

Sharing information across patient subgroups to draw conclusions from sparse treatment networks

Theodoros Evrenoglou¹  | Silvia Metelli¹ | Johannes-Schneider Thomas² |
 Spyridon Siafis² | Rebecca M. Turner³ | Stefan Leucht² | Anna Chaimani¹

¹Center of Research in Epidemiology and Statistics (CRESS-U1153), Université Paris Cité, INSERM, Paris, France

²Department of Psychiatry and Psychotherapy, School of Medicine, Technical University of Munich, Munchen, Germany

³MRC Clinical Trials Unit, University College London, London, UK

Correspondence

Theodoros Evrenoglou, Université Paris Cité, Research Center of Epidemiology and Statistics (CRESS-U1153), INSERM, Paris, France; Hôpital Hôtel-Dieu, 1 Place du Parvis Notre-Dame, 75004 Paris.
 Email: tevrenoglou@gmail.com

Funding information

French National Research Agency, Grant/Award Number: ANR-22-CE36-0013-01; UK Medical Research Council, Grant/Award Number: MC_UU_00004/06

Abstract

Network meta-analysis (NMA) usually provides estimates of the relative effects with the highest possible precision. However, sparse networks with few available studies and limited direct evidence can arise, threatening the robustness and reliability of NMA estimates. In these cases, the limited amount of available information can hamper the formal evaluation of the underlying NMA assumptions of transitivity and consistency. In addition, NMA estimates from sparse networks are expected to be imprecise and possibly biased as they rely on large-sample approximations that are invalid in the absence of sufficient data. We propose a Bayesian framework that allows sharing of information between two networks that pertain to different population subgroups. Specifically, we use the results from a subgroup with a lot of direct evidence (a dense network) to construct informative priors for the relative effects in the target subgroup (a sparse network). This is a two-stage approach where at the first stage, we extrapolate the results of the dense network to those expected from the sparse network. This takes place by using a modified hierarchical NMA model where we add a location parameter that shifts the distribution of the relative effects to make them applicable to the target population. At the second stage, these extrapolated results are used as prior information for the sparse network. We illustrate our approach through a motivating example of psychiatric patients. Our approach results in more precise and robust estimates of the relative effects and can adequately inform clinical practice in presence of sparse networks.

KEYWORDS

informative priors, mixed treatment comparisons, sharing information, sparse data

1 | INTRODUCTION

Network meta-analysis (NMA) has become an essential tool for the comparative effectiveness research of healthcare interventions because it allows integrating all available information on a specific disease and usually provides estimates of relative effects with the highest possible precision (Chaimani et al., 2019; Nikolakopoulou et al., 2018). However, networks

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *Biometrical Journal* published by Wiley-VCH GmbH.

of interventions with limited data that fail to provide useful conclusions may arise under certain situations. A recent empirical study found 92 (7.4%) published NMAs that included more treatments than the number of studies in a sample of 1236 NMAs of randomized controlled trials (RCTs) with at least four interventions (Veroniki et al., 2021). These numbers suggest that the phenomenon of “sparse” networks, namely, networks with limited direct evidence in the form of few direct comparisons and/or few—1 or 2—studies per comparison, is not very rare in the literature. For example, this is often the case for “sensitive” subgroups of the population that cannot be easily included in trials, such as children, elder patients, or individuals with multimorbidity.

Results from such networks of interventions are accompanied with substantial uncertainty not only in the estimation of the relative treatment effects but also in the plausibility of the underlying assumptions because their formal evaluation is impossible (Dias et al., 2010). In addition, the large-sample approximations on which NMA models typically rely upon fail in the presence of only a handful of studies per comparison. Hence, lack of robustness and potentially limited reliability are common issues that might be encountered when analyzing sparse treatment networks (Yoon et al., 2021). To avoid these issues, it might seem reasonable to wait until more studies become available for the outcome(s) of interest. However, the pace of trial production in these contexts is usually very slow and the chance to obtain new study results shortly is low.

The use of external evidence in the form of informative prior distributions has been suggested previously in meta-analysis, either to achieve more accurate estimation of the heterogeneity parameter (Röver et al., 2019; Turner et al., 2019) or for the incorporation of non-randomized evidence (Efthimiou et al., 2017). Here, we introduce a new framework that allows sharing information between two networks of treatments that pertain to different subgroups of the population. Our framework refers to the case when only aggregate data are available; in the presence of individual participant data, other approaches may be used, such as population adjustment methods (Phillippo et al., 2020). Specifically, we use the results from a subgroup with a lot of direct evidence (i.e., many direct comparisons and many studies per comparison) forming a “dense” network to construct informative priors for the relative effects in the target subgroup forming a sparse network.

Information sharing across population subgroups is also common outside the context of NMA. For example, similar approaches are used in the so-called “basket trials,” where information is borrowed from large baskets to facilitate the estimation of treatment effects in smaller baskets (Parmar et al., 2017; Renfro & Sargent, 2017; Turner et al., 2022) involving subpopulations that are challenging to include in clinical trials. The idea of extrapolating information from a population with a lot of data to another with limited evidence has been discussed previously in an FDA report (U.S. Department of Health & Human Services Food & Drug Administration, 2016) that used adult data to draw conclusions for children for the case of individual studies. Extending these ideas to the meta-analytic setting would allow us to obtain more robust conclusions as individual trials are often underpowered (Salanti et al., 2008).

