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

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. 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

2. 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

Methods for Staging disease activity In ulcerative colitis; A comparison of endoscopy and Histology (MESIAH)

Narrative Summary:

It is essential tests accurately gauge inflammation for ulcerative colitis. Endoscopy is the best test, but calprotectin (stool test) is an alternative. This study will use data from clinical trials to identify:
- calprotectin thresholds for endoscopic inflammation
- if calprotectin can monitor response to medication
- factors influencing calprotectin results
- Combine symptoms, blood and calprotectin to create a model to predict endoscopic inflammation. Calprotectin and the proposed model are inexpensive, less invasive and more accessible than endoscopy; potentially improving patient outcomes. Findings will inform design and interpretation of therapeutic drug trials.

Scientific Abstract:

Background: Endoscopic and histological remission for ulcerative colitis (UC) is associated with successful steroid tapering; reduced hospitalisation, surgery and cancer. This highlights the need for robust markers to stage inflammation in response to treatment. Endoscopy is the gold standard, but is invasive, expensive and can delay treatment decisions due to availability. Objective: To identify the utility of non-invasive markers of disease activity of UC. Specifically: thresholds of calprotectin for endoscopic activity; serial calprotectin results to determine treatment response; and ultimately, create a model that combines markers (symptom, serum markers, calprotectin) to predict endoscopic activity. Study Design: This study will use prospectively collected data from therapeutic clinical trials available on the YODA and Vivli platforms to address the objectives and validate findings. Participants: patients enrolled in therapeutic clinical trials with established UC. Main Outcome: endoscopic remission (mayo endoscopic subscore = 0/1) and histological remission (absence of neutrophils). Statistical analysis: agreement statistics (Fleiss’ and Cohen’s Kappa), statistics for diagnostic accuracy (sensitivity, specificity, positive predictive value, negative predictive value, receiver operator curve analysis), regression analysis, decision curve analysis and predictive modelling experiments.

Brief Project Background and Statement of Project Significance:

Ulcerative colitis (UC) is a chronic inflammatory condition predominantly affecting young patients (Ordás et al., 2012). The prevalence is rising, especially in newly industrialised nations (GBD, 2020). This will directly impact delivery of care across healthcare systems including resource poor settings.

Robust methods of monitoring disease activity aid physicians when planning treatment escalation. Patients achieving mucosal healing have reduced rates of colectomy, colorectal cancer (Flores et al., 2017); and improved quality of life (Knowles et al., 2018). A ‘treat to target’ (T2T) approach has been proposed with focus on achieving mucosal healing (mayo endoscopic score = 0/1) (Peyrin-Biroulet et al., 2015). There are challenges delivering this strategy using endoscopy. Highlighting the need for non-invasive markers of activity; existing methods include symptom indexes and faecal calprotectin (FC).
At best symptoms indexes differentiate patients in remission (score of 0 for stool frequency and rectal bleeding on PRO2, sensitivity = 36%, specificity = 96% for remission); and are weakly correlated with endoscopic activity ($r = 0.44, p<0.001$). Patients can have asymptomatic endoscopic disease activity; or elevated symptom indexes (increased stool frequency) with quiescent endoscopic disease (Dulai et al., 2020; Falvey et al., 2015; Narula et al., 2019).

A cut-off of 50 for FC has been proposed to identify endoscopic remission (M. H. Mosli et al., 2015; Rokkas et al., 2018). Studies have demonstrated that calprotectin is directly correlated with endoscopic activity (Baron score), this has not been replicated consistently (Schoepfer et al., 2013). Methodological differences in study design provide uncertainty in the use of calprotectin to replace endoscopy. Despite establishing the optimum cut-off of 50, studies included in meta-analyses identified thresholds for endoscopic remission ranging from 6 to 400 (M. H. Mosli et al., 2015; Rokkas et al., 2018). Furthermore, meta-analyses have demonstrated moderate/high levels of heterogeneity (I² greater than 70%). Likely attributable to use of different assays, timing of stool collection, differing endoscopic scoring systems and inclusion of paediatric and adult patients.

