Algorithmic parameterization of mixed treatment comparisons

**Gert van Valkenhoef · Tommi Tervonen ·
Bert de Brock · Hans Hillege**

Abstract Mixed Treatment Comparisons (MTCs) enable the simultaneous meta-analysis (data pooling) of networks of clinical trials comparing ≥ 2 alternative treatments. Inconsistency models are critical in MTC to assess the overall consistency between evidence sources. Only in the absence of considerable inconsistency can the results of an MTC (consistency) model be trusted. However, inconsistency model specification is non-trivial when multi-arm trials are present in the evidence structure. In this paper, we define the parameterization problem for inconsistency models in mathematical terms and provide an algorithm for the generation of inconsistency models. We evaluate running-time of the algorithm by generating models for 15 published evidence structures.

G. van Valkenhoef (✉) · H. Hillege
Dept. of Epidemiology, University Medical Center Groningen, P.O. Box 30.001, 9700 RB Groningen, The Netherlands
e-mail: g.h.m.van.valkenhoef@rug.nl

H. Hillege
e-mail: j.l.hillege@epi.umcg.nl

G. van Valkenhoef · B. de Brock
Faculty of Economics and Business, University of Groningen, Groningen, The Netherlands

B. de Brock
e-mail: e.o.de.brock@rug.nl

T. Tervonen
Econometric Institute, Erasmus University Rotterdam, Rotterdam, The Netherlands
e-mail: tervonen@ese.eur.nl

1 Introduction

Meta-analysis refers to statistical methods that summarize evidence from multiple studies (most commonly: clinical trials). Traditional meta-analysis (Hedges and Vevea 1998; Normand 1999) has focused on pairwise comparisons of treatments based upon summary measures of relative effect as reported in the original studies. Several models to simultaneously compare more than two treatments have recently appeared (Sutton and Higgins 2008), also leading to reported applications of the methodology (see Salanti et al. 2008b). Such simultaneous comparisons are called Mixed Treatment Comparisons (MTCs), or network meta-analyses. MTCs allow for the use of both direct and indirect evidence for comparisons, and to calculate the rank-probabilities of a set of alternative treatments with regard to a single evaluation criterion.

An MTC is implemented as a Bayesian hierarchical model and estimated using Markov Chain Monte Carlo (MCMC) simulation (Lu and Ades 2006; Salanti et al. 2008a). As in pairwise meta-analysis, the goal is to combine evidence from multiple studies in order to derive a best estimate of the relative effect of treatments. MTC extends pairwise meta-analysis by simultaneously estimating the relative effects of all possible pairs of the included treatments. Normally consistency is assumed, i.e., that direct and indirect evidence are in agreement. For example, if we have different studies comparing treatments a versus b, b versus c and a versus c, we add the constraint $d_{ac} = d_{ab} + d_{bc}$

to the model (Lu and Ades 2006), where the d_{xy} are relative effects. This assumption of consistency does not necessarily hold and needs to be tested. To do this, an inconsistency model is formulated by relaxing the consistency constraint by introducing an Inconsistency Factor (ICF): $d_{ac} = d_{ab} + d_{bc} + w_{abca}$. Evidence can only be inconsistent if there are closed loops in the evidence structure (Lumley 2002): the ICF w_{abca} corresponds to the loop $abca$. Evidence consistency can be tested by individually assessing the null hypothesis that $w_C = 0$ for each ICF w_C (Salanti et al. 2008a), and further comparison of consistency and inconsistency models can be based on global goodness of fit (Lu and Ades 2006).

No general formula or algorithm exists for evidence structures with multi-arm trials (trials with three or more arms—i.e., treatment groups) to determine the consistency equations that must be relaxed with ICFs to achieve correct model parameterization (Lu and Ades 2006; Salanti et al. 2008a). In addition, baseline treatments have to be chosen for the individual studies, which can prove to be problematic in the presence of multi-arm trials (Lu and Ades 2006). The absence of an algorithmic solution causes MTC model construction to be error prone and only applicable by experts in Bayesian modeling. Thus, MTC model generation would enable wider adoption of MTC and should allow greater confidence in the correctness of subsequently published MTCs. In this paper, we formally define the model generation problem for MTC inconsistency models and provide an algorithmic solution.

The remainder of the paper is structured as follows. First, an overview of MTC models and a mathematical formulation of their evidence structure is given in Sect. 2. Then, we give a precise definition of the parameterization problem as the problem of finding the spanning tree of the evidence structure that maximizes the number of ICFs, the Inconsistency Degree (ICD), while satisfying the constraint that every relevant parameter must be informed directly by at least one trial (Sect. 3). An algorithmic solution to the problem is given in Sect. 4. We give a detailed example of how parameterization is done in Sect. 5. In Sect. 6, we evaluate the feasibility of our algorithm on a number of published evidence structures. Finally, in Sect. 7, we discuss our results.

2 Mixed treatment comparison models

The Bayesian hierarchical model for an MTC evidence structure is specified following the general formulation in Lu and Ades (2006), which in turn extends that by Higgins and Whitehead (1996). We shall only introduce the concepts that are relevant to the parameterization problem, and refer the interested reader to Lu and Ades (2006) for a full discussion. The evidence structure for any MTC consists of a

number of studies, that together determine an undirected evidence graph in which the treatments are the vertices and the available comparisons are the edges. Since a trial S_i provides evidence for all possible comparisons between the included treatments $T(S_i)$, each study can be considered to provide a fully connected evidence graph $G(S_i) = (T(S_i), E(S_i))$. Here, $E(S_i)$ represents the estimates of relative effects that can be made based on the trial data. So a two-arm trial is a pair, a three-arm trial a triangle, a four-arm trial a fully connected 4-treatment graph, and so on.

Denote by $S = \{S_1, \ldots, S_n\}$ the set of n studies included in the MTC. The evidence graphs $G(S_i)$, $S_i \in S$ form an evidence structure, as illustrated in Figs. 1(a) and 1(b). These figures introduce two hypothetical examples that will be used throughout the paper to illustrate the introduced concepts. Structure I consists of two overlapping three-arm trials and one two-arm trial, while structure II has only one three-arm trial and two two-arm trials. The union of the individual study evidence graphs forms the MTC evidence graph:

Definition 1 (Evidence graph) The graph $G(S)$ of all comparisons made in at least one of the trials in S is defined as:

$$G(S) = (T(S), E(S)) = \left(\bigcup_{S_i \in S} T(S_i), \bigcup_{S_i \in S} E(S_i) \right).$$

For example, the evidence structures I and II have the same evidence graph, shown in Fig. 1(c). A graph corresponding to an MTC problem has to be connected. If it is not, S must be decomposed into two or more independent problems, corresponding to connected subgraphs of $G(S)$ that can be analyzed separately. Given the (connected) evidence graph $G(S)$, every edge in $E(S)$ becomes an effect parameter in the MTC model:

Definition 2 Given the MTC problem S and an arbitrary ordering $\prec$ on $T(S)$ (e.g., alphabetical order of treatments), the set of effect parameters $D(S)$ is given by:

$$D(S) = \{d(\{x, y\}) \mid \{x, y\} \in E(S)\}$$

where $d(\cdot)$ identifies a unique parameter with the set $\{x, y\}$:

$$d(\{x, y\}) = \begin{cases} d_{xy} & \text{if } x \prec y, \\ d_{yx} & \text{if } y \prec x. \end{cases}$$