In this manuscript, we propose a two-stage approach where at the first stage, we synthesize the data from the dense network using a hierarchical NMA model. This includes a location parameter that shifts the distribution of the relative effects to make them applicable to the population of the sparse network. We also add a scale parameter that allows us to further downweight the external data and to reduce their influence. At the second stage, the results from the first stage are used as prior information in the analysis of the sparse network. We inform the location and scale parameters either through the data or using expert opinion. We illustrate our approach through an example examining the relative effects of antipsychotic treatments in two subgroups of patients: a target subgroup of children adolescents (CAs) that forms a rather sparse network and a subgroup of chronic adult patients with acute exacerbation of schizophrenia, referred here as “general patients” (GP), which forms a dense network.

2 | MOTIVATING DATASET

Our work is motivated by a recent article investigating the differences in the treatment effects among several subgroups of schizophrenia patients (Leucht et al., 2022). The aim of the article was to compare the subgroup of GP, defined as chronic patients with an acute exacerbation of positive symptoms, with more specific subgroups of patients: CA, first-episode patients, treatment-resistant patients, patients with negative symptoms, patients under substance abuse, and elder patients. In contrast to the dataset for GP, all these specific subgroups were informed by very limited direct evidence, and therefore, the original authors, who acknowledged the problems of analyzing sparse networks, only used pairwise meta-analysis. To illustrate our approach, we use the data from two of the above subgroups: CA and GP. The former is considered the target subgroup that forms a sparse network and the latter is the only available subgroup with abundant direct evidence (forming a dense network) that can be safely used to elicit informative priors for CA.

subgroup P_1 . Each study i provides the observed mean (change score/endpoint) y_{it_k} for the treatment $t_k \in T_a$ of arm k with $k = 1, \dots, K_i$ and K_i the total number of arms in i . For simplicity, in what follows we write the y_{it_k} as y_{ik} , its variance sd_{ik}^2 , and the respective “true” mean is denoted as θ_{ik} . We consider treatment $j = 1$ as the common network reference for both P_1 and P_2 and the relative effects μ_{1j} ($j > 1, j \in T_a$) as the basic parameters. For every study, we arbitrarily choose the treatment of arm $k = 1$ (t_1) as the study-specific baseline treatment.

3.2 | Models for sharing information between two population subgroups

Different models stem from different assumptions about the relation between the two subgroups P_1 and P_2 . We start with the description of the standard hierarchical NMA model that in the present setting synthesizes the two subgroups in a naïve way, namely, as if they were equivalent. Then, we move to more plausible NMA models in which we incorporate a location parameter acknowledging the difference between P_1 and P_2 and a scale parameter aiming to increase the uncertainty of the studies in subgroup P_2 .

3.2.1 | Naïve synthesis

In this approach, the two subgroups are combined together as if all participants were coming from the same population P . In other words, this approach makes the strong assumption that P_1 and P_2 are equivalent, namely, $P_1 \equiv P_2 \equiv P$. Hence, for every study $i = 1, \dots, N$ and treatment $t_k \in T_a$, the observed means follow a normal distribution

$$y_{ik} \sim N(\theta_{ik}, sd_{ik}^2). \quad (1)$$

Then, the underlying SMD between every treatment t_2, t_3, \dots, t_{K_i} versus the baseline treatment t_1 is

$$\delta_{i,1k} = \frac{\theta_{ik} - \theta_{i1}}{sd_i^{pooled}}, (k > 1), \quad (2)$$

with $sd_i^{pooled} = \sqrt{\sum_{k=1}^{K_i} \frac{n_{ik} sd_{ik}^2}{n_{ik} - K_i}}$ the between-arm standard deviation. Under the random-effects assumption, the underlying relative effects are assumed to follow a multivariate normal distribution

$$\delta_i \sim N_{K_i-1}(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \quad (3)$$

where $\boldsymbol{\mu} = (\mu_{12}, \dots, \mu_{1K_i})$ is the vector of the summary relative effects and $\boldsymbol{\Sigma}$ the between-study variance–covariance matrix with entries τ^2 (the common heterogeneity variance across comparisons) in the diagonal and $\frac{\tau^2}{2}$ in the off-diagonal. Finally, the transitivity assumption for every pair of treatments $j, l \in T_a$ ($j, l > 1$) implies that

$$\mu_{jl} = \mu_{1l} - \mu_{1j}.$$

Non-informative prior distributions are usually employed for every μ_{1j} and τ , such as $N(0, 10000)$ and the half-normal $HN(1)$, respectively.

Apart from the strong assumption about the similarity of the two subgroups, a further problem in the present setting is that the final results for the joint population P will be dominated by the dense network (P_2) rather than from the target subgroup P_1 . The same would apply if the results of the dense network were used directly as prior information for P_1 . To avoid this issue, in the proposed approach, we extrapolate the results from P_2 to P_1 , acknowledging differences between the two subgroups, before forming informative prior distributions for P_1 .

3.2.2 | Using the external subgroup P_2 to construct informative priors for the target subgroup P_1

This is a two-stage approach where at the first stage, we extrapolate the results from the dense network of subgroup P_2 to P_1 , and at the second stage, we use predictions from this extrapolation to form prior distributions for the analysis of P_1 .

First stage: Extrapolating external results to the target population subgroup

In contrast to the naïve synthesis, the main assumption here is that the two subgroups have some underlying differences in the treatment effects that can be considered as a different location of their outcome distributions. To incorporate some uncertainty about this assumption, in Equation (1), we introduce a scale parameter $w_i \in (0, 1]$ that inflates the variance of each y_{ik} in subgroup P_2 for each $i = n_1 + 1, n_1 + 2, \dots, N$ (Baker & Jackson, 2013). This modifies Equation (1) into

$$y_{ik} \sim N\left(\theta_{ik}, \frac{sd_{ik}^2}{w_i}\right). \quad (4)$$

Equation (4) implies that the parameters w_i only affect the variance of the the mean scores y_{ik} and not the true study-specific SMDs that are still the same as in Equation (2). Assuming that $\beta_{1j} = \mu_{1j}^{P_2} - \mu_{1j}^{P_1}$ for each $\{1, j\}, j \in T_c$, the location parameters β that aim to shift the original distribution of the SMDs in P_2 toward the distribution of P_1 are then added in Equation (3):

$$\delta_i \sim N_{K_i-1}(\mu^{P_2} - \beta, \Sigma), \quad (5)$$

with $\beta = (\beta_{12}, \dots, \beta_{1K_i})$ being the vector of the “true” differences between the comparison-specific SMDs of the two subgroups. The extrapolated means $\mu^* = \mu^{P_2} - \beta$ for population P_2 in Equation (5) are then considered similar to those expected for population P_1 . Note that Equation (5) could be equivalently written as $\delta_i \sim N_{K_i-1}(\mu^{P_2} + \beta, \Sigma)$, if we had assumed that $\beta_{1j} = \mu_{1j}^{P_1} - \mu_{1j}^{P_2}$.

By fitting this modified NMA model, we obtain the extrapolated SMD estimates $\hat{\mu}_{jl}^*$ for every pair of treatments $j, l \in T_c$. Subsequently, the predictive distributions (Higgins et al., 2009) of μ_{jl}^{new} , namely, $N(\mu_{jl}^*, (\tau^{P_2})^2)$ are used at the second stage as prior distributions for P_1 . Note that the parameter $(\tau^{P_2})^2$ represents the heterogeneity across the studies in P_2 .

Second stage: NMA of the target subgroup using informative prior distributions for relative effects

Here, the model of Section 3.2.1 is used for $i = 1, \dots, n_1$. The key difference is that for the μ_{1l} s, we use as prior distribution $\mu_{1l}^{P_1} \sim N(\hat{\mu}_{1l}^{new}, \text{var}(\hat{\mu}_{1l}^{new}))$ where $\hat{\mu}_{1l}^{new}$ and $\text{var}(\hat{\mu}_{1l}^{new})$ are the posterior mean and variance of μ_{1l}^{new} , respectively. The use of informative priors is expected to improve the precision in the final NMA estimates without dominating the analysis. This is because the priors are constructed at the first stage by extrapolating the results of the dense network to those expected for the target sparse network.

3.2.3 | Informing the location and the scale parameters

Constructing prior distributions for β s using the data

To obtain an estimate of the difference in the outcome distributions between the two subgroups, we first use the estimates $u_{1j}^{P_1}$ and $u_{1j}^{P_2}$ obtained from separate pairwise meta-analyses for each population and each pair of treatments $\{1, j\}, j \in T_c$. To allow the estimation of heterogeneity for every comparison in P_1 , here we assume a common comparison-specific heterogeneity across the two subgroups $(\sigma_{1j}^{P_1})^2 = (\sigma_{1j}^{P_2})^2 = \sigma_{1j}^2$. Using the difference of these estimates $d_{1j} = u_{1j}^{P_2} - u_{1j}^{P_1}$, we construct prior distributions for the parameters β_{1j} , namely,

$$\beta_{1j} \sim N(d_{1j}, \text{var}(d_{1j})). \quad (6)$$

For treatment comparisons not evaluated in P_2 , we use a non-informative normal prior $N(0, 10000)$.

Constructing prior distributions for β s using expert opinion

Elicitation of expert opinion can be undertaken with several methods such as face-to-face interviews, software tools, or questionnaires. The pooled change score of each treatment $j \in T_c$ in P_2 can be used as a reference for the experts who need to provide an “estimate” for the change score in P_1 . Suppose that x_{hj} is the change score and sd_{hj} the standard deviation provided by the expert $h = 1, 2, \dots, H$ for treatment $j \in T_c$. Let also γ_h denote the overall confidence of each expert to

provided opinions. To combine the estimates provided by different experts, we can then use the following meta-analysis model:

$$x_{hj} \sim N\left(c_{hj}, \frac{sd_{hj}^2}{\gamma_h}\right), \quad (7)$$

$$d_{hj} \sim N\left(\xi_j^{P_1}, \sigma^2\right), \quad (8)$$

with $d_{hj} = \frac{c_{hj}}{\text{med}(sd_i^{\text{pooled}})}$, $\forall i, h, j$ the standardized values of c_{hj} . Given that the pooled standard deviation cannot be elicited by experts, the median across the studies in the data is used. Note that standardization is not necessary when all studies in the data use the same scale. Then, the pooled SMDs can be obtained as $u_{1j}^{P_1} = \xi_j^{P_1} - \xi_1^{P_1}$ and the priors for the parameter β_{1j} is constructed as in Equation (6).

Prior distributions for the scale parameter

Dividing the variances in Equation (4) by w_i can be seen as a special case of the power prior method where a specific power is chosen for the likelihood of each study (Chen & Ibrahim, 2000; Efthimiou et al., 2017; Ibrahim et al., 2015). Values of w_i close to 0 reflect a serious downweight of the evidence coming from the external subgroup P_2 , whereas values close to 1 imply a mild downweight. Typical choices of prior distributions for the scale parameter can be the beta or the uniform distribution. The parameters of these distributions should be chosen to reflect the prior beliefs regarding the specific characteristic according to which the downweight takes place. For example, a Beta(3, 3) can be used for cases where a moderate downweight needs to apply for the external information. This distribution places its mass around 0.5 with 95% range equal to [0.15, 0.84] thus reflecting more uncertainty about the magnitude of our downweighting. Left-skewed beta priors (e.g., Beta(1, 6)) can be used for weight values close to 0, while on the other hand, right-skewed beta priors (e.g., Beta(6, 1)) can be used for weight values close to 1. The previous downweighting schemes can be achieved with other types of prior distributions. For example, a Unif(0.4, 0.6), a Unif(0.1, 0.3) or a Unif(0.8, 1) can be used for moderate, serious, or mild downweight, respectively. Finally, it is generally not recommended to use priors for the scale parameter that can allow for values larger than 1. This could further increase the impact of external data and enable their dominance in the final NMA estimates.