In summary, endoscopy is the gold standard to stage UC activity, (Lamb et al., 2019; Maaser et al., 2019; Takeuchi et al., 2015); disadvantages include: cost; inconvenience(bowel preparation, time off work); complications (perforation, bleeding) and delay in appointments with subsequent delays in treatment. Studies of alternative markers (e.g. FC) have been limited by their methodology. We propose data collected through therapeutic trials will allow us to accurately determine the utility of non-invasive markers, through larger sample size and greater methodological consistencies.

Specific Aims of the Project:

1. Determining agreement between endoscopic and a novel histological definition of remission (Personal Communication; B. Hayee, 2020 - forthcoming peer reviewed consensus statement)
2. Identification of a endoscopically diagnosed (upper) threshold for relapse/flare based on faecal calprotectin results
3. Evaluating the increase in diagnostic accuracy from knowledge of an index calprotectin measurement (i.e. the role of longitudinal measurements in diagnosing response to treatment)
4. Identification of covariates that are associated with variation in the faecal calprotectin results for individuals with endoscopically confirmed remission.
5. Developing a prediction model for endoscopic remission based on features collected routinely in clinical practice
6. Rasch analysis [psychometric validation] of the Mayo Tool

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

- New research question toe examine accuracy of diagnostic methods based on secondary analysis of clinical trial data

Research Methods

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

Analyses will be performed using data from the therapeutic trials for vedolizumab (via the Vivli platform: NCT02497469 and NCT00783718) and golimumab (via the YODA platform: NCT00488631 and NCT00487539). All patients with confirmed ulcerative colitis enrolled in the mentioned therapeutic trials will be included in analyses for our proposed study.

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

Endoscopic Remission: Defined based on the mayo endoscopic score (MES). Two thresholds will be used: 0, and less than or equal to 1.

The MES is categorised as:
- Mayo 0 - Normal mucosa / inactive disease
- Mayo 1 - Mild erythema, decreased vascular pattern and friability
- Mayo 2 - Marked erythema, lack of vascular pattern, friability and erosions
- Mayo 3 - Spontaneous bleeding and ulceration
Histological remission: absence of neutrophils

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

Faecal calprotectin - continuous outcome from assay (week 0, post-induction review alongside endoscopy, and final follow-up)

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

Gender - Binary category (Male/Female)

Age - continuous (years)

Disease History - Location; Montreal/Paris classification (Ordinal); duration (years/months)

Race - Categories defined in the original studies

Smoking status - Binary (yes/no)

Concurrent medical conditions - categorised by the presence or absence of the condition for each individual.

Medication history - presence or absence of a prescription (30 days prior to consent), or current prescription for each individual.

height - Continuous in cm

weight - Continuous in kg

Extraintestinal manifestations of disease – presence/absence of ‘accepted’ extraintestinal features of IBD on physical examination at screening

Laboratory results - all as continuous variables

**Statistical Analysis Plan:**

A more extensive description of the statistical methods that we will employ have been detailed in the attached statistical analysis plan, which provides links to specific references which demonstrate that the applicants have previously applied these methods in other settings, and are both aware of and involved in the development of best practice guidelines for carrying out the described analyses.

This proposal is based on using three core sets of methods:

1) Agreement statistics (incl. Fleiss’ and Cohen’s Kappa) to explore the concordance between endoscopic and histological definitions of remission. This will be based on the 94 individuals with paired histological sampling form the PURSUIT trial. Agreement will be tested at the 0-week time point, and the re-randomisation (post-induction) time-point. Given the sample size this will be explicitly framed as a pilot study to inform future prospective applications to generate a larger databank.

