Transforming Population Health Management: The value of integrating socio-economic data into predictive models

August 2015
Introduction

The promise of population health management is to keep entire populations of patients generally well and out of the hospital. It's not just a better way of doing business and a better way to care for people. It's a health plan survival necessity under a radically restructuring health care system, especially with the Patient Protection and Affordable Care Act rewarding practitioners for value over volume when it comes to providing and paying for patient care services. To make the promise of population health a reality, at-risk health care organizations need to know details about their patients that go beyond medical records or pharmacy claims; to keep people out of inpatient facilities, those health care organizations must analyze data from multiple sources—familiar and not so familiar, clinical and non-clinical—and then push the results out to front line care managers.

Recent innovation in the realm of predictive modeling and population stratification is the incorporation of socio-economic attributes—from public records data, primarily—into existing stores of big data derived from traditional health care sources. It’s the next great leap forward that will power population health management. Now, the model developers at LexisNexis® have accomplished what scientists across the country have been attempting. They’ve created a Stress Index model to augment the firm’s predictive analytics options, representing the first meaningful industry integration of medical and pharmacy claims information with socio-economic data from outside sources. With the integration of socio-economic data, the LexisNexis Stress Index accurately predicts patients at a high risk of experiencing the onset or exacerbation of health conditions brought on by high levels of stress, stress levels that can’t be detected with medical or claims data alone.

Indeed, LexisNexis data scientists have confirmed that the Stress Index Model performs better than models that don’t include socio-economic data at identifying individuals with high stress, and thus a greater likelihood of requiring behavioral or social programs or interventions. The innovative model allows at-risk health care organizations to better identify patients they can contact to better manage their stress or their external stressors through programs available to them as part of their benefits packages or through wraparound behavioral or mental health care services.

The future, in other words, is here.
New Market Realities Require At-risk Health Care Organizations to Maximize Data Opportunities

At-risk health care organizations do not have the option to ignore the power of using external data sources to drive the predictive analytics that make population health management effective. The scientific and anecdotal evidence of the powerful link between stress and physical and mental health is far too compelling; at-risk health care organizations simply cannot attempt to do battle with stress without every analytical tool at their disposal.

- A recent article on the WebMD website notes that “stress that continues without relief can lead to a condition called distress—a negative stress reaction. Distress can lead to physical symptoms including headaches, upset stomach, elevated blood pressure, chest pain, and problems sleeping. Research suggests that stress also can bring on or worsen certain symptoms or diseases.” Specifically, the article points out, some 43% of all adults “suffer adverse health effects from stress.”

- A detailed article on the Working Mother website cites a well-known number from the Ft. Worth, Texas-based American Institute of Stress: “Stress-related absenteeism, lost productivity, legal and insurance fees, and employee turnover cost U.S. companies more than $300 billion per year.” Additional points the article makes include these:
  - “Stress is responsible in some way for a whopping 70% to 90% percent of doctor visits, either directly or by aggravating an ongoing medical condition,” it says, quoting Cynthia Ackrill MD, who chairs the Workplace Stress Board for the American Institute of Stress.
  - “Recent data show American private-sector workers take almost four times as many days away from work yearly—a median of 31 days—because of stress, anxiety and related issues than they do for all other nonfatal injuries and illnesses,” it adds, quoting Roxanne Conrad, Director of Work Life Products and Services for ComPsych, an employee assistance program provider.

Better Data Is Imperative to Understanding Risk

The key to taming the stress beast is knowing which members of an at-risk health care organization’s patient population are most likely to encounter it. And that requires using the most powerful data resources available.
Until today, hospital billing information, lab data, and medical and pharmacy claims data mostly told you what happened to a patient in the past and, thus, what might happen in the future. But in a changing health care environment, that’s no longer adequate. All data are not created equal; at-risk health care organizations must use the best information available, and not just what’s traditionally been put to use. Now, non-clinical socio-economic data from public sources—financial, environmental, and legal, for example—must be added to paint more complete pictures of patients and how aspects of their environment can impact their mental and physical health, and thus their costs to your organization. The following information provides an assessment of the relative value a data source may have today to improve predictive models to find hidden health risk.

- Medical and pharmacy claims have been used for decades by payers to understand cost trends and, to the extent possible, quality outcomes. And a majority of members in a health plan—80% to 90%—will have claims data; new enrollees and non-users will not. In most cases, some data cleansing is required due to gaps in data processing or collection, but, overall, claims data is available in a very structured and standardized format, so it is reliable. Claims data is the benchmark to compare the next three categories of data.

- Lab results would be very beneficial in predicting future risk, but they’re not always readily available to link with claims data. For ambulatory services and holistic management of patients, labs would paint a very narrow picture on their own without claims or clinical data. With their time-sensitive nature, along with their lack of availability, lab results have limited potential to improve a risk prediction’s accuracy for revealing hidden health risk today in comparison to claims data.

- Health Risk Assessment (HRA) data is also not readily available, particularly in comparison to claims data—but that’s changing quickly. HRA data could be helpful as an early warning for potential risk of new enrollees; however, due to its limited availability—as of this paper’s publication—and the fact that HRA data is self-reported, it is considered to have questionable potential to improve the accuracy of a prediction. HRA data tells you something about missed/unknown diagnoses, family history, other lifestyle factors that may not be available anywhere else; therefore, it is widely believed that it will have value—but it will take some time to further refine how it compares to the use of other data sources.

- The final category is the most compelling. Public and proprietary data that reveals socio-economic factors and other community influencers is available and used across multiple industries today to fight crime, locate missing people, and assess risk. The use of this data in health care has been very limited, even though there is a relatively large amount of data available on individuals and the data is very reliable in its format and structure. Its potential to improve the accuracy in predicting risk and revealing hidden or unknown trends is extremely high.
The value of integrating socio-economic data into predictive models

LexisNexis Maintains a Database of Unique Data

When you know how to look for useful information, the models you design can reveal new or hidden trends that are invisible—or difficult to predict—with basic clinical claims data only. The key to integrating socio-economic data with existing clinical information is knowing which available datasets actually help a model better identify exactly the types of patients it’s designed to identify. Data scientists need to have vast volumes of data, on a surprisingly broad set of topics, and need to know just how much weight each aspect of the data should have in determining which individual patients to pinpoint.

Toward that end, LexisNexis maintains more than two petabytes of unique data; the records it manages contain billions of public and proprietary documents. Included are, for example, some 9 billion name/address combinations and about 4 billion property records. Several aspects of a person’s address may provide powerful insights into that person’s stress risk, such as whether the person rents or owns a home and whether he or she resides in an urban or rural setting—and of potentially particular interest may be the amount of crime in the person’s neighborhood. If you know what you’re looking for, those records can shine a light on previously undetectable stress risks in time to manage them most effectively.

LexisNexis First to Quantify Effects of Stressors on Health

The LexisNexis difference is knowing what to look for in public data—even when no one else has been bold enough to define it and quantify its effect on a patient’s health. In this case, “it” was stress.

“The biggest challenge in creating the Stress Index was defining and measuring stress,” explains Ogi Asparouhov PhD, Chief Data Scientist. “Typical or common stressors that come to mind are money—or lack of it—work situations, the economy, personal safety, relationships, and family responsibilities. However, quantifying the effect of those stressors on physical health is another challenge.” So the clinical team created a model that incorporates socio-economic data to help reveal how stress contributes to the overall risk of an individual member. “We know what happens from a medical claims perspective,” Dr. Asparouhov explains, “but what about daily life? We wanted to hold a magnifying glass over what’s happening on a personal basis.”
For example, LexisNexis scientists developing the Stress Index looked carefully at how income affects stress. Lower income is generally linked to higher stress, but it’s not only low-income patients who are affected by it. Annual salary increases can also impact a person’s stress level. LexisNexis data scientists discovered that people generally suffer less stress if their annual increases are high than if they’re low; that sounds intuitive on the surface, but small differences in salary increases can drive the Stress Index from a low figure to one that’s off the proverbial charts. Similarly, when LexisNexis scientists quantified the effect of neighborhood crime, they found a strong correlation between data trends in publicly available current address burglary information and the stress a patient suffers.

Another important finding: Trauma is correlated to stress level. That’s incredibly significant for the payer community because trauma is regarded by many in health care predictive analytics as an unpredictable event. But if anything can be related to it, it is stress level, and LexisNexis scientists determined that the two trend together “across the board.” More stress generally, possibly, equals less attention to risks, and that can equal a higher propensity for trauma. No other model in the health care industry can target such a rise in the potential for trauma.

“We try to go beyond the obvious,” Dr. Asparouhov says, “things other people can’t tackle because there’s nothing out there that says exactly what they are, nothing that’s easy to create a regression model around. We’re willing to go after the less obvious, which is harder to do.” The final version of the LexisNexis Stress Index model integrates dozens of socio-economic attributes with 750 medical attributes. Socio-economic data tends to provide better-quality information that the traditional data sources—those mentioned previously—used in predictive models. While medical and pharmacy claims data is voluminous, readily available and reliable, other sources, such as Health Risk Assessment data, are often less available, less reliable, and incomplete. Using vast and reliable public record data for socio-economic predictors has been shown to add significant value in identifying patients with high stress and hidden future costs.

The LexisNexis Stress Index Is Better—It Works

The Stress Index has the ability to capture some of the hidden future high-cost complications and events that aren’t identified by other models, making it especially useful for high-risk members who have near-term future savings potential that can’t be captured by other standard industry predictions. LexisNexis scientists have found that several outcomes that the health care industry perceives as dependent on socio-economic factors do in fact trend very well with severity on the Stress Index scale—including ER visits, admissions and readmissions, mental health costs, and acute costs.
The Stress Index is used alone or in combination with other predictions to better “see” a member and his or her potential for avoidable risk. Care managers examining a patient’s record can use the LexisNexis Stress Index just like other predictions available in LexisNexis products—Total Forecasted Cost, Compliance Risk, Motivation and others—that can help them better design and better customize care management programs for patients who are going to have higher stress. A care manager in an asthma disease management program, for example, now knows one of his patients is also under tremendous stress. That alerts him to the potential for myriad categories of higher costs in time to offer additional interventions that can keep the patient healthier and lower the payer’s costs.

With richer and more reliable data, care management teams can stratify higher-risk patients for inclusion in wraparound services. For example, high-stress patients may need wraparound behavioral or social services plus their chronic or disease management program interventions. Low-stress individuals, then, could be stratified specifically into disease-based programs that don’t require significant wraparound services.

The LexisNexis Stress Index changes the conversation. Many people who may want to engage in their own care simply have other life stressors that prevent them from doing so. Care managers can now be prepared ahead of time to ask questions designed to better understand a member and what may challenge him or her to receive care. The care manager’s day-to-day employment of the Stress Index could, thus, include specific assessments for those behavioral or social programs, including moving patients out of programs they’re already in; a specifically tailored interview style; or notification of the availability of employee assistance programs. Indeed, care managers would have the information they need to move their patients with the highest stress into programs that target it first, making secondary a focus on their specific medical diseases.

About the Author

Dr. Ogi Asparouhov is Chief Data Scientist and innovator at LexisNexis, one of the leading U.S. health care analytics companies. He has spent more than three decades inventing and developing solutions to address a variety of pressing industry issues across clinical analytics, big data, risk adjustment and health care predictive modeling.

He is an expert in the creation of risk assessment indices; identification of segments, clusters, patterns, and outliers; classification and regression, using both industry standard and custom built predictive modeling and data mining tools.
LexisNexis Population Health Management Analytics

LexisNexis offers a continuum of clinical solutions that address population health and provider performance management. The goal of our unique continuum is to empower payers wherever they find themselves in terms of technology and analytic sophistication. Our population health analytics continuum starts with the availability of socioeconomic health risk data attributes sold for ingestion into existing client software and platforms to a full turnkey-workflow solution, LexisNexis Population Health Monitor, which aggregates and analyzes data to enable care managers and better inform care plans.

This white paper is provided solely for general informational purposes and presents only summary discussions of the topics discussed. This white paper does not represent legal advice as to any factual situation; nor does it represent an undertaking to keep readers advised of all relevant developments. Readers should consult their attorneys, compliance departments and other professional advisors about any questions they may have as to the subject matter of this white paper.

LexisNexis does not utilize clinical or socioeconomic data for the purposes of determining or validating health insurance premiums.

For more information:
Call 866.396.7703 or visit
www.lexisnexis.com/risk/healthcare

1 The Effects of Stress on Your Body. © 2014 WebMD LLC.

2 The Real Cost...of Stress. http://www.workingmother.com/content/real-costof-stress

LexisNexis® Population Health Monitor and the Stress Index is not provided by “consumer reporting agencies,” as that term is defined in the federal Fair Credit Reporting Act (15 U.S.C. § 1681, et seq.) (FCRA) and do not constitute “consumer reports,” as that term is defined in the FCRA. Accordingly, Population Health Monitor and the Stress Index may not be used in whole or in part as a factor in determining eligibility for credit, insurance, employment or another purpose in connection with which a consumer report may be used under the FCRA. Due to the nature of the origin of public record information, the public records and commercially available data sources used in reports may contain errors. Source data is sometimes reported or entered inaccurately, processed poorly or incorrectly, and is generally not free from defect. This product or service aggregates and reports data, as provided by the public records and commercially available data sources, and is not the source of the data, nor is it a comprehensive compilation of the data. Before relying on any data, it should be independently verified.

LexisNexis and the Knowledge Burst logo are registered trademarks of Reed Elsevier Properties Inc., used under license. Other products and services may be trademarks or registered trademarks of their respective companies. Copyright © 2015 LexisNexis. All rights reserved. NXR11018-00-0515-EN-US