eAppendix A Remarks on the sampling model

Combining data across multiple data sources requires some sort of sampling model that underlies the observed data and specifies the relationships between the data sources. For example, in "conventional" random effects meta-analysis models [54], the true study effects (e.g., the study-specific population-averaged treatment effects underlying each published study) are assumed to be independently sampled from an infinite population of study effects (e.g., [4, 54]).

In this paper we operate under a sampling model that relates to the sampling model we proposed for generalizability and transportability analyses that involve a single randomized trial using non-nested trial designs [20]. Specifically, we assume that the research question defines an infinite superpopulation of individuals that can be stratified in several non-overlapping groups: (1) the non-overlapping groups of individuals who belong in the population underlying each trial s in the collection $\mathcal{S} = \{1, \ldots, m\}$; that is, individuals who receive care or services in each trial's recruitment centers, who meet that trial's eligibility criteria, and who would be invited and agree to participate in the trial; and (2) individuals who belong to the target population of substantive interest, $S = 0$, and who are not invited to participate in any trial or who would decline if invited. Individuals with $S = 0$ may also be a well-defined subset of individuals who were not invited to participate in any trial or who were invited to participate but declined; that is, individuals with $S = 0$ do not need to exhaust the set of individuals not participating in any trial.

The relative size of subsets of the superpopulation defined by different S values is determined by the scientific question, as well as the trial eligibility criteria, individuals' access to centers recruiting patients into different trials, and individuals' preferences (e.g., desire to participate in a trial). This infinite superpopulation [55] is a convenient fiction [29], but it seems to us to be somewhat more plausible than the infinite population of study effects invoked by random effects meta-analyses.

We assume that the data are generated by random sampling of individuals from the superpopu-

lation, stratified by S, with sampling fractions are (1) constant for individuals from the same subset of the superpopulation, $S = s$, for $s \in \{0, 1, \ldots, m\}$; (2) possibly variable across groups defined by different S values; (3) and unknown to the investigators. Constancy of the sampling fraction for individuals with $S = s$ for every $s \in \{0, 1, \ldots, m\}$ is a reflection of the frequently entertained assumption that individuals participating in a trial can be viewed as a simple random sample from the population underlying the trial; as well at the assumption that the sample of the target population is representative of that population. Variation of the sampling fractions across S values reflects differences in recruitment practices across trials as well as the inability to obtain data from all individuals from the target population. The sampling fractions are unknown to the investigators because they depend on complex mechanisms that drive the design and conduct of randomized trials, as well as the underlying population of each trial.

The sampling step induces a biased sampling model [18] in the following sense: in the data, the ratios $\frac{n_j}{n} = \frac{n_j}{n_0 + \sum_{s \in \mathcal{S}} n_s}$, for $j \in \{0, 1, \ldots, m\}$, do not reflect the population probabilities of belonging to the subset of the population with $S = j$, because the sampling fraction from each subset is unknown (and possibly variable between subsets). As a technical condition, we require that as $n \longrightarrow \infty$, $\frac{n_j}{n} \longrightarrow \pi_j > 0$, for $j \in \{0, 1, \ldots, m\}$. Nevertheless, under the biased sampling model, the limiting values, π_j, are not equal to the superpopulation probabilities $\Pr[S = j]$.

In effect, we view the data as generated by stratified random sampling of the actual population with sampling fractions that only depend on S but are unknown to the investigators. Arguments analogous to those in [19] and [20] show that the identifiability of the causal quantities of interest is unaffected by biased sampling.

eAppendix B Identification of the effect of treatment assignment

In this eAppendix, we collect results about the identification of average treatment effects of treatment assignment (intention-to-treat average treatment effects) in the target population.

