*Research Synthesis Methods* (2025), **0**: 1–14
doi:10.1017/rsm.2025.10026

RESEARCH ARTICLE

# Trials and triangles: Network meta-analysis of multi-arm trials with correlated arms

Gerta Rücker and Guido Schwarzer

Institute of Medical Biometry and Statistics, Medical Faculty and Medical Center—University of Freiburg, Freiburg, Germany
**Present address:** Stefan-Meier-Straße 26, D-79104 Freiburg, Germany.

**Corresponding authors:** Gerta Rücker; Email: gerta.ruecker@uniklinik-freiburg.de

**Received:** 11 January 2025; **Revised:** 6 June 2025; **Accepted:** 10 June 2025

**Keywords:** consistency; correlated arms; multi-arm studies; network meta-analysis (NMA)

## Abstract

For network meta-analysis (NMA), we usually assume that the treatment arms are independent within each included trial. This assumption is justified for parallel design trials and leads to a property we call consistency of variances for both multi-arm trials and NMA estimates. However, the assumption is violated for trials with correlated arms, for example, split-body trials. For multi-arm trials with correlated arms, the variance of a contrast is not the sum of the arm-based variances, but comes with a correlation term. This may lead to violations of variance consistency, and the inconsistency of variances may even propagate to the NMA estimates. We explain this using a geometric analogy where three-arm trials correspond to triangles and four-arm trials correspond to tetrahedrons. We also investigate which information has to be extracted for a multi-arm trial with correlated arms and provide an algorithm to analyze NMAs including such trials.

### Highlights

### What is already known?

The contrasts in parallel-arm trials, included in a meta-analysis, show consistent variances. The approach used by R package *netmeta* so far relied on variance consistency.

### What is new?

Variance consistency can be violated in trials with correlated arms.

### Potential impact for RSM readers

We provide a method for analyzing trials with inconsistent variances that is implemented in R package *netmeta*.

## 1. Introduction

Network meta-analysis (NMA), a method that generalizes standard pairwise meta-analysis to a network of three or more treatments, has become an important tool for evidence synthesis in health care.[1,2] Bayesian and frequentist methods have been proposed to analyze such networks. Both contrast-based and arm-based methods are available.[3]

---

This article was awarded Open Data and Open Materials badges for transparent practices. See the Data availability statement for details.
**Abbreviations:** NMA, network meta-analysis; SMD, standardized mean difference.

Multi-arm trials come with a particular challenge for the analysis of NMA. In a multi-arm trial, the pairwise contrasts involve the same treatments, thus creating a dependency of the treatment effects and also their variances. The standard contrast-based NMA method accounts for this dependency in an $n$-arm trial by using only the $n - 1$ contrasts to a baseline treatment and deriving the treatment effects of all other contrasts from these baseline contrasts (also called basic parameters). In the framework of the graph-theoretical method, implemented in R package *netmeta*, we had developed an alternative adjustment method by reducing the weights of all contrasts of a multi-arm trial by certain factors.[4–6] This method is based on the observation that the variances in a multi-arm trial in parallel design (and also the variances of the network estimates) build a metric on the set of treatments and therefore obey the triangle inequality. This also follows from the analogy between NMA and electrical circuit theory, where this metric is called the resistance distance.[7–9] In the NMA context, we suggest calling this property variance consistency.

The variance-adjusting method relies on the assumption that the arms within a multi-arm trial are independent (though the contrasts become dependent). However, this assumption is violated for trials with correlated arms, for example, split-body trials. For multi-arm trials with correlated arms, the variance of a contrast is not the sum of the arm-based variances, but comes with a correlation term. This may lead to violations of variance consistency, and the inconsistency of variances may even propagate to the NMA estimates. Knowing the variances of the baseline contrasts is then not sufficient for knowing all variances, and the method so far used in R package *netmeta* did not necessarily lead to correct results or did not even work at all.

Su and Tu (2018) introduced three NMA methods to include studies with correlated treatment arms, a data augmentation approach, an adjusting variance approach, and a reducing weight approach, the latter based on *netmeta*.[10] They did not explicitly discuss variance inconsistency.

We present three examples, one with fictitious data and two with real data, in Section 2 of this article. Another fictitious example is given in the online Supporting Information on Zenodo, see Data Availability Statement. In Section 3, we first clarify the correspondence between three-arm trials and triangles and then describe the specialties of trials where treatment arms are correlated. We show that variance consistency is not guaranteed for correlated arms and present a method that works also in the case of variance inconsistency. In Section 4, we apply this method, implemented in the R package *netmeta*, to the examples, followed by a discussion in Section 5.

## 2. Data

### 2.1. Fictitious data example

In Table 1. we give a fictitious example of correlated arm data for individual participants. We assume four individuals, each having three different treatments (1, 2, 3) such that each individual $i$ provides three measurements $x_i^{(1)}, x_i^{(2)}, x_i^{(3)}$ of a continuous outcome. To compare these measurements at the individual level, we consider the three pairwise differences $d_i^{(12)} = x_i^{(1)} - x_i^{(2)}, d_i^{(13)} = x_i^{(1)} - x_i^{(3)}$, $d_i^{(23)} = x_i^{(2)} - x_i^{(3)}$. Averaging these at the trial level provides three treatment effects (TE, contrasts) $d^{(12)}, d^{(13)}, d^{(23)}$ with standard errors (SE) and sampling variances as given in Table 1.