4 | APPLICATION

4.1 | Implementation

We consider throughout CA being the target population subgroup P_1 and GP the external subgroup P_2 . We applied in total nine models six of which refer to the combinations of the different prior distributions for the location (β_{1j}) and scale parameters (w_i).

1. Naïve synthesis model for CA and GP with non-informative priors.
2. NMA for CA with informative priors from GP.
 - a. data-based prior distributions for the location parameters with
 - (i) no downweight for any GP study ($w_i = 1$) using only treatments in T_c (No DW)
 - (ii) moderate downweight (i.e., $w_i \sim \text{Beta}(3, 3)$) to high risk of bias (RoB) GP studies using only treatments in T_c (RoB DW). The risk of bias was assessed according to the Cochrane's RoB tool (Sterne et al., 2019) and in total 16 studies were rated at high RoB for GP.
 - (iii) no downweight for any GP study evaluating treatments in T_c and moderate downweight (i.e., $w_i \sim \text{Beta}(3, 3)$) for those evaluating at least one treatment in $T_a - T_c$ (NCT DW). This means that here we use the full network of GP with the 34 treatments and we downweight all studies that contain at least one treatment that has never been evaluated in CA (107 studies).
 - b. prior distributions based on expert opinion for the location parameter combined the three above (i–iii) possibilities for the scale parameters.
3. NMA for CA with non-informative priors (i.e., “standard” NMA).
4. Pairwise meta-analysis for CA with non-informative priors.

All the analyses were conducted using the `rjags` (22) package through the R statistical software (R version 4.0.3, 2020-10-10) (23). For all models, we ran two chains in parallel, performed 50,000 iterations and discarded the first 10,000 samples of each chain. We checked the chains convergence using the Gelman–Rubin criterion (Brooks & Gelman, 1998) with a value below 1.1 to indicate failure of convergence. We also visually checked the Markov chains using trace plots to inspect the sampling behavior and assess mixing across chains and convergence. The R code and the data used in this article are available in the Appendix and can also be accessed at: https://github.com/TEvrenoglou/codes_sharing_information

4.2 | Elicitation of expert's opinion

We prepared a questionnaire and circulated it to psychiatrists with experience in treating schizophrenia for CA and GP (available in Appendix 2). We asked each expert to provide an estimate of the “expected” mean score reduction in the PANSS scale from baseline to endpoint for each drug for CA given the pooled mean score reduction obtained from the data for GP. We further asked them to provide a measure of uncertainty around the given expected mean scores (a) in the form of standard deviation and (b) in a 10-point scale of their confidence. We obtained in total 22 responses which we averaged for each drug using the meta-analysis model given in Equations (7)–(8). We used a non-informative $N(0, 10000)$ and a weakly informative $HN(1)$ prior for the means ξ_j^{T-1} , $j \in T_c$ and the parameter σ , respectively.

4.3 | Results

The estimates for the basic comparisons of the CA network are depicted in Figure 2 and for all relative comparisons in Online Appendix 1, Tables 10–17. As expected, for most comparisons, the use of informative priors leads to a substantial improvement in the precision of the relative effects. The naïve synthesis model provides the most precise results but relies on a strong assumption that is likely to be implausible. Interestingly, CA results from the standard NMA model (with non-informative priors) tend to be less precise than the respective direct estimates. This occurs often in sparse networks because indirect comparisons as well as heterogeneity are estimated with large uncertainty. Overall, standard NMA most often does not give any insight for comparisons without direct evidence as it yields very large credible intervals. For some of these comparisons (e.g., Haloperidol vs. Placebo), the NMA models with informative priors result in more conclusive relative effect estimates. In terms of point estimates, results appear generally robust across the different models with only few exceptions, such as the extreme cases of Fluphenazine and Trifluoperazine. Those two are very old drugs for which the evidence is outdated and sparse in both GP and CA networks. The latter implies that the informative prior can dominate in the analysis and yield to estimate closer to the GP. Overall, point estimates appeared to be more robust for the basic comparisons for which direct evidence is available in the network of CA. Finally, no convergence issues were identified across the chains. The corresponding trace plots are available in Appendix 1, Figures 8–13.

Only small differences can be observed among the six models with informative priors, suggesting that the approach of informing the location and the scale parameters does not affect materially the final predictions; this is possibly due to the huge amount of data in the GP network. For a few comparisons, the data-based approach for β provides to some degree different relative effects than those obtained from the expert opinion approach, implying that the experience from clinical practice does not fully agree with the available data. Those comparisons might need to be prioritized for the design of new trials. Of course, when the full GP network was used and only studies comparing treatments in $T_a - T_c$ were downweighted some extra precision was gained. Such an approach might be more useful when the external network is not as dense as in the present dataset.