Furthermore, for directed edges (x, y), we define a signed function, that takes into account the direction of (x, y) relative to the parameter $d(\{x, y\})$:

$$d((x, y)) = \begin{cases} d(\{x, y\}) & \text{if } x \prec y, \\ -d(\{x, y\}) & \text{if } y \prec x. \end{cases}$$
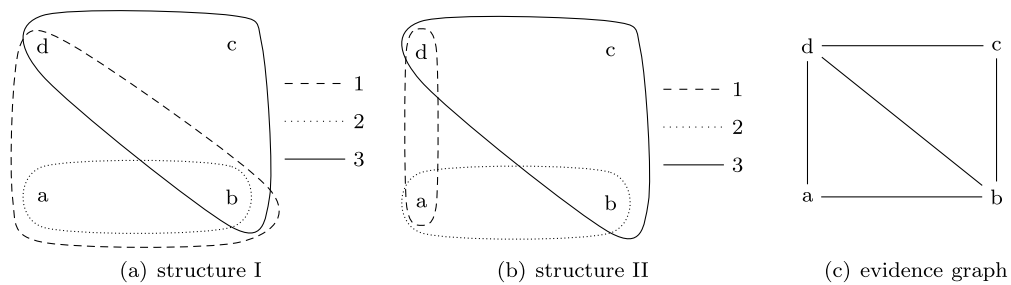
(a) structure I (b) structure II (c) evidence graph

Fig. 1 Two examples of evidence structures used throughout the paper. Evidence structure I (**a**) contains three trials, where trial 1 compares a, b, and d, trial 2 compares a and b, and trial 3 compares b, c, and d. Evidence structure II (**b**) also contains three trials, but differs in that trial 1 compares only a and d. Both structures have an identical evidence graph (**c**)

For example, if we adopt alphabetical ordering for $\prec$, we have

$$d((a, b)) = d_{ab},$$

$$d((b, a)) = -d_{ab},$$

meaning we do not have to worry about the direction in which the evidence graph is traversed.

2.1 Study level effects

We first discuss how the relative effects at the study level are parameterized in terms of the parameters $D(S)$. For each $S_i \in S$, for each treatment $t \in T(S_i)$, we have a certain absolute effect μ_{it}. The way these absolute effects are defined depends on the type of model, and is not important for the current discussion; Salanti et al. (2008a) gives the formulation for both dichotomous and continuous data. Now, because we are interested in the *relative* effects, we choose a *baseline* treatment $x \in T(S_i)$. The baseline effect μ_{ix} is then a random variable for which we assume some prior distribution $\pi(\mu_{ix})$. For every other treatment $u \in T(S_i)$, $u \neq x$ the treatment effect is:

$$\mu_{iu} = \mu_{ix} + \delta_{ixu},$$

where δ_{ixu} is the random effect of treatment u relative to x. The distribution for the random effects is:

$$\begin{pmatrix} \delta_{ixu} \\ \vdots \\ \delta_{ixw} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} d((x, u)) \\ \vdots \\ d((x, w)) \end{pmatrix}, \Sigma \right),$$

where Σ is an appropriately defined variance-covariance matrix. For a full discussion of the study level absolute and relative effects, see Lu and Ades (2006), Salanti et al. (2008a).

From the definition of the relative effects, it is clear that they are transitive (Lu and Ades 2009): if $u, v, w \in T(S_i)$ are distinct treatments, then

$$\delta_{iuv} = \delta_{iuw} + \delta_{iwv}. \tag{1}$$

Based on this we conjecture that if a triangle of treatments is included in precisely the same set of studies, then this relation also holds for the estimates of the effect sizes:

Conjecture 1 (Internal consistency) Write $f((u, v), S')$ for the estimate of $d((u, v))$ based solely on the studies $S' \subset S$. Then if $u, v, w \in T(S_i)$; $\forall S_i \in S'$,

$$f((u, v), S') = f((u, w), S') + f((w, v), S').$$

Note that this implies that each of the studies in S' has at least three arms.

2.2 Consistency models

Normally conclusions are drawn under the assumption of evidence consistency. Basically, this is a generalization of the conjecture in the sense that we assume that it holds regardless of the supporting studies:

Definition 3 (Consistency) Let $u, v, w \in T(S)$ be distinct treatments, then assuming consistency,

$$d((u, v)) = d((u, w)) + d((w, v)).$$

This can be justified by assuming exchangeability of the study level relative effects and taking expectations on both sides of (1) (Lu and Ades 2009). The consistency assumption is essential, as it models the relationships between treatment contrasts and allows the model to borrow strength across the evidence structure (Lu and Ades 2009). More generally, a consistency equation can be written for any cycle and reference effect, as shown by the following lemma and corollary.

Lemma 1 *Given the evidence graph $G(S)$, let (w_1, w_n) be any pair of vertices of $G(S)$ and $p = (w_1, \ldots, w_n)$ a path of length $n - 1$ between them, $n > 2$ (see Appendix). Then,*

under the assumption of consistency,

$$d((w_1, w_n)) = \sum_{i=1}^{n-1} d((w_i, w_{i+1})). \qquad (2)$$

Proof (by induction) If $p = (u, w, v)$, then the lemma is just a restatement of the assumption. Now let the lemma hold for (w_2, w_n) and $p' = (w_2, \ldots, w_n)$, i.e.,

$$d((w_2, w_n)) = \sum_{i=2}^{n-1} d((w_i, w_{i+1})).$$

Then for $p = (w_1, w_2, \ldots, w_n)$, we get

$$d((w_1, w_n)) = d((w_1, w_2)) + d((w_2, w_n))$$
$$= \sum_{i=1}^{n-1} d((w_i, w_{i+1})).$$

The first equality by the consistency assumption and the second by the induction hypothesis. $\qquad \square$

Corollary 1 (Consistency relation) *Given the evidence graph $G(S)$, a (simple) cycle $C \subseteq E(S)$, then if we take any edge $\{u, v\} \in C$, $u \prec v$ as a reference effect, we can write a consistency equation as follows*:

$$d_{uv} = d((u, v)) = \sum_{i=1}^{n-1} d((w_i, w_{i+1})),$$

where $(w_1, \ldots, w_n)$ is the directed path from $w_1 = u$ to $w_n = v$ consisting of the edges $(C - \{u, v\})$.

Thus, in a consistency model, the consistency relation defines d_{uv} completely in terms of the other comparisons in the cycle. In that case, d_{uv} is called a *functional* parameter (Lu and Ades [2006]). For each functional parameter there must be a cycle in which it is the only one, otherwise a circular definition would result. Moreover, each cycle should have at least one functional parameter, or we do not assume full consistency. The right hand side parameters are called *basic* parameters and are defined through suitable distributions. The division of parameters into basic and functional ones is not arbitrary; it has previously been stated that the basic parameters should form a spanning tree (Lu and Ades [2006]). This is proven by the following theorem.

Theorem 1 (Basic parameters) *If we divide the parameter edges $E(S)$ into a set of basic parameters E_b and a set of functional parameters E_f, such that $E_b \cup E_f = E(S)$ and $E_b \cap E_f = \emptyset$, the basic parameters form a spanning tree $G_b = (T_b, E_b)$ of the evidence graph $G(S)$.*

Proof It is sufficient to show (Appendix) that

1. G_b is a connected graph,
2. G_b is acyclic,
3. $T_b = T(S)$.

Proof of 1. Assume G_b is not connected. Then, since $G(S)$ is connected, there is an edge $e \in E(S)$, $e \notin E_b$ that connects two vertices not connected in G_b. Since $e = (u, v)$ does not correspond to a basic parameter, it must be a functional parameter. Hence, there must be a simple directed path from u to v in G_b, and therefore G_b must be connected.

Proof of 2. If $C \subseteq E_b$ is a cycle in G_b, then for an edge $e = (u, v) \in C$, $u \prec v$, Corollary 1 lets us write a consistency equation in terms of the other (basic) parameters in the cycle. Thus, if there would be a cycle in G_b, we would not be assuming full consistency.

Proof of 3. From the proof to the first part and $E_b \cup E_f = E(S)$. $\qquad \square$

Corollary 2 *The functional parameters E_f are the non-tree edges (Appendix) corresponding to the spanning tree G_b, and the (simple) cycle C created by adding $e \in E_f$ to G_b generates a consistency relation, as considered in Corollary 1.*

The theorem and corollary imply that for any valid parameterization of $G(S)$, we will have $\dim(G(S)) = |T(S)| - 1$ basic parameters and $\text{nul}(G(S)) = |E(S)| - |T(S)| + 1$ functional parameters (Appendix). For example, in Fig. 2 we show a spanning tree of the evidence graph in Fig. 1(c). Specifically, it is made up of $|T(S)| - 1 = 3$ basic parameters:

$$E_b = \{\{a, b\}, \{a, d\}, \{d, c\}\},$$

and thus there are $5 - 3 = 2$ functional parameters:

$$E_f = \{\{b, d\}, \{b, c\}\}.$$

Corresponding to the first functional parameter, $\{b, d\}$, we identify the cycle $badb$. This implies that:

$$d((b, d)) = d((b, a)) + d((a, d)),$$

and for the second functional parameter we get:

$$d((b, c)) = d((b, a)) + d((a, d)) + d((d, c)).$$

2.3 Inconsistency models

The assumption of consistency does not necessarily hold and should be tested. Inconsistency can only occur if there are closed loops in the evidence structure (Lumley [2002]). An inconsistency relation is obtained by expanding a consistency relation with an ICF, e.g., for a loop $abca$, we add w_{abca} (Lu and Ades [2006]):

$$d_{ac} = d_{ab} + d_{bc} + w_{abca}, \qquad (3)$$

Fig. 2 Choosing a spanning tree (**b**) for an evidence graph (**a**) determines the partition into basic and functional parameters. The spanning tree induces a set of fundamental cycles, (**c**) and (**d**), that determine the equations that define the functional parameters

(a) Evidence graph   (b) Spanning tree   (c) Cycle I   (d) Cycle II
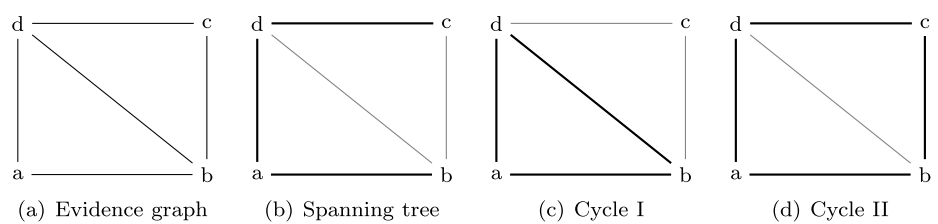
for which we again assume some distribution (Lu and Ades 2006). If multi-arm trials are included, some of the comparisons may be informed by only multi-arm trials, and evidence within a multi-arm trial is consistent by definition. For example, if we replace trials of a versus b, b versus c and a versus c with a three-arm trial a versus b versus c, the inconsistency model would not include w_{abca}. As we show in Sect. 3, the choice of basic parameters determines the number of ICFs that are included in the model.

3 Problem definition

The parameterization of an evidence structure requires partition of the parameters into *basic* and *functional* parameters, as presented in Theorem 1. However, when an inconsistency model is constructed for an evidence structure with multi-arm trials, the choices of the spanning tree and the individual study baselines are not arbitrary.

3.1 Spanning tree selection

If S contains only two-arm studies, then we may choose any spanning tree of $G(S)$ (Lu and Ades 2006; Salanti et al. 2008a). By contrast, when multi-arm studies are present, the choice of spanning tree is not arbitrary as some contrasts may be informed by only multi-arm trials, and the measurements within a multi-arm trial can not be inconsistent. For a cycle to potentially be inconsistent, it must be supported by at least three independent sources of evidence (Lu and Ades 2006), which is formalized below in Theorem 2. To be able to do this, we introduce the concept of the *partition* of a cycle into comparisons with their supporting studies and the operation of *reduction*, which allows us to simplify a partition.

Definition 4 (Elementary partition of C) Let C be a directed cycle in $G(S)$, represented by its set of (directed) edges. The elementary partition of C is (P, r), where $P = \{e \mid e \in C\}$, and $r(e) = \{S_i \in S \mid e \in E(S_i)\}$.

Note that there are, for each (undirected) cycle, two possible elementary partitions, depending on the direction in

which the cycle is traversed. Again using the evidence structure of Fig. 1(a), an elementary partition of the cycle $abcda$ is (P, r), where:

$$P = \{(a, b), (b, c), (c, d), (d, a)\},$$

$$r((a, b)) = \{1, 2\},$$

$$r((b, c)) = r((c, d)) = \{3\},$$

$$r((d, a)) = \{1\}.$$

Given an elementary partition (P, r) of an evidence cycle C, the inconsistency equation is given by:

$$w_C = F(P, r) = \sum_{(u, v) \in P} f((u, v), r((u, v))), \qquad (4)$$

with $f((u, v), S')$ defined as in Conjecture 1. This is a generalization of (3). Based on the conjecture, it seems that if two adjacent comparisons have the same set of supporting studies, we should be able to simplify the equation. We call this reducing the partition:

Definition 5 (Reduction) Let (P, r) be a partition of C and $e_k = (u, w_1), \ldots, e_l = (w_n, v) \in P$ a sequence of pairwise adjacent edges, such that $r(e_k) = \cdots = r(e_l)$. Then we may reduce this partition to (P', r'); where $P' = (P - \{e_k, \ldots, e_l\}) \cup \{(u, v)\}$, and

$$r'(e) = \begin{cases} r(e), & \text{if } e \neq (u, v), \\ r(e_k), & \text{if } e = (u, v). \end{cases}$$

Note that the numbering of the e_i is arbitrary and that for any specific reduction step, we can always choose the numbering scheme such that the reduced sequence $e_k, \ldots, e_l$ does not contain the subsequence e_n, e_1. Given this definition, it is natural to think of two edges e_i and e_j as *independent* if $r(e_i) \neq r(e_j)$ (and dependent otherwise). For convenience, we will call any pair (P, r) that was obtained from the elementary partition of C by (repeated) application of reduction, a partition of C. For example, the previously discussed elementary partition (P, r) of the cycle $abcda$ can be reduced, because $r(b, c) = r(c, d) = \{3\}$. We get (P', r'),

where

$$P' = \{(a, b), (b, d), (d, a)\},$$

$$r'((a, b)) = \{1, 2\},$$

$$r'((b, d)) = \{3\}, r'((d, a)) = \{1\}.$$

The following lemma shows that the inconsistency equation (4) is preserved under reduction of partitions:

Lemma 2 *Assume* (P, r) *is a partition of the cycle* C*, and* (P', r') *is obtained from* (P, r) *by a single reduction step. Then* $F(P, r) = F(P', r')$*.*

Proof Let $e_1, \ldots, e_k$ be the edges reduced to e', then $r'(e') = r(e_1) = \cdots = r(e_k)$. Now, if P has n edges:

$$F(P, r) = \sum_{i=1}^{n} f(e_i, r(e_i))$$

$$= \sum_{i=1}^{k} f(e_i, r(e_i)) + \sum_{i=k+1}^{n} f(e_i, r(e_i))$$

$$= f(e', r'(e')) + \sum_{i=k+1}^{n} f(e_i, r'(e_i))$$

$$= F(P', r').$$

Here, $f(e', r'(e')) = \sum_{i=1}^{k} f(e_i, r(e_i))$ holds by the same induction argument used for Lemma 1, but this time using Conjecture 1. $\square$

