

WHITE PAPER

Social determinants
of health fuel deeper
understanding of health risks
and ability to proactively
impact health outcomes

JUNE 2017

Predictive analytics organizations are recognizing the value of integrating non-clinical data, primarily socioeconomic data on social determinants of health from outside sources, into the analytics that feed a patient's care plan. LexisNexis® is a leader in identifying, integrating and applying data on these social risk factors to create predictive models that can improve risk stratification and population healthcare management initiatives.

The reason for focusing on social determinants of health data

The implementation of the Affordable Care Act in 2010 boosted the development and use of Big Data and analytics across healthcare. Suddenly, the surge of new health system enrollees made it imperative for predictive analytics organizations to locate and employ new data sources that were capable of assessing and predicting risk without the use of medical history data. Since then, the continuing shift to value-based delivery and payment models has only intensified the need to be able to proactively predict and address potential health risks before they occur or worsen.

Prior to utilizing socioeconomic data, the healthcare outcome predictive modeling industry relied on a strong mixture of four essential components: clinical, healthcare, mathematical and computer knowledge. The industry's first 25 years—the “risk adjustment methodology epoch” spanning 1980 to 2005—were mostly dominated by a medical/healthcare approach, including Diagnosis-Related Groups (DRGs), Diagnostic Cost Groups (DxCs) and Episode Risk Groups (ERGs). All predictors were infused with medical terminology and meaning.



Risk Adjustment Methodology Epoch

Clinical • Healthcare • Mathematical • Computer Knowledge

Ancillary information sources mainly stressed medical information, such as lab data and health risk assessment (HRA) questionnaires that primarily queried patients about diseases and drugs. The first HRA questions without medical context were focused on style of living, including exercise and physical activities, drinking habits and level of education. A problem quickly arose: HRA information was not available consistently in all medical records, what was available was not necessarily reliable since the data was self-reported and quickly became outdated, and the medical information contained within them was correlated with the diseases and drugs already available in the medical and Rx claims records.

There were also no new sources of medical information primed to add predictive power to the currently available medical knowledge. The only medical source that had not been thoroughly explored was the genomics data, which required significant investment—and 20 to 30 years. The conclusion was inescapable: improvement in the accuracy of healthcare outcome predictive models required new sources of information that were easily available on a huge percentage of the population, but that were not correlated with already-available medical sources. The choices were:

- socioeconomic data
- lifestyle data
- consumer survey data and
- unstructured data such as text mining and social media analytics.

After assessing these options for availability, reliability, human experience working with such data and cost, integration of medical and pharmacy data with socioeconomic data soon emerged as the next logical step of the healthcare outcome predictive modeling industry. This data had the ability to improve the accuracy of predictions for as many people as possible.



Incorporation of socioeconomic data

The most common types of social determinants of health used in healthcare predictive modeling

LexisNexis has extensively tested the validity of hundreds of data attributes created from public and proprietary records. Validated socioeconomic attributes have been found to correlate to health outcomes including cost, hospitalizations, 30-day readmissions, emergency room visits, medication adherence, motivation and stress. Key data correlated to health outcomes includes level of education, address stability, income, taxes, properties, assets, change in housing costs and more. Correlated derogatory attributes include arrests, liens, bankruptcies, evictions, accidents, evidence of unpaid financial obligations, high-interest loan applications, relatives' information, business associates' information and data on neighborhood burglaries and other crimes.

Trends in those attributes can be more important than the values themselves. For example, a serious reduction in income could emerge as a harbinger of high stress, with all its associated health consequences. A woman's last name change during the last 24 months could be a sign of a future pregnancy associated with getting married and starting a family or of a period of high stress related to getting divorced.

Social determinants of health outcomes



Neighborhood

Housing ♦ Transportation ♦ Safety
Parks ♦ Playgrounds ♦ Walkability



Economic Stability

Employment ♦ Income ♦ Expenses
Debt ♦ Medical Bills ♦ Support



Education

Literacy ♦ Language ♦ Early Childhood
Vocational Training ♦ Higher Education



Community

Social Integration ♦ Support Systems
Engagement ♦ Discrimination



Food

Hunger ♦ Access to Healthy Options



Health Care System

Coverage ♦ Availability ♦ Linguistics
and Cultural Competency ♦ Quality

OUTCOMES

Mortality ♦ Morbidity ♦ Life Expectancy ♦ Expenditures ♦ Health Status ♦ Limitations



To predict risk related to these social determinants of health, LexisNexis leverages more than 65 billion records from over 10,000 different public and proprietary sources, including:

- ◆ 19.2 billion consumer records
- ◆ 10.2 billion unique name and address records
- ◆ 5 billion property records
- ◆ 45 million active US business entities
- ◆ 44 million business contact records
- ◆ 278 million unique cell phone numbers
- ◆ 1.5 billion bankruptcy records monitored monthly
- ◆ 5 billion motor vehicle registrations
- ◆ 1.1 billion vehicle title records
- ◆ 517 million criminal records

To tap into the most valuable information for the customer, LexisNexis Data Scientists employed a variety of statistical methods and visualization techniques to conduct a comprehensive evaluation of the predictive power of different attributes that touch on important healthcare outcomes—based on a 24-month medical and pharmacy claim history and more than 400 socioeconomic attributes. The resultant observations are stored in the Enhanced Attribute Selection (EASE) framework, a proprietary software tool that provides an efficient way to recalculate predictive power of socioeconomic attributes, rank and select top attributes, and create reports that illustrate the value of those attributes in predicting healthcare outcomes. This data can be used for predicting health risk in conjunction with—or in the absence of—other risk adjustment models, improving the accuracy of existing models or predicting risk for individuals who have no or limited clinical history. The resulting healthcare predictive models can be used to better coordinate proactive care management to improve health outcomes, reduce costs and improve member or patient satisfaction and retention.