### 2.2. Real data example 1

We use data from a 28 week prospective double-blind phase I trial in patients with alopecia universalis investigating hair regrowth where each of 16 patients was treated with each of four treatments in different areas of the scalp (2% tofacitinib, 1% ruxolitinib, clobetasol dipropionate 0.005%, and placebo).[11] We had no access to the individual patient data. Here, we use data of two patients from Table 1 of the publication, here reproduced in Table 2. Treatment started at visit 2 (week 0) and ended at visit 8 (week 12), with two follow-up visits at week 18 and week 24. The outcome is a hair regrowth

***Table 1.*** *A trial with four participants and three correlated arms.*[a]

| | Measurement | | | Contrasts | | |
|---|---|---|---|---|---|---|
| Individual | $x^{(1)}$ | $x^{(2)}$ | $x^{(3)}$ | $d^{(12)}$ | $d^{(13)}$ | $d^{(23)}$ |
| 1 | 0 | −1 | 0 | 1 | 0 | −1 |
| 2 | 0 | 0 | 2 | 0 | −2 | −2 |
| 3 | 0 | 1 | 4 | −1 | −4 | −3 |
| 4 | 0 | −1 | −1 | 1 | 1 | 0 |
| TE | | | | 0.25 | −1.25 | −1.50 |
| SE | | | | 0.553 | 1.280 | 0.745 |
| Variance | | | | 0.306 | 1.639 | 0.556 |

[a] TE = treatment effect; SE = standard error.

score with integer values from 1 (worst) to 4 (best). We added the average score from visit 2 to visit 10 to the table (last column).

### 2.3. *Real data example 2*

The second real data example is a small NMA provided by a pseudonymous user of R package *netmeta*.[1] We use the last, corrected, version of this data set that is given in Table 3. The effect measure is Cohen's d with a pooled standard deviation. This would guarantee variance consistency for multi-arm trials in a parallel-arm design, but not in a design with correlated arms.

## 3. Methods

First, we consider trials with independent arms and refer to a geometric analogy between trials and triangles that has been pointed out by Lu and Ades.[12]

### 3.1. *Independent arms*

Consider a three-arm trial comparing treatments A, B, C in a parallel-arm design and denote by $\mathrm{SE}_{AB}$, $\mathrm{SE}_{AC}$, and $\mathrm{SE}_{BC}$ the standard errors of the three possible comparisons. We construct the triangle built using these standard errors as side lengths. If the three arms include different individuals, they are independent, and the variance of each contrast is the sum of the variances of two arm-based responses, for example, for treatments $A$ and $B$,

$$\mathrm{SE}_{AB}^2 = s_A^2 + s_B^2, \tag{1}$$

where $s_A^2$ and $s_B^2$ denote the sampling variances of the arm-based responses $y_A$ and $y_B$, which could be means or log odds.

We now consider (without loss of generality) the angle $\gamma$ between the AC and BC edges. Applying the law of cosines to $\gamma$ we obtain

$$\cos\gamma = \frac{\mathrm{SE}_{AC}^2 + \mathrm{SE}_{BC}^2 - \mathrm{SE}_{AB}^2}{2\,\mathrm{SE}_{AC}\,\mathrm{SE}_{BC}} \tag{2}$$

---

[1] https://github.com/guido-s/netmeta/issues/17.

***Table 2.*** *Data from two patients on four treatments for alopecia from Bokhari and Sinclair (two participants).*[11][a]

| Patient | Treatment | Visit 1 Screening | Visit 2 Week 0 | Visit 3 Week 2 | Visit 4 Week 4 | Visit 5 Week 6 | Visit 6 Week 8 | Visit 7 Week 10 | Visit 8 Week 12 | Visit 9 Week 18 | Visit 10 Week 24 | Average (visits 2 to 10) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| P1 | Ruxolitinib | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 4 | 3 | 1 | 1.625 |
| | Tofacitinib | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 4 | 2 | 1 | 1.500 |
| | Clobetasol | 1 | 2 | 1 | 1 | 3 | 3 | 3 | 4 | 2 | 1 | 2.250 |
| | Placebo | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 4 | 2 | 1 | 1.875 |
| P2 | Ruxolitinib | 1 | 1 | 1 | 2 | 2 | 2 | 1 | 3 | 3 | 3 | 2.125 |
| | Tofacitinib | 1 | 1 | 1 | 2 | 2 | 2 | 1 | 3 | 3 | 3 | 2.125 |
| | Clobetasol | 1 | 1 | 1 | 2 | 2 | 2 | 1 | 4 | 3 | 3 | 2.250 |
| | Placebo | 1 | 1 | 1 | 2 | 2 | 2 | 1 | 3 | 3 | 3 | 2.125 |

[a] The entries correspond to a four-point categorical hair regrowth scale where 1 is worst and 4 is best.

***Table 3.*** *An NMA of three studies. Study Batch10 has correlated arms.*

| Study | trt1 | trt2 | smd | se.smd |
|---|---|---|---|---|
| Batch7 | A | Control | −1.8305545 | 0.3316958 |
| Batch9 | B | Control | −0.8826485 | 0.2240514 |
| Batch10 | B | C | 1.2587150 | 0.2132226 |
| Batch10 | B | Control | −0.6531304 | 0.1700709 |
| Batch10 | C | Control | −1.9118454 | 0.2739463 |

which we may rewrite using (1) as

$$\cos\gamma = \frac{s_A^2 + s_C^2 + s_B^2 + s_C^2 - s_A^2 - s_B^2}{2\sqrt{(s_A^2 + s_C^2)(s_B^2 + s_C^2)}} = \frac{s_C^2}{\sqrt{(s_A^2 + s_C^2)(s_B^2 + s_C^2)}}.$$