Transporting from an individual study

Proposition 1. *Under conditions A1 through A5 in the main text, the average treatment effect in the target population, comparing treatments z and z' in $\mathcal{Z}$,* $\mathrm{E}[Y^z - Y^{z'}|S=0]$*, using data from study $s^* \in \mathcal{S}$ and the target population, is identified by the following functional of the observed data distribution:*

$$\psi(z, z', s^*) \equiv \mathrm{E}\left[\mathrm{E}[Y|X, S=s^*, Z=z] - \mathrm{E}[Y|X, S=s^*, Z=z']\middle|S=0\right]. \tag{B.1}$$

Proof:

$$\begin{aligned}
&\mathrm{E}[Y^z - Y^{z'}|S=0] \\
&\quad = \mathrm{E}\left[\mathrm{E}[Y^z - Y^{z'}|X, S=0]\middle|S=0\right] \\
&\quad = \mathrm{E}\left[\mathrm{E}[Y^z - Y^{z'}|X, S=s^*]\middle|S=0\right] \\
&\quad = \mathrm{E}\left[\mathrm{E}[Y^z|X, S=s^*] - \mathrm{E}[Y^{z'}|X, S=s^*]\middle|S=0\right] \\
&\quad = \mathrm{E}\left[\mathrm{E}[Y^z|X, S=s^*, Z=z] - \mathrm{E}[Y^{z'}|X, S=s^*, Z=z']\middle|S=0\right] \\
&\quad = \mathrm{E}\left[\mathrm{E}[Y|X, S=s^*, Z=z] - \mathrm{E}[Y|X, S=s^*, Z=z']\middle|S=0\right] \\
&\quad \equiv \psi(z, z', s^*).
\end{aligned}$$

□

Proposition 2. *Under positivity conditions A3 and A5,*

$$\psi(z,z',s^*) = \frac{1}{\Pr[S=0]}\mathrm{E}\left[\left(\frac{I(Z=z)}{\Pr[Z=z|X,S=s^*]} - \frac{I(Z=z')}{\Pr[Z=z'|X,S=s^*]}\right)\frac{I(S=s^*)Y\Pr[S=0|X,I(S\in\{0,s^*\})=1]}{\Pr[S=s^*|X,I(S\in\{0,s^*\})=1]}\right]. \tag{B.2}$$

Proof: We begin from the left-hand-side, using the definition of $\psi(z,z',s^*)$,

$$\begin{aligned}
\psi(z,z',s^*) &\equiv \mathrm{E}\left[\,\mathrm{E}[Y|X,S=s^*,Z=z] - \mathrm{E}[Y|X,S=s^*,Z=z']\middle|S=0\right] \\
&= \mathrm{E}\left[\mathrm{E}\left[\frac{I(Z=z)Y}{\Pr[Z=z|X,S=s^*]}\middle|X,S=s^*\right] - \mathrm{E}\left[\frac{I(Z=z')Y}{\Pr[Z=z'|X,S=s^*]}\middle|X,S=s^*\right]\middle|S=0\right] \\
&= \mathrm{E}\left[\mathrm{E}\left[\left(\frac{I(Z=z)}{\Pr[Z=z|X,S=s^*]} - \frac{I(Z=z')}{\Pr[Z=z'|X,S=s^*]}\right)Y\middle|X,S=s^*\right]\middle|S=0\right] \\
&= \mathrm{E}\left[\mathrm{E}\left[\left(\frac{I(Z=z)}{\Pr[Z=z|X,S=s^*]} - \frac{I(Z=z')}{\Pr[Z=z'|X,S=s^*]}\right)\frac{I(S=s^*)Y}{\Pr[S=s^*|X]}\middle|X\right]\middle|S=0\right] \\
&= \frac{1}{\Pr[S=0]}\mathrm{E}\left[I(S=0)\,\mathrm{E}\left[\left(\frac{I(Z=z)}{\Pr[Z=z|X,S=s^*]} - \frac{I(Z=z')}{\Pr[Z=z'|X,S=s^*]}\right)\frac{I(S=s^*)Y}{\Pr[S=s^*|X]}\middle|X\right]\right] \\
&= \frac{1}{\Pr[S=0]}\mathrm{E}\left[\Pr[S=0|X]\,\mathrm{E}\left[\left(\frac{I(Z=z)}{\Pr[Z=z|X,S=s^*]} - \frac{I(Z=z')}{\Pr[Z=z'|X,S=s^*]}\right)\frac{I(S=s^*)Y}{\Pr[S=s^*|X]}\middle|X\right]\right] \\
&= \frac{1}{\Pr[S=0]}\mathrm{E}\left[\mathrm{E}\left[\left(\frac{I(Z=z)}{\Pr[Z=z|X,S=s^*]} - \frac{I(Z=z')}{\Pr[Z=z'|X,S=s^*]}\right)\frac{I(S=s^*)Y\Pr[S=0|X]}{\Pr[S=s^*|X]}\middle|X\right]\right] \\
&= \frac{1}{\Pr[S=0]}\mathrm{E}\left[\left(\frac{I(Z=z)}{\Pr[Z=z|X,S=s^*]} - \frac{I(Z=z')}{\Pr[Z=z'|X,S=s^*]}\right)\frac{I(S=s^*)Y\Pr[S=0|X]}{\Pr[S=s^*|X]}\right] \\
&= \frac{1}{\Pr[S=0]}\mathrm{E}\left[\left(\frac{I(Z=z)}{\Pr[Z=z|X,S=s^*]} - \frac{I(Z=z')}{\Pr[Z=z'|X,S=s^*]}\right)\frac{I(S=s^*)Y\Pr[S=0|X,I(S\in\{0,s^*\})=1]}{\Pr[S=s^*|X,I(S\in\{0,s^*\})=1]}\right],
\end{aligned}$$

