

Principal Investigator

First Name: Adam
Last Name: Kapelner
Degree: PhD
Primary Affiliation: Queens College, CUNY
E-mail: kapelner@gmail.com
Phone number: 7189975817
Address: 65-30 Kissena Blvd, Kiely Hall Suite 604

City: Flushing
State or Province: NY
Zip or Postal Code: 11367
Country: United States

2017-1466

General Information

Key Personnel (in addition to PI): **First Name:** Adam
Last name: Kapelner
Degree: PhD
Primary Affiliation: Department of Mathematics, Queens College, CUNY
SCOPUS ID:

First Name: Alina
Last name: Levine
Degree: BA
Primary Affiliation: Division of Healthcare Policy and Research, Weill Cornell Medical College
SCOPUS ID:

First Name: Zachary
Last name: Cohen
Degree: MA
Primary Affiliation: Department of Psychology, University of Pennsylvania
SCOPUS ID:

First Name: Rob
Last name: DeRubeis
Degree: PhD
Primary Affiliation: Department of Psychology, University of Pennsylvania
SCOPUS ID:

Are external grants or funds being used to support this research?: External grants or funds are being used to support this research.

Project Funding Source: MQ Foundation, London UK

How did you learn about the YODA Project?: Internet Search

 [coi_rob.pdf](#) [coi_zach.pdf](#) [coi_adam.pdf](#) [coi_alina.pdf](#)

Certification

Certification: All information is complete; I (PI) am responsible for the research; data will not be used to support litigious/commercial aims.

Data Use Agreement Training: As the Principal Investigator of this study, I certify that I have completed the YODA Project Data Use Agreement Training

Associated Trial(s): [NCT01106677 - A Randomized, Double-Blind, Placebo and Active-Controlled, 4-Arm, Parallel Group, Multicenter Study to Evaluate the Efficacy, Safety, and Tolerability of Canagliflozin in the Treatment of Subjects With Type 2 Diabetes Mellitus With Inadequ](#)

[NCT00968812 - A Randomized, Double-Blind, 3-Arm Parallel-Group, 2-Year \(104-Week\), Multicenter Study to Evaluate the Efficacy, Safety, and Tolerability of JNJ-28431754 Compared With Glimepiride in the Treatment of Subjects With Type 2 Diabetes Mellitus](#)

[NCT01137812 - A Randomized, Double-Blind, Active-Controlled, Multicenter Study to Evaluate the Efficacy, Safety, and Tolerability of Canagliflozin Versus Sitagliptin in the Treatment of Subjects With Type 2 Diabetes Mellitus With Inadequate Glycemic Con](#)

What type of data are you looking for?: Individual Participant-Level Data, which includes Full CSR and all supporting documentation

Research Proposal

Project Title

The use of clinical experimental data to build personalized medicine models and the evaluation of their effectiveness and statistical significance.

Narrative Summary:

In medical practice, when more than one treatment option is viable, there is little systematic use of individual patient characteristics to estimate which treatment option is most likely to result in a better patient outcome. For instance, some diabetes patients may have better outcomes on metformin than on insulin (or vice-versa).

It would be valuable to have a way of (1) sorting these two types of patients using a statistical model (2) estimating how clinically impactful the model will be when it is used to determine treatments for future patients. This type of system is not available presently and it would be of tremendous use to clinicians and 29.1 million people with Diabetes.

Scientific Abstract:

Background: Physicians use clinical experience when deciding between two effective treatments to give to a patient. However, it has been shown that data driven modeling can outperform clinician personal decisions when it comes to choosing best treatment option.

Objective: Using data from clinical trials each comparing two or more treatments, we'll create sorting rules that will use patient covariates to assign future patients to treatments and will test how well these rules perform when compared to random or uniform allocation.

Study Design: We employ a custom study design because our deliverables are (1) the quantified personalization advantage over business-as-usual and (2) its confidence interval and (3) its significance level. To do so, we make use of standard tools such as regression, imputation (if there is missing data) and the bootstrap.

Participants: The participants in our study will be people with diabetes who have participated in clinical trials comparing different treatments.

Main Outcome Measures: The main outcome measure is drop in A1C, but our significance test will be based on the

difference between the expected drop in A1C for future patients when our allocation rule is employed vs. when random or uniform allocation is used.

Statistical Analysis: Our sorting rule will be based on a simple linear model using first order interactions and our significance test will be done using a bootstrap approximation to the estimator for allocation rule advantage.

Brief Project Background and Statement of Project Significance:

Background

Consider an individual seeking medical treatment for a set of symptoms. After a diagnosis, suppose a medical practitioner has two treatment options, neither of which is known to be superior for all patients. How does the practitioner choose which treatment to administer?

Sometimes practitioners will select a treatment based informally on personal experience. Other times, practitioners may choose the treatment that their clinic or peers recommend. If the practitioner happens to be current on the literature of published RCTs, the studies' "superior" treatment may be chosen.

