Miscellaneous fields clinical data request

# Scientific Abstract \* Provide a structured abstract for the proposed study, using the following sub-headings: Background; Objective; Study Design; Participants; Main Outcome Measure(s); and Statistical Analysis. 1600 characters

Background

Using evidence synthesis, estimates of efficacy from clinical trials can be applied to standard treatment comparator (natural history) event rates from observational data to estimate effectiveness in target populations.

In order to extend this approach to estimate treatment effectiveness in people with additional secondary diseases (comorbidity), we need to determine whether treatment efficacy is similar in people with and without comorbidity.

Objective

To estimate the variation in efficacy by comorbidity within clinical trials, and summarise this for different drug-classes and wider groupings of related drug classes.

Study design

Meta-analysis

Participants

Trials of drugs used to prevent and treat long-term medical conditions (38 trials from the YODA repository, ~200 trials from other repositories)

Main outcome measures

Outcomes common to trials of specific drug-classes, eg HbA1c in diabetes trials.

Statistical analysis

For each outcome, in each trial, we will model main effects and interactions with treatment allocation for age, sex and the 6 commonest comorbidities. We will use generalized linear models, or Cox regression as appropriate to the outcome.

In subsequent analyses, using the coefficients and variance-covariance matrix from these models as the dependent variables, we will estimate the mean (and variance) of each comorbidity-treatment interaction for specific drug-classes, and across drug-classes for wider related drug-groupings.

We will explore the effect of different scales and transformations on the extent of heterogeneity.

# Narrative summary. Provide a plain English summary of the proposed study that is suitable for a general or lay audience, clarifies the design, and explains the relevance of the research project to science and public health. 700 characters

Randomised controlled trials provide the best evidence concerning treatment effectiveness. Trial results can be applied to “target” populations, for example; if a diabetes trial shows that a drug reduces blindness by 20%, and other data shows that 8 out of 100 people with diabetes develop blindness, then 2 fewer people are expected to develop blindness per 100 treated.

We would like to use a similar approach for a new target, people with additional (co-morbid) diseases. To make sure this is valid, we will examine clinical trial data from YODA, and other clinical trial repositories to see if trial participants with and without different co-morbid diseases experience similar treatment benefits.

# Scientific Abstract \*

Provide a structured abstract for the proposed study, using the following sub-headings: Background; Objective; Study Design; Participants; Main Outcome Measure(s); and Statistical Analysis. 1600 characters

# Brief Project Background and Statement of Project Significance \*

Provide a brief summary of the research project’s background, including a clear description of the project’s **significance** and how information gained from this work will be used to **create or materially enhance generalizable scientific and/or medical knowledge to inform science and public health**. Provide references to prior work on the topic if applicable at the end of the proposal.

**3200 characters**

## Background

Multimorbid patients are less likely than those with single diseases to receive recommended drugs, even where there are no specific contraindications.[6–11] This difference may represent under treatment or may have arisen from uncertainty among clinicians as to the applicability of clinical guidelines, which rarely provide specific advice on managing multimorbidity.[3,12,13]

Where multimorbidity alters the (baseline) natural history of diseases, the effects of treatment are likely to differ. For example, while there is strong evidence that the benefits of dual antiplatelet therapy (DAPT) following myocardial infarction (versus a single antiplatelet) outweigh the risks,[14,15] this may not be true for patients with comorbid chronic obstructive pulmonary disease (COPD). Cardiovascular mortality is commoner in COPD than the general population, favouring DAPT.[16] However, non-cardiovascular mortality is also higher,[17] favouring single-antiplatelet therapy because of competing risks. Glucose and intensive risk factor control in diabetes[18,19] and anticoagulant use in atrial fibrillation[20] provide similar examples, where the net overall treatment benefits are uncertain for multimorbid patients.

Neither clinical trials nor observational studies fully address this problem.[21] Trialists do not report results stratified by combinations of important comorbidities.[12] Nor is it feasible to conduct trials sufficiently large to allow such comparisons for all indications.[5,21] Administrative healthcare databases are sufficiently large. However, while estimates of treatment effectiveness obtained from observational studies are generally comparable to those from clinical trials,[22] they have in several cases led to inaccurate conclusions.[23–25]

There is therefore a need to find alternative approaches to produce estimates of net overall treatment benefits for patients with multimorbidity.

In a Wellcome Trust funded project, we intend to develop and validate an approach using Bayesian Evidence Synthesis to estimate treatment effectiveness in multimorbidity. As with standard evidence synthesis, we plan to apply estimates of treatment efficacy from clinical trials to observational data representative of the target population. However, in most evidence synthesis it is assumed that estimates of relative treatment efficacy (RTE) from clinical trials can be applied, unmodified, to the target population. It is not known whether this assumption is valid in applying efficacy estimates to patients with multimorbidity.

In a limited number of large individual-patient meta-analyses (such as of aspirin for primary and secondary prevention) RTE was similar regardless of cardiovascular risk factors, age and sex.[27,28] However, for most drugs and conditions empirical evidence is lacking as to whether and how RTE differs between patients with and without comorbidity, especially for comorbidities which are not established causes or complications of the target disease.