where the last step uses the fact that $\dfrac{\Pr[S=0|X]}{\Pr[S=s^*|X]} = \dfrac{\Pr[S=0|X,I(S\in\{0,s^*\})=1]}{\Pr[S=s^*|X,I(S\in\{0,s^*\})=1]}$.

□

Remark. *The derivation for Proposition 2 does not use any conditions that involve potential outcomes and thus holds whether or not $\psi(z,z',s^*)$ has a causal interpretation.*

Transporting the entire collection of trials

Proposition 3. *If conditions B1 through B5 from the main text, hold for every trial $s \in \mathcal{S}$, then the average treatment effect in the target population, comparing treatments z and z' in $\mathcal{Z}$,* $\mathrm{E}[Y^z - Y^{z'}|S = 0]$*, equals the following functional of the observed data distribution:*

$$\phi(z, z') \equiv \mathrm{E}\left[\tau(z, z'; X)\middle|S = 0\right], \tag{B.3}$$

where $\tau(z, z'; X)$ is the common (across trials) observed outcome mean difference comparing treatments z and z' conditional on X.

Proof: If conditions *B4* and *B5*, hold for every trial $s \in \mathcal{S} = \{1, \ldots, m\}$ then, as shown in the main text,

$$\mathrm{E}[Y^z - Y^{z'}|X, S = 1] = \ldots = \mathrm{E}[Y^z - Y^{z'}|X, S = m] = \mathrm{E}[Y^z - Y^{z'}|X, I(S \in \mathcal{S}) = 0].$$

Under conditions 1 through 3, the above result implies,

$$\mathrm{E}[Y|X, S = 1, Z = z] - \mathrm{E}[Y|X, S = 1, Z = z'] = \ldots = \mathrm{E}[Y|X, S = m, Z = z] - \mathrm{E}[Y|X, S = m, Z = z'].$$

Using the notation, $\tau(z, z'; X)$ to denote the common (across trials) observed outcome mean difference comparing treatments z and z' conditional on X, and Proposition 1, we obtain

$$\phi(z, z') = \mathrm{E}\left[\tau(z, z'; X)\middle|S = 0\right]. \tag{B.4}$$

□

Proposition 4. *Under positivity conditions B3 and B5,*

$$\phi(z,z') = \frac{1}{\Pr[S=0]}\mathrm{E}\left[\omega(z,z';X,S)\frac{Y\Pr[S=0|X]}{\Pr[I(S\in\mathcal{S})=1|X]}\right], \tag{B.5}$$

where

$$\omega(z,z';X,S) = \left(\frac{I(Z=z)}{\Pr[Z=z|X,S,I(S\in\mathcal{S})=1]} - \frac{I(Z=z')}{\Pr[Z=z'|X,S,I(S\in\mathcal{S})=1]}\right) I(S\in\mathcal{S}).$$

