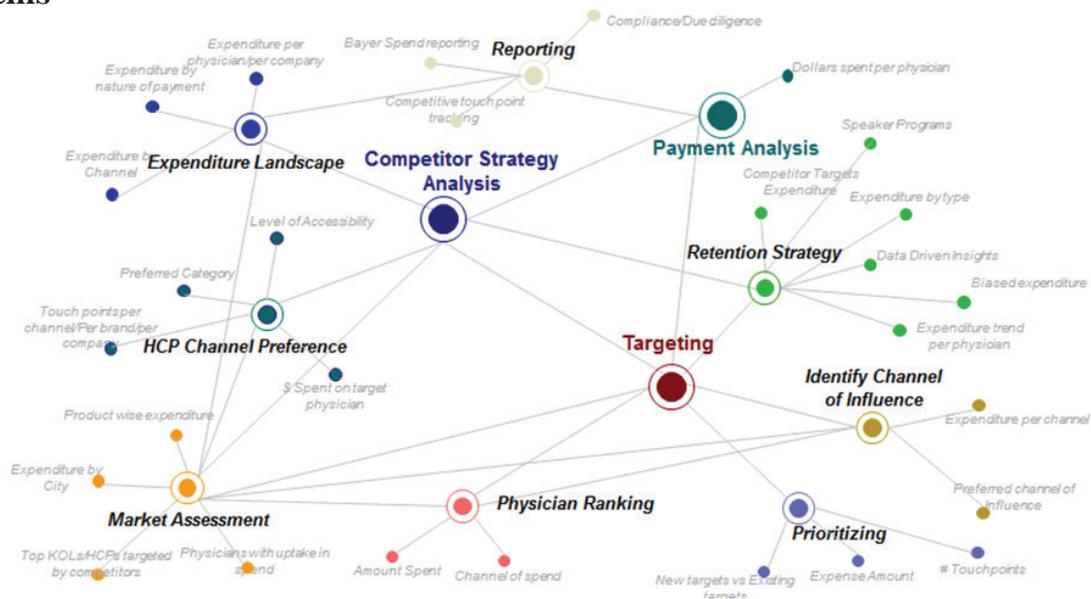
prescription data. This posed a challenge to the analytics team to explore other data sources.

In this case study, *CMS Open Payments data[6]* was explored to identify new potential targets for the sales force. The HCP universe for the market (including the company and defined competitors) was established as per the business rules identified earlier. Their level of contact across commercial channels (*Speaker Program – Speaker, Speaker Program – Attendee/Trainings, Speaker Training and In-Service Activities/Others)* was studied to funnel down to HCPs who could be targeted by reps. This was overlaid on the existing call plan to exclude HCPs who were already present in the plan. HCPs who participated in competitor clinical trials, or were called historically, were also excluded from the analysis (Figure 6).

The existing sales force target list and their planned annual touchpoints were validated. 3,108 additional physicians were identified for the cardiopulmonary sales and marketing team. A recommendation was made to add 450 of

them as potential targets (8% of the existing targets) to the call plan. Since there were no historical calls made to additional targets to measure promotional responsiveness, tier status was assigned based on comparison of targets per HCP/claims per HCP against current target list. (Figure 7)

Conclusion

This paper covered examples of how sales and marketing teams have used *CMS data[6]* to improve HCP targeting by identifying channel preferences for existing and potential new targets. There is more untapped potential in the CMS Open *Payments data[6]* to answer questions in sales and marketing space as represented in the universe of the problems (Figure 8). Over time, the richness and accuracy of data will further improve, providing companies the opportunity to monitor existing and newly acquired targets over time. Given that there is no cost associated with the data, if analytical rigor is appropriately applied, pharmaceutical companies can only derive more consumable insights, or at the least, directional insights. It's still a win-win situation.

About the Authors

Rahul Anand, *Engagement Manager, Mu Sigma has more than 5 years of experience in driving data-driven decision making in sales & marketing analytics for major pharmaceutical companies across multiple therapeutic areas.*

Duggan Collier, *Deputy Director, Commercial Analytics, Bayer is an accomplished business analytics executive with 20+ years of proven experience supporting global pharmaceutical sales operations through data analytics, business analysis, business intelligence and reporting.*

Yan Jiang, *Director, Business Insights, Rare Diseases, Bayer is an experienced market analytics, insight, and strategy professional with focus on commercialization of innovative biomedicine, medical devices, and pharmaceuticals for small biotechnology and large, multinational pharmaceutical companies.*

Janardhan Vellore, *Director, Commercial Analytics, Bayer is an analytics leader experienced in advanced sales, marketing and payer analytics and commercial ROI analyses leading forecasting efforts, brand planning, driving productivity initiatives, to mention a few of his responsibilities for supporting specialty drugs across the various stages of the product lifecycle.*