Due to the independence of the arm-based responses the covariance between two contrasts $d_{AC} = y_A - y_C$ and $d_{BC} = y_B - y_C$ is given by

$$\mathrm{Cov}(d_{AC}, d_{BC}) = \mathrm{Cov}(y_C, y_C) = \mathrm{Var}(y_C) = s_C^2.$$

Thus we have

$$\cos\gamma = \frac{\mathrm{Cov}(d_{AC}, d_{BC})}{\sqrt{\mathrm{Var}(d_{AC})\mathrm{Var}(d_{BC})}} = \mathrm{Cor}(d_{AC}, d_{BC}).$$

Note that because of $s_C^2 \geq 0$, it follows $\cos\gamma \geq 0$ (that is, $0 \leq \gamma \leq 90°$). This means that all angles in our triangle are acute. Using (2) we also see that, equivalently,

$$\mathrm{SE}_{AC}^2 + \mathrm{SE}_{BC}^2 > \mathrm{SE}_{AB}^2 \tag{3}$$

holds which is the triangle inequality for the variances. We note that another triangle inequality is also valid for the standard errors. We can use equation (3) to see that

$$\mathrm{SE}_{AC}^2 + \mathrm{SE}_{BC}^2 + 2\mathrm{SE}_{AC}\mathrm{SE}_{BC} > \mathrm{SE}_{AB}^2 + 2\mathrm{SE}_{AC}\mathrm{SE}_{BC} > \mathrm{SE}_{AB}^2 \tag{4}$$

$$(\mathrm{SE}_{AC} + \mathrm{SE}_{BC})^2 > \mathrm{SE}_{AB}^2 \tag{5}$$

$$\mathrm{SE}_{AC} + \mathrm{SE}_{BC} > \mathrm{SE}_{AB}. \tag{6}$$

This confirms that a three-arm trial can be represented by a triangle with the standard errors as side lengths, as noted by Lu and Ades who referred to the relation in (6) as second-order consistency.[12]

For a summary, see Box 1.

**BOX 1 Trials and triangles**

A trial with three independent arms corresponds to an acute-angled triangle, the side lengths representing the standard errors of the contrasts and the angles representing the (non-negative) correlations between contrasts. The triangle inequality holds for both the standard errors and the variances. The same is true for every subset of three arms of a multi-arm trial in general. For example, a four-arm trial could be represented by a tetrahedron, with the four subsets of three arms building the faces of the tetrahedron, as shown in Figure 1.

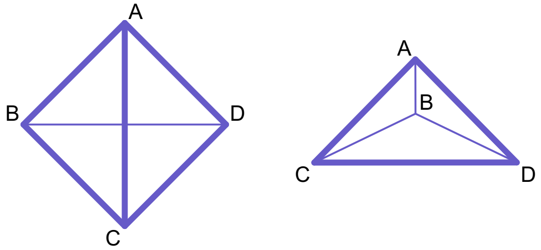

***Figure 1.*** *Visualizations of a four-arm trial as a tetrahedron.*

*Note:* Vertices represent the treatments, edges the comparisons, the four faces the four three-arm subtrials. Left panel: All faces acute. Right panel: All faces obtuse.

### *3.1.1. Adjusting the standard errors of multi-arm trials in NMA*

In NMA, it is necessary to account for the within-trial correlation of contrasts in a multi-arm trial. This can be done in two ways. The standard method is to represent an $n$-arm trial by $n-1$ basic parameters and their covariance matrix. An alternative method, implemented in the R package *netmeta*, is to use all $\binom{n}{2} = n(n-1)/2$ comparisons and adjust (that is, inflate) their variances by certain factors in a way that corresponds to reversing the NMA process for this trial. Thus, when including the trial in an NMA, the inflated variances are reduced to their original values.[4] For trials with independent arms, it has been shown that both methods are equivalent.[5]

## ***3.2. Trials with correlated arms***

If the treatments are not provided to different individuals in independent arms, but each individual receives all three (or more) treatments (e.g., in a cross-over trial), the above-mentioned conditions of independence are not fulfilled. There are a number of different terms for this type of trial, for example, split-mouth trial (in dentistry), split-body trial (frequent in dermatology), intra-individual/intra-subject trial, or within-person/within-patient/within-subject trial, self-control trial, or non-parallel arm trial. In many of these trials, the unit of randomization is a part of the body, for example, a tooth, an eye, or a lesion, such that randomization can take place within an individual.[13] We want to investigate the validity of the triangle inequality for the variances of the comparisons if arms are correlated.

### *3.2.1. Methods of analysis*

The optimal method to analyze a single multi-arm trial with correlated arms is a mixed effects model with fixed effects for the treatments and a random effect for individuals. Such a model does not use the pairwise contrast variances, but produces identical variances for all contrasts to a given baseline such that the resulting triangle becomes equilateral. This is not a problem for including the results in an NMA. It is different if pairwise contrasts are analyzed separately within the patients. We assume a multi-arm trial with at least three treatments $(i, j, k, \dots)$. Generalizing (1), we have for the standard errors of the contrasts

$$\mathrm{SE}_{ij}^2 = s_i^2 + s_j^2 - 2\rho_{ij} s_i s_j \tag{7}$$

$$\mathrm{SE}_{ik}^2 = s_i^2 + s_k^2 - 2\rho_{ik} s_i s_k \tag{8}$$

$$\mathrm{SE}_{jk}^2 = s_j^2 + s_k^2 - 2\rho_{jk} s_j s_k \tag{9}$$

with certain correlation coefficients $\rho_{ij}, \rho_{ik}, \rho_{jk}$. For the covariance of a pair of contrasts with a treatment in common we find

$$\mathrm{Cov}(d_{ij}, d_{ik}) = s_i^2 - \rho_{ij} s_i s_j - \rho_{ik} s_i s_k + \rho_{jk} s_j s_k$$

and for the covariance of a pair of contrasts without a treatment in common (possible if there are at least four treatments) we have

$$\mathrm{Cov}(d_{ij}, d_{kl}) = \rho_{ik} s_i s_k - \rho_{il} s_i s_l - \rho_{jk} s_j s_k + \rho_{jl} s_j s_l.$$