Proof: First, it is easy to see that for each trials $s \in \mathcal{S}$,

$$\begin{aligned}
&\mathrm{E}[Y|X,S=s,Z=z] - \mathrm{E}[Y|X,S=s,Z=z'] \\
&\quad = \mathrm{E}\left[\left(\frac{I(Z=z)}{\Pr[Z=z|X,S=s]} - \frac{I(Z=z')}{\Pr[Z=z'|X,S=s]}\right) Y \Big| X, S=s\right] \\
&\quad = \mathrm{E}\left[\left(\frac{I(Z=z)}{\Pr[Z=z|X,S,I(S\in\mathcal{S})=1]} - \frac{I(Z=z')}{\Pr[Z=z'|X,S,I(S\in\mathcal{S})=1]}\right) Y \Big| X, S=s\right].
\end{aligned} \tag{B.6}$$

From the derivation of Proposition 3,

$$\mathrm{E}[Y|X,S=1,Z=z] - \mathrm{E}[Y|X,S=1,Z=z'] = \ldots = \mathrm{E}[Y|X,S=m,Z=z] - \mathrm{E}[Y|X,S=m,Z=z'],$$

which, combined with the result in (B.6), implies that

$$\begin{aligned}
&\mathrm{E}\left[\left(\frac{I(Z=z)}{\Pr[Z=z|X,S,I(S\in\mathcal{S})=1]} - \frac{I(Z=z')}{\Pr[Z=z'|X,S,I(S\in\mathcal{S})=1]}\right) Y \Big| X, S=1\right] = \ldots \\
&\quad = \mathrm{E}\left[\left(\frac{I(Z=z)}{\Pr[Z=z|X,S,I(S\in\mathcal{S})=1]} - \frac{I(Z=z')}{\Pr[Z=z'|X,S,I(S\in\mathcal{S})=1]}\right) Y \Big| X, S=m\right].
\end{aligned} \tag{B.7}$$

Let

$$\mathrm{E}\left[\left(\frac{I(Z=z)}{\Pr[Z=z|X,S,I(S\in\mathcal{S})=1]}-\frac{I(Z=z')}{\Pr[Z=z'|X,S,I(S\in\mathcal{S})=1]}\right)Y\Big|X,I(S\in\mathcal{S})=1\right]$$

denote the quantity equal to all terms in the above chain of equations. Because this quantity is equal to each of the conditional mean differences $\mathrm{E}[Y|X,S=s,Z=z]-\mathrm{E}[Y|X,S=s,Z=z']$, for every $s\in\mathcal{S}$, it has to be that

$$\tau(z,z';X)=\mathrm{E}\left[\left(\frac{I(Z=z)}{\Pr[Z=z|X,S,I(S\in\mathcal{S})=1]}-\frac{I(Z=z')}{\Pr[Z=z'|X,S,I(S\in\mathcal{S})=1]}\right)Y\Big|X,I(S\in\mathcal{S})=1\right].$$

Using the derivation of Proposition 3, we obtain

$$\begin{aligned}\phi(z,z')&=\mathrm{E}[\tau(z,z';X)|S=0]\\&=\mathrm{E}\left[\mathrm{E}\left[\left(\frac{I(Z=z)}{\Pr[Z=z|X,S,I(S\in\mathcal{S})=1]}-\frac{I(Z=z')}{\Pr[Z=z'|X,S,I(S\in\mathcal{S})=1]}\right)Y\Big|X,I(S\in\mathcal{S})=1\right]\Big|S=0\right].\end{aligned}$$

Because

$$\begin{aligned}&\mathrm{E}\left[\mathrm{E}\left[\left(\frac{I(Z=z)}{\Pr[Z=z|X,S,I(S\in\mathcal{S})=1]}-\frac{I(Z=z')}{\Pr[Z=z'|X,S,I(S\in\mathcal{S})=1]}\right)Y\Big|X,I(S\in\mathcal{S})=1\right]\Big|S=0\right]\\&=\frac{1}{\Pr[S=0]}\mathrm{E}\left[I(S=0)\,\mathrm{E}\left[\left(\frac{I(Z=z)}{\Pr[Z=z|X,S,I(S\in\mathcal{S})=1]}-\frac{I(Z=z')}{\Pr[Z=z'|X,S,I(S\in\mathcal{S})=1]}\right)Y\Big|X,I(S\in\mathcal{S})=1\right]\right]\\&=\frac{1}{\Pr[S=0]}\mathrm{E}\left[\Pr[S=0|X]\,\mathrm{E}\left[\left(\frac{I(Z=z)}{\Pr[Z=z|X,S,I(S\in\mathcal{S})=1]}-\frac{I(Z=z')}{\Pr[Z=z'|X,S,I(S\in\mathcal{S})=1]}\right)Y\Big|X,I(S\in\mathcal{S})=1\right]\right]\\&=\frac{1}{\Pr[S=0]}\mathrm{E}\left[\mathrm{E}\left[\left(\frac{I(Z=z)}{\Pr[Z=z|X,S,I(S\in\mathcal{S})=1]}-\frac{I(Z=z')}{\Pr[Z=z'|X,S,I(S\in\mathcal{S})=1]}\right)\frac{I(S\in\mathcal{S})Y\Pr[S=0|X]}{\Pr[I(S\in\mathcal{S})=1|X]}\Big|X\right]\right]\\&=\frac{1}{\Pr[S=0]}\mathrm{E}\left[\left(\frac{I(Z=z)}{\Pr[Z=z|X,S,I(S\in\mathcal{S})=1]}-\frac{I(Z=z')}{\Pr[Z=z'|X,S,I(S\in\mathcal{S})=1]}\right)\frac{I(S\in\mathcal{S})Y\Pr[S=0|X]}{\Pr[I(S\in\mathcal{S})=1|X]}\right],\end{aligned}$$

we conclude that

$$\phi(z,z')=\frac{1}{\Pr[S=0]}\mathrm{E}\left[\left(\frac{I(Z=z)}{\Pr[Z=z|X,S,I(S\in\mathcal{S})=1]}-\frac{I(Z=z')}{\Pr[Z=z'|X,S,I(S\in\mathcal{S})=1]}\right)\frac{I(S\in\mathcal{S})Y\Pr[S=0|X]}{\Pr[I(S\in\mathcal{S})=1|X]}\right],$$

which completes the proof. □

Remark. *The derivation for Proposition 4 does not use any conditions that involve potential outcomes and thus holds whether or not $\phi(z, z')$ has a causal interpretation.*

eAppendix C Addressing non-adherence in the randomized trials

In this section we sketch an approach to transporting per-protocol effects from a collection of randomized trials to a target population. We only consider the simple case of a binary adherence indicator (all-or-nothing adherence) and an outcome measured at the end of the study, because the strategy for combining adjustments for non-adherence and transportability in this simple case readily extends to more complicated non-adherence patterns (multiple time periods, more than two treatment recommendations, etc.), using well-known methods for the analysis of the randomized trials with non-adherence [56, 57].

C.1 Setup and notation

We begin by introducing some notation to represent non-adherence and define per-protocol effects. In each trial S, information is collected on baseline covariates X, treatment assignment Z, post-treatment assignment covariates L, treatment received A, and the outcome Y. As in the main text, lower case letters denote realizations of the corresponding random variables. The set of possible assigned treatments is $\mathcal{Z}$; the set of possible received treatments is $\mathcal{A}$. Each specific pair of assigned and received treatment is denoted as (z, a), $z \in \mathcal{Z}$, $a \in \mathcal{A}$.