References

[1] Quintiles. Empowering healthcare research with realworld data [Internet]. Available from: http://www.quintiles.com/library/brochures/empowering-healthcare-research-with-realworld-data#pdf

[2] Symphony Health. PHAST Prescription Sales Sheet [Internet]. Available from: http://symphonyhealth.com/wp-content/uploads/2017/01/phast_prescription_sales_sheet.pdf

[3] Centers for Medicare & Medicaid Services. DUA – Limited Data Sets [Internet]. Available from: https://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/Privacy/DUA_-_LDS.html

[4] Centers for Medicare & Medicaid Services. DUA – Requesting CMS's Limited Data Set (LDS) Files [Internet]. Available from: https://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/LimitedDataSets/downloads/RevisedLDSInstructions.pdf

[5] Centers for Medicare & Medicaid Services. NPI Registry [Internet]. Available from: https://npiregistry.cms.hhs.gov/

[6] Centers for Medicare & Medicaid Services. Open Payments Data in Context [Internet]. Available from: https://www.cms.gov/OpenPayments/About/Open-Payments-Data-in-Context.html

[7] Centers for Medicare & Medicaid Services. Natures of Payment [Internet]. Available from: https://www.cms.gov/OpenPayments/About/Natures-of-Payment.html

[8] Centers for Medicare & Medicaid Services. NPI Files download [Internet]. Available from: http://download.cms.gov/nppes/NPI_Files.html

How a Lack of Data Integration Can Hurt Patient Services

Anshul Agarwal, Principal, ZS

Abstract: What are the consequences of an uninformed decision? When it comes to patients and their journeys toward health, the stakes are very high. Pharmaceutical companies struggle to fully understand each patient: Patient data is limited in access, and companies must gather disparate information across multiple sources to paint a full picture of the patient journey. While many data-savvy pharmaceutical companies understand the value of new data sources, such as electronic medical records and specialty pharmacy data, those information sources don't tell the whole story. Data integration is the key—and the future—for pharmaceutical companies looking to create collaborative patient services teams that can design effective treatment plans and customized patient experiences, resulting in improved outcomes for patients. With improved insight, companies are able to start, retain and take patients through the entire journey with better success.

Recently, Bill was diagnosed with non-Hodgkin lymphoma and prescribed a new medication. Shortly after, the reimbursement hub team calls Bill to welcome him and delivers a complimentary supply of his prescription. The next week, Bill receives a second welcome call, but this time it's from the specialty pharmacy team saying that his prescription has been approved and will be delivered in a few days. However, Bill has now received a second complimentary prescription, and his commercial shipment is on the way.

The next day, Bill receives another phone call. This time it's from a nurse educator on the pharmaceutical company's manufacturing team, and she discusses the medication's potential side effects and how to manage them. One week later, Bill receives two more compliance calls before his next shipment of the drug. Needless to say, these redundant calls from multiple patient services groups and the abundance of medication leave Bill feeling confused about his main point of contact and unsure of his required dosage.

This is a hypothetical example, of course, and Bill is a hypothetical patient, but the experience unfortunately is a common one. Patients often are stepping into unknown territory when they start a new treatment, and their inherent uncertainty or confusion often is exacerbated—rather than alleviated—by the "helpful information" that they receive from providers, payers and pharmaceutical companies. Put simply, a lack of coordination and data integration among pharmaceutical patient services teams can result in a disjointed, negative patient experience.

From the pharmaceutical company's perspective, the situation illustrated in that hypothetical example also is riddled with problems. Without the complete view of the patient shared among patient services teams, those teams are unable to design effective services that work together to assist the patient along his entire journey—resulting in inefficiencies, wasted spending and missed opportunities.

An incomplete view of the patient journey also inhibits the ability to measure a patient services effort's success: The nurse educator, for example, thinks that her follow-up phone call lead Bill to stay on his medication for the duration of his therapy, but the specialty pharmacist, who has been making regular compliance and adherence calls, thinks that he is the reason that Bill has stayed on his medication. Moreover, this inefficiency could lead to a liability risk. Patient services teams working in isolation could miscommunicate or deliver mixed messages regarding a critical element of a patient's therapy, such as dosage adjustment.

To improve the patient experience—an imperative for all of healthcare—it's key to understand the patient journey in its entirety, from diagnosis to recovery. This holistic view helps patient services teams work together across the organization to design effective treatment plans and customized patient experiences, resulting in improved outcomes for patients.

Solving the Data Problem
Most pharmaceutical companies are struggling to use data and technology to glean the necessary insights. Access to patient data is limited, and the data that is available exists in silos, so companies must gather disparate data sets from across multiple sources. However, now that pharmaceutical companies are starting to identify patient-level data—largely thanks to the recent maturation of certain data sources, like electronic medical records, claims data and specialty pharmaceutical data—technology has improved as a result. Data providers are better enabled to share patient-level data, and an influx in data platform options now help pharmaceutical companies combine data sources and view data more securely.