Combining equations (7) to (9) in a suitable way—for example, adding equations (7) and (8) and subtracting equation (9)—provides

$$\mathrm{Cov}(d_{ij}, d_{ik}) = (\mathrm{SE}_{ij}^2 + \mathrm{SE}_{ik}^2 - \mathrm{SE}_{jk}^2)\,/\,2 \tag{10}$$

$$\mathrm{Cov}(d_{ij}, d_{kl}) = (\mathrm{SE}_{il}^2 - \mathrm{SE}_{ik}^2 + \mathrm{SE}_{jk}^2 - \mathrm{SE}_{jl}^2)\,/\,2 \tag{11}$$

meaning that all covariances and correlations between contrasts, also for non-adjacent edges, are determined by the contrast-based variances. This is plausible when thinking of the triangle analogy: all angles are determined by the side lengths of all edges.

While NMA of independent arms always leads to variance consistency, corresponding to acute-angled triangles for all subsets of three trials, this is not necessarily the case for correlated arms, as we will see in our examples. The variance inflation method, implemented in R package *netmeta*, is based on the assumption of independent arms (corresponding to acute triangles) and does not work for a trial that corresponds to an obtuse-angled triangle. In this case, it results in a negative estimate for an arm-based variance, leading to an error message. We cannot ignore the off-diagonal elements of the covariance matrix for trials with correlated arms. The right panel of Figure 1, showing a tetrahedron with all faces obtuse, represents a four-arm trial with correlated arms.

### *3.2.2. The rank of the covariance matrix*

Let us denote the variance-covariance matrix, as derived using (10) and (11), by **C**. As defined here, **C** is fully determined by the $\binom{n}{2}$ standard errors (or variances) of all comparisons in a multi-arm trial with $n$ treatments. **C** is also, alternatively, determined by the variances of all $n-1$ comparisons to an arbitrarily chosen baseline treatment plus the $\binom{n-1}{2}$ covariances between them, likewise resulting in $\binom{n}{2}$ entries, see Appendix A for a proof that the covariance matrix has rank at most $n-1$.

This can also be illustrated using our geometric analogy. We may think of the baseline treatment as a node of the geometric polyhedron (for example, a tetrahedron for a four-arm trial). If we know (i) all side lengths from this baseline node to its $n-1$ neighbors (corresponding to $n-1$ standard errors), and (ii) all $\binom{n-1}{2}$ angles between each pair of these edges (corresponding to the correlations), the polyhedron is completely determined.

### *3.2.3. A method for correlated arms*

For multi-arm trials with independent arms, we may still use the variance inflation method described above to obtain a diagonal matrix **W** with adjusted weights. For multi-arm trials with potentially correlated arms, we need a more general method. Instead of adjusting, we use the Moore–Penrose generalized inverse (also known as pseudoinverse) of the covariance matrix **C**,[14]

$$\mathbf{W} = \mathbf{C}^{+}$$

which is not a diagonal matrix. We then apply the graph-theoretical method as usual, but with a block-diagonal matrix consisting of blocks such as **W** for multi-arm trials, instead of a diagonal matrix of weights, see Appendix B. In Appendix C, we discuss the special case where the rank is less than $n-1$.

If we enter the full covariance matrix for a multi-arm trial into the NMA equations, the effective variances and covariances of this trial agree with the original variances and covariances. Correctly, the variances are neither inflated beforehand, nor reduced afterwards by the NMA: providing the full covariance matrix replaces the adjustment for multi-arm studies. We have implemented the more general method in R package *netmeta*, version 3.2-0, to allow including multi-arm trials with correlated

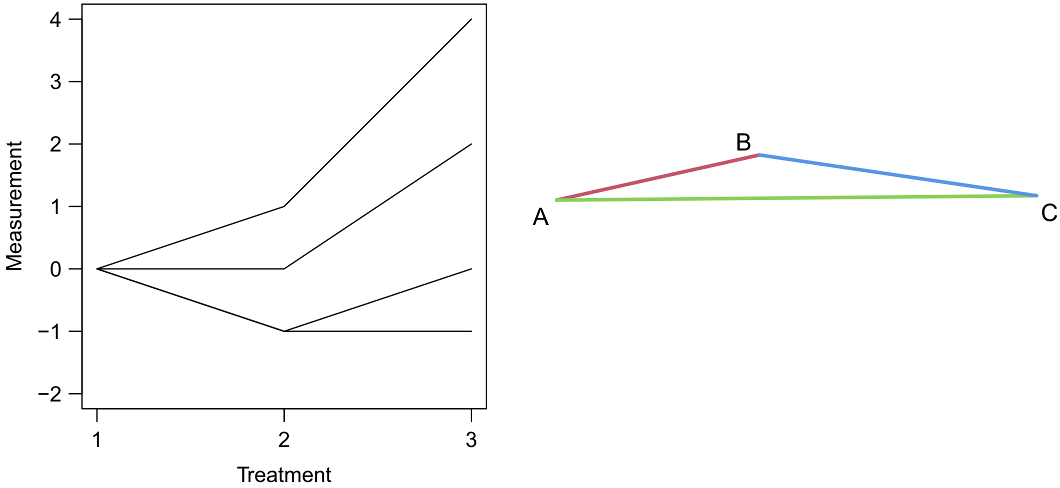

***Figure 2.*** *Visualization of the data in Table 1.*

*Note:* Left: Individual participant data, represented as a line chart. Right: Obtuse-angled triangle, side lengths representing the standard errors.

arms, even in case of variance inconsistency. In the R script in the Supporting Information on Zenodo, we demonstrate the use of the new argument "correlated."

