Principal Investigator

First Name: Dennis
Last Name: Shung
Degree: MD MHS PhD
Primary Affiliation: Yale School of Medicine
E-mail: dennis.shung@yale.edu
State or Province: CT
Country: USA

General Information

Key Personnel (other than PI):
First Name: Jasjeet
Last name: Sekhon
Degree: PhD
Primary Affiliation: Yale University
SCOPUS ID:
Requires Data Access? Unknown

First Name: Dennis
Last name: Shen
Degree: PhD
Primary Affiliation: University of California-Berkeley
SCOPUS ID:
Requires Data Access? Unknown

First Name: Siddharth
Last name: Singh
Degree: MD
Primary Affiliation: University of California-San Diego
SCOPUS ID:
Requires Data Access? Unknown

First Name: Badr
Last name: Al-Bawardy
Degree: MD
Primary Affiliation: Yale School of Medicine
SCOPUS ID:
Requires Data Access? Unknown

First Name: Loren
Last name: Laine
Degree: MD
Primary Affiliation: Yale School of Medicine
SCOPUS ID:
Requires Data Access? Unknown

First Name: Kisung
Last name: You
Degree: PhD
Primary Affiliation: Yale School of Medicine
SCOPUS ID:
Requires Data Access? Unknown

First Name: Rohan
Last name: Alur
Degree: PhD
Primary Affiliation: Massachusetts Institute of Technology
SCOPUS ID:
Requires Data Access? Unknown

First Name: Theo
Last name: Saarinaen
Degree: BA/ BS/BSc
Primary Affiliation: University of California, Berkeley
SCOPUS ID:
Requires Data Access? Unknown

First Name: Mauro
Last name: Giuffre
Degree: MD
Primary Affiliation: Yale School of Medicine
SCOPUS ID:
Requires Data Access? Unknown

First Name: Bahar
Last name: Ardestani
Degree: MD
Primary Affiliation: Yale New Haven Hospital
SCOPUS ID:
Requires Data Access? Unknown

First Name: Simone
Last name: Kresevic
Degree: MA / MS / MSc
Primary Affiliation: Yale University
SCOPUS ID:
Requires Data Access? Unknown

Are external grants or funds being used to support this research?: No external grants or funds are being used to support this research.
How did you learn about the YODA Project?: Colleague

Conflict of Interest

https://yoda.yale.edu/wp-content/uploads/2022/03/COI_FORM_MG.pdf
https://yoda.yale.edu/wp-content/uploads/2019/11/coi_form_ba.pdf
https://yoda.yale.edu/wp-content/uploads/2018/01/coi_form_ds.pdf
https://yoda.yale.edu/wp-content/uploads/2019/08/coi_form_js.pdf
https://yoda.yale.edu/wp-content/uploads/2017/05/coi_form_ll.pdf
https://yoda.yale.edu/wp-content/uploads/2016/05/coi_form_ss.pdf
https://yoda.yale.edu/wp-content/uploads/2017/09/coi_form_ky.pdf
https://yoda.yale.edu/wp-content/uploads/2019/12/coi_form_ra.pdf
https://yoda.yale.edu/wp-content/uploads/2014/10/coi_form_ts.pdf
https://yoda.yale.edu/wp-content/uploads/2022/03/COI_FORM_SK.pdf
https://yoda.yale.edu/wp-content/uploads/2022/03/COI_FORM_MG.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

1. NCT00036439 - C0168T37 - A Randomized, Placebo-controlled, Double-blind Trial to Evaluate the Safety and Efficacy of Infliximab in Patients With Active Ulcerative Colitis
2. NCT00096655 - C0168T46 - A Randomized, Placebo-controlled, Double-blind Trial to Evaluate the Safety and Efficacy of Infliximab in Patients With Active Ulcerative Colitis
3. NCT00487539 - C0524T17 - A Phase 2/3 Multicenter, Randomized, Placebo-controlled, Double blind Study to Evaluate the Safety and Efficacy of Golimumab Induction Therapy, Administered Subcutaneously, in Subjects with Moderately to Severely Active Ulcerative Colitis
4. NCT01551290 - CR018769; REMICADEUCO3001 - A Phase 3, Multicenter, Randomized, Double-Blind, Placebo-Controlled Study Evaluating the Efficacy and Safety of Infliximab in Chinese Subjects With Active Ulcerative Colitis
5. NCT00488631 - C0524T18 - A Phase 3 Multicenter, Randomized, Placebo-controlled, Double-blind Study to Evaluate the Safety and Efficacy of Golimumab Maintenance Therapy, Administered Subcutaneously, in Subjects With Moderately to Severely Active Ulcerative Colitis
6. NCT00488774 - C0524T16 - A Phase 2/3 Multicenter, Randomized, Placebo-controlled, Double-blind Study to Evaluate the Safety and Efficacy of Golimumab Induction Therapy, Administered Intravenously, in Subjects With Moderately to Severely Active Ulcerative Colitis

