**EFFICACY AND SAFETY OF THERAPIES FOR INDUCTION AND MAINTENANCE OF REMISSION IN CROHN’S DISEASE: PROTOCOL FOR A NETWORK META-ANALYSIS**

**Corresponding author**

1. Morris Gordon

**Authors:**

1. Vasiliki Sinopoulou
2. Gordon W. Moran
3. Aderonke Ajiboye
4. Sudheer Vuyyuru
5. Shellie J. Radford

Contributorship Statement:

An original draft for this protocol manuscript was made by MG with VS and then finalized by the wider team.

Co-Authors agreed on the review questions, approved the protocol and contributed to the final version of this manuscript.

**Introduction**

Crohn’s disease (CD) is a complex immune mediated inflammatory disease of the intestine, affecting approximately 1 in 300 patients in countries with high prevalence such as the United States of USA and the United Kingdom and is expected to increase in future. (1)

Conventional therapies such as corticosteroids, immunomodulators and surgery played a central role in the management of IBD before the development of advanced targeted therapies. The management principles since then have been transforming, since approval of the first biologic therapy targeting tumour necrosis factor (TNF), infliximab, in early 2000s. Since then various classes of biologics namely anti-integrins, anti-IL-12/23p49, anti-IL23p19 and oral small molecules (JAKi) have been developed in the last two decades and many of them have been approved. With the availability of multiple treatment options, the choice of therapy can be challenging in clinical practice. While head-to-head blinded controlled trials are the ideal method for comparing the efficacy of different drugs, practical limitations make it challenging to conduct multiple trials encompassing all available treatment options.

A network meta-analysis (NMA) is where multiple treatments are compared using both direct comparisons of interventions within randomized controlled trials (RCTs) and indirect comparisons across trials based on a common comparator (i.e., placebo). In other words, if compound A is compared with compound B in one trial, and the same compound B is compared with compound C in another trial, indirect information can be obtained from compound A versus compound C under the assumption of transitivity.

NMA is often understood to allow ranking of therapies, but there are significant limitations in this approach and goal. The main opportunities for NMA analysis in this context are much broader than simply ranking a top therapy and include:

* To allow interventions studied in multiple standalone comparisons to be combined in a single node and therefore enhance the precision of estimates for such interventions.
* To allow borderline therapies to be considered with greater precision.
* To rank therapies with consideration of certainty, ensuring the interpretation of both elements of the data together
* To produce subgroup or sensitivity networks to clarify the effect of clinical and method factors on findings.

There is a need for NMAs that examines and compares the efficacy and safety of advanced therapies for CD and immunomodulators.

This protocol describes the steps that will be followed to produce an NMA for induction of remission and another for maintenance of remission.

**Methods**

**Evidence selection**

Types of studies: Randomized controlled trials (RCTs) for induction or maintenance of remission in CD.

Types of participants: Trials enrolling adults (>18 y.o. or above) with diagnosed CD will be considered for inclusion.

Types of interventions: Advanced and immunomodulator therapies including, TNF-α inhibitors, anti-integrins, IL23/IL12 antagonists, JAK inhibitors, and biosimilars. All administration routes and dosages will be considered. Studies on advanced therapies where participants received more than 50% immunomodulators will be considered combination interventions (advanced therapy + immunomodulator). Corticosteroids treatments will not be considered.

**Types of outcome measures**

Dichotomous outcomes will be considered for inclusion.

Induction of remission:

* Clinical remission, as defined by the studies
* Clinical response, as defined by the studies
* Endoscopic remission, as defined by the studies
* Withdrawals due to adverse events (WAEs)
* Serious adverse events (SAEs)
* Total adverse events (TAEs)

Maintenance of remission:

* Clinical relapse, as defined by the studies.
* Loss of clinical response, as defined by the studies.
* Endoscopic relapse, as defined by the studies.
* Withdrawals due to adverse events (WAEs)
* Serious adverse events (SAEs)
* Total adverse events (TAEs)

**Thresholds for outcomes**

For each of the included outcomes, the threshold will be pre-defined as per the BSG GDG guidelines’ predetermined thresholds. These will ensure interpretation is against this a priori defined framework.

* The minimum threshold for a small difference to be defined (lower than this would be ‘trivial’)
* The minimum threshold for a moderate difference to be defined (lower than this would be ‘small’)
* The minimum threshold for a large difference to be defined (lower than this would be ‘moderate’ and all above this would be ‘large’)

**Search methods for identification of studies**

We will use a search strategy designed and checked by an information specialist with Cochrane expertise.

We will search EMBASE, MEDLINE, CENTRAL, PubMed (excluding MEDLINE), ClinicalTrials.gov and WHO ICTPR. We will place no restrictions on language of publication. As complementary search methods, we will carefully check Cochrane systematic reviews on immunomodulators for eligible studies.