Lemma 3 *If a partition* (P, r) *of* C *contains* $k > 1$ *independent pairs of adjacent edges* (e_i, e_j)*, then there is a reduced partition* (P', r') *composed of* k *adjacently independent edges, such that* $F(P, r) = F(P', r')$*.*

Proof By the previous lemma, a single reduction step will preserve $F(P, r)$, thus so will repeated reduction. It remains to be shown that there is a reduction with exactly k edges, each independent of its adjacent edges. To see this, number the edges $e_1, \ldots, e_n$ so that $r(e_1) \neq r(e_n)$. Create a strictly increasing index list $i(1), \ldots, i(k-1)$ so that $r(e_{i(j)}) \neq r(e_{i(j)+1}); \forall_{1 \leq j \leq k-1}$. Then if we set $i(k) = n$, this list enumerates all independent pairs of adjacent edges. We can reduce $e_1, \ldots, e_{i(1)}$ to e'_1, $e_{i(1)+1}, \ldots, e_{i(2)}$ to e'_2 and so on, until $e_{i(k-1)+1}, \ldots, e_{i(k)}$ to e'_k. Then (P', r') with $P' = \{e'_1, \ldots, e'_k\}$ and $r'(e'_j) = r(e_{i(j)})$ consists of k adjacent independent edges. The reduction is unique up to the numbering of the e'_j. $\square$

The lemma leads to a simple test of when an inconsistency can occur in an evidence cycle, as given in the following theorem. We make the distinction between *potentially*

inconsistent, which is a property of the evidence structure, and *actually* inconsistent, which depends additionally on the data. A cycle is potentially inconsistent if we can devise data so that it becomes actually inconsistent.

Theorem 2 (Inconsistency cycle) *Let* C *be a cycle of length* n *and suppose that the elementary partition* (P, r) *of* C *has* m *independent pairs of adjacent edges. Then,* C *is potentially inconsistent iff* $m \geq 3$*.*

Proof Case I ($m < 3$) The first possibility is that all studies include the complete set of vertices in C ($m = 0$), and through internal consistency we have:

$$F(P, r) = \sum_{e \in P} F(e, r(e)) = 0.$$

From Lemma 3, if $m = 1$, C is not a cycle. If $m = 2$, we derive, using Lemma 3:

$$F(P, r) = F(\{(u, v), (v, u)\}, r);$$

$$r((u, v)) = R_1, \qquad r((v, u)) = R_2.$$

Thus, $F(P, r) \neq 0$ reduces to

$$f((u, v), R_1) + f((v, u), R_2) \neq 0,$$

$$f((u, v), R_1) \neq f((u, v), R_2),$$

which is just inter-study heterogeneity.

Case II ($m \geq 3$) Using Lemma 3, reduce the elementary partition to a partition where each pair of adjacent edges is independent. Since there are at least three distinct sets of supporting studies, the equations cannot be reduced as was done for $m < 3$, and this gives us sufficient freedom to choose data so that $w_C \neq 0$ without reducing to heterogeneity. Thus, C is potentially inconsistent. $\square$

In evidence structure II (repeated in Fig. 3(a)), if we consider the cycle $bcdb$, each of the comparisons is supported only by study 3, and hence $m = 0$, so according to the theorem, $bcdb$ is *not* potentially inconsistent. On the other hand, each of the comparisons in $abda$ is supported by a different study, so $m = 3$, making this cycle potentially inconsistent. The same holds for the longer cycle $abcda$, in which (b, c) and (c, d) are both supported by study 3, (d, a) by study 1 and (a, b) by study 2, also giving $m = 3$.

It would appear that this theorem allows us to count the number of inconsistency cycles for a given spanning tree. However, although the fundamental cycles for any spanning tree are independent, some may reduce to the same set of linear equations. An example of this is shown in Fig. 3, where the cycles $abda$ and $abcda$ discussed previously have the
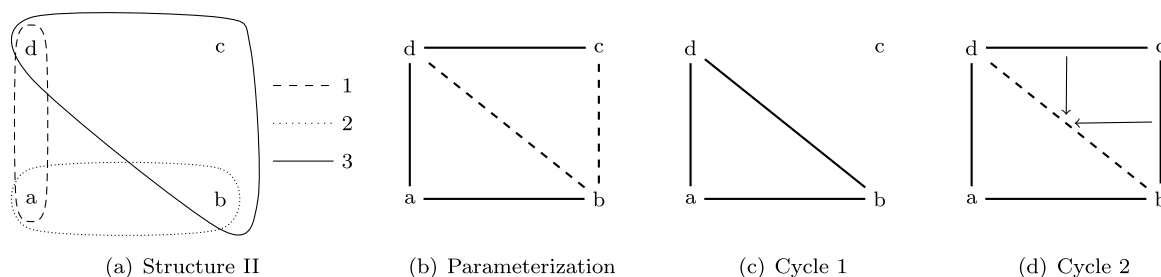
Fig. 3 An evidence structure (**a**) and spanning tree (**b**) in which two fundamental cycles reduce to the same set of equations. The *bcd* path in (**d**) collapses (through reduction) and leaves the same cycle as in (**c**), because (b, c), (b, d) and (c, d) are all supported only by study 3

same reduced partition, namely (P, r) with:

$$P = \{(a, b), (b, d), (d, a)\},$$

$$r((a, b)) = \{2\}, \qquad r((b, d)) = \{3\}, \qquad r((d, a)) = \{1\}.$$

Hence, to count the number of inconsistencies, we should count the number of distinct reduced partitions among the fundamental cycles. Moreover, any cycles that reduce to the same set of linear equations should be assigned the same ICF. The following definitions and lemma make this notion precise:

Definition 6 Let $g((u, v)) = (u, v)$ or $g((u, v)) = (v, u)$ be a one-one correspondence $P_1 \to P_2$. Then the partitions (P_1, r_1) and (P_2, r_2) are equivalent if

$$r_1(e) = r_2(g(e)); \quad \forall e \in P_1.$$

Lemma 4 *Let (P_1, r_1) and (P_2, r_2) be equivalent partitions under g. Then $F(P_1, r_1) = F(P_2, r_2)$ or $F(P_1, r_1) = -F(P_2, r_2)$, for $g((u, v)) = (u, v)$ or $g((u, v)) = (v, u)$ respectively.*

Proof Assuming $g((u, v)) = (u, v)$, we have

$$f((u, v), r_1((u, v))) = f(g((u, v)), r_2(g((u, v)))),$$

and thus

$$F(P_1, r_1) = \sum_{e \in P_1} f(e, r_1(e))$$

$$= \sum_{g(e) \in P_2} f(g(e), r_2(g(e))) = F(P_2, r_2),$$

where the second equality holds because g is a one-one correspondence. If $g((u, v)) = (v, u)$, we have that

$$f((u, v), r_1((u, v))) = -f((v, u), r_2((v, u)))$$

$$= -f(g((u, v)), r_2(g((u, v)))),$$

since $d((u, v)) = -d((v, u))$. Then clearly

$$F(P_1, r_1) = -F(P_2, r_2). \qquad \square$$

Definition 7 (*S*-equivalence) Two cycles C_1 and C_2 are *S*-equivalent ($C_1 \sim_S C_2$) iff their maximally reduced elementary partitions (in the evidence structure S) are equivalent.

