Discretion and Guidelines: Evidence from Warfarin Administration

Jason Abaluck, Leila Agha and David Chan

August 15, 2016

Extended Abstract

A large literature in health economics focuses on whether too much money is being spent: whether we are on the ``flat of the curve'' in terms of the marginal return to additional spending (Garber and Skinner 2008). Physicians typically ask instead how a given dollar can be spent wisely: which patients should receive various tests and treatment? Some recent evidence in health economics takes the latter approach and finds potentially large efficiency gains from physicians more effectively allocating tests and treatments to patients with the largest benefits (Currie and MacLeod 2013; Abaluck, et al. 2014). Advisory boards and medical experts are increasingly promoting the use of clinical risk scores to guide treatment decisions.

This paper studies a popular risk score based guideline. We first ask how adoption of the existing guideline impacts physician prescribing behavior and patient outcomes. Next, we develop a methodology to estimate heterogeneous causal treatment effects and then apply the estimates we recover to characterize optimal guidelines. We conduct simulations comparing status quo treatment decisions to counterfactual strict adherence to existing and optimal guidelines as well as discretionary adherence in which physicians integrate information not observable in medical records into their treatment decisions.

Specifically, we study the decision of whether to prescribe anticoagulation, typically the drug warfarin, to treat patients with atrial fibrillation. Atrial fibrillation is an irregular heartbeat that increases stroke risk and causes 450,000 hospitalizations and 99,000 deaths each year (Ott et al. 1997). Anticoagulation can reduce stroke risk but increases hemorrhage risk. Through alternative prescribing patterns, could physicians achieve fewer strokes for the same number of bleeds? During our study period, a treatment guideline called the CHADS2 score is introduced. Using a large database of clinical health records from the Veterans' Health Administration, we construct the CHADS2 score for each patient with atrial fibrillation, record the treatment choice, identify stroke and bleed outcomes, and observe whether the physician noted the CHADS2 score in their patient record.

Guidelines like the CHADS2 score are constructed to predict risk among untreated patients, and then recommend treatment for patients with the greatest stroke risk. There are several reasons to think that this approach might be flawed. First, treatment effects need not be proportional to risk. Second, the CHADS2 score may omit important variables, and the limited set of conditions belonging to the score may not have been selected systematically. Third, doctors consider patient characteristics that are often difficult to record objectively in medical record data.

Our approach addresses each of these problems. We use a jackknife instrumental variables strategy based on patient quasi-random assignment to physicians to estimate treatment effects. We combine this approach with machine learning methods to estimate how treatment effects vary with patient characteristics. Finally, we embed all of this in a generalized Roy selection model. This estimation procedure recovers, for every set of patient characteristics, how the number of strokes prevented and bleeds induced vary as physicians move from treating 0% to 100% of patients with those characteristics. We can thus determine the fraction of patients doctors should treat with each set of observables given their private information about patients' returns to treatment.

We find that when physicians adopt the CHADS2 score, they become more likely to prescribe anticoagulation to patients with a high score and less likely to prescribe to patients with a low score. Stroke rates decrease among high score patients. By the end of our sample, 65% of physicians have noted using the CHADS score. Second, we find that our machine learning model selects quite different variables than physician experts have selected as being most important to include in clinical guidelines. Preliminary results suggest that an optimal guideline could reduce the number of strokes in the population by 11% to 20% without increasing the current rate of hemorrhage.