**Data collection and analysis**

We will carry out data collection and analysis according to the methods recommended in the Cochrane Handbook for Systematic Reviews of Interventions. (2)

**Selection of studies**

Two authors will independently screen the titles and abstracts. All will be screened in duplicate independently and disagreements solved by a third senior author. Full reports will be obtained for studies deemed potentially eligible and the same screening process will be followed. Multiples reports of the same RCT will be merged. RCTs reported as one study in one publication will be considered as separate RCTs.

**Data extraction**

Pairs of reviewers will independently extract data on outcome data and effect modifiers and disagreements will be resolved by discussion with a third senior author. We will attempt to contact authors or sponsors of the studies for unclear or missing data and information.

**Assessment of risk of bias in included studies**

Two authors will independently assess risk of bias in the included studies using the Cochrane Risk of Bias 1 tool, outlined in the Cochrane Handbook for Systematic Reviews of Interventions. (3). Disagreements will be resolved by discussion with a third senior author.

**Measures of treatment effect**

We will express treatment effect as risk ratios (RR) with corresponding 95% confidence intervals (CI). The participant will be the unit of analysis.

**Dealing with missing data**

We will perform analyses on a modified intention-to-treat basis. For safety outcomes, patients with missing or unclear withdrawal data were considered as withdrawals due to adverse events. In maintenance trials which report remission/response rates instead of relapse/loss of response we will invert the reported remission and response rates to infer relapse and loss of response.

**Assessment of heterogeneity**

To evaluate the presence of clinical heterogeneity, we will examine trial and trial population characteristics across all eligible trials that compared each pair of interventions. We will assess the presence of clinical heterogeneity within each pairwise comparison by comparing these characteristics.

**Assessment of transitivity across treatment comparisons**

We will assess the assumption of transitivity by comparing the distribution of potential effect modifiers across the different pairwise comparisons. In this context we expect that the transitivity assumption will hold assuming the common treatment used to compare different drugs indirectly is similar when it appears in different trials, and all pairwise comparisons do not differ with respect to the distribution of effect modifiers.

We will evaluate the assumption of transitivity epidemiologically by comparing the clinical and methodological characteristics of sets of studies from the various treatment comparisons.

**Assessment of statistical heterogeneity and inconsistency**

In standard pairwise meta-analyses we will estimate different heterogeneity variances for each pairwise comparison. In the network meta-analysis, we will assume a common estimate for the heterogeneity variance across the different comparisons.

We will assess statistically the presence of heterogeneity within each pairwise comparison using the I2 statistic and its 95% CI. (4) We will base the assessment of statistical heterogeneity in the entire network on the magnitude of the heterogeneity variance parameter (Tau2) estimated from the network meta-analysis models. We will compare the magnitude of the heterogeneity variance with the empirical distribution as derived by Turner 2012. (5) We will also estimate a total I2 statistic value for heterogeneity in the network as described in Higgins 2022. We will consider downgrading the certainty of the evidence for inconsistency where I2 is greater than 60%.

**Assessment of statistical inconsistency**

We will use global and local approaches to evaluate the statistical agreement between the various sources of evidence in a network of interventions (consistency) to complement the evaluation of transitivity. To evaluate the presence of inconsistency locally we will use the loop-specific approach. This method evaluates the consistency assumption in each closed loop of the network separately as the difference between direct and indirect estimates for a specific comparison in the loop (inconsistency factor). Then, the magnitude of the inconsistency factors and their 95% CIs can be used to infer the presence of inconsistency in each loop. We will assume a common heterogeneity estimate within each loop. To check the assumption of consistency in the entire network we will use the 'design-by-treatment' model as described by Higgins 2012. (6) This method accounts for different sources of inconsistency that can occur when studies with different designs (two-arm trials versus three-arm trials) give different results as well as disagreement between direct and indirect evidence. Using this approach, we will infer the presence of inconsistency from any source in the entire network based on a Chi2 test.

All analyses will be run with R statistical package (R Development Core Team) and the netmeta library. (7)

**Assessment of reporting biases**

Most reporting biases are minimized by using an inclusive search strategy. We plan to investigate publication bias using a funnel plot if there are 10 or more studies. The magnitude of publication bias will be determined by visual inspection of the asymmetry of the funnel plot. In addition, we will test funnel plot asymmetry by performing a linear regression of intervention effect estimate against its standard error, weighted by the inverse of the variance of the intervention effect estimate. (8)

**Direct comparisons of treatment effects**

We will combine data from individual trials for meta-analysis when the interventions, patient groups and outcomes are sufficiently similar (as determined by consensus). A random-effect model will be used to pool data. (9)

**Indirect and network comparisons**

We will initially generate and assess the network diagrams to determine if a network meta-analysis is feasible. Then we will perform the network meta-analysis on all outcomes within a frequentist framework using multivariate meta-analysis.