Each of these approaches can sometimes lead to improved outcomes, but each also can be badly flawed. For example, in a variety of clinical settings, "craft lore" has been demonstrated to perform poorly, especially when compared to even very simple statistical models (Dawes, 1979). It follows that each of these "business-as-usual" treatment allocation procedures can in principle be improved if patient characteristics related to how well an intervention performs are taken into account.

These patient features can be used to construct a "personalized medicine model" (Chakraborty & Moodie, 2013). The need for personalized medicine is by no means a novel idea. As noted as early as Bernard (1865), "the response of the average patient to therapy is not necessarily the response of the patient being treated". Since then, there has been much work done on finding moderators i.e patient characteristics that relate to differential drug response (Gail & Simon, 1985, Silvapulle, 2001, Dusseldorp & Van Mechelen, 2014); on building multiple-stage experiments using patient characteristics (Murphy, 2003, James, 2004); on model selection given the patient characteristics (Gunter et al., 2011) and much more. Our previous work, (Kapelner et al., 2014) provided a framework for evaluating such procedures by answering these outstanding questions on personalized medicine models:

How well do these models perform on future subjects?

How much advantages so these models provide to patients compared to the "naive strategies for allocating treatments currently used by medical practitioners?

How confident can one be about these estimates of patient "improvement"

Statement of Project Significance

Our work is of paramount importance as we seek to create personalized treatments for widely-used drugs that treat the world's most pernicious diseases. The models we use are built with the RCT data you provide and evaluated for performance and significance using our open-source software that has been in use around the world over for a couple of years already.

Our procedure is general, but we choose to begin with diabetes. NCT01106677, NCT01137812, and NCT00968812 are clinical trials that compared canagliflozin to either glimepiride or sitagliptin (in combination with metformin for two of them). After our work is done, we will know which patients should be given each drug and how much better they are expected to respond to this personalization. Such models will also provide insight into the inner-workings of these drugs that has the potential to spawn future research.

Specific Aims of the Project:

Our project aims to (1) create diabetes treatment allocation rules and (2) test whether they perform significantly better than random allocation or "best allocation" (every patient is given the treatment that performed best on average in clinical trials). Our project will help physicians decide which treatment to give a specific patient, when deciding between multiple treatments.

We aim for a completely different type of analysis from the analysis conducted in the requested trials. Generally speaking, the requested trials studied the average differences between treatment regimens. They randomized diabetes patients into multiple groups, gave each group a different treatment, and assessed whether or not the

averages of the endpoints in each group were significantly different.

Our analysis instead will find the average improvement with certain allocation rules. Therefore, we are not validating the original results. A treatment that is found to be better on average is not necessarily better for each patient and a treatment that is on average not found to be better or worse than another treatment may be more effective for certain patients. Our project will allow physicians to assign a patient to a treatment depending on which treatment will most likely work best for that patient.

What is the purpose of the analysis being proposed? Please select all that apply. Preliminary research to be used as part of a grant proposal
Other

Research Methods

Data Source and Inclusion/Exclusion Criteria to be used to define the patient sample for your study:

It is likely the vast majority of patient-level data will satisfy our broad requirements that are at the study level only.

- (1) Randomization into at least two treatment conditions (the Diabetes drug regimens)
- (2) sample size of at least 500

Main Outcome Measure and how it will be categorized/defined for your study:

The main outcome measure of our study will be the same as the primary endpoint from the trials from which we are requesting data. For our analysis of all three trials, we will be interested in the change in HbA1c at either 26 weeks or 52 weeks (depending on the study). We will be calculating the difference between the value of different allocation rules, and this value will be defined as the expected decrease in HbA1c when using this particular rule.

Main Predictor/Independent Variable and how it will be categorized/defined for your study:

In each of our analyses, treatment allocation is our main independent variable. For study NCT01106677, patient treatment will be either sitagliptin in combination with or canagliflozin both in combination with metformin and in study NCT01137812 treatment groups will be either canagliflozin or sitagliptin. For study NCT00968812 patient treatment will be either canagliflozin or glimepiride. For modeling purposes, the treatment allocation will be defined as a binary dummy variable.

Other Variables of Interest that will be used in your analysis and how they will be categorized/defined for your study:

The other variables of interest in our study are the patient level characteristics. These are our potential moderators inducing heterogeneous treatment effects. Likely many of the baseline variables that were collected in these studies fall into this category. The following is an example list of baseline variables that may affect treatment performance that could be included in our analysis.

