TAKING ISSUE

The Promise of Large, Longitudinal Data Sets

In a study in this issue, Gören and colleagues used longitudinal data to assess whether providers followed guidelines when prescribing antipsychotics. This is an advance over studies that have assessed the quality of care with cross-sectional data. The point prevalence of antipsychotic polypharmacy may be of interest when assessing the quality of care. However, starting a patient with schizophrenia on antipsychotic polypharmacy immediately, rather than reserving it until other treatments have failed, has different implications.

Using longitudinal data to assess quality makes sense because clinical care is longitudinal. It consists of a series of treatment decisions made by the physician and patient over time as more information is gathered. "Doc, that medication affected my sex drive" is something we take into account when deciding next steps, along with the extent to which symptoms were reduced by the most recent drug tried and the drug tried before that. In essence, we make a series of increasingly informed trial-and-error decisions. Guidelines have become more longitudinal as treatment options have multiplied and are being applied in a stepped fashion.

However, large longitudinal data sets can be quite complex to analyze, and treatment pathways become numerous after just a few decision nodes. Gören and colleagues attempted to address this problem by collapsing pathways and simplifying their analyses and conclusions. The complexities of analyzing longitudinal treatment choices and pathways will be further compounded as data sets become richer and administrative and pharmacy databases, such as those used in this study, are combined with genetic data, medical record data, and patients' reports of outcomes.

Luckily, new approaches to analyzing longitudinal "big data" resources—machine learning and data-mining approaches—are also being developed. Such approaches may allow us to go beyond determining whether treatment recommendations are being followed to actually improving the recommendations. (For example, the recommendation for clozapine use after two failed antipsychotic trials is a "best guess," and patients might be better off with clozapine after one or three failed trials.)

Already, large data sets and data-mining techniques have been used to identify previously unknown drug interactions (such as the combination of paroxetine and pravastatin causing high blood sugar) and to explore ways to improve longitudinal treatment decisions for major depression. "Big data" and sophisticated analytic approaches may soon help us make better treatment decisions.—MARCIA VALENSTEIN, M.D., M.S., *U.S. Department of Veterans Affairs Center for Clinical Management Research and Department of Psychiatry, University of Michigan, Ann Arbor*