Pharmaceutical companies need to go into data integration with clear business goals in mind. Companies need to assess whether they're aiming to improve patient services overall, or if there's a particular team within patient services that needs to become more successful. Once they've identified their business goals, companies can get started with their data integration strategy by addressing these six key elements:

1. **Master data management:** Implementing an MDM system is the first step to conforming multiple sources of data at the patient level. To do this, the MDM system uses a unified patient ID—such as first name, ZIP code, etc.—that facilitates integration. During initial data processing, an MDM system also will limit the use of sensitive patient information. Clean, standardized data is a core foundation for integrated data, so this step is crucial.

2. **Developing a "secure enclave" for protected health information:** HIPAA regulations limit access to secure data, but within a secure platform, data scientists can work to integrate and enrich patient data, and patient services teams can leverage the integrated, de-identified data to improve their programs and support services. All teams "win" within this environment, and confidential patient information still remains confidential.

3. **Cultivating an efficient ecosystem:** Stakeholders within a pharmaceutical company must have access to an efficient ecosystem that allows them to work within clearly defined rules for leveraging data. This ecosystem allows those who are unable to view patient-identified data to continue to weigh in on strategy decisions with the patient services teams. This is an important step for those on the business side, keeping them connected and able to make changes on a future business strategy.

4. **Integrating non-pharmaceutical data sources:** To build a more complete view of the patient, companies need to get a view of the patient beyond her disease, and third-party data sources can help. While data such as media consumption and lifestyle habits might not seem relevant for pharmaceutical companies, information like this can help patient services teams devise the best way to engage with a patient to keep her on therapy, for example. For companies that want to continue to improve their patient services offerings, this is an important step because they can begin to identify commonalities among patients on the whole versus segmenting patients based solely on observations on symptoms and diagnosis.

5. **Building an ability to make predictions based on small data sets:** The more that data sources are integrated and filtered, the thinner the patient information becomes, so patient services teams need to feel comfortable gleaning insights from a smaller patient pool. There is no single, complete data set, so predictions become a necessity.

6. **Collaboration:** Integration of patient data is only one piece of the puzzle. Companies need to think about how to leverage this integrated view efficiently. It's key to incorporate change management at this step, ensuring that collaboration and communication occurs, and making sure that everyone involved with the design of a patient services plan is working together.

While advancements in technology and data availability finally make integration a possibility, companies need to understand the potential—and long-standing—challenges that come along with it. These key elements of a data integration strategy involve contracting with multiple vendors to provide and integrate patient data—a time- and effort-intensive endeavor. Additionally, because ownership of patient data lies across several functional teams within an organization, aligning them on an overall objective and solution poses another challenge. And finally, this undertaking requires significant monetary investment and executive buy-in.

These challenges, however, are nothing new, and companies should be prepared to handle them as they start their journey.

Getting It Right
One large biopharmaceutical company has already started expanding and improving its view of patients' journeys through data integration. The company aimed to maintain and increase market share, and with a number of competitor launches planned, the company knew that it needed to start integrating data and developing patient-focused analytics that would help uncover insights related to four key areas:

• **Patient adherence and compliance, and reasons for discontinuations:** Getting the data is one thing, but being able to connect it with the patient journey is another. Consider the earlier hypothetical patient, Bill: The patient services team could use historical patient data to identify characteristics that increase the likelihood of discontinuation. If Bill is deemed likely to discontinue, the team could give him a call and offer the support he needs.

• **The effectiveness of patient services being offered and the impact on patient retention:** If patient services aren't being offered at the right time in the patient journey, then they're useless. With visibility into all of the services that a patient is receiving and the point in

time in which he is receiving them, the patient services team can also assess the relationship between these services and determine their true impact—and how they could be optimized.

- **Potential new customer segmentations based on patient's groups and prior treatment preferences:** At this level, companies utilize underlying patient-level data but aggregate it at the physician level: The idea is to use type of patients and treatment selection by patient type to segment the customers.

- **Details of patients' journeys and the identification of risk areas:** Connecting patient data sources together allows for a better understanding of your patient, which should shape your messaging and patient services offerings. A more complete view of the patient could reveal, for example, that Bill has a comorbidity, allowing the patient services team to readjust their messaging and provide more specific support.

In order for the company to accomplish these complex goals, it aimed to develop an integrated information management solution that connected multiple patient-level data sources. The goal was to connect data from specialty pharmaceuticals, claims, lab, patient services, HUB and electronic medical records.