Assessment of inconsistency was performed using the node-splitting method (Dias et al., 2010). Inference regarding the presence of inconsistency was based on Bayesian p -values, and on the visual inspection of the posterior densities of direct and indirect evidence, Bayesian p -values were calculated as $2 \times \min\{P, 1 - P\}$ where P is the probability that the difference between direct and indirect comparisons is positive. Additionally, the different antipsychotics and placebo were ranked across all models using the surface under the cumulative ranking curve (SUCRA) (Salanti et al., 2011) method. Overall, the respective posterior densities of the direct and indirect estimates appeared to be quite similar across the different comparisons and models and no p -value indicated lack of consistency (Online Appendix 1, Figures 2–6, Tables 2–9). In terms of ranking, the results appear to be generally robust across models including the naïve synthesis and the standard NMA model. Clozapine was always ranked as the most effective antipsychotic. Finally, Olanzapine, Risperidone, and Molindone were ranked in second, third, and fourth positions, respectively, in most of the fitted models (Appendix 1, Figure 7).

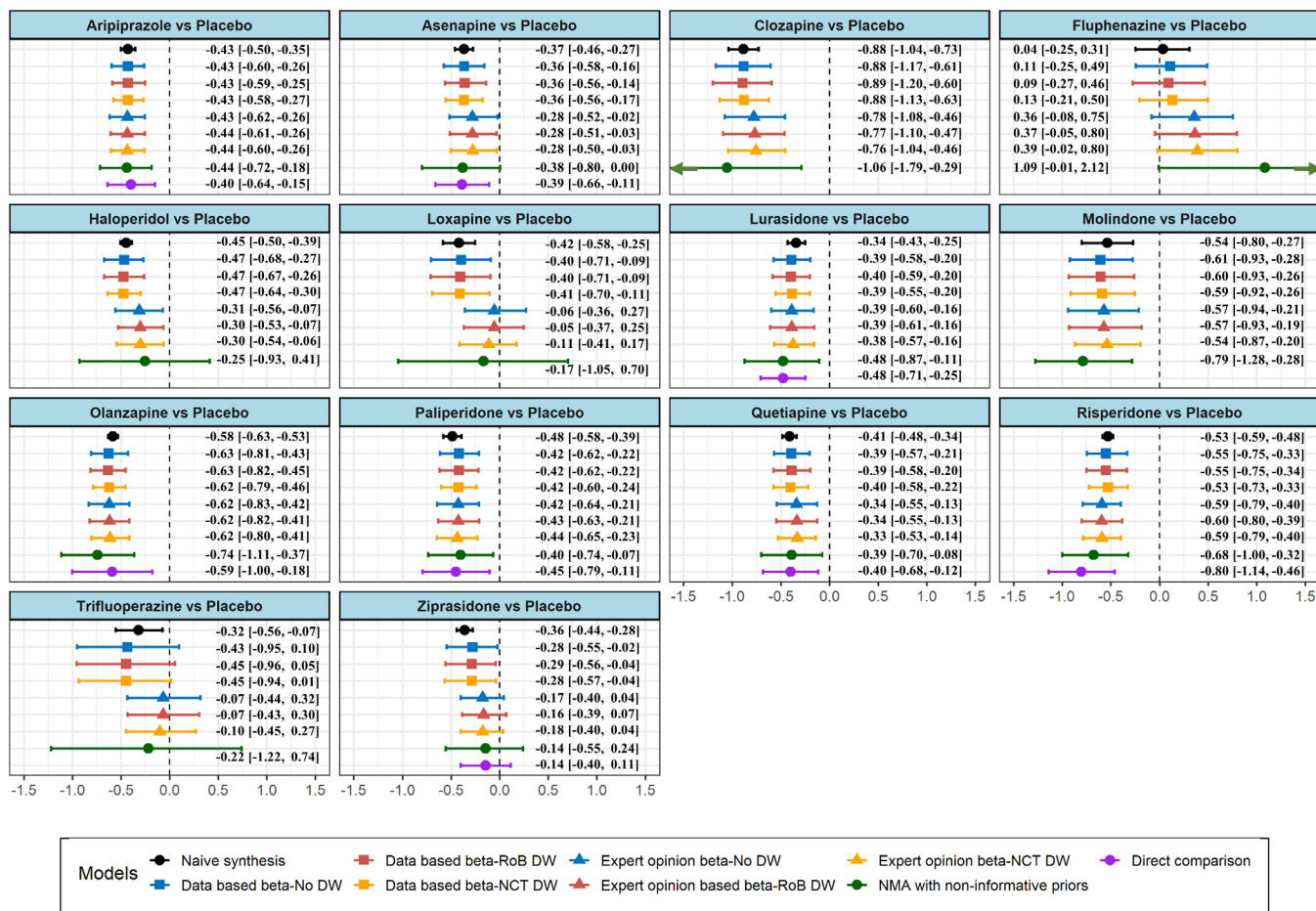


FIGURE 2 Results from the analysis of the network defined for the population of CA. The outcome of interest is the reduction of the overall schizophrenia symptoms and the effect measure is the standardized mean difference. DW, downweight; RoB, risk of bias; NCT, non-common treatment.

5 | DISCUSSION

In this paper, we present a Bayesian framework for NMA of sparse networks aiming to improve the performance of the relative effect estimates in terms of precision and reliability. To this end, we borrow information using external data from a dense network that targets a different subgroup of the population. We model the differences between the two populations with a two-stage approach. At the first stage, we analyze only the dense network, and we extrapolate its results to the sparse network though the incorporation of a location parameter in the distribution of the summary relative effects. Then, we use the extrapolated results to construct informative prior distributions for the target population subgroup. Similar approaches have been used previously at the level of primary studies (Dunne et al., 2011; Dunoyer, 2011; Hampson et al., 2014; Rhodes et al., 2015). We further introduce a scale parameter that inflates the variance of the studies in the dense network to reflect the potential uncertainty regarding the assumed relationship between the two population subgroups (Hlavin et al., 2016; Wadsworth et al., 2018).