We are now interested in counterfactual outcomes under joint intervention to assign treatment z and enforce the receipt of treatment a; we denote these counterfactual outcomes as $Y^{z,a}$. The average treatment effect comparing two different joint interventions, (z, a) and (z', a'), in the target population is $\mathrm{E}[Y^{z,a} - Y^{z',a'} | S = 0]$. For example, suppose that we are evaluating a binary treatment; then, in our setup, the most interesting causal contrast is arguably the per-protocol effect in the target population, $\mathrm{E}[Y^{z=1,a=1} - Y^{z'=0,a'=0} | S = 0]$.

C.2 Identifiability conditions for transporting inferences from a single trial

We assume that the following identifiability conditions hold for some trial $s^* \in \mathcal{S}$:

C1. Consistency of potential outcomes: For every individual i in the trials or target population, for every $z \in \mathcal{Z}$ and every $a \in \mathcal{A}$, if $Z_i = z$ and $A_i = a$, then $Y_i^{z,a} = Y_i$.

C2. Sequential conditional exchangeability for assignment and receipt of treatment in the trial: for trial $s^* \in \mathcal{S}$, each treatment assignment $z \in \mathcal{Z}$, and each treatment received $a \in \mathcal{A}$, we have that $\mathrm{E}[Y^{z,a}|X = x, S = s^*, Z = z] = \mathrm{E}[Y^{z,a}|X = x, S = s^*]$ for each x with $f(x, S = s^*) > 0$; and $\mathrm{E}[Y^{z,a}|X = x, S = s^*, Z = z, L = l, A = a] = \mathrm{E}[Y^{z,a}|X = x, S = s^*, Z = z, L = l]$ for each x and l with $f(x, S = s^*, Z = z, l) > 0$.

C3. Sequential positivity of the probability of treatment assignment and receipt of treatment in the randomized trial: for trial $s^* \in \mathcal{S}$ we have that $\Pr[Z = z|X = x, S = s^*] > 0$ for every x with $f(x, S = s^*) > 0$; and for each treatment $a \in \mathcal{A}$, $\Pr[A = a|X = x, S = s^*, Z = z, L = l] > 0$ for every x and l with $f(x, S = s^*, Z = z, L = l) > 0$.

C4. Conditional exchangeability in measure for the per-protocol effect: for trial s^*, for each pair of treatment assignments $z \in \mathcal{Z}$ and $z' \in \mathcal{Z}$, and each pair of treatments received $a \in \mathcal{A}$ and $a' \in \mathcal{A}$,

$$\mathrm{E}[Y^{z,a} - Y^{z',a'}|X = x, S = 0] = \mathrm{E}[Y^{z,a} - Y^{z',a'}|X = x, S = s^*]$$

for every x with $f(x, S = 0) > 0$.

C5. Positivity of the probability of participation in the trial: $\Pr[S = s^*|X = x] > 0$ for every x with $f(x, S = 0) > 0$.

Conditions *C1* through *C3* are "standard" conditions used for the analysis of randomized trials with non-adherence [29]. Condition *C4* is a per-protocol effect version of condition *A4* for intention-

to-treat effects given in the main text; and condition *C5* is the same as condition *A5* given in the main text.

C.3 Identification of per-protocol effects in the target population

We will now show that under conditions *C1* through *C5* per-protocol effects in the target population, $\mathrm{E}[Y^{z,a} - Y^{z',a'}|S = 0]$, are identifiable using randomized trial data on $(X, S = s^*, Z, L, A, Y)$ from the trial s^*, and data on $(X, S = 0)$ from the sample of the target population.