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

Research Proposal

Project Title

Precision Medicine for Patients with Ulcerative Colitis Using Synthetic Controls

Narrative Summary:

Biologic therapies are the cornerstone of treatment for moderate to severe ulcerative colitis. With the proliferation of different inflammatory targets and separate Phase III and Phase IV trials for each therapy, study of comparative efficacy and rates of adverse events are necessary to evaluate how to use the different biologic therapies given different mechanisms of action, delivery mechanism, and risk of adverse events. The synthetic control method can be used on a patient level to estimate comparative effectiveness to biologic therapies and safety profile as measured by adverse events using data from Phase III and Phase IV clinical trials.

Scientific Abstract:

Scientific Abstract:
Background
Previous efforts to use Phase III clinical trial data for ulcerative colitis to evaluate the comparative effectiveness of biologic therapies include network meta-analyses and propensity score studies.1-4 Clinical trial data with synthetic controls can be used to quantify the efficacy of specific treatments

and the risk for adverse events for individual patients.
Objective: Validate the performance in quantifying treatment effect and adverse event risk of multiple biologic therapies for patients with ulcerative colitis using the synthetic control method.
Study Design: Match overlapping covariates between the two import sources (e.g., Phases III/IV clinical trials and RWD). For each outcome measure of interest, identify the predictive covariates using feature importance algorithms (e.g., regression-based models such as LASSO or decision-tree based models such as random forest). Validation is performed by holding out a test set of patients (i.e., obfuscate their outcome measures) and apply the method to the remaining patients on the predictive covariates to recreate the test set (ground truth) outcomes. Counterfactual outcome measures will be predicted by applying the method to all patients of interest on the predictive covariates.
Participants: Patients enrolled in ACT-1, -2, PURSUIT-SC and PURSUIT-M, and UNIFI
Main Outcome Measures: PRO-2 defined remission without physician global assessment (PGA) or Mayo endoscopy score
Secondary: Endoscopic outcomes, adverse events
Statistical Analysis: The baseline comparisons will be nave methods (mean, median), parametric methods (multivariate logistic regression covariate adjustments), non-parametric methods (nearest neighbor matching methods using various distance metrics, e.g., l2 or l1), and random forest (mapping from covariates to outcomes).
For the primary outcome variables that are continuous, average prediction error (e.g., mean-squared-error) will be calculated across all patients in the test set and then compared to the above baseline methods.
For the secondary outcome variables that are binary (e.g. adverse events), the area under the receiver operating characteristic curve will be calculated for our method and compared to other baseline methods using the DeLong non-parametric test. Robustness tests will be performed with the following strategy: changing the proportion of the dataset used for training and evaluate the average prediction error.

Brief Project Background and Statement of Project Significance:

The absence of multiple head-to-head trials for biologic therapies in the treatment of ulcerative colitis leaves a knowledge gap regarding comparative efficacy. Synthetic controls is a widely used method in econometrics that have been used in clinical applications to quantify treatment effects and adverse event rate in cohorts of patients using real-world data. 5-7 Synthetic controls can be used to construct a synthetic comparison group of ?patients? most similar to patients studied for treatment efficacy in a Phase III Clinical Trial and adverse event rate studied in a Phase IV clinical trial. By creating multiple synthetic comparisons, we can quantify the treatment effect and adverse event rate of different interventions studied across different Phase III/IV clinical trials. We will apply the method to UC patients undergoing induction therapy across multiple Phase III trials to quantify treatment effect and comparative efficacy as well as Phase IV trials to quantify risk of adverse events for different biologic therapies.