We proposed two different approaches for informing the location parameters; one based on the data and one based on expert opinion. The former is easier to implement but the data are used in both stages: once to form the priors and once to perform the NMA. This is a drawback of the data-based approach as it might introduce some correlation that is ignored. In addition, this approach requires that a certain number of direct comparisons are available in both networks; otherwise, the use of non-informative priors would be necessary for several β s at the first stage. The prior elicitation approach is general as it does not rely on the amount of direct evidence. However, results from such an approach are always subjective to some degree and handling disagreements between different experts might be challenging. Here, we requested the necessary information from the psychiatrists through a questionnaire and some of them raised concerns that

it was difficult for them to make a suggestion. Other ways to elicit information from experts exist. For example, Turner et al. used a three-stage elicitation approach where the experts were interviewed through 1:1 meetings or by telephone (Turner et al., 2022). Alternatively, information on these parameters may be obtained by using, when available, external data, such as from observational studies, electronic health records, or other networks pertaining to populations with similar characteristics.

In our motivating example, the two approaches for constructing priors for the location parameters were generally in agreement except the comparison between Loxapine and Placebo where, in contrast to the data-based approach, the analyses based on the expert opinion may not support the presence of an important effect. In all other comparisons, despite some differences in the magnitude of the estimated summary effects, the two approaches yielded similar conclusions. Regarding the scale parameters, we applied no downweight for trials considered more relevant to our target population and moderate downweight for studies that were considered less relevant or reliable. The results remained robust across the different downweighting schemes that we used. However, the use of Beta(3, 3) prior for the w_i s can be considered a conservative choice as it is centered at 0.5 and provides similar mass to both the left and right tails allowing for both strong and mild downweight, for example, $w_i \approx 0$ or $w_i \approx 1$, respectively. Sensitivity analysis could further consider left-skewed Beta distributions (e.g., Beta(1, 8)) that would assign weights mostly close to 0 and right-skewed Beta's (e.g., Beta(8, 1)) that would assign weights mostly close to 1. Compared to the standard NMA with non-informative priors, our two-stage approach provided substantially more precise estimates. The lack of robustness of the standard NMA model for this dataset is also deduced from the fact that almost always it yielded less precise results than pairwise meta-analysis. As expected, the most precise results were obtained from the naïve synthesis model. Nevertheless, this model does not account for any differences across the two population subgroups.

In the presence of few studies, estimation of the heterogeneity is challenging (Turner et al., 2019). In our analysis, we used a weakly-informative half normal prior distribution for the heterogeneity parameter (Röver et al., 2021). However, a sensitivity analysis on the prior distribution for heterogeneity could be considered as in the absence of sufficient data, the estimation of heterogeneity might be affected by the choice of the prior. For example, using informative priors for heterogeneity could possibly further increase the precision of our approach. Usually, these are based on empirical distributions that depend on the type of outcome and treatment comparisons (Rhodes et al., 2015; Turner et al., 2012, 2015). Alternatively, more sophisticated options that allow the incorporation of external data in the estimation of the heterogeneity parameter may be applied (Turner et al., 2019).

A limitation of our work is that we did not explore under which conditions the introduced location parameters reduce or increase the bias of the summary estimates. This would require to evaluate the proposed Bayesian framework through a simulation study. Nevertheless, data generation in this setting might be challenging as it would require construction of the differences between the two populations based on clinically meaningful baseline characteristics. This merits a lot of further consideration and can form a follow-up article. In addition, such a study could only evaluate the data-based approach for the priors as it would be impossible to obtain expert opinion for many simulated datasets.

Our proposed method was built upon a Bayesian framework as this offers a straightforward and intuitive way to incorporate external information in the analysis through informative prior distributions. Alternatively, a two-stage frequentist framework could be used where the prior information for the location parameters could be incorporated through pseudodata as in Rhodes et al. (Rhodes et al., 2016). However, this approach uses meta-regression to fit an NMA model, and therefore, it is directly applicable only to NMA of two-arm studies (Rhodes et al., 2016). A further possibility may be to model the differences in the two populations through inconsistency parameters using the random inconsistency design-by-treatment interaction model proposed by Jackson et al. (2014). Such an approach would assume different inconsistency parameters for populations P_1 and P_2 but the same for studies within each population. A potential future simulation study could investigate similarities and differences in the estimation between the three approaches.

Overall, the proposed framework offers a reliable approach for assisting the analysis of sparse networks. Of course, its application requires close collaboration between statisticians and experienced clinicians to ensure that sharing of information between different population subgroups is clinically meaningful. Finally, performing NMA in cases of sparse networks will always be a challenging procedure and the best approach will likely be context-dependent.

ACKNOWLEDGMENTS

Theodoros Evrenoglou received funding by the French National Research Agency under the project ANR-22-CE36-0013-01. Rebecca Turner was supported by the UK Medical Research Council (Programme MC_UU_00004/06). We are thankful to the following psychiatrists for providing expert opinion: Arango Celso, Carpenter Will, Crespo Facorro Benedicto, Davidson Michael, Dazzan Paola, Gerhard Gruender, Hasan Alkomet, Heres Stephan, Hui Wu, Jauhar Sameer, Juckel

Georg, McGorry Pat, Murray Robin, Musil Richard, Peter Natalie, Samara Myrto, Stip Emmanuel, Takeuchi Hiroyoshi, Weiser Mark, and Zhu Yikang.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

All data are available in the Supplementary Material and on the following public GitHub repository: https://github.com/TEvrenoglou/codes_sharing_information