The overall solution design was created with the help of the IT team, but a crucial component was the support and input from the business side. Compliance and legal teams were integrated into the process to identify and address patient data security concerns while still meeting business needs. The teams also were leveraged to help manage contract negotiations with different data vendors to

ensure that patient-level data could be shared and connected.

Once the team designed the solution and secured the data inputs, the company embarked on developing the central component to the plan: a centralized patient de-identification and mastering solution. The team also worked to create a data warehouse capable of business intelligence, which connected patient-level data sources and housed business rules to create longitudinal patient journeys that facilitated analytics.

With integrated patient data now at their fingertips, the biopharmaceutical company was able to garner insight into patients who needed help. There was an uptake in patients with specific comorbidities, for example. The company also achieved visibility into risk areas that could be proactively managed through the right messaging, as well as the ability to design analytics-driven content for the market and for patients. Most importantly, the company was able to accomplish its goal of increasing market share in spite of growing competition, and it took a big step forward in its effort to become a patient-centric organization.

The Path Ahead
Once a company has integrated and optimized its patient data, how else can it make the data work for it? Turn that data into actionable insights that drive change. Here's a four-step framework for getting started:

1. **Engage:** Determine which patients will receive the most value out of a patient services program. Integrated data will provide greater visibility into historical patient adherence and compliance, as well as patient impact. Then, based on the collected data and patient experiences, companies begin to take that information

and segment future patients based on what they've learned about past patients, and then predict patient needs based on these segments. With these predictions, a company can design a stronger patient services program for the future. For example, based on the segmented patient data, companies can design a patient services program to provide unique support to patients with comorbid diseases, or they can adjust their programs to suit the needs of patients receiving private, in-home care. It's important to keep in mind, though, that much of the data required is "softer" data, like lifestyle habits and credit card usage, which is completely different than data from the medical ecosystem, and it might be more difficult to capture and merge.

2. **Get:** Incorporating newer data sources, such as social media or third-party data, will give a better glimpse into patient affinity and preferences. After integrating patient data, personalizing the experience for patients becomes a very real possibility, and technology plays a role. Social media, for example, can be a very effective tool for patient awareness. With the increasing number of people on social networks, the pharmaceutical industry needs to develop better social listening skills, following other industries' lead. This information will help map out the appropriate patient services—financial assistance, co-pay support, samples/free products, etc.—that best meet individual patients' needs in the early stages of the patient journey.

3. **Start:** Providing the appropriate onboarding services—treatment and disease education, peer support, injection/medication administration training—will help set patients up for success from the start. Interactive apps are now being used

to onboard patients faster: Humira offers an instant benefit verification app, and Pfizer has an app that includes information on all of its products. These apps not only offer a new way of onboarding patients faster and more easily, but also serve as data sources that can result in advanced analytics and help design a better patient experience. Companies should beware, though, of spending too much time designing these services and not enough time ensuring that patients are adopting them. Implementation support is key.

4. **Stay:** Lastly, ensuring that patients are well-equipped for long-term success is a key component of patient services. In addition to adherence and treatment reminders, community support, and lifestyle/treatment support, the movement toward the Internet of Things—in which everyday objects send and receive data—and connected health will help take adherence programs to the next level. Sensors and wearable devices will not only ensure adherence, but also help improve outcomes by providing the ability to continuously monitor patient progress and react in real time to various events. Here are a few examples:
   - *Gaming and loyalty:* Companies like HealthPrize and MangoHealth both have apps that leverage points and rewards to encourage adherence, much like airline miles or hotel points.
   - *Digital coaching:* Fit4D, CareSpeak, LegacyHealth and many others use telephonic and virtual coaches to support their patients in ongoing adherence.
   - *Peer connections:* Apps and programs are starting to create peer connections outside of the program, itself.

Because of this continued advancement in technology, though, companies need to

ensure that they're agile enough to adapt and refine their approach in lockstep with changes to patient adherence.

While these futuristic scenarios may seem far off for some companies, there are many out there that are already experimenting and seeing results. For pharmaceutical companies that embrace integrated patient data, a new world of insightful analytics opens up the possibility for improved patient services, an increase in adherence and an uptake in patient retention. Keep in mind, though, that how a company integrates data and uses it to improve its patient services function is a journey, not a flip of the switch. It isn't as easy as simply buying a technology and installing it. Therefore, it's imperative to get started now. It takes time to build up the right skills internally so that the entire organization can feel comfortable working with new and improved data sets, but the payoff is worth it: Integrated patient data is the foundation for creating stronger support for patients.

About the Author

Anshul Agarwal *is a Principal in ZS's San Francisco office. He has over 11 years of sales and marketing consulting experience in sales force design, distribution channel strategy design, patient analytics and information strategy design. He leads the global oncology and specialty therapeutics (distribution channel) strategy design and information management vertical at ZS. He's helped numerous clients design/execute a patient analytics and data management strategy.*