## 4. Results

We apply the method to our examples.

### *4.1. Fictitious data example*

We see that the variances in Table 1 violate the triangle inequality, as $1.639 > 0.306 + 0.556$. This is visualized in the left panel of Figure 2: the variance of comparison 1 vs. 3 is greater than the sum of the variances of the comparisons 1 vs. 2 and 2 vs. 3. The right panel shows the obtuse triangle that belongs to these data. An NMA reproduces the results in Table 1. R code for analyzing this data set and creating Figure 2 is provided as Supporting Information on Zenodo.

### *4.2. Real data example 1*

We consider the hair regrowth score, averaged across visits 2 to 10 (that is, all visits after treatment onset), given in the last column of Table 2. To demonstrate that the phenomenon of inconsistent variances exists in these data, we build means of pairwise score differences between treatments within each patient. The results are given in Table 4. As there are four treatments, we have six pairwise comparisons. The last three columns give the mean difference, its standard error, and the sampling variance for each pair of treatments. It is easily checked that the triangle inequality is violated for the variances of each of the four possible triangles. Moreover, all triangles are degenerate in the sense that all nodes are lying in one line. For example, consider the triangle Tofacitinib–Clobetasol–Placebo: the sum of the last two standard errors in Table 4 gives the standard error of the fourth comparison. See Appendix C and the R code provided in the Supporting Information on Zenodo. Figure 3 shows the results as a forest plot, suggesting a slight superiority of clobetasol over placebo, but no other differences between treatments.

***Table 4.*** *Pairwise within-patient differences with mean and standard error. Data from Bokhari and Sinclair (two participants).*[11]

| Comparison | Patient | Difference | MD[a] | SE[b] | Variance |
|---|---|---|---|---|---|
| Ruxolitinib - Tofacitinib | P1 | 0.125 | | | |
| Ruxolitinib - Tofacitinib | P2 | 0.000 | 0.0625 | 0.0625 | 0.0039 |
| Ruxolitinib - Clobetasol | P1 | −0.625 | | | |
| Ruxolitinib - Clobetasol | P2 | −0.125 | −0.3750 | 0.2500 | 0.0625 |
| Ruxolitinib - Placebo | P1 | −0.250 | | | |
| Ruxolitinib - Placebo | P2 | 0.000 | −0.1250 | 0.1250 | 0.0156 |
| Tofacitinib - Clobetasol | P1 | −0.750 | | | |
| Tofacitinib - Clobetasol | P2 | −0.125 | −0.4375 | 0.3125 | 0.0977 |
| Tofacitinib - Placebo | P1 | −0.375 | | | |
| Tofacitinib - Placebo | P2 | 0.000 | −0.1875 | 0.1875 | 0.0352 |
| Clobetasol - Placebo | P1 | 0.375 | | | |
| Clobetasol - Placebo | P2 | 0.125 | 0.2500 | 0.1250 | 0.0156 |

[a] MD = mean difference.
[b] SE = standard error.

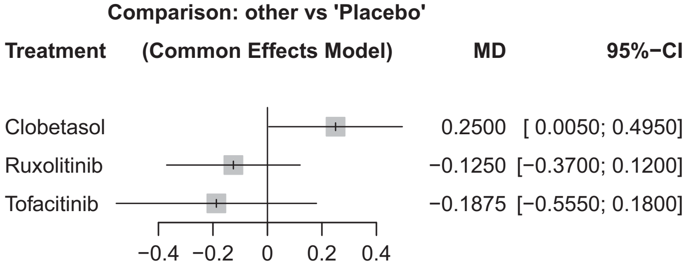

***Figure 3.*** *Forest plot of the hair growth trial results.*

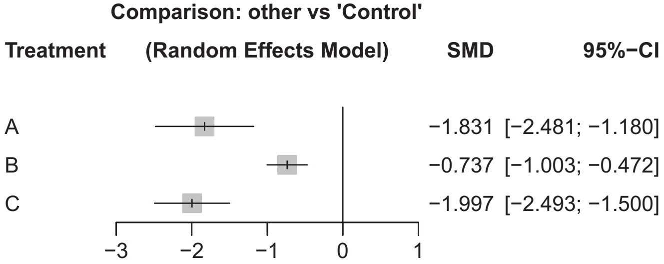

***Figure 4.*** *Forest plot of the second real data example, produced with netmeta.*

### *4.3. Real data example 2*

We analyzed the second real data example using two methods, our method and the method implemented in the R package *metafor*.[15] We show the results for the random effects model as standardized mean differences (SMD) compared to treatment “Control” in Figures 4 (*netmeta*) and 5 (*metafor*). They are in agreement. The full R code, also for generating the covariance matrices, is provided in the supporting information for this article on Zenodo.