Therefore, there is a specific need to determine how commonly RTE differs by age, sex and comorbid diseases for different drug-classes.

# Specific Aims of the Project \*

Provide a concise description of the specific aims of the project, including the study objectives and the specific hypotheses to be evaluated.

1300 characters

## Aim

Therefore, in this proposal we aim to examine and quantify the variation in relative treatment efficacy by comorbidity. We intend to do so by examining such variation within clinical trials, for a range of drug-classes and indications.

## Objectives

### Stage One – Produce trial-level summary estimates

1. Assign drugs to broad drug-classes based on 5-character ATC codes
2. Assign related drug-classes to wider drug-groupings
3. Identify comorbidities within individual clinical trials according to the Clinical Classification Software scheme[29]
4. For each trial estimate interactions between treatment allocation and age, sex and selected comorbidities

### Stage Two – Analyse trial-level aggregated estimates

1. In Bayesian hierarchical generalized linear models, use the estimates from 4 **alongside estimates from clinical trials obtained from other repositories** to estimate average drug-class-level and drug-grouping level comorbidity-treatment interaction
2. Summarise 5 as a probability distribution for use in subsequent evidence synthesis; either as off-the-shelf informative priors for meta-analyses, or as inputs to probabilistic sensitivity analyses in which treatment effects are modelled

# Data Source and Inclusion/Exclusion Criteria to be used to define the patient sample for your study \* 1200 characters

The purpose of this project is to produce summaries, for and across drug-classes, of the covariate-treatment interactions for different outcomes within drug-groupings formed of clinically-related drug-classes, for subsequent use by ourselves and others either as informative priors, or in probabilistic sensitivity analyses. The covariates of interest being comorbidities, age, sex and socio-economic status.

Consequently, trials were chosen to include drugs used to treat or prevent long term medical (and, as these are a common issue in the elderly, urological) conditions. Both short and long term indications were included (eg short term post MI therapy). We have excluded trials where the indication was for neoplastic, infectious, affective, psychotic or developmental disorders. Topical eye and primary prevention trials in the general adult population were also excluded. Small trials, and those with highly restrictive inclusion criteria were also excluded.

This led to us identifying 38 relevant trials in the YODA repository. Covariate-treatment interaction estimates from which will be analysed alongside estimates from similar trials obtained from other repositories.

# Main Outcome Measure and how it will be categorized/defined for your study \*1200 characters

As our aim, is to ‘survey’ many trials to determine how much variation in treatment effect by comorbidity is plausible within different clinically-meaningful groupings, we intend to obtain measures which are similar across multiple trials. Therefore, within each indication we will select endpoints common to trials for that indication. For example, the majority of trials for rheumatoid arthritis include the American College of Rheumatology Criteria and Disease Activity Score endpoints.

In order to simplify comparisons, we will scale continuous outcomes and will also use log-transformation in order to examine heterogeneity by comorbidity on the relative scale. After examining trial documentation and datasets, but prior to performing any comparisons by treatment allocation, we will with our steering group agree on the final outcomes, transformations etc.

While we believe that this approach is suitable for our analysis, we recognise that it is not likely to be optimal for estimating the main effect for each trial. For that reason, we will not report the main effects (ie marginal across sub-groups) from individual trials.

# Main Predictor/Independent Variable and how it will be categorized/defined for your study \* 1200 characters

We will examine covariate-treatment interactions for each trial. The covariates included will be age, sex, socio-economic status and comorbid disease.

Age, sex and socio-economic status

For each drug grouping we will also categorise age into equal width bands. Sex will be defined as per the study documentation. We will define socio-economic status separately using educational attainment and income as proxies, collapsing the reported categories into a three-level variable.

Comorbidities

Comorbidities will be categorised using the Agency for Healthcare Research and Quality clinical classifications system (CCS). [30] We will attempt to define in the trial datasets comorbid conditions from 94 of the CCS groups.

Trial-specific operational definitions of each comorbidity will be defined using a combination of demographic, past medical history, lifestyle (eg smoking) and drug variables as well as information from trial protocols (ie inclusion and exclusion criteria). We anticipate that it will frequently not be possible to arrive at an operational definition for many of the comorbidities.

Each operational trial definition will be finalised prior to analysing outcome data.

# Other Variables of Interest that will be used in your analysis and how they will be categorized/defined for your study 1200 characters

# Statistical Analysis Plan \*Describe how you will analyze the requested clinical trial data, including descriptive, bivariate and multivariable analyses, and any other planned advanced analyses (such as propensity score methods, Kaplan-Meier or Cox modeling approaches, non-parametric testing). 4000 characters

Stage one

For each outcome, we will model main effects and interactions with treatment allocation for age, sex, socio-economic status (in 3-levels) and the 6 commonest comorbidities for that drug-grouping. Where a particular trial does not have a comorbidity variable defined we will record that variable as missing.

