Managing the health needs of groups of patients, rather than one patient at a time, can improve health outcomes. However, better methods of predicting health risks are needed to manage healthcare costs. The current scientific literature on “risk stratification in population health management of chronic disease” agrees that predicting risks for groups of patients can improve outcomes, but doesn’t agree on the best methods for risk prediction and doesn’t agree on whether it costs less. The current scientific literature on “data mining for predictive healthcare analytics” agrees that risk predictions need to be interpretable by doctors and nurses, rather than “black box” techniques that don’t give explanations for their predictions. This study used interpretable methods, combined with simple but powerful techniques such as breaking down costs and using charts to show patient data, to improve prediction of health risks for a group of patients.

This study looked at the medical histories of 10,000 mid-Missouri Medicare and Medicaid patients between 2012 and 2014. Mathematical techniques were used to identify patients at different levels of risk for high healthcare needs. The changes in patient risk level over time were measured, and the differences between patients with different risk levels were analyzed. The high-risk and very-high-risk patients made up only 21% of all the patients but went on to incur 43% of hospital charges for the group. Patients in the most expensive sub-group of the most very-high-risk were nearly twice as costly as very-high-risk patients on average. These new techniques predicted the most expensive 5% of patients with 84% accuracy.

All the strategies used in this study, from the simplest to the most sophisticated, produced useful insights. By predicting the small number of patients who will incur the majority of healthcare expenses in ways that doctors and nurses can interpret, these methods can support healthcare providers to focus on preventive care more effectively. These models, and similar models developed by combining different mathematical techniques, could improve health outcomes, delivery, and costs.