the YODA PROJECT | Forging a unified scientific community

Principal Investigator

First Name: Joanna
Last Name: Harton
Degree: MS (Candidate for PhD)
Primary Affiliation: University of Pennsylvania Department of Biostatistics, Epidemiology, and Informatics
E-mail: joanna.harton@gmail.com
Phone number:
Address:
428 Guardian Drive Room 108/109-C
City: Philadelphia
State or Province: Pennsylvania
Zip or Postal Code: 19104
Country: United States of America

General Information

Key Personnel (in addition to PI):
First Name: Joanna
Last name: Harton
Degree: MS
Primary Affiliation: University of Pennsylvania

First Name: Nandita
Last name: Mitra
Degree: PhD
Primary Affiliation: University of Pennsylvania

First Name: Rebecca
Last name: Hubbard
Degree: PhD
Primary Affiliation: University of Pennsylvania

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: NIH
How did you learn about the YODA Project?: Colleague

Conflict of Interest

https://yoda.yale.edu/system/files/yoda_project_coi_form_for_data_requestors_2019_mitra.pdf
https://yoda.yale.edu/system/files/yoda_project_coi_form_for_data_requestors_2019_hubbard.pdf
https://yoda.yale.edu/system/files/yoda_project_coi_form_for_data_requestors_2019_harton_0.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

.misc-fixes { display: none; } #admin-region { z-index: 9999999; } #admin-menu { z-index: 99999999; }

1. NCT00638690 - COU-AA-301 - A Phase 3, Randomized, Double-Blind, Placebo-Controlled Study of Abiraterone Acetate (CB7630) Plus Prednisone in Patients With Metastatic Castration-Resistant Prostate Cancer Who Have Failed Docetaxel-Based Chemotherapy
2. NCT02236637 - 212082PCR4001 - A Prospective Registry of Patients With a Confirmed Diagnosis of Adenocarcinoma of the Prostate Presenting With Metastatic Castrate-Resistant Prostate Cancer

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

Research Proposal

Project Title

Data-Adaptive Weighting of Real-World and Randomized Controls Using Propensity Scores: Creating a Hybrid Control Arm

Narrative Summary:

Clinical trials with a hybrid control arm, an arm with both randomized patients and real-world data on standard clinical care patients, are lower in cost. But, due to stringent trial inclusion criteria and differences in care between trials and community practice, control patients will likely have better outcomes than real-world ones. Our new method to analyze these trials controls for bias and error. We weight each real-world subject by their similarity to the randomized controls. Hence patients that better resemble randomized controls count more and dissimilar subjects less. We compare our approach to existing ones via simulations and apply these methods to a study using real-world data.

Scientific Abstract:

Background: Clinical trials with a control arm constructed from both trial patients and real-world data on patients receiving standard care have the potential to decrease the cost of randomized trials. However, due to stringent trial inclusion criteria and differences in care quality between trials and community practice, randomized control patients will likely have superior outcomes compared to their real-world counterparts.
Objective: For our proposed approach, each real-world subject is weighted by a function of the propensity score reflecting their similarity to the randomized controls while randomized subjects receive full weight. This weighting allows for real-world patients that more closely resemble randomized controls to have a larger contribution to the likelihood while dissimilar subjects are discounted.
Study Design: This is a hybrid control arm study where subjects in the disease registry who meet the incl/exclusion criteria for the clinical trial will be added to the clinical trial patients to demonstrate results using several different statistical methods.
Participants: All subjects have metastatic castration-resistant prostate cancer: NCT02236637, the external data source, and NCT00638690, a clinical trial. This will serve as a real-world example to supplement the simulations that have been conducted to evaluate the proposed method.
Outcome Measure: The main outcome of interest is overall survival.
Statistical Analysis: Survival models will be fit using the power prior, the normalized power prior, the commensurate prior, and the proposed method, DAW

Brief Project Background and Statement of Project Significance:

The goal of this project is to introduce a new statistical method, data-adaptive weighting (DAW) to the larger statistical community. The concept of DAW is to incorporate electronic health record (EHR) data into clinical trial data in order to increase power or decrease the sample size needed to achieve a given power. The development of these types of methods has become more important in the past few years as EHR databases become more widespread with entire health systems keeping records for patients in a central location. While there are existing methods that combine EHR and trial data, they weight each subject within the EHR data with a common weight, which does not account for the heterogeneity seen in EHR data. Furthermore, several methods rely on a researcher deciding a priori how similar s/he thinks the EHR data is to the clinical trial data and uses that weight in

the analysis, requiring several sensitivity analyses to be performed as well. Data-adaptive weighting uses a function of the propensity score, which is defined here as the probability that a subject will be on-trial as opposed to in the EHR, given their observed covariates, to determine the similarity between each subject present in the EHR and the trial subjects.