Examples of attributes included in two social determinants of health predictive risk scores

LexisNexis used Socioeconomic Health Attributes to create Socioeconomic Health Scores that calculate risk based on an individual's predicted future total cost of healthcare and based on a patient's predicted readmission risk. Below are examples of how attributes related specifically to property correlate to these health outcomes. Hundreds of other attributes were also integrated in the scores.

Case 1:



Prediction of Future Total Cost

The Address Recent Economic Trajectory Index, which indicates the property type and value trajectory for an input address, reveals that individuals who downsize houses more typically have higher healthcare costs, less education, more applications for high-interest loans and lower income than people who move up. This trend is simply not captured by information available in medical and pharmacy claims, but is factored into the LexisNexis Socioeconomic Health Score: Total Cost Risk Score.

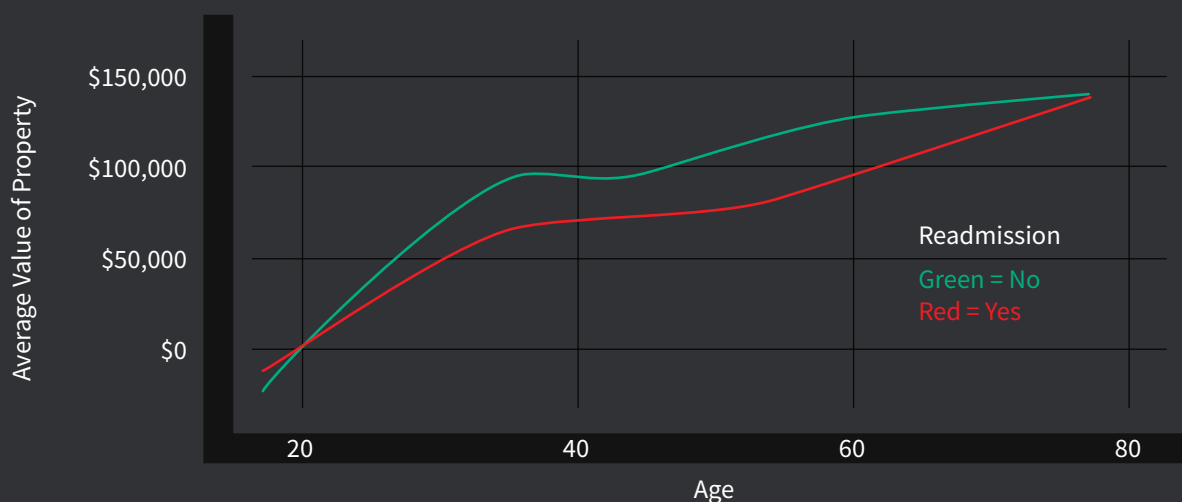
Index	Count	Age	Per Member Per Month Healthcare Cost	Level of Education	Applications for High-Interest Loans	Income
Moved downscale	119,545	49.1	\$480	0.19	0.55	\$37,473
Moved down	465,376	49.6	\$472	0.25	0.24	\$48,243
Stayed	357,048	49.6	\$455	0.33	0.15	\$61,590
Moved up	115,758	51.0	\$449	0.37	0.06	\$87,377
Moved upscale	16,028	52.1	\$459	0.34	0.06	\$104,537
Stayed upscale	8,164	51.8	\$456	0.38	0.05	\$95,637

Case 2:



Prediction of Readmission Likelihood

Property characteristics also correlate to readmission, as shown in the graph below. For patients ages 25-75, those who own properties with higher tax assessed values are less likely to be readmitted to the hospital within 30 days compared to patients of the same age who own properties with lower assessed values. This trend is also not captured by medical and pharmacy claims, nor would it be evident based on age alone. It is factored into the LexisNexis Socioeconomic Health Score: Readmission Risk Score.



Conclusion: Using social determinants of health data sources to stratify degree of health risk is a real game-changer in healthcare predictive analytics

We believe the industry should remain focused on socioeconomic data because that data is reliable and can be consistently accessed when sourced from public and proprietary records. Low correlation exists with already-available medical data which means that social determinants of health can identify pockets of high health risk hidden in medical data, which does not account for an individual's social, economic or environmental factors. The revelations go beyond the

available information of just age and gender, allowing for more precise risk stratification. Using socioeconomic data also provides the opportunity to identify trends over time for relevant attributes.

It is becoming increasingly apparent that data derived from social determinants of health should be integrated into the prediction of healthcare outcomes such as those affected by stress, motivation, medication adherence, readmissions and disease complications. All are extremely important in the population health management arena, and social determinants of health data could be a game-changer in improving healthcare outcomes and curbing rising costs.

For more information, call 866.396.7703 or visit
lexisnexis.com/risk/healthcare



Health Care

About LexisNexis® Risk Solutions

At LexisNexis Risk Solutions, we believe in the power of data and advanced analytics for better risk management. With over 40 years of expertise, we are the trusted data analytics provider for organizations seeking actionable insights to manage risks and improve results while upholding the highest standards for security and privacy. Headquartered in metro Atlanta, LexisNexis Risk Solutions serves customers in more than 100 countries and is part of RELX Group plc, a world-leading provider of information and analytics for professional and business customers across industries. For more information, please visit www.lexisnexis.com/risk.

Our healthcare solutions combine proprietary analytics, science and technology with the industry's leading sources of provider, member, claims and public records information to improve cost savings, health outcomes, data quality, compliance and exposure to fraud, waste and abuse.

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 and the Knowledge Burst logo are registered trademarks of RELX Inc. Copyright © 2017 LexisNexis. NXR12100-00-0617-EN-US