2) Prediction modelling – we intend to explore both regression (i.e. prediction of the faecal calprotectin), and dichotomous classification tasks (i.e. prediction of remission, as defined by a mayo score cut-off). We intend to benchmark a series of classical statistical models and machine learning methods, using k-fold cross validation, to determine the optimal approach based on out-of-sample errors. Multiple imputation by chained equations will be used for missing data handling. Fractional polynomials will be sued to transform potentially useful non-linear relationships identified from a random 10% hold-out sample. We believe that the sample sizes of the studies in question satisfy current thinking around minimum requirements (see Riley et al., 2020, BMJ). A full description of the proposed computational experiments is outline the attached statistical analysis plan document, including how we intend to handle the issue of variable selection.
3) Rasch Analysis – we will use pre-existing packages in R (e.g. eRM) to fit a Rasch model to the Mayo scale data. The purpose will be to determine whether the ordinal presentation of the Mayo has a corresponding interval scale based on the hypothesis that the score relates to a latent trait of ‘remission/disease activity’. Confirmation to the assumptions of the Rasch model will be checked using a chi-squared test, and post-hoc changes to the Rasch-andrich thresholds will be justified based on domain expertise and carried out in sequence to determine their impact on the fit statistics. Alongside these novel Rasch-specific measures, we will also report the cronbach’s alpha as a measure on reliability in both the original format, and following post-hoc correction of misfit.

Software Used:

R

Project Timeline:

Data Access = January - March 2021
Analysis completion date = September 2021
Dissemination (manuscript preparation and submission of draft for publication = October-November 2021
Results Reported back to YODA project = December 2021

Dissemination Plan:

Each of the above objectives represents a unique manuscript that we intend to publish in a gastroenterology or general medical journal as appropriate. Whilst working towards publication in a peer-reviewed journal, we may submit abridged versions of the results to relevant conferences. Moreover, where appropriate our results will be submitted to members of the British Society for Gastroenterology IBD sub-committee for consideration in circumstances where we feel they might consider an interim update to their recently published IBD clinical guidelines.

Bibliography:

Crohn’s and Colitis, 13(2), 144–164. https://doi.org/10.1093/ecco-jcc/jiy113
ulcerative colitis. Cochrane Database of Systematic Reviews, 5.
https://doi.org/10.1002/14651858.CD011256.pub2
(2015). C-reactive protein, fecal calprotectin, and stool lactoferrin for detection of endoscopic activity in
symptomatic inflammatory bowel disease patients: A systematic review and meta-analysis. American Journal
of Gastroenterology, 110(6), 802–819. https://doi.org/10.1038/ajg.2015.120
Endoscopic Appearance of Ulcerative Colitis: A Systematic Review and Meta-analysis. Clinical Gastroenterology
and Hepatology, 17(3), 411-418.e3. https://doi.org/10.1016/j.cgh.2018.06.015
380, 1606–1609. https://doi.org/http://dx.doi.org/10.1016/S0140-6736(12)60150-0
Peyrin-Biroulet, L., Sandborn, W., Sands, B. E., Reinisch, W., Bemelman, W., Bryant, R. V., D’Haens, G., Dotan,
Bowel Disease (STRIDE): Determining Therapeutic Goals for Treat-to-Target. American Journal of
Gastroenterology, 110(9), 1324–1338. https://doi.org/10.1038/ajg.2015.233
disease endoscopic activity: A diagnostic accuracy meta-analysis. Journal of Gastrointestinal and Liver Diseases,
Schoepfer, A. M., Beglinger, C., Straumann, A., Safroneeva, E., Romero, Y., Armstrong, D., Schmidt, C.,
Trummeler, M., Pittet, V., & Vavricka, S. R. (2013). Fecal calprotectin more accurately reflects endoscopic activity of
ulcerative colitis than the lichtiger index, C-reactive protein, platelets, hemoglobin, and blood leukocytes.
Inflammatory Bowel Diseases, 19(2), 332–341. https://doi.org/10.1097/MIB.0b013e3182810066
Takeuchi, Y., Hanafusa, M., Kanzaki, H., Ohta, T., Hanaoka, N., Yamamoto, S., Higashino, K., Tomita, Y., Uedo,
N., Ishihara, R., & Iishi, H. (2015). An alternative option for “resect and discard” strategy, using magnifying narrow-

Supplementary Material:

https://yoda.yale.edu/sites/default/files/statistical_analysis_plan.pdf