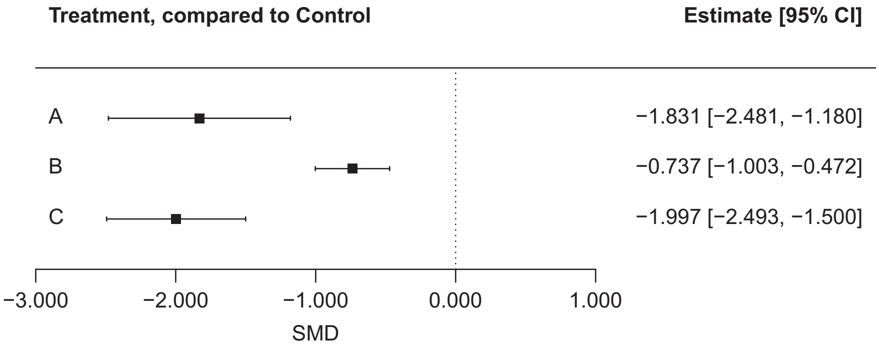

***Figure 5.*** *Forest plot of the second real data example, produced with metafor.*

## 5. Discussion

This article was primarily inspired by a message from a pseudonymous user of the R package *netmeta* who reported a negative treatment arm variance for a three-arm trial, see our real data example 2. In our past experience, negative variances in multi-arm trials were typically caused by an inadequate variance imputation or data errors. However, in this case, we identified the correlation of treatment arms as the culprit. After finding a mathematical solution for correlated arm data which is described in Section 3.2, we accordingly extended R package *netmeta*. Users can now easily specify trials with correlated arms as "correlated" to obtain correct results.

For the usual case of multi-arm studies with independent arms we know three approaches: the standard approach (i), using only $n-1$ baseline contrasts (basic parameters) for an $n$-arm study, leading to a full-rank $(n-1) \times (n-1)$ covariance matrix; the variance-adjusting approach (ii) from *netmeta*, using a $\binom{n}{2} \times \binom{n}{2}$ diagonal matrix of adjusted weights for all contrasts; and the approach introduced in this article (iii), likewise using all contrasts, but, instead of a diagonal matrix of adjusted weights, utilizing the full covariance matrix and its pseudoinverse. For independent arms, all three approaches are equivalent. In particular, this means that our new approach is completely general. To demonstrate this, we provide an example in the R script in the Supporting Information of this article on Zenodo.

For the case discussed in this article (correlated arms), the standard approach (i) works, but to provide a full correlation matrix more information is needed than only the basic contrasts. The variance-adjusting approach (ii) does not work, which is why we introduced the novel approach (iii) that replaces the diagonal matrix with the full covariance matrix of all contrasts.

Technically, treatment effects and standard errors of all pairwise comparisons of a multi-arm trial are used in *netmeta*. It is not necessary to know arm-based variances $s_i^2, s_j^2, s_k^2$ or any correlations between arms $(\rho_{ij}, \rho_{ik}, \rho_{jk})$ for deriving equations (10) and (11) that provide the full covariance matrix also for the correlated case. Particularly, it is not necessary to impute "typical" values for correlations such as 0.25 or 0.5, or to assume a constant correlation. Notably, this also holds for multi-arm studies with more than three arms. However, it is rarely the case that variances are presented for all comparisons. Missing variances can, in principle, be derived from known correlations, but correlations are provided in primary studies even less frequently. We point out that this type of inconsistency is disguised if only the contrasts to a chosen baseline are used in the analysis. Therefore, it is due to the special approach of *netmeta* that inconsistency of variances within a trial becomes apparent.

If present within one or more multi-arm trials in an NMA, variance inconsistency may also propagate to the NMA, at least in theory. At present, we are not aware of a real-data example where this happened. Instead, we provide a toy example (an NMA with four three-arm trials) in Appendix D and the R script in the Supporting Information of this article on Zenodo.

Su and Tu (2018) provided three methods for imputing unreported correlations that could be valuable.[10] Their reducing weight approach is based on *netmeta*. Su and Tu presented their data examples in an arm-based format, such that it depends on the observed (or assumed) correlations whether variances become inconsistent or not. Su and Tu did not encounter variance inconsistency in any of their examples. In a recent publication, Liu et al. (2025) discuss within-study correlations in NMA, but do not seem to refer to correlated arms.[16]

Examples of inconsistent variances appear to occur rarely or are at least rarely reported. We found it difficult to find a "good" real data example: the pseudonymous user's example has, to the best of our knowledge, not been published, and the hair growth example is very artificial, as discussed in Section 2.3 and Appendix C. The main reason for the apparent lack of examples is that most authors report results of studies with correlated arms not contrast-wise, but arm-wise and without providing information on correlations. An exception is Table 2 in Farook et al. who used a mixed model, and there was no inconsistency.[17]

A reviewer pointed out that another reason for inconsistent variances might be the use of the SMD as an effect measure. While this holds for Hedges' g with different group sample sizes or standard deviations for the contrasts, it is not a problem for Cohen's d with a single pooled standard deviation for all contrasts, as implemented in function `pairwise()` of R package *meta* which is used in a typical workflow of *netmeta*. The real data example 2 uses Cohen's d with a single pooled standard deviation for the three-arm study.

The estimation method in *netmeta* originated from the analogy between NMA and electrical circuits.[4] For electrical circuits, the triangle inequality for resistances, which corresponds to variance consistency, is always fulfilled. As we have seen, this does not extend to NMA in general: it can be violated if the treatment arms are correlated. We know of no analog in physics for correlated treatment arms. This makes our adjusting variance (or reducing weights) approach inappropriate for the examples described in this article. However, our new method, implemented in R package *netmeta*, version 3.2-0, overcomes this limitation and can be recommended for the NMA of multi-arm studies with correlated data.