Using conditions *C4* and *C5*, and taking expectations over the conditional distribution of X in the target population, we have

$$\begin{aligned}
&\mathrm{E}[Y^{z,a} - Y^{z',a'}|S = 0] \\
&\quad = \mathrm{E}\big[\,\mathrm{E}[Y^{z,a} - Y^{z',a'}|X, S = 0]\big|S = 0\big] \\
&\quad = \mathrm{E}\big[\,\mathrm{E}[Y^{z,a} - Y^{z',a'}|X, S = s^*]\big|S = 0\big] \\
&\quad = \mathrm{E}\big[\,\mathrm{E}[Y^{z,a}|X, S = s^*] - \mathrm{E}[Y^{z',a'}|X, S = s^*]\big|S = 0\big].
\end{aligned} \tag{C.1}$$

To complete the identification analysis, note that the only terms involving counterfactual quantities in the last expression of display (C.1) are the trial-specific conditional potential outcome means $\mathrm{E}[Y^{z,a}|X, S = s^*]$, for every $z \in \mathcal{Z}$ and every $a \in \mathcal{A}$. Using conditions *C1* through *C3*, we have

$$\begin{aligned}
\mathrm{E}[Y^{z,a}|X, S = s^*] &= \mathrm{E}[Y^{z,a}|X, S = s^*, Z = z] \\
&= \mathrm{E}\big[\,\mathrm{E}[Y^{z,a}|X, S = s^*, Z = z, L]\big|X, S = s^*, Z = z\big] \\
&= \mathrm{E}\big[\,\mathrm{E}[Y^{z,a}|X, S = s^*, Z = z, L, A = a]\big|X, S = s^*, Z = z\big] \\
&= \mathrm{E}\big[\,\mathrm{E}[Y|X, S = s^*, Z = z, L, A = a]\big|X, S = s^*, Z = z\big].
\end{aligned} \tag{C.2}$$

Define $\theta(z, a, s^*; X) \equiv \mathrm{E}\big[\,\mathrm{E}[Y|X, S = s^*, Z = z, L, A = a]\big|X, S = s^*, Z = z\big]$ for each $z \in \mathcal{Z}$ and

$a \in \mathcal{A}$. Combining the results from displays (C.1) and (C.2), we have

$$\mathrm{E}[Y^{z,a} - Y^{z',a'}|S = 0] = \mathrm{E}\Big[\theta(z, a, s^*; X) - \theta(z', a', s^*; X)\Big|S = 0\Big],$$

which establishes the identifiability of per-protocol effects when transporting inferences from a randomized trial to a target population.

C.4 Transporting inferences from a collection of trials

Suppose that the identifiability conditions *C1* through *C5* hold for every trial $s \in \mathcal{S}$. Then, we would have

$$\mathrm{E}[Y^{z,a} - Y^{z',a'}|X, S = 1] = \ldots = \mathrm{E}[Y^{z,a} - Y^{z',a'}|X, S = m] = \mathrm{E}[Y^{z,a} - Y^{z',a'}|X, S = 0]. \tag{C.3}$$

Using the results from the previous section, we obtain

$$\theta(z, a, 1; X) - \theta(z', a', 1; X) = \ldots = \theta(z, a, m; X) - \theta(z', a', m; X) = \mathrm{E}[Y^{z,a} - Y^{z',a'}|X, S = 0]. \tag{C.4}$$

Note that, similar to our results in the main text, the chain of equalities

$$\theta(z, a, 1; X) - \theta(z', a', 1; X) = \ldots = \theta(z, a, m; X) - \theta(z', a', m; X) \equiv \lambda(z, a, z', a'; X), \tag{C.5}$$

is an observed data implication of assuming transportability of the per-protocol effects across the collection $\mathcal{S}$, and we use the notation $\lambda(z, a, z', a'; X)$ for that common (across trials) quantity.

Using the law of total expectation, and the identification result for a single trial, we obtain the

following identification result for transporting the entire collection of trials

$$\mathrm{E}[Y^{z,a} - Y^{z',a'}|S=0] = \mathrm{E}[\lambda(z,a,z',a';X)|S=0].$$

Obtaining a weighting re-epression of these results is also instructive. For the single trial $s^* \in \mathcal{S}$, and every $z \in \mathcal{Z}$

$$\begin{aligned}
&\theta(z,a,s^*;X) \\
&\quad \equiv \mathrm{E}\Big[\,\mathrm{E}[Y|X,S=s^*,Z=z,L,A=a]\Big|X,S=s^*,Z=z\Big] \\
&\quad \mathrm{E}\left[\frac{I(Z=z,A=a)}{\Pr[Z=z|X,S=s^*]\Pr[A=a|X,S=s^*,Z=z,L]}Y\Big|X,S=s^*\right] \\
&\quad \mathrm{E}\left[\frac{I(Z=z,A=a)}{\Pr[Z=z|X,S,I(S\in\mathcal{S})=1]\Pr[A=a|X,S,Z=z,L,I(S\in\mathcal{S})=1]}Y\Big|X,S=s^*,I(S\in\mathcal{S})=1\right],
\end{aligned} \tag{C.6}$$

and consequently,

$$\begin{aligned}
&\theta(z,a,s^*;X) - \theta(z',a',s^*;X) \\
&\quad = \mathrm{E}\left[\left(\frac{I(Z=z,A=a)}{\Pr[Z=z|X,S,I(S\in\mathcal{S})=1]\Pr[A=a|X,S,Z=z,L,I(S\in\mathcal{S})=1]}\right.\right. \\
&\quad \left.\left. - \frac{I(Z=z',A=a')}{\Pr[Z=z'|X,S,I(S\in\mathcal{S})=1]\Pr[A=a'|X,S,Z=z',L,I(S\in\mathcal{S})=1]}\right)Y\,\middle|\,X,S=s^*,I(S\in\mathcal{S})=1\right].
\end{aligned} \tag{C.7}$$

It follows that, if all trials $s \in \mathcal{S}$ are transportable to the target population, then

$$\begin{aligned}
\lambda(z,a,z',a';X) = \mathrm{E}\left[\left(\frac{I(Z=z,A=a)}{\Pr[Z=z|X,S,I(S\in\mathcal{S})=1]\Pr[A=a|X,S,Z=z,L,I(S\in\mathcal{S})=1]}\right.\right. \\
\left.\left. - \frac{I(Z=z',A=a')}{\Pr[Z=z'|X,S,I(S\in\mathcal{S})=1]\Pr[A=a'|X,S,Z=z',L,I(S\in\mathcal{S})=1]}\right)Y\,\middle|\,X,I(S\in\mathcal{S})=1\right].
\end{aligned}$$

Finally, using the above result, the identification result for the collection of trials becomes

$$\mathrm{E}[Y^{z,a} - Y^{z',a'}|S=0] = \mathrm{E}\Bigg[\ \mathrm{E}\Bigg[\Bigg(\frac{I(Z=z,A=a)}{\Pr[Z=z|X,S,I(S\in\mathcal{S})=1]\Pr[A=a|X,S,Z=z,L,I(S\in\mathcal{S})=1]}$$
$$-\frac{I(Z=z',A=a')}{\Pr[Z=z'|X,S,I(S\in\mathcal{S})=1]\Pr[A=a'|X,S,Z=z',L,I(S\in\mathcal{S})=1]}\Bigg)Y\Bigg|X,I(S\in\mathcal{S})=1\Bigg]\Bigg|S=0\Bigg],$$

or equivalently,

$$\mathrm{E}[Y^{z,a} - Y^{z',a'}|S=0] = \frac{1}{\Pr[S=0]}\,\mathrm{E}\Bigg[\Bigg(\frac{I(Z=z,A=a)}{\Pr[Z=z|X,S,I(S\in\mathcal{S})=1]\Pr[A=a|X,S,Z=z,L,I(S\in\mathcal{S})=1]}$$
$$-\frac{I(Z=z',A=a')}{\Pr[Z=z'|X,S,I(S\in\mathcal{S})=1]\Pr[A=a'|X,S,Z=z',L,I(S\in\mathcal{S})=1]}\Bigg)\frac{I(S\in\mathcal{S})Y\Pr[S=0|X]}{\Pr[I(S\in\mathcal{S})=1|X]}\Bigg].$$

eAppendix D Code to implement the estimators

We have provided `R` code that can be modified to implement the methods in a new dataset and a simulated dataset that illustrates the input data structure. The code is available at https://github.com/serobertson/TransportingMultipleTrials.