By this definition the cycles *abda* and *abcda* shown in Fig. 3 and discussed above are *S*-equivalent. This means that if we assign the inconsistency factor w to *abda*:

$$d((b, d)) = d((b, a)) + d((a, d)) + w,$$

we should assign the same one to *abcda*:

$$d((b, c)) = d((b, a)) + d((a, d)) + d((d, c)) + w.$$

Note that according to Lemma 4, the direction in which we go around the cycle matters. In the above case the equivalence is due to $g((u, v)) = (u, v)$, so we use $+w$. If we traverse it in the other direction, $g((u, v)) = (v, u)$, we should use $-w$:

$$d((c, b)) = d((c, d)) + d((d, a)) + d((a, b)) - w.$$

Definition 8 (Inconsistency degree) For an evidence structure S and spanning tree G_b, let $\mathbf{C} = \mathbf{C}(G(S), G_b)$ be the set of fundamental cycles. Then, $\mathbf{C}/\sim_S$ is the set of equivalence classes under $\sim_S$ in $\mathbf{C}$. The ICD of G_b is the number of equivalence classes that contain inconsistency cycles:

$$\mathrm{icd}(S, G_b) = \sum_{X \in \mathbf{C}/\sim_S} \mathrm{icc}(S, C); \quad C \in X,$$

where $C \in X$ may be chosen arbitrarily and

$$\mathrm{icc}(S, C) = \begin{cases} 1, & \text{if } C \text{ is an inconsistency cycle,} \\ 0, & \text{otherwise.} \end{cases}$$

To clarify the meaning of the quotient set $\mathbf{C}/\sim_S$, consider again the situation of Fig. 3. Since both cycles are equivalent, we have

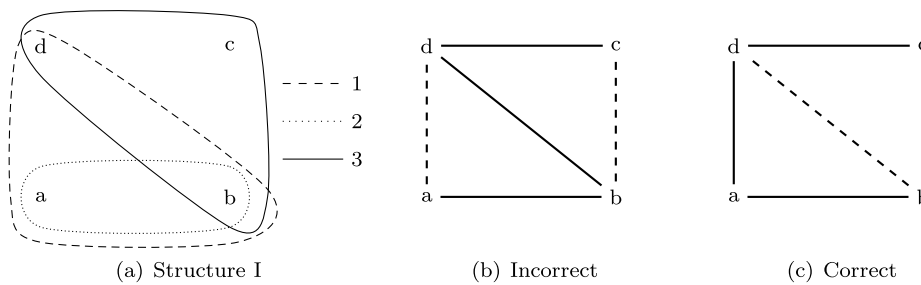$$\mathbf{C}/\sim_S = \{\{abda, abcda\}\}.$$

(a) Structure I   (b) Incorrect   (c) Correct

Fig. 4 An evidence structure in which a non obvious choice of ICFs is required to arrive at the correct ICDF. The structure (**a**) contains three trials, where trial 1 compares a, b and d, trial 2 compares a and b and trial 3 compares b, c and d. The subfigures give an incorrect (**b**) and a correct (**c**) parameterization

On the other hand, for the evidence structure in Fig. 4(a) and the spanning tree of Fig. 4(b), the cycles $abda$ and $bcdb$ are clearly not equivalent, so then

$$\mathbf{C}/\sim_S= \{\{abda\}, \{bcdb\}\}.$$

Definition 8 allows us to count the inconsistency degree of a spanning tree. For example, consider the evidence structure in Fig. 4(a). Clearly, both Figs. 4(b) and 4(c) are parameterized so that the $\dim(G(S)) = 3$ basic parameters (solid edges) form a spanning tree, and the remaining $\mathrm{nul}(G(S)) = 2$ edges become the functional parameters. This implies that the Inconsistency Degrees of Freedom (ICDF) is at most 2, the number of functional parameters. The spanning tree G_1 in Fig. 4(b) has two fundamental cycles, namely $abda$ and $bdcb$. The first cycle is partitioned into $e_1 = (a, b)$, $e_2 = (b, d)$ and $e_3 = (d, a)$, with support $r(e_1) = \{1, 2\}$, $r(e_2) = \{1, 3\}$ and $r(e_3) = \{1\}$. Since the sets of supporting studies are all distinct, Theorem 2 leads us to conclude that $\mathrm{icc}(abda) = 1$. For the latter cycle, we have $e_1 = (b, d)$, $e_2 = (d, c)$ and $e_3 = (c, b)$, for which the supporting studies are $r(e_1) = \{1, 3\}$, $r(e_2) = \{3\}$ and $r(e_3) = \{3\}$. Thus, using Theorem 2, this reduces to heterogeneity on (b, d), so $\mathrm{icc}(bdcb) = 0$. Hence, in this parameterization $\mathrm{icd}(S, G_1) = 1$.

Now, consider the tree G_2 in Fig. 4(c), with fundamental cycles $abda$ and $abcda$. We already know that $\mathrm{icc}(abda) = 1$. The partition of $abcda$ reduces to $e'_1 = (a, b)$, $e'_2 = (b, d)$, and $e'_3 = (d, a)$ with $r'(e'_1) = \{1, 2\}$, $r'(e'_2) = \{3\}$, and $r'(e'_2) = \{1\}$. All three edges are independent, and hence $\mathrm{icc}(abcda) = 1$. Moreover, the partitions of $abda$ and $abcda$ are not equivalent and thus $\mathrm{icd}(S, G_2) = 2$, the maximum possible. Hence, the choice of spanning tree determines the ICD:

Lemma 5 *The ICD* $\mathrm{icd}(S, G_b)$ *depends on the chosen spanning tree* G_b.

Theorem 3 (Spanning tree selection problem) *To parameterize the model correctly, we need to find a spanning tree* G_b *that maximizes* $\mathrm{icd}(S, G_b)$. *Then,* $\mathrm{icdf}(S) = \mathrm{icd}(S, G_b)$.

Proof $\mathrm{icd}(S, G_b)$ determines the number of independent inconsistency factors in the model. It has previously been shown that the ICFs $\mathbf{w}$ under one parameterization can be represented as linear combinations of the ICFs $\mathbf{w}'$ under another (Lu and Ades 2006), assuming equal icd. However, from Lemma 5, not all spanning trees result in the same icd. Therefore, in order to be able to express any ICF as a linear combination of the chosen ICFs, a maximal set of independent ICFs must be chosen. $\square$

Only one ICF should be created for each equivalence class of inconsistency cycles. Thus, whereas Lu and Ades (2006) claim that each ICF corresponds to exactly one functional parameter, actually each ICF may correspond to several. Their assertion that not every functional parameter need correspond to an ICF is confirmed by our work.

3.2 Baseline selection

The individual studies have to be parameterized in such a way that every comparison for which there is direct evidence (and which can be inconsistent) is expressed in the parameterization of at least one trial (Lu and Ades 2006). Again, this problem occurs only for multi-arm trials. For example, consider the structure in Fig. 5(a); we might parameterize trial 1 with b as the baseline and trial 3 with c as the baseline (Fig. 5(b)), having δ_{1ba}, δ_{1bd}, δ_{2ab}, δ_{3cb}, and δ_{3cd} as study parameters. Consider the cycle $abda$, where we have the inconsistency relation $d_{ad} = d_{ab} + d_{bd} + w_{abda}$. Now, d_{ad} is not informed directly by any of the study parameters δ_{ixy}, and hence the choice of d_{ad} is free, meaning that w_{abda} is also unconstrained. Hence, given this parameterization of the individual studies, the ICF w_{abda} cannot be estimated. A correct choice of baselines, covering all edges, is given in Fig. 5(c).