We are not aware of any study which has examined the effect of different scales on the extent of heterogeneity in treatment-comorbidity interactions between patients. Therefore, for continuous measures we will model outcomes on both the absolute and relative (log-transformed) scale, and for event outcome data we will model outcomes on the log-odds ratio log-rate ratio scales and log-risk ratio as well as log-hazard ratios scales. We will estimate hazard ratios using Cox regression, and for all other models will use Generalized Linear Models with appropriate link functions and error distributions.

Analyses will be conducted in R. We will obtain the coefficients and variance covariance matrices from each of these models. For simplicity, missingness will be treated using complete case analysis. We will however record and report, for each operational comorbidity definition, the proportion of participants for whom sufficient data was missing to prevent a participant being defined as having a diagnosis.

Summary of outputs from stage one

In summary, therefore, the following results, all of which are trial-level aggregates/summaries, will be obtained from the above analyses:-

1. The proportion of participants with each combination of comorbidities

2. The proportion of participants with missing data included in the definition of each comorbidity

3. The coefficients and variance-covariance matrix of the regression models described above.

Stage Two

All subsequent modelling will be conducted on aggregated summary-level data.

In the second part of a two-step approach we will model the comorbidity-treatment interactions across multiple trials using the parameters described above as the dependent variables.

This will be done using Bayesian hierarchical generalized linear models. The main effects for treatments and covariates will have independent priors for each trial, while the priors for the covariate-treatment interactions will have hyper-priors, to allow sharing of information. We have successfully fit such models on simulated data.

For each comorbidity-treatment interaction we will report the mean effect for each drug-class and drug grouping, along with the between drug-class variance. We will first model each comorbidity in turn, after which we will simultaneously model multiple comorbidities. We will separately examine associations for datasets comprising trials where a trial drug is compared with placebo, an agent from a different drug-class, and an agent from the same drug-class.

An assumption of this modelling approach is that comorbidity-treatment interactions are exchangeable for trials within drug-classes, and for drug-classes within wider drug groupings. This assumption makes use of the structure inherent in the anatomic therapeutic classification system; thereby allowing borrowing of information across drug-classes (eg antithrombotic agents).

If there is no similarity in treatment-comorbidity interactions within our drug-groupings, the predicted comorbidity-treatment interaction for an unknown drug-class will have a wide uncertainty interval, indicating that little can be learned about one drug class comorbidity-treatment interaction based on knowledge about such an interaction for another drug-class within that grouping.

These models will be fit in the JAGS package. Alternative model specifications will be explored and compared using the deviance information criterion (DIC) and the Bayesian information criterion (BIC). Model convergence and autocorrelation will be assessed using diagnostic plots as well as summaries such as the Gelman-Rubin statistic.

# Project Timeline \*

# Provide an estimation of key milestone dates for the proposed study, including anticipated project start date, analysis completion date, date manuscript drafted and first submitted for publication, and date results reported back to the YODA Project. Please note: if your data request is approved, the Data Use Agreement allows for access for a 12 month period, with the possibility of an extension.1300 characters

# Dissemination Plan \*

# Provide a description of anticipated products and target audience(s), including expectation for study manuscript(s) and potentially suitable journals for submission of the completed research project. 1300 characters

The purpose of this project is to produce summaries, for and across drug-classes, of the covariate-treatment interactions for different outcomes within drug-groupings formed of clinically-related drug-classes, for subsequent use by ourselves and others either as informative priors, or in probabilistic sensitivity analyses. The covariates of interest being comorbidities, age, sex and socio-economic status.

We reviewed trials listed under suitable drug-classes (Cholinesterase Inhibitors, Anticonvulsants, Antipsoriatics, Antirheumatic agents, Colony-stimulating factors, and Diabetes Related) or diseases (Arthritis, Psoriatic, Arthritis, Rheumatoid, Atrial Fibrillation, Colitis, Ulcerative, Crohn's Disease, Dementia, Diabetes Mellitus, Type 2, Epilepsy, Migraine Disorders, Partial Seizure Disorder, Psoriasis, Sacroiliac Joint Dysfunction, Seizures and Spondylitis, Ankylosing).

Using metadata from clincialtrials.gov for the 90 trials identified, we excluded 32 smaller trials (< 300 participants) and two phase 2 trials. Of the remaining 56 trials, 13 were excluded (eg because these were cancer trials or trials of psychiatric drugs (other than anti-dementia drugs).

None of the remaining 43 trials had restrictive age criteria (maximum ages were 65 or older). 5 were excluded because of restrictive exclusion criteria such that patients with comorbidities would be very unlikely to be recruited. An additional 3 trials had moderately restrictive criteria, and will be excluded in a sensitivity analysis.

In order to obtain trials likely to include at least some participants with one or more comorbid diseases, we selected phase 3 and phase 4 trials for which people aged over 60 were not explicitly excluded (as recorded on clinicaltrials.gov). We selected trials with at least 300 participants to avoid problems with very small numbers, and parallel group designs to avoid additional complexity.

We will use the intention to treat rather than the per protocol population in order to maintain the instrument. Where a trial has more than 2 arms we will select the most extreme comparison – eg a placebo arm versus the highest dose arm.