Specific Aims of the Project:

- Validate the synthetic controls method in quantifying treatment effect of biologic therapies from clinical trial data for patients with UC
- Validate the synthetic controls method in quantifying adverse event rate of biologic therapies from clinical trial data for patients with UC

Research Methods

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

Phase III/IV trials of biologic therapies for Ulcerative Colitis

Primary and Secondary Outcome Measure(s) and how they will be categorized/defined for

your study:

PRO-2 defined remission without physician global assessment (PGA) or Mayo endoscopy score. The PRO-2 is the patient-reported outcome score, which ranges from 0 to 6. This raw score will be used as the continuous outcome after the treatment course. The Mayo endoscopy score is an ordinal score based on endoscopist assessment from 0 to 3. This raw score will also be used as the continuous outcome after the treatment course.

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

Age, Gender, Race (primary race)/Ethnicity (non-Hispanic or Hispanic), Medical Conditions and Co-morbidities (ICD-9 or ICD-10 codes for specific diagnoses), Concomitant Medications (Medications listed upon recruitment to the trial with dosages), Surgical History (ICD-9 or ICD-10 codes for specific surgical diagnoses in the past, including endoscopic evaluation), Medication History (Previous biologic or immunomodulating treatments in particular), Previous treatment of IBD (Previous treatment history with specific formulations as by medication history), Smoking History (quantified by estimated pack-years), Vital Signs (systolic blood pressure, heart rate, diastolic blood pressure, oxygen saturation, respiratory rate), Laboratory Values (C-reactive protein, sedimentation rate, fecal calprotectin, basic metabolic panel, hepatic function panel, complete blood count, international normalized ratio, prothrombin time, albumin)

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

n/a

Statistical Analysis Plan:

The baseline comparisons will be nave methods (mean, median), parametric methods (multivariate logistic regression covariate adjustments), non-parametric methods (nearest neighbor matching methods using L2 distance between covariates), and random forest (mapping from covariates to outcomes).
For the primary outcomes that are continuous values, average prediction error will be calculated across all patients in the test set and then compared to the above baseline methods.
For the secondary outcomes that are binary values (e.g. adverse events), the area under the receiver operating characteristic curve will be calculated for our method and compared to other baseline methods using the DeLong non-parametric test.
Robustness tests will be performed with the following strategy: we will begin with 50% training, with five-fold cross-validation across the training set and 50% test set. In order to evaluate if a smaller proportion of the full dataset could be sufficient for quantifying treatment effect and adverse event risk, we will trial 25% training with five-fold cross-validation across the training set and 75% test set and 10% training with five-fold cross-validation across the training set and 90% test set. Our evaluation metrics will be the average prediction error for primary outcomes and AUROC for secondary outcomes for our method and comparison methods.

Project Timeline:

Perform analysis on Phase III and IV Trials 2022-2023
Prepare manuscript by July 2023
Publication by January 2024

Dissemination Plan:

Publication in Nature Medicine, Nature Digital Medicine, Lancet Digital Health

Bibliography:

1. Singh S, Murad MH, Fumery M, et al. First- and Second-Line Pharmacotherapies for Patients With Moderate to Severely Active Ulcerative Colitis: An Updated Network Meta-Analysis. Clin Gastroenterol Hepatol 2020;18:2179-2191.e6.
2. Lasa JS, Olivera PA, Danese S, et al. Efficacy and safety of biologics and small molecule drugs for patients with moderate-to-severe ulcerative colitis: a systematic review and network meta-analysis. The Lancet Gastroenterology & Hepatology 2022;7:161-170.
3. Burr NE, Gracie DJ, Black CJ, et al. Efficacy of biological therapies and small molecules in moderate to severe ulcerative colitis: systematic review and network meta-analysis. Gut 2021:gutjnl-2021-326390.
4. Singh S, Fumery M, Sandborn WJ, et al. Systematic review with network meta-analysis: first- and second-line pharmacotherapy for moderate-severe ulcerative colitis. Aliment Pharmacol Ther 2018;47:162-175.
5. Ko Y-A, Chen Z, Liu C, et al. Developing a synthetic control group using electronic health records: Application to a single-arm lifestyle intervention study. Preventive medicine reports 2021;24:101572-101572.
6. Chen Z, Zhang H, Guo Y, et al. Exploring the feasibility of using real-world data from a large clinical data research network to simulate clinical trials of Alzheimer’s disease. NPJ Digit Med 2021;4:84.
7. Sagami S, Nishikawa K, Yamada F, et al. Post-marketing analysis for biosimilar CT-P13 in inflammatory bowel disease compared with external data of originator infliximab in Japan. J Gastroenterol Hepatol 2021;36:2091-2100.