The YODA Project
Research Proposal Review

The following page contains the final YODA Project review approving this proposal.

The YODA Project
Research Proposal Review - Final
(Protocol #: 2023-5251)

Reviewers:

- ☐ Nihar Desai
- ☒ Cary Gross
- ☐ Harlan Krumholz
- ☒ Richard Lehman
- ☒ Joseph Ross
- ☐ Joshua Wallach

Review Questions:	**Decision:**
1. Is the scientific purpose of the research proposal clearly described?	Yes
2. Will request create or materially enhance generalizable scientific and/or medical knowledge to inform science and public health?	Yes
3. Can the proposed research be reasonably addressed using the requested data?	Yes, or it's highly likely
4. Recommendation for this data request:	Approve

Comments:

No additional comments

The YODA Project
Research Proposal Review

Revisions were requested during review of this proposal.

The following pages contain the original YODA Project review and

the original submitted proposal.

The YODA Project
Research Proposal Review - Revisions Requested
(Protocol #: 2023-5251)

Reviewers:

- ☐ Nihar Desai
- ☒ Cary Gross
- ☐ Harlan Krumholz
- ☒ Richard Lehman
- ☒ Joseph Ross
- ☐ Joshua Wallach

Review Questions: **Decision:**

1. Is the scientific purpose of the research   No
   proposal clearly described?

2. Will request create or materially enhance   Yes
   generalizable scientific and/or medical
   knowledge to inform science and public
   health?

3. Can the proposed research be reasonably   Yes, or it's highly likely
   addressed using the requested data?

4. Recommendation for this data request: Not Approve

Comments:

It is still unclear how disease remission is going to be defined (different measures were used in the trial)? Will remission be defined as a score >170? >180? >190?

The YODA Project
Research Proposal Review - Revisions Requested
(Protocol #: 2023-5251   )

Reviewers:

☐ Nihar Desai

☒ Cary Gross

☐ Harlan Krumholz

☒ Richard Lehman

☒ Joseph Ross

☐ Joshua Wallach

Review Questions:	**Decision:**
1. Is the scientific purpose of the research proposal clearly described?	No
2. Will request create or materially enhance generalizable scientific and/or medical knowledge to inform science and public health?	Yes
3. Can the proposed research be reasonably addressed using the requested data?	Yes, or it's highly likely
4. Recommendation for this data request:	Not Approve

Comments:

It is still unclear how disease remission is going to be defined (different measures were used in the trial)? Will remission be defined as a score >170? >180? >190?

Principal Investigator

First Name: Saeed
Last Name: Ghodsi
Degree: PhD
Primary Affiliation: University of California, Los Angeles
E-mail: saeedghodsi93@gmail.com
Phone number: 3232758358
Address:

City: Los Angeles
State or Province: CA
Zip or Postal Code: 90064
Country: United States

General Information

Key Personnel (in addition to PI):
  First Name: Reza
Last name: Ahmadi
Degree: PhD
Primary Affiliation: University of California, Los Angeles
SCOPUS ID: 54379482100

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?: Scientific Publication

Conflict of Interest

https://yoda.yale.edu/system/files/saeed_ghodsi_-_coi.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. 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

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

Research Proposal

Project Title

Efficient Learning of Continuous-Time Hidden Markov Models with Discrete-Time Irregular Observations for Healthcare Intervention Planning

Narrative Summary:

Disease progression models provide a mechanism for understanding and predicting the impact of interventions on the health state of patients. CT-HMMs have recently attracted attention, as they are able to handle the complexities of real-world data. The main contribution of this research project is to propose a CT-HMM disease progression model, which incorporates the effect of interventions, and to present an efficient approach for learning the parameters of this model based on the EM algorithm (https://escholarship.org/content/qt3gz0c7qx/qt3gz0c7q). We believe Ulcerative Colitis is an appropriate disease for demonstrating the advantages of our model.

Scientific Abstract:

Background: Disease progression models provide a mechanism for understanding and predicting the impact of interventions on the health state of patients. In this study, we develop a statistical disease progression model that is capable of handling the shortcomings of the existing methods.
Objective: Provide a statistical framework for tracking disease progression as a function of treatment.
Primary and Secondary Outcome Measure(s): Statistical error of the model in predicting ulcerative colitis remission.
Statistical Analysis: We present a Continuous-Time Hidden Markov Model for tracking the course of the disease (https://escholarship.org/content/qt3gz0c7qx/qt3gz0c7qx.pdf).

Brief Project Background and Statement of Project Significance:

The availability of vast amounts of healthcare data has inspired an increasing interest in data-driven healthcare intervention planning methods. Disease progression models provide a mechanism for understanding and predicting the impact of interventions on the health state of patients. Most traditional Markovian state-transition models perform poorly on real-world data since they are incapable of capturing complexities such as unobservability of the underlying health state and irregularity of observation times. Moreover, most of the existing frameworks are unable to explicitly model the effect of interventions on disease progression. CT-HMMs have recently attracted attention, as they are able to handle these complexities. The main contribution of this research is to propose a CT-HMM disease progression model, which incorporates the effect of interventions, and to present an efficient approach for learning the parameters of this model based on the EM algorithm.

Specific Aims of the Project:

We'll track the progression of the disease in response to the treatment over time. The model is supposed to provide guidelines for predicting disease remission.

What is your Study Design?:

Methodological research

What is the purpose of the analysis being proposed? Please select all that apply.

Develop or refine statistical methods

Research Methods

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

We include patients who have been diagnosed with Ulcerative Colitis. The statistical model is capable of incorporating covariates (e.g. age, gender, etc.) naturally.

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

We will measure the statistical error of the model in predicting disease remission.

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

We're using a Markvian model, which tracks the disease state over time. Therefore, we do not have a classical regression analysis with independent variables.

Statistical Analysis Plan:

We have already developed the statistical model as well as a statistical inference algorithm for fitting data into the model. The mathematical details are available online at
https://escholarship.org/content/qt3gz0c7qx/qt3gz0c7qx.pdf
Software Used:
R
Project Timeline:

We expect the analysis to take 6-12 months depending on the possible complexities that arise while working with real-world data. The mathematical results are ready as mentioned earlier.

Dissemination Plan:

We plan to publish the work in healthcare - machine learning or healthcare - AI journals.

Bibliography:

https://escholarship.org/content/qt3gz0c7qx/qt3gz0c7qx.pdf