

EHR Data, Claims Data Can Successfully Predict Hospitalizations

Christopher Jason

Researchers at Atrius Health in Massachusetts found that the use of EHR data is an accurate way to predict hospitalizations with a diverse group of patients.

Predictive tools using EHR data, claims data, and combined data all possessed similar predictive value when identifying potential hospitalizations in a six month period, [according](#) to a study in the *American Journal of Managed Care*.

Predicting hospitalization among a diverse group of patients is a difficult process, but can be useful for determining provider workflows and allocating hospital resources. The use of [EHR data](#) is an accurate way to predict hospitalization.

"The healthcare system generates, collects, and stores a tremendous amount of data during the course of a patient's clinical encounter, with one study finding an average of more than 200,000 individual data points available during a single hospital stay," the authors wrote.

"These data are used to monitor a patient's progress, coordinate care among all members of the healthcare team, and provide documentation for billing and reporting activities. Although the use of data for these purposes has been long-standing, the availability of these data has increased substantially."

Researchers analyzed EHR data adult patients seen at Atrius Health, a large multi-specialty group in Massachusetts, from June 2013 to November 2015. To get a [broad sample size](#), they selected patients with different demographics, medications, clinical dosages, and prior utilization. Some were insured under Medicare, Medicaid, and commercial contracts.

"Data sets capable of linking EHR and claims data at the patient level remain uncommon," wrote the authors. "We hypothesized that when combined, these two data sources would complement each other and lead to stronger prediction than that observed previously."

From there, researchers developed three different types of models to predict hospitalization within six months, including EHR-only, claims-only, or both EHR-and-claims data.

Overall, 185,388 patients were included for analysis. With this large and broad sample size, a variety of predictive characteristics were observed such as age, prior healthcare utilization, and risk level.

Throughout the study, researchers were able to develop a risk score that accurately projected hospitalization in the following six months. Using the area under the receiver operating curve, a measure used to determine prediction model accuracy, the researchers found that using only EHR data, only claims data, or combined data sources were nearly all equally accurate.

EHR-only and claims-only data both received a 0.84 score, while the combined claims and EHR data received a 0.846 score.

"Although our results suggest some utility to combining EHR and claims data to inform predictive model creation, we find that even in scenarios in which only EHR or claims data are available, strong performance can be achieved provided that a diverse collection of variable types is represented," explained the authors.

"The risk prediction score was also found to be well calibrated in those less likely to be hospitalized in the next 6 months, but it did become less accurate among those at higher risk of hospitalization," they continued. "The model tended to overestimate the likelihood of hospitalization in those with higher than 30% predicted risk, likely owing to the small number of patients demonstrating such high risk."

One major limitation was that all the data was derived from one health system without an external center to authenticate results. Still, researchers do believe that any health system could apply these same methods to adjust the model for its own patients and system of care.

"We believe that our model approach is a meaningful step toward identifying patients at highest risk of hospitalization," the authors concluded. "Tying the model to care interventions that are likely to modify the risk of hospitalization represents a promising area for future research."