Thus, in addition to choosing the basic parameters correctly, the study baselines must be chosen so that at least one study provides direct evidence where needed. That is, every cycle C for which $\mathrm{icc}(C) = 1$, all $|C|$ parameters should have direct evidence, while for cycles where $\mathrm{icc}(C) = 0$,

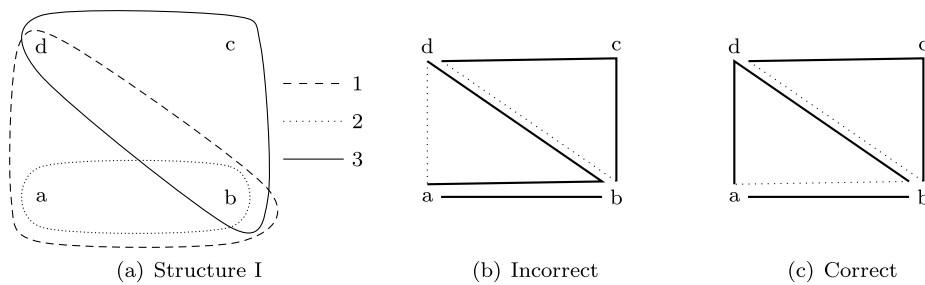(a) Structure I   (b) Incorrect   (c) Correct

Fig. 5 An evidence structure in which the choice of study baselines is not arbitrary. The subfigures (**b**)–(**c**) give an incorrect and a correct choice of baselines. The *solid* edges connect to the study baseline. In (**b**) baselines are 1:*b*, 2:*a*, 3:*c* and in (**c**) 1:*d*, 2:*a*, 3:*c*. The *dotted* edges are not connected to the baseline for that study and are thus not informed by direct evidence from that study

only $|C| - 1$ need direct evidence. In case of an equivalence class of inconsistency cycles, only one of the cycles needs all $|C|$ parameters to have direct evidence. This is formalized as follows:

Definition 9 (Evidence cover constraint) Let $X \in \mathbf{C}/\sim_S$ be an equivalence class of cycles. Let every cycle $C \in X$ be represented by its edge-set. Define the indicator function φ_X that is 1 if the direct evidence constraint is satisfied by the edge set E:

$$\varphi_X(E) = \begin{cases} 1, & \text{if } \forall_{C \in X}(|C \cap E| \geq |C| - 1) \wedge \\ & \exists_{C \in X}(\text{icc}(C) = 0 \vee |C \cap E| = |C|), \\ 0, & \text{otherwise.} \end{cases}$$

The individual studies are parameterized by choosing a baseline, and the effect parameters are the relative effects of all other treatments compared to the chosen baselines:

Definition 10 (Baseline study graph) Given a study S_i and a baseline $b_i \in T(S_i)$, the baseline study edge set is:

$$E(S_i, b_i) = \{\{b_i, x\} \mid x \in (T(S_i) - b_i)\}.$$

And the baseline study graph is given by $G(S_i, b_i) = (T(S_i), E(S_i, b_i))$.

Any given choice of baselines results in a graph representing the relative effects supported by at least one source of direct evidence:

Definition 11 (Baseline evidence graph) Given a baseline assignment $B = \{(S_i, b_i) \mid S_i \in S\}$, the baseline evidence graph is:

$$G(S, B) = (T(S), E(S, B)) = \left(T(S), \bigcup_{S_i \in S} E(S_i, b_i)\right).$$

Thus, the baseline selection problem is to find the baseline assignment that simultaneously satisfies the evidence cover constraint (Definition 9) for all equivalence classes of fundamental cycles:

Definition 12 (Baseline selection problem) Given S and a spanning tree G_b for S, the baseline selection problem is to find a baseline assignment B, that satisfies the constraint

$$\varphi_X(E(S, B)) = 1; \quad \forall X \in \mathbf{C}(G(S), G_b)/\sim_S .$$

3.3 Parameterization problem

Together, the problems of maximizing the ICDF and selecting the baselines form the parameterization problem:

Definition 13 (Parameterization problem) To choose a spanning tree G_b of S that maximizes $icd(S, G_b)$, while allowing a solution B to the baseline selection problem.

4 The algorithm

With the problem precisely defined, we developed a naive, inefficient algorithm that is sufficiently fast to solve all problem instances encountered in practice. An open source implementation is available from http://drugis.org/mtc.

The baseline selection sub-problem (Definition 12) is solved through an exhaustive search over the space of possible assignments, as is shown in Algorithm 1. Before the search, an arbitrary baseline is assigned for the two-arm studies since either baseline will cover all included comparisons (lines 3–4). For the multiple arm studies, all possible baseline choices are constructed (lines 5–6). Then, these are combined with the two arm study assignments (line 9) to construct the space A of possible baseline assignments. Then, an exhaustive search over the space A is performed (lines 10–14). As soon as a valid baseline assignment is found, the search is terminated. Note that in practice the set A is not constructed beforehand, but the space of baseline assignments is explored with e.g. a depth-first search (Cormen et al. 2001).

Algorithm 1 Find-baseline-assignment, procedure to find a baseline assignment satisfying a certain goal condition.

Input: Evidence structure S, goal condition $\varphi(\cdot)$
Output: Baseline assignment, or undefined if none exists
1: $b \leftarrow \emptyset, p \leftarrow \emptyset$
2: **for** $S_i \in S$ **do**
3:     **if** $|T(S_i)| = 2$ **then**
4:         $b \leftarrow b \cup \{(S_i, \text{some-element-of}(T(S_i)))\}$
5:     **else**
6:         $p \leftarrow p \cup \{S_i \times T(S_i)\}$
7:     **end if**
8: **end for**
9: $A \leftarrow p_1 \times \cdots \times p_n \times b_1 \times \cdots \times b_m$
10: **for** $B \in A$ **do**
11:     **if** $\varphi(B)$ **then**
12:         **return** B
13:     **end if**
14: **end for**
15: **return** undefined

The algorithm to solve the full parameterization problem is described as pseudocode in Algorithm 2. We start by trying to solve the baseline selection problem for the maximally constrained case in which all edges need direct evidence. The indicator procedure required in find-baseline-assignment for checking whether all edges are covered is presented in Algorithm 3. If no solution exists, any solution G_b to the parameterization problem will have $\text{icd}(S, G_b) < \text{nul}(G(S))$ (see Appendix).

Then, we use the standard algorithm presented in Gabow and Myers (1978) to iterate over all spanning trees of the evidence graph $G(S)$ (Definition 1). For each generated tree g, we determine the ICD $\text{icd}(S, g)$ (Definition 8). The procedure for computing ICD is given in Algorithm 4. If $\text{icd}(S, g)$ is greater than the largest so far, we determine whether there is a solution to the baseline selection problem for this tree. In this case the find-baseline-assignment requires an indicator procedure for checking whether the parameterization satisfies the baseline selection constraints (Definition 12); this one is presented in Algorithm 5. If there exists a solution to the baseline selection problem, we record this spanning tree and its ICD as the best so far. We stop if for the best tree so far $\text{icd}(S, g) = k$, the maximum possible, or if all spanning trees have been enumerated. For difficult problems this will be intractable, since there may be exponentially many spanning trees, and if the evidence structure has lower than maximal ICDF all of them have to be enumerated. However, it seems that most real-world problems are easy, as is shown by the computational tests in Sect. 6.

Using an exhaustive search to identify a baseline selection solves the baseline selection problem. Since the spanning tree search also (potentially) generates all possible spanning trees and maximizes the ICD taking into account

Algorithm 2 Parameterization of a mixed treatment comparison model as finding the spanning tree that maximizes the ICD while having a valid baseline assignment.

Input: Evidence structure S
Output: Solution (G_b, B) to the parameterization problem
1: best-$g \leftarrow$ undefined, best-$b \leftarrow$ undefined
2: $b \leftarrow$ undefined
3: full-$b \leftarrow$ find-baseline-assignment(S, φ_S)
4: **if** defined(full-b) **then**
5:     $k \leftarrow \text{nul}(G(S))$
6: **else**
7:     $k \leftarrow \text{nul}(G(S)) - 1$
8: **end if**
9: **for** $g \in$ gabow-myers$(G(S))$ **do**
10:     **if not** defined(best-g) **or** $\text{icd}(S, g) > \text{icd}(S, \text{best-}g)$ **then**
11:         **if** defined(full-b) **then**
12:             $b \leftarrow$ full-b
13:         **else**
14:             $b \leftarrow$ find-baseline-assignment$(S, \varphi_{S,g})$
15:         **end if**
16:         **if** defined(b) **then**
17:             best-$g \leftarrow g$
18:             best-$b \leftarrow b$
19:         **end if**
20:     **end if**
21:     **if** $\text{icd}(S, \text{best-}g) = k$ **then**
22:         **return** (best-g, best-b)
23:     **end if**
24: **end for**
25: **return** (best-g, best-b)

Algorithm 3 φ_S, indicator procedure for checking whether all edges are covered. $E(S_i, b_i)$ is the baseline study graph (Definition 10).