ORCID

Theodoros Evrenoglou  <https://orcid.org/0000-0003-3336-8058>

REFERENCES

- Baker, R. D., & Jackson, D. (2013). Meta-analysis inside and outside particle physics: Two traditions that should converge? *Research Synthesis Methods*, 4(2), 109–124. <https://doi.org/10.1002/jrsm.1065>
- Brooks, S. P., & Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7(4), 434–455.
- Chaimani, A., Caldwell, D. M., Li, T., Higgins, J. P., & Salanti, G. (2019). Undertaking network meta-analyses. In: Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., Welch, V. A. (Eds.), *Cochrane handbook for systematic reviews of interventions* (pp. 285–320). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781119536604.ch11>
- Chen, M. H., & Ibrahim, J. G. (2000). Power prior distributions for regression models. *Statistical Science*, 15(1), 46–60. <https://doi.org/10.1214/ss/1009212673>
- Dias, S., Welton, N. J., & Caldwell, D. M. (2010). Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Statistics in Medicine*, 29(7–8), 932–944. <https://doi.org/10.1002/sim.3767>
- Dunne, J., Rodriguez, W. J., Murphy, M. D., Beasley, B. N., Burckart, G. J., Filie, J. D., Lewis, L. L., Sachs, H. C., Sheridan, P. H., Starke, P., & Yao, L. P. (2011). Extrapolation of adult data and other data in pediatric drug-development programs. *Pediatrics*, 128(5), e1242–e1249. <https://doi.org/10.1542/peds.2010-3487>
- Dunoyer, M. (2011). Accelerating access to treatments for rare diseases. *Orphanet Journal of Rare Diseases*, 10(7), 475–476. <https://doi.org/10.1038/nrd3493>
- Efthimiou, O., Mavridis, D., Debray, T. P. A., Samara, M., Belger, M., Siontis, G. C. M., Leucht, S., & Salanti, G. (2017). Combining randomized and non-randomized evidence in network meta-analysis. *Statistics in Medicine*, 36(8), 1210–1226. <https://doi.org/10.1002/sim.7223>
- Hampson, L. V., Whitehead, J., Eleftheriou, D., & Brogan, P. (2014). Bayesian methods for the design and interpretation of clinical trials in very rare diseases. *Statistics in Medicine*, 33(24), 4186–4201. <https://doi.org/10.1002/sim.6225>
- Higgins, J. P. T., Thompson, S. G., & Spiegelhalter, D. J. (2009). A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172, 137–159. <https://doi.org/10.1111/j.1467-985X.2008.00552.x>
- Hlavin, G., Koenig, F., Male, C., & Posch, M., & Bauer, P. (2016). Evidence, eminence and extrapolation. *Statistics in Medicine*, 35(13), 2117–2132. <https://doi.org/10.1002/sim.6865>
- Ibrahim, J. G., Chen, M. H., Gwon, Y., & Chen, F. (2015). The power prior: Theory and applications. *Statistics in Medicine*, 34(28), 3724–3749. <https://doi.org/10.1002/sim.6728>
- Jackson, D., Barrett, J. K., Rice, S., White, I. R., & Higgins, J. P. T. (2014). A design-by-treatment interaction model for network meta-analysis with random inconsistency effects. *Statistics in Medicine*, 33(21), 3639–3654. <https://doi.org/10.1002/sim.6188>
- Law, M., Alam, N., Veroniki, A. A., Yu, Y., & Jackson, D. (2019). Two new approaches for the visualisation of models for network meta-analysis. *BMC Medical Research Methodology [Electronic Resource]*, 19(1), 61. <https://doi.org/10.1186/s12874-019-0689-9>
- Leucht, S., Chaimani, A., Krause, M., Schneider-Thoma, J., Wang, D., Dong, S., Samara, M., Peter, N., Huhn, M., Priller, J., & Davis, J. M. (2022). The response of schizophrenia subgroups to different antipsychotic drugs: A systematic review and meta-analysis. *Lancet Psychiatry*, 9, 884–893. [https://doi.org/10.1016/S2215-0366\(22\)00304-2](https://doi.org/10.1016/S2215-0366(22)00304-2)
- Nikolakopoulou, A., Mavridis, D., Furukawa, T. A., Cipriani, A., Tricco, A. C., Straus, S. E., Siontis, G. C. M., Egger, M., & Salanti, G. (2018). Living network meta-analysis compared with pairwise meta-analysis in comparative effectiveness research: Empirical study. *British Medical Journal*, 360, k585. <https://doi.org/10.1136/bmj.k585>
- Parmar, M. K., Sydes, M. R., Cafferty, F. H., Choodari-Oskooei, B., Langley, R. E., Brown, L., Phillips, P. P., Spears, M. R., Rowley, S., Kaplan, R., James, N. D., Maughan, T., Paton, N., & Royston, P. J. (2017). Testing many treatments within a single protocol over 10 years at MRC Clinical Trials Unit at UCL: Multi-arm, multi-stage platform, umbrella and basket protocols. *Clinical Trials*, 14(5), 451–461. <https://doi.org/10.1177/1740774517725697>
- Phillippo, D. M., Dias, S., Ades, A. E., & Welton, N. J. (2020). Assessing the performance of population adjustment methods for anchored indirect comparisons: A simulation study. *Statistics in Medicine*, 39(30), 4885–4911. <https://doi.org/10.1002/sim.8759>
- Plummer, M. (2015). rjags: Bayesian Graphical Models using MCMC. R package version 4–10. <https://CRAN.R-project.org/package=rjags>
Published online 2015