**Relative treatment ranking**

We will estimate the cumulative probabilities for each treatment being at each possible rank and obtained a treatment hierarchy using the surface under the cumulative ranking curve (SUCRA); the larger the SUCRA the higher its rank among all available agents. (10)

**Subgroup analyses**

* Patients naïve to advanced treatments (>50% of all participants being naïve) versus patients that have failed advanced treatments previously (>50% of all participants being not naïve)

**Sensitivity analyses**

* Removal of studies where the population is mixed regarding the use of immunosuppressants (>20% of all participants on concomitant immunosuppresants)
* For the maintenance NMA, removal of studies with mixed populations of patients in a state of both clinical remission and clinical response at baseline randomization

**Summary of findings and assessment of the certainty of the evidence**

We will assess the certainty of the evidence using the GRADE approach as outlined in the GRADE handbook in order to assess the certainty of the body of evidence relating to each outcome for all comparisons. (11) We will use the GRADE working group’s approach for rating the certainty of the network meta-analysis effect estimates for all the comparisons and all outcomes. (12) We will appraise the certainty of the direct, indirect, and network evidence sequentially.

First, we will assess the certainty of the direct evidence (where available) for a given outcome, and rate the evidence using the standard GRADE approach based on consideration of: trial design limitations (risk of bias); inconsistency; imprecision; indirectness and publication bias. (13) In this approach, the direct estimates are rated for risk of bias, inconsistency, indirectness and publication bias; followed by the indirect estimates are rated based on the lowest ratings of the direct comparisons forming the most dominant loop and intransitivity; and finally, the network estimates are rated based on imprecision, incoherence, and the rating of the direct or indirect estimate that contributes the most.

At each of these stages, review authors will independently appraise the certainty ratings for the direct, indirect and network evidence. We will resolve disagreements between authors through discussion and consultation. We will rate the certainty of network evidence for each outcome as ‘high’, ‘moderate’, ‘low’ or ‘very low’ in accordance with the GRADE approach.

* High certainty: we are very confident that the true effect lies close to that of the effect.
* Moderate certainty: we are moderately confident in the effect estimate; the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
* Low certainty: our confidence in the effect estimate is limited; the true effect may be substantially different from the estimate of the effect.
* Very low certainty: we have very little confidence in the effect estimate; the true effect is likely to be substantially different from the estimate of effect.

**References**

1. Ng SC, Shi HY, Hamidi N, et al. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. Lancet Lond Engl. 2017;390(10114):2769-2778. doi:10.1016/S0140-6736(17)32448-0

2. Cumpston M, Li T, Page MJ, Chandler J, Welch VA, Higgins JP, Thomas J. Updated guidance for trusted systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions. Cochrane Database Syst Rev. 2019;10:Ed000142.

3. Cohen J. Statistical power analysis for the behavioral sciences: Academic press; 2013.

4. Mills HL, Higgins JPT, Morris RW, Kessler D, Heron J, Wiles N, et al. Detecting Heterogeneity of Intervention Effects Using Analysis and Meta-analysis of Differences in Variance Between Trial Arms. Epidemiology. 2021;32(6):846-54.

5. Turner RM, Davey J, Clarke MJ, Thompson SG, Higgins JP. Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews. Int J Epidemiol. 2012;41(3):818-27.

6. Higgins JP, Jackson D, Barrett JK, Lu G, Ades AE, White IR. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. Res Synth Methods. 2012;3(2):98-110.

7. Shim SR, Kim SJ, Lee J, Rücker G. Network meta-analysis: application and practice using R software. Epidemiol Health. 2019;41:e2019013.

8. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. Bmj. 1997;315(7109):629-34.

9. DerSimonian R, Laird N. Meta-analysis in clinical trials. Controlled Clinical Trials. 1986;7(3):177-88.

10. Salanti G, Ades AE, Ioannidis JPA. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. Journal of Clinical Epidemiology. 2011;64(2):163-71.

11. Puhan MA, Schünemann HJ, Murad MH, Li T, Brignardello-Petersen R, Singh JA, et al. A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis. Bmj. 2014;349:g5630.

12. Brozek JL, Canelo-Aybar C, Akl EA, Bowen JM, Bucher J, Chiu WA, et al. GRADE Guidelines 30: the GRADE approach to assessing the certainty of modeled evidence-An overview in the context of health decision-making. J Clin Epidemiol. 2021;129:138-50.

**Funding statement**

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests statement**

None to declare.