Input: Set B of pairs (S_i, b_i): (study, baseline)
Output: **true**, if all edges are covered, otherwise **false**
1: $E \leftarrow \bigcup_{(S_i, b_i) \in B} E(S_i, b_i)$
2: **return** $E = E(S)$

whether there is a solution to the baseline selection problem, the algorithm outlined here solves the parameterization problem (Definition 13).

5 Example

As a full example of finding a correct parameterization for an evidence structure, we consider a network of treatments for smoking cessation therapy comparing (*a*) nicotine replacement therapy, (*b*) bupropion, (*c*) varenicline and (*d*) placebo or no treatment (Wu et al. 2006). The outcome of

Algorithm 4 icd, procedure for computing the ICD. $C / \sim_S$ is the set of equivalence classes of cycles (Definition 8), and $C(\cdot, \cdot)$ is the set of fundamental cycles (Appendix).

Input: Evidence structure S, Spanning tree g
Output: $\mathrm{icd}(S, g)$
1:   $C \leftarrow C(G(S), g)$
2:   $\mathrm{icd} \leftarrow 0$
3:   **for** $X \in (C / \sim_S)$ **do**
4:      $\mathrm{icd} \leftarrow \mathrm{icd} + \mathrm{icc}(\text{some-element-of}(X))$
5:   **end for**
6:   **return** icd

Algorithm 5 $\varphi_{S,g}$, indicator procedure for checking whether the parameterization satisfies the baseline selection constraints (Definition 12). $E(S_i, b_i)$ is the baseline study graph (Definition 10), $C / \sim_S$ is the set of equivalence classes of cycles (Definition 8), and $C(\cdot, \cdot)$ is the set of fundamental cycles (Appendix).

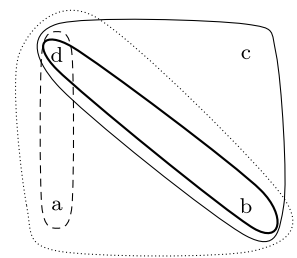**Input:** Set B of pairs (S_i, b_i): (study, baseline)
Output: true, if Definition 12 is satisfied, otherwise **false**
1:   $E \leftarrow \bigcup_{(S_i, b_i) \in B} E(S_i, b_i)$
2:   $C \leftarrow C(G(S), g)$
3:   **for** $X \in (C / \sim_S)$ **do**
4:     **for** $C \in X$ **do**
5:       **if** $|C \cap E| < |C| - 1$ **then**
6:         **return false**
7:       **end if**
8:     **end for**
9:     **if** $\mathrm{icc}(\text{some-element-of}(X)) = 0$ **then**
10:      **return true**
11:     **end if**
12:     **for** $C \in X$ **do**
13:       **if** $|C \cap E| = |C|$ **then**
14:         **return true**
15:       **end if**
16:     **end for**
17:   **end for**
18:   **return false**

Fig. 6 The evidence structure for the outcome 'smoking cessation after 12 months' from Wu et al. (2006). a = nicotine replacement therapy, b = bupropion, c = varenicline, and d = placebo or no treatment

interest is smoking cessation at 12 months. For this outcome there are 78 studies with 4 different treatment comparisons: S_{ad} (66 studies), S_{bd} (6 studies), S_{abd} (3 studies) and S_{bcd} (3 studies). The evidence structure is shown in Fig. 6, having $|T(S)| = 4$ treatments and $|E(S)| = 5$ comparisons. Thus, any correct parameterization will have $|E_b| = 4 - 1 = 3$ basic and $|E_f| = 5 - 3 = 2$ functional parameters.

The first step in the algorithm is to try and find a baseline selection that covers all edges. Given this structure, that is easy, e.g. a for the S_{ad} studies, b for S_{bd}, a for S_{abd} and c for S_{bcd} will suffice. We could also have chosen different baselines for studies of the same type, but that is not neces-

sary here. Then, we set $k = |E_f| = 2$, meaning we will try to find an ICD equal to the number of functional parameters.

Now, we start iterating over the spanning trees of the evidence graph. Say the first spanning tree we are given is

$$g_1 = \{\{b, a\}, \{b, c\}, \{b, d\}\}.$$

Then the cycles to evaluate are $badb$ and $bcdb$. For $badb$ we get the partition $P = \{(b, a), (a, d), (d, b)\}$ with $r((b, a)) = S_{abd}$, $r((a, d)) = S_{ad} \cup S_{abd}$ and $r((d, b)) = S_{bd} \cup S_{bcd}$. This partition cannot be reduced any further, and there are 3 distinct sets of studies, so

$$\mathrm{icc}(S, badb) = 1.$$

For $bcdb$, both (b, c) and (c, d) are supported only by S_{bcd}, and thus we have only 2 sets of supporting studies, so

$$\mathrm{icc}(S, bcdb) = 0.$$

The cycles are not equivalent, so they fall into two separate classes, and we get $\mathrm{icd}(S, g_1) = 1$. This is the best so far, so we store g_1.

The ICD of g_1 is $1 < k = 2$, so we continue iterating over the spanning trees. The second spanning tree might be:

$$g_2 = \{\{b, a\}, \{a, d\}, \{d, c\}\},$$

having fundamental cycles $bdab$ and $bcdab$. We already know that $\mathrm{icc}(S, bdab) = 1$, and we also recognize that $bcdab$ is a basically a longer version of $bdab$, so we will have to evaluate whether they are equivalent. We also recall that (d, c) and (c, b) are only supported by S_{bcd} and hence reduce to (d, b) with $r'((d, b)) = S_{bcd}$. The other two comparisons, (b, a) and (a, d) cannot be reduced. Thus we get a reduction for $bcdab$ that has the same comparisons as $bdab$, but a different set of supporting studies for (b, d): S_{bcd} for $bcdab$ and $S_{bd} \cup S_{bcd}$ for $bdab$. Hence $bdab$ and $bcdab$ are not equivalent in S: $bdab \not\sim_S bcdab$. Moreover, we also get

$$\mathrm{icc}(S, bcdab) = 1,$$

so that $\mathrm{icd}(S, g_2) = 2 = k$. Hence we have identified g_2 with the full baseline assignment identified earlier as the solution to the parameterization problem.

The example structure is structure number 18 in Table 1, and our implementation of the algorithm actually evaluates

Table 1 Performance of our algorithm on evidence structures from Salanti et al. (2008b). Structures are listed in the same order as Fig. 1 in Salanti et al. (2008b), omitting the first three; the first column (ref) gives the reference number in that paper. v is the number of included treatments, e is the number of comparisons, s is the number of functional parameters and n is the number of studies. The x_j indicate the number of different *types* of j-arm studies, e.g. if we have 2 ab studies and 3 bc studies, $x_2 = 2$. For the results, i is the ICD of the solution, N is the number of evaluated spanning trees and t is the time taken (seconds)

Ref	Structure							Result		
	v	e	s	n	x_2	x_3	x_4	i	N	t
18	4	5	2	78	2	2	0	2	4	0.3
19	4	5	2	34	5	0	0	2	1	0.2
20	4	4	1	12	3	1	0	1	1	0.1
21	4	4	1	10	3	1	0	1	1	0.1
22	4	5	2	21	5	0	0	2	1	0.1
23	10	21	12	43	16	6	0	11	1	3.8
24	9	14	6	54	12	3	0	6	1	0.5
25	7	16	10	22	13	4	0	10	1	0.5
26	9	16	8	18	10	3	1	7	1	1.0
3	6	9	4	14	7	1	0	4	13	0.4
27	7	8	2	25	8	0	0	2	1	0.2
28	16	22	7	34	15	4	0	6	1	0.9
29	7	8	2	10	8	1	0	2	1	0.1
30	8	10	3	14	9	1	0	3	1	0.2
31	10	12	3	14	10	2	0	3	1	0.2

four spanning trees before it finds the correct one, rather than the two shown here.