**Acknowledgements.** We thank the pseudonymous user of *netmeta* who made us aware of the problem with correlated arms.

**Author contributions.** G.R. developed the methods, wrote the first draft, and revised the manuscript. G.S. implemented the methods and revised the manuscript. Both authors read and approved the final version of the manuscript.

**Competing interest statement.** The authors declare that no competing interests exist.

**Data availability statement.** The R script and datasets supporting this work are openly available on Zenodo at https://doi.org/10.5281/zenodo.15707313.

**Funding statement.** The authors declare that no specific funding has been received for this article.

## References

[1] Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Res. Synth. Methods* 2012;3(2): 80–97.

[2] Ades AE, Welton NJ, Dias S, Phillippo DM, Caldwell DM. Twenty years of network meta-analysis: Continuing controversies and recent developments. *Res. Synth. Methods* 2024;15(5): 702–727. https://doi.org/10.1002/jrsm.1700.

[3] White IR, Turner RM, Karahalios A, Salanti G. A comparison of arm-based and contrast-based models for network meta-analysis. *Stat. Med.* 2019;38(27): 5197–5213. https://doi.org/10.1002/sim.8360.

[4] Rücker G. Network meta-analysis, electrical networks and graph theory. *Res. Synth. Methods* 2012;3(4): 312–324.

[5] Rücker G, Schwarzer G. Reduce dimension or reduce weights? Comparing two approaches to multi-arm studies in network meta-analysis. *Stat. Med.* 2014;33: 4353–4369. https://doi.org/10.1002/sim.6236.

[6] Balduzzi S, Rücker G, Nikolakopoulou A, et al. netmeta: An R package for network meta-analysis using frequentist methods. *J. Stat. Softw.* 2023;106(2): 1–40.

[7] Klein DJ, Randić M. Resistance distance. *J. Math. Chem.* 1993;12(1–4): 81–95.

[8] Balaban A, Klein D. Co-authorship, rational Erdos numbers, and resistance distances in graphs. *Scientometrics* 2002;55(1): 59–70.

[9] Xiao WJ, Gutman I. Resistance distance and Laplacian spectrum. *Theor. Chem. Acc.* 2003;110(4): 284–289.

[10] Su YX, Tu YK. Statistical approaches to adjusting weights for dependent arms in network meta-analysis. *Res. Synth. Methods* 2018;9(3): 431–440.

[11] Bokhari L, Sinclair R. Treatment of alopecia universalis with topical Janus kinase inhibitors—A double blind, placebo, and active controlled pilot study. *Int. J. Dermatol.* 2018;57(12): 1464–1470.

[12] Lu G, Ades AE. Modeling between-trial variance structure in mixed treatment comparisons. *Biostatistics* 2009;10(4): 792–805.

[13] Leducq S, Dugard A, Allemang-Trivalle A, Giraudeau B, Maruani A. Design and methodological issues of within-person (split-body) randomized controlled trials evaluating a topical treatment: A systematic review. *Dermatology* 2023;239(5): 720–731. https://doi.org/10.1159/000530149

[14] Albert AE. *Regression and the Moore-Penrose Pseudoinverse*, Mathematics in Science and Engineering. Academic Press; 1972. ISBN: 0-12-048450-1.

[15] Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* 2010;36(3): 1–48. https://doi.org/10.18637/jss.v036.i03.

[16] Liu YL, Zhang B, Chu H, Chen Y. Network meta-analysis made simple: A composite likelihood approach. *Res. Synth. Methods*. 2025;16(2): 272–290. https://doi.org/10.1017/rsm.2024.12.

[17] Farook F, Alrumi A, Aldalaan K, et al. The efficacy of manual toothbrushes in patients with fixed orthodontic appliances: a randomized clinical trial. *BMC Oral Health* 2023;23(1): 315.

## Appendix

In the Appendix, we give additional details.

### A. The rank of the covariance matrix

We choose treatment $i = 1$ as baseline. That is, we assume to know $\text{SE}^2_{1j}$ for all $j = 2, \dots, n$ and also, see (10), $\text{Cov}(d_{1j}, d_{1k}) = (\text{SE}^2_{1j} + \text{SE}^2_{1k} - \text{SE}^2_{jk})/2$ for all $j, k = 2, \dots, n$, building an $(n-1) \times (n-1)$ submatrix of $\mathbf{C}$ with $\binom{n}{2}$ potentially different entries.

We show that any other row (or column) of $\mathbf{C}$ is obtained as a linear combination of the first $n-1$ columns (rows). Consider, for example, entry $((1j), (kl))$. We have by (11)

$$\begin{aligned}\text{Cov}(d_{1j}, d_{kl}) &= (\text{SE}^2_{1l} - \text{SE}^2_{1k} + \text{SE}^2_{jk} - \text{SE}^2_{jl})/2\\ &= (\text{SE}^2_{1l} - \text{SE}^2_{jl})/2 - (\text{SE}^2_{1k} - \text{SE}^2_{jk})/2\\ &= (\text{SE}^2_{1j} + \text{SE}^2_{1l} - \text{SE}^2_{jl})/2 - (\text{SE}^2_{1j} + \text{SE}^2_{1k} - \text{SE}^2_{jk})/2\\ &= \text{Cov}(d_{1j}, d_{1l}) - \text{Cov}(d_{1j}, d_{1k})\end{aligned}$$

which are terms from the top left submatrix, assumed as known. Thus, starting from the top left submatrix, we can derive all entries in the top rows of the remaining columns and in the leftmost columns of the remaining rows. Finally, by the same process, we also can fill in all remaining entries. This result means that the variance-covariance matrix $\mathbf{C}$ has a rank of at most $n-1$.