Age
Sex
Race
Time from diagnosis
Weight
Systolic blood pressure
Diastolic blood pressure
History of Hypertension (Y/N)
BMI
Fasting blood glucose
HbA1c
Fasting insulin
Current Smoking Status
Past Smoking Status
HDL cholesterol
LDL cholesterol
Total Triglycerides
Triacylglycerols carbon number

Triacylglycerols double bond number
Prior Treatments
Additional Diseases / Comorbidities

Note that the more such possible moderating variables, the better as we will be able to fit better personalized medicine models.

Statistical Analysis Plan:

The analyses have three steps.

Step 1.

Here we first build a model that attempts to capture heterogenous treatment effects. This is similar to discovering “qualitative interactions” (e.g. Silvapulle, 2001). This model we call a “personalized medicine model” as it allows us to predict a new patient’s endpoint on both treatment alternatives and allows a practitioner to estimate the better of the two treatment alternatives.

To build this model we make use of standard tools. If the endpoint is continuous, we will use OLS regression; incidence, logistic regression; survival, the suite of the popular survival modeling techniques. If there is missing data, we will use multiple imputation. We will return to that discussion in Step 3. To induce heterogeneity, we will use first order interactions with the treatment allocation. Time permitting, we can examine more elaborate interaction models, even machine learning techniques.

We would like to stress that we do not need the personalized medicine model to be “true” or for its assumptions to be “true” in any absolute sense. This is a divergence from classical statistics that does indeed necessitate the model to be true and assumptions met for valid inference. We are only assessing if the model is useful and we turn to this now.

Step 2.

If we were to use this personalized medicine model in the future, i.e. use it to predict personalized treatments to new patients, how well would we do? This requires a definition of “how well” and necessitates a competitor for comparison purposes.

“How well” we will define to be the average outcome of the patient that is administered a treatment based on the personalized medicine model’s recommendation minus the average outcome for the competitor. The competitor we define in two ways (a) randomly allocating the two treatments and (b) always administering the treatment which does better on average in the RCT data. Thus “how well” is measured in the native units of the endpoint in the RCT data and thus it is interpretable e.g. “our personalized medicine model lowers fasting serum glucose by 12.7mg/dL on average OR lowers incidence of heart attacks by 15.2% on average OR increases survival by 1.9 years on average”. These numbers express the “advantage” of employing the personalized medicine model in the real world.

How are we able to estimate the personalized medicine model advantages in the future? Here, we employ 10-fold cross validation, an out-of-sample validation procedure that gives honest results of future performance (Hastie et al., 2013).

Note that “better models” will have greater personalization advantages. This is the main reason that the model does not need to be “true”, only “useful” in our context.

Step 3

Step 2 provides a point estimate to the personalization advantage. However, we would like to know the uncertainty in this estimate (confidence intervals) and whether it is statistically significantly different from zero, indicating a stable advantage of the personalized medicine model (hypothesis testing).

In order to provide confidence intervals and hypothesis testing, we make use of the bootstrap. We bootstrap estimates of the out-of-sample advantage metric. This is a rather elaborate, computationally expensive procedure, but it is asymptotically valid.

The open-source software implementation of all three steps is already available within the R package Personalized Treatment Evaluator currently available on CRAN (as package "PTE"). And thus, we can hit the ground running in your secure data access system.

Project Timeline:

We can start our proposed research program immediately upon access to the data. The analysis will take at most one month and writing a short paper will take at most two months. This timeline is comfortably within the limit of the 12-month access period.

Dissemination Plan:

We anticipate highly impactful, original and well-articulated results. Thus, we plan to publish in a top journal such as NEJM or JAMA, etc i.e. prestigious journals that have not yet seen quantitative results concerning personalized medicine a la our research program.

Bibliography:

Bernard, Claude (1865). *Introduction à l'étude de la médecine expérimentale*. Paris.

Chakraborty, B. and Moodie, E. E. M. (2013). *Statistical Methods for Dynamic Treatment Regimes*. Springer, New York.

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, 34(7):571-582.

Dusseldorp, E. and Van Mechelen, I. (2014). Qualitative interaction trees: a tool to identify qualitative treatment-subgroup interactions. *Statistics in medicine*, 33(2):219-37.

Gail, M. and Simon, R. (1985). Testing for qualitative interactions between treatment effects and patient subsets. *Biometrics*, 41(2):361-72.

Gunter, L., Zhu, J., and Murphy, S. (2011). Variable selection for qualitative interactions in personalized medicine while controlling the family-wise error rate. *Journal of Biopharmaceutical Statistics*, 21(6):1063-1078.

Hastie, T., Tibshirani, R., and Friedman, J. H. (2013). *The Elements of Statistical Learning*. Springer Science, Tenth printed edition.

Kapelner, A., Bleich, J., Cohen, Z. D., DeRubeis, R. J. & Berk, R. A. (2014) *Inference for Treatment Regime Models in Personalized Medicine*. arXiv

Silvapulle, M. J. (2001). Tests against qualitative interaction: Exact critical values and robust tests. *Biometrics*, 57(4):1157-65.