6 Evaluation of the running-time

A review of published evidence networks (Salanti et al. 2008b) identified 18 different networks in the literature. Three of those were star-shaped, and have a trivial solution to the parameterization problem. For the other 15 networks, we extracted the evidence structure from the original papers and evaluated the running time of our algorithm, as well as the ICDF of each structure and the number of spanning trees that were generated before a solution was found. The results are summarized in Table 1, and we give the exact evidence structures in an online supplement. There are three structures with only two-arm trials, the remaining 12 have at least one three-arm trial. There is one structure that includes a four-arm trial.

All of the evidence structures were parameterized within four seconds (on a 3 GHz processor), which is negligible compared to the time usually taken by the MCMC simulation used to estimate the models. The longest time taken was on structure 23, which contains the largest number (6) of distinct types of three-arm trials. Only three structures had

non-maximal ICDF (23, 26, 28), namely $s - 1$, one less than the number of functional parameters. Note that if the ICDF would be $< s - 1$, our algorithm would need to enumerate all spanning trees to terminate. In only two cases more than one spanning tree needed to be explored. The number of distinct spanning trees the evidence graph had varied between three (structures 21 and 22) to 13611 (structure 23). All running times of > 0.5 seconds were observed for structures with non-maximal ICDF. In all three cases, this reflects a failed exhaustive baseline search for full evidence cover.

7 Discussion

In this paper, we defined the parameterization problem for MTC evidence structures and we provided an algorithm which can be used for automated model generation for MTC. We refine previous work (Lu and Ades 2006) on identifying the ICDF by giving a precise problem definition, and point out the additional problem of equivalent cycles. An open source implementation of the algorithm is available (http://drugis.org/mtc). Although the worst-case complexity of our algorithm is exponential, it seems that real-world problems can be solved quickly. We evaluated running time of the algorithm with evidence structures from the literature, and all were solved within four seconds on a standard PC.

Future work should aim to develop more efficient algorithms, and further investigate the relationship between the spanning tree and baseline selection problems. In this paper, we took the pragmatic approach of defining the combined problem as finding the maximal ICD for which a baseline selection can be derived. The question remains whether there may be evidence structures for which the optimal spanning tree does not have an associated baseline assignment, and what would be the implications for the MTC method. There also seems to be a certain redundancy in the cycles *bdab* and *bcdab* of Fig. 6 discussed in Sect. 5, even though they are not equivalent according to Definition 7. This is correct since the w-factors associated with these cycles are not provably equal. However, the w-factors should differ only by heterogeneity on the (b, d) comparison. Future work should address whether and, if so, how this should be incorporated in the parameterization of the evidence structure.

Appendix: Definitions from graph theory

Definition 14 (Spanning tree—Gabow and Myers 1978) In a connected, undirected graph G, a *spanning tree* G_s is a subgraph having a unique simple path (a path containing each vertex at most once) between any two vertices of G. If G has t vertices, G_s has $\dim(G) = t - 1$ edges.

Definition 15 (Fundamental cycle set—Deo et al. 1982) The fundamental cycle set of a connected, undirected graph $G = (T, E)$ with respect to a spanning tree $G_s = (T, F)$ is generated from the set $E' = E \backslash F$, as follows:

$$\mathbf{C}(G, G_s) = \{C(G, G_s, e) \mid e \in E'\}, \quad \text{with}$$

$$C(G, G_s, \{u, v\}) = \text{path}(v, u) \cup \{\{u, v\}\},$$

where $\text{path}(v, u)$ gives the (unique simple) path from v to u in G_s. The size of the set of non-tree edges, $\text{nul}(G) = |E'| = |E| - |T| + 1$ is called the *nullity* of G, and determines the number of *independent* cycles in G. The set $\mathbf{C}(G, G_s)$ consists of independent cycles and $|\mathbf{C}(G, G_s)| = \text{nul}(G)$, which means that the set of fundamental cycles is also a *cycle basis* of G.

Definition 16 (Path) A path is a sequence of directed edges, such that the target of each edge connects to the source of the next one: $p = ((w_1, w_2), (w_2, w_3), \ldots, (w_{n-1}, w_n))$ is a path of length $n - 1$, as counted by the number of edges. Often, the path p is conveniently written as $(w_1, w_2, \ldots, w_n)$, which should be read as shorthand for the longer notation.

References

Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C.: Introduction to Algorithms, 2nd edn. MIT Press, Cambridge (2001)

Deo, N., Prabhu, G.M., Krishnamoorthy, M.S.: Algorithms for generating fundamental cycles in a graph. ACM Trans. Math. Softw. **8**(1), 26–42 (1982). doi:10.1145/355984.355988

Gabow, H.N., Myers, E.W.: Finding all spanning trees of directed and undirected graphs. SIAM J. Comput. **7**(3), 280–287 (1978). doi:10.1137/0207024

Hedges, L.V., Vevea, J.L.: Fixed- and random-effects models in meta-analysis. Psychol. Methods **3**(4), 486–504 (1998). doi:10.1037/1082-989X.3.4.486

Higgins, J.P.T., Whitehead, A.: Borrowing strength from external trials in a meta-analysis. Stat. Med. **15**(24), 2733–2749 (1996). doi:10.1002/(SICI)1097-0258(19961230)15:24<2733::AID-SIM562>3.0.CO;2-0

Lu, G., Ades, A.E.: Assessing evidence inconsistency in mixed treatment comparisons. J. Am. Stat. Assoc. **101**(474), 447–459 (2006). doi:10.1198/016214505000001302

Lu, G., Ades, A.: Modeling between-trial variance structure in mixed treatment comparisons. Biostatistics **10**(4), 792–805 (2009). doi:10.1093/biostatistics/kxp032

Lumley, T.: Network meta-analysis for indirect treatment comparisons. Stat. Med. **21**(16), 2313–2324 (2002). doi:10.1002/sim.1201

Normand, S.: Meta-analysis: formulating, evaluating, combining, and reporting. Stat. Med. **18**(3), 321–359 (1999). doi:10.1002/(SICI)1097-0258(19990215)18:3<273::AID-SIM19>3.0.CO;2-7

Salanti, G., Higgins, J.P.T., Ades, A.E., Ioannidis, J.P.A.: Evaluation of networks of randomized trials. Stat. Methods Med. Res. **17**(3), 279–301 (2008a). doi:10.1177/0962280207080643

Salanti, G., Kavvoura, F.K., Ioannidis, J.P.A.: Exploring the geometry of treatment networks. Ann. Intern. Med. **148**(7), 544–553 (2008b)

Sutton, A.J., Higgins, J.P.T.: Recent developments in meta-analysis. Stat. Med. **27**(5), 625–650 (2008). doi:10.1002/sim.2934

Wu, P., Wilson, K., Dimoulas, P., Mills, E.J.: Effectiveness of smoking cessation therapies: a systematic review and meta-analysis. BMC Public Health **6**, 300 (2006). doi:10.1186/1471-2458-6-300