### B. Estimation method for the non-degenerate case

Restricting the consideration to the multi-arm trial of interest, we consider an $n$-arm trial with $m = \binom{n}{2}$ treatment effects (contrasts). First, we define the edge-vertex incidence matrix $\mathbf{B}$ with $m = \binom{n}{2}$ rows, corresponding to ordered comparisons, and $n$ columns, corresponding to the treatments. In the row corresponding to comparison $(i, j)$, $i < j$, $\mathbf{B}$ has an entry 1 in column $i$ and $-1$ in column $j$; all other entries are zero.

Based on $\mathbf{B}$ and $\mathbf{W} = \mathbf{C}^+$ we obtain the hat matrix $\mathbf{H}$ by

$$\mathbf{H} = \mathbf{B}(\mathbf{B}^\top \mathbf{W} \mathbf{B})^+ \mathbf{B}^\top \mathbf{W}.$$

Treatment effects within a multi-arm trial are consistent and therefore linearly dependent. Let $\mathbb{S}$ be the $(n-1)$-dimensional consistent subspace of $\mathbb{R}^m$ which is spanned by the columns of $\mathbf{B}$.[5] $\mathbf{H}$ is the projection onto $\mathbb{S}$. If $\mathbf{C}$ and $\mathbf{W} = \mathbf{C}^+$ have maximal rank $n-1$, $\mathbf{H}$ has the same rank and is the projection onto $\mathbb{S}$.

## C. The degenerate case

The hair growth example shows that the covariance matrix, and thus the Laplacian matrix, its pseudoinverse, and $\mathbf{H}$ can have rank less than $n-1$, in this case even 1. We call this the degenerate case. It can be shown that this may happen if we have very few observations (such as only two patients in the hair growth data). It turns out that for rank 1 the covariance matrix has the form

$$\mathbf{C} = \mathbf{s}\,\mathbf{s}^\top,$$

where $\mathbf{s}$ is the vector of standard errors. $\mathbf{C}$ has rank 1 because all columns (or all rows) are by definition proportional to each other, i.e., to $\mathbf{s}$. If the rank is 1, all eigenvalues except the first (largest) must be zero. The first eigenvalue therefore equals the sum of eigenvalues, which is the trace of the matrix, $\lambda_1 = \sum s_i^2$. The principal eigenvector is $\mathbf{s}$ due to

$$\mathbf{C}\mathbf{s} = \mathbf{s}\,\mathbf{s}^\top\mathbf{s} = \sum s_i^2\,\mathbf{s} = \lambda_1\mathbf{s}.$$

The matrix $\mathbf{C}^+$ and the hat matrix $\mathbf{H}$, derived using $\mathbf{C}^+$, have the same principal eigenvector as $\mathbf{C}$. Therefore, the vector of observed treatment effects is a fixed point of $\mathbf{H}$ only if it is proportional to $\mathbf{s}$, the vector of standard errors (where deviations of some signs are possible, depending on the direction of the edges). If this is not the case, $\mathbf{H}$ does not map the consistent treatment effects onto themselves.

Technically, this means that we need an adapted estimation method for the degenerate case. To obtain the correct treatment effect estimates, our implementation uses $\mathbf{W} = (\mathrm{diag}(\mathbf{s}))^{-1}$ instead of $\mathbf{W} = \mathbf{C}^+$, while still using $\mathbf{C}$ to derive the standard errors.

## D. A network meta-analysis with inconsistent variances

This fictitious example demonstrates that variance inconsistency can propagate from individual trials to NMA. We assume a network of four treatments A, B, C, D based on four three-arm trials. The data are given in Table D1.

We can easily confirm that all trials in Table D1 except ACD have inconsistent variances (column SE$^2$), as $1 + 1 < 2.89$; $1 + 42.25 < 47.61$. The last column shows the variances of the resulting NMA estimates, and they are also inconsistent: $0.364 + 0.364 < 0.754$; $0.364 + 13.502 < 14.019$. R code to reproduce this result is found in the Supporting information on Zenodo.

***Table D1.*** *Fictitious data of an NMA of four three-arm studies, all with correlated arms.*

| Study | Treat1 | Treat2 | TE | SE | $SE^2$ | NMA variance |
|---|---|---|---|---|---|---|
| ABC | A | B | 0 | 1.0 | 1.00 | 0.364 |
| | A | C | 0 | 1.7 | 2.89 | 0.754 |
| | B | C | 0 | 1.0 | 1.00 | 0.364 |
| ABD | A | B | 0 | 1.0 | 1.00 | 0.364 |
| | A | D | 0 | 6.5 | 42.25 | 13.502 |
| | B | D | 0 | 6.9 | 47.61 | 14.019 |
| ACD | A | C | 0 | 1.7 | 2.89 | 0.754 |
| | A | D | 0 | 6.5 | 42.25 | 13.502 |
| | C | D | 0 | 6.5 | 42.25 | 13.502 |
| BCD | B | C | 0 | 1.0 | 1.00 | 0.364 |
| | B | D | 0 | 6.9 | 47.61 | 14.019 |
| | C | D | 0 | 6.5 | 42.25 | 13.502 |

*Note*: TE = treatment effect; SE = standard error; NMA variance = variance of the NMA estimate.

**Cite this article:** Rücker, G, Schwarzer, G Trials and triangles: Network meta-analysis of multi-arm trials with correlated arms. *Research Synthesis Methods*. 2025: 1–14. https://doi.org/10.1017/rsm.2025.10026