Specific Aims of the Project:

The objective of this study is to evaluate the performance of existing methods that incorporate EHR data with clinical trial data (power prior with several alpha values, normalized power prior, commensurate prior) in addition to the proposed method, DAW, under a variety of simulation scenarios that vary trial size, proportion of the trial that is treated, size of the EHR data available, treatment effect, "on-trial" effect, and confounding strength.

What is the purpose of the analysis being proposed? Please select all that apply.
Develop or refine statistical methods
Research on clinical trial methods

Research Methods

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

NCT00638690 for the trial data and NCT02236637 for the external data source. The inclusion/exclusion criteria that have been applied to the trial data will be applied to the external data source as best as they can in order to obtain a pool of external controls. The trial data will be used in its entirety.

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

The main outcome measure will be overall survival if there are enough events present or progression-free survival if that is recorded and there are not enough events with overall survival.

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

The main predictor will be an indicator for whether the subject is on-trial or in the EHR database and whether the subject received abiraterone acetate plus prednisone or prednisone only.

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

Continuous covariates: age, PSA level (ng/mL), hemoglobin (g/dL), alkaline phosphatase (U/L), time from initial diagnosis to beginning of study
Categorical covariates: geographic location (country-level), ECOG performance status (<=2 vs >2), Gleason score (<=6 vs. >6), M-stage (Mx and M0 vs all others), presence of metastases (number and/or location), comorbidities/comorbidity score, treatment history at study start

Statistical Analysis Plan:

Several Bayesian statistical methods will be performed with the data requested as a 'real-world' example to complement the simulation studies performed in the main section of the paper. The power prior, developed by Ibrahim and Chen in 2000 is a method that raises the likelihood for the entirety of the available EHR data to a power, also known as the alpha value, that is selected by the researcher and bounded by 0 and 1. An alpha value of 0 corresponds to not incorporating any information from the EHR data and an alpha value of 1 corresponds to fully pooling the EHR and trial data together. Another method is the normalized power prior, which was developed by Duan and Ye in 2008 and is similar to the regular power prior but with a normalizing constant present that allows for the alpha value to be estimated from the data instead of set by the researcher. The commensurate prior was developed by Hobbs, et al. in 2011 and uses a prior that quantifies a distribution for the true hazard based upon the hazard in the external data set. The proposed method, DAW, weights each external subject by their inverse probability weight (IPW), if the estimand of interest is the average treatment effect, or by their inverse odds weight (IOW), if the estimand of interest is the average treatment effect for those on-trial. Both IPW and IOW are functions of the propensity score that quantifies the probability of a subject being on-trial given their observed covariates.

Software Used:
RStudio
Project Timeline:

The project has already begun and simulation studies are completed. Manuscript writing aside from the real-data component will be completed by the end of October and real-data analysis as well as the finalization of the manuscript will be completed by early December 2020. Results will be reported back to the YODA project when the manuscript is submitted for publication at the end of 2020.

Dissemination Plan:

The project will result in a scientific paper with a target audience of the larger statistical community. The manuscript will be submitted to Biopharmaceutical Statistics for their special issue on real-world evidence.

Bibliography:

Duan, Yuyan and Keying Ye (2008). "Normalized power prior Bayesian analysis". In: The University of Texas at San Antonio, College of Business Working Paper Series.

Hobbs, Brian P, Bradley P Carlin, et al. (2011). "Hierarchical commensurate and power prior models for adaptive incorporation of historical information in clinical trials". In: Biometrics 67.3, pp. 1047–1056.

Hobbs, Brian P, Daniel J Sargent, and Bradley P Carlin (2012). "Commensurate priors for incorporating historical information in clinical trials using general and generalized linear models". In: Bayesian analysis (Online) 7.3, p. 639.

Ibrahim, Joseph G and Ming-Hui Chen (2000). "Power prior distributions for regression models". In: Statistical Science 15.1, pp. 46–60