- R Core Team. (2016). R: A Language and Environment for Statistical Computing. Published online 2016. <https://www.R-project.org/>
- Renfro, L. A., & Sargent, D. J. (2017). Statistical controversies in clinical research: Basket trials, umbrella trials, and other master protocols: a review and examples. *Annals of Oncology*, 28(1), 34–43. <https://doi.org/10.1093/annonc/mdw413>
- Rhodes, K. M., Turner, R. M., & Higgins, J. P. T. (2015). Predictive distributions were developed for the extent of heterogeneity in meta-analyses of continuous outcome data. *Journal of Clinical Epidemiology*, 68(1), 52–60. <https://doi.org/10.1016/j.jclinepi.2014.08.012>
- Rhodes, K. M., Turner, R. M., White, I. R., Jackson, D., Spiegelhalter, D. J., & Higgins, J. P. T. (2016). Implementing informative priors for heterogeneity in meta-analysis using meta-regression and pseudo data. *Statistics in Medicine*, 35(29), 5495–5511. <https://doi.org/10.1002/sim.7090>
- Röver, C., Bender, R., Dias, S., Schmid, C. H., Schmidli, H., Sturtz, S., Weber, S., & Friede, T. (2021). On weakly informative prior distributions for the heterogeneity parameter in Bayesian random-effects meta-analysis. *Research Synthesis Methods*, 12(4), 448–474. <https://doi.org/10.1002/jrsm.1475>
- Röver, C., Wandel, S., & Friede, T. (2019). Model averaging for robust extrapolation in evidence synthesis. *Statistics in Medicine*, 38(4), 674–694. <https://doi.org/10.1002/sim.7991>
- Salanti, G., Ades, A. E., & Ioannidis, J. P. A. (2011). Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: An overview and tutorial. *Journal of Clinical Epidemiology*, 64(2), 163–171. <https://doi.org/10.1016/j.jclinepi.2010.03.016>
- Salanti, G., Higgins, J. P. T., Ades, A. E., & Ioannidis, J. P. A. (2008). Evaluation of networks of randomized trials. *Statistical Methods in Medical Research*, 17(3), 279–301. <https://doi.org/10.1177/0962280207080643>
- Sterne, J. A. C., Savović, J., Page, M. J., Blencowe, N. S., Boutron, I., Cates, C. J., Cheng, H.-Y., Corbett, M. S., Eldridge, S. M., Emberson, J. R., Hernán, M. A., Hopewell, S., Hróbjartsson, A., Junqueira, D. R., Jüni, P., Kirkham, J. J., Lasserson, T., Li, T., ... Higgins, J. P. T. (2019). RoB 2: A revised tool for assessing risk of bias in randomised trials. *Bmj*, 366, 14898. <https://doi.org/10.1136/bmj.14898>
- Turner, R. M., Davey, J., Clarke, M. J., Thompson, S. G., & Higgins, J. P. (2012). Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews. *International Journal of Epidemiology*, 41(3), 818–827. <https://doi.org/10.1093/ije/dys041>
- Turner, R. M., Domínguez-Islas, C. P., Jackson, D., Rhodes, K. M., & White, I. R. (2019). Incorporating external evidence on between-trial heterogeneity in network meta-analysis. *Statistics in Medicine*, 38(8), 1321–1335. <https://doi.org/10.1002/sim.8044>
- Turner, R. M., Jackson, D., Wei, Y., Thompson, S. G., & Higgins, J. P. T. (2015). Predictive distributions for between-study heterogeneity and simple methods for their application in Bayesian meta-analysis. *Statistics in Medicine*, 34(6), 984–998. <https://doi.org/10.1002/sim.6381>
- Turner, R. M., Turkova, A., Moore, C. L., Bamford, A., Archary, M., Barlow-Mosha, L. N., Cotton, M. F., Cressey, T. R., Kaudha, E., Lugemwa, A., Lyall, H., Mujuru, H. A., Mulenga, V., Musiime, V., Rojo, P., Tudor-Williams, G., Welch, S. B., Gibb, D. M., ... Ford, D. (2022). Borrowing information across patient subgroups in clinical trials, with application to a paediatric trial. *BMC Medical Research Methodology*, 22(1), 49. <https://doi.org/10.1186/s12874-022-01539-3>
- U.S. Department of Health and Human Services Food and Drug Administration. (2016). Leveraging existing clinical data for extrapolation to pediatric uses of medical devices—Guidance for industry and Food and Drug Administration staff draft guidance.
- Veroniki, A. A., Tsokani, S., Zevgiti, S., Pagkalidou, I., Kontouli, K.-M., Ambarcioglu, P., Pandis, N., Lunny, C., Nikolakopoulou, A., Theodoros, P., Chaimani, A., Straus, S. E., Brian, H., Tricco, A. C., Mavridis, D., & Salanti, G. (2021). Do reporting guidelines have an impact? Empirical assessment of changes in reporting before and after the PRISMA extension statement for network meta-analysis. *Systematic Reviews*, 10(1), 246. <https://doi.org/10.1186/s13643-021-01780-9>
- Wadsworth, I., Hampson, L. V., & Jaki, T. (2018). Extrapolation of efficacy and other data to support the development of new medicines for children: A systematic review of methods. *Statistical Methods in Medical Research*, 27(2), 398–413. <https://doi.org/10.1177/0962280216631359>
- Yoon, J. H., Dias, S., & Hahn, S. (2021). A method for assessing robustness of the results of a star-shaped network meta-analysis under the unidentifiable consistency assumption. *BMC Medical Research Methodology [Electronic Resource]*, 21(1), 113. <https://doi.org/10.1186/s12874-021-01290-1>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Evrenoglou, T., Metelli, S., Thomas, J.-S., Sifis, S., Turner, R. M., Leucht, S., & Chaimani, A. (2024). Sharing information across patient subgroups to draw conclusions from sparse treatment networks. *Biometrical Journal*, 66, 2200316. <https://doi.org/10.1002/bimj.202200316>