Social Determinants of Health

And Predictive Modeling

SOHAYLA PRUITT
Director
Product Management
Health systems must devise new ways to adapt to an aggressively changing reimbursement environment, competitive pressures, and economic realities. This requires finding new ways to reduce readmissions, control emergency room utilization, manage chronic conditions, prevent medication nonadherence, and manage a patient throughout his or her entire clinical episode to achieve bundled payment success.

To address these challenges, many organizations are turning to real-time predictive analytics, which hold promise in being able to predict a patient’s risk, prescribe appropriate interventions, and ensure the allocation of resources that align with the needs of the individual patient. However, they struggle to harness the clinical data they already have access to, and often find external data — a key ingredient to an accurate and actionable predictive program — too costly and complicated to integrate. But it need not be.

Social Determinants of Health

According to recent research, clinical information from a care encounter contributes only about 10% of what drives population health (Figure 1). Social determinants of health, including health behaviors and physical environment, account for approximately 80%. A recent IOM report underscores this knowledge gap and recommends that social data be incorporated within EHRs, but it also acknowledges the difficulty of integrating such sources external to the care environment.

External person-, household-, and geographic level data sets can be costly to acquire, keep up to date, and require a number of full time specialty resources and expensive software in order to ensure the data sets are appropriately related to the patients’ clinical record, and accessible by predictive models.

Given these challenges, how does a health system become better aware of the SDH that impact their patients’ clinical outcomes, utilization, and cost of care?
The Forecast Health Solution

Forecast Health’s next generation analytics help healthcare organizations improve patient and population health by identifying high-risk patients with unprecedented accuracy, further identifying the subset of patients who are potentially “impactable”, and providing patient-specific guidance to mitigate this risk. We have overcome the challenges of exposing thousands of SDH variables to our predictive models, and we are able to achieve a predictive accuracy that is twice as precise as other models.

We have roughly 4000 person- and household-specific data variables, which span a number of SDH topics. Through master data management and automated processes, we enrich each patient’s clinical record with information, such as car ownership, debt, credit risk, and discretionary income.

“"We provide health systems with access to thousands of SDH variables, without the expense of data acquisition, processing, specialized software and personnel."”

We also collect a number of geospatial data sets which help characterize the physical environment in which the patient resides, works, or frequents. Additionally, the socio-economic, lifestyle, and behavioral patterns of the neighborhoods in which each patient resides are included. The clinical EHR data is enriched to compute the relationship between the patient’s geocoded address and thousands of features in their physical environment; to include their driving distance to their PCP, the nearest hospital, park, fitness facility, or fast food restaurant, as well as the block group neighborhood statistics related to income, socio-economic status, education, consumer spending, and customer segmentation.

By merging this information with clinical and claims-level data, we are able to create more accurate, reliable, and actionable model results. We are also able to identify for health systems a much smaller subset of SDH variables that are predictive of a patient’s risk of poor health outcomes, utilization, and cost. In doing so, we provide the benefit of having access to thousands of SDH variables, without the expense of data acquisition, processing, specialized software, and personnel. Only those statistically significant SDH variables that we identify as relevant to a health system’s population are accessed in our risk algorithms in order to provide real time intervention information within a health system’s electronic medical record system.

Actionable, Not Just Visual

With increasing awareness of the importance of social determinants of health, multiple vendors have started claiming that they offer such capabilities. It is, however, important to distinguish between our solutions and those available in the marketplace.

Some claim to raise awareness of SDH through attractive map visualizations within business intelligence tools. Clinical phenomenon can be overlaid alongside features such as: business points, parks, and U.S. Census measures. Figure 2 is an example of a visualization published by ESRI, where dark red ZIP codes represent high percentages of low birth weight births, and green dots represent the location of OBGYN providers.3

While it can be initially useful to visualize data outside the confines of rows and columns to discover emerging patterns, it is difficult to translate the information into prescriptive action or to identify new patients with similar influences.
In the case of Figure 2, the graphic visually implies that ZIP codes with higher rates of low birth rate babies are correlated with being located further away from OBGYN clinics. Clinicians might see such graphics and ask any one of the following questions: Is the issue proximity to OBGYN locations or issues with transportation? Are lower birth rates correlated with median household income also? Maybe it’s not distance to OBGYN or income that correlates, but what if it’s a combination of factors? How do we use this when I’m seeing a patient from one of these low birth rate ZIP codes?

While Forecast Health offers a number of similar visualizations, we take the additional steps needed for healthcare systems to make this information clinically actionable. Such steps include the enrichment of the clinical information with thousands of discrete person-, household-, and socio-geographic- level data elements which enable the data to be amenable to advanced predictive modeling capabilities capable of analyzing thousands of data variables to identify the ones that are statistically significant and contribute to a particular phenomenon. This allows clinicians the ability to identify specific patients who are at risk, understand their clinical and SDH risk drivers, and take recommended action based on such information.

The Highest Levels of Granularity

Although other vendors claim to analyze SDH information alongside clinical information, they generally; use only highly aggregate data sources at the county-, city-, ZIP-level of geography. While data sources at these levels are readily available and inexpensive, they are not particularly useful in predicting a patient’s risk or determining the best way to intervene. Many health systems serve populations that are incredibly diverse within any given geography. For this reason, Forecast Health uses only the most granular person and household specific information.
Seamless, Interoperable, and Scalable

There are a number of academic medical centers who have shown value in using socio-geographic data to explore links between health outcomes and the social determinants of health. However, these efforts are often grant funded for a specific cohort of patients and are not scaled to the entire patient population. The methods used to investigate SDH on clinical outcomes are often ad-hoc, manual, inconsistent, and time-consuming; therefore, the results are not easily integrated back into the EMR. These methods also make it difficult to incorporate thousands of SDH variables for scientific selection using advanced machine learning methods. Instead, the SDH variables considered are often hand-selected based on what the investigator thinks may have influence, simply because there are not sufficient resources to get all the data needed to do a more advanced factor selection process.

Forecast Health solutions allow for thousands of SDH variables to be factored in alongside the clinical information in order to identify the unique clustering of factors that influence a patient’s risk and their likelihood to be impactable. Our solutions then integrate seamlessly back into a health system’s EMR, where they only license and pay for the SDH variables needed to achieve good outcomes, reduced utilization, and lower costs.

Conclusion

Given the potential impact that social, demographic, lifestyle, behavioral, and physical factors have on the health of a population, Forecast Health’s solution will assist health systems with integration of this information at the point of care — within each patient’s EHR, without having to take on the burden of data acquisition, or implementing processes to keep accurate, and amenable by predictive models. Furthermore, the use of this information within clinical decision support tools embedded within the EMR or care management platform, allows this information to be a part of the clinical workflow and to be translated into clinical, prescriptive action.

About the Author

SOHAYLA PRUITT — With 16+ years of experience in predictive analytic and geospatial solutions within healthcare, homeland security, law enforcement, and marketing, Sohayla leads our predictive analytics development. Most recently, she led the Health Intelligence team at Duke Medicine, and deployed analytics within the Epic EMR, using both clinical and socio-geographic information. She’s Epic-certified in several Epic data models and applications.

REFERENCES:

