

Indirect Treatment Comparison Without Network Meta-Analysis: Overview of Novel Techniques

K. Jack Ishak, PhD, Executive Director, Statistics and Senior Research Scientist

BACKGROUND

There is generally a paucity of evidence about the relative effectiveness of a new treatment and its competitors. And yet, this is a critical consideration in reimbursement decisions as well as in the planning of future research. In the absence of head-to-head studies, comparative evidence is derived through indirect comparisons, relying on common comparators to link data from trials of the various treatments of interest. That is, treatments A and B, which were compared to treatment C in their respective trials, can be indirectly compared to each other by contrasting effects of A vs. C to that of B vs. C. Network meta-analysis or mixed/indirect treatment comparison

(MTC) is the standard technique used for this purpose. This approach is broadly used and accepted by the research community as well as health technology assessment agencies, in part because it can incorporate data from all competing treatments in a therapeutic area, thus reflecting the totality of evidence that is available.

In some cases, however, MTCs may not be able to produce the comparisons of interest (i.e., when common comparators were not available), or may be subject to limitations (e.g., heterogeneity between trials) affecting the reliability of the results. Two alternative approaches—simulated treatment comparisons (STCs)¹ and Matching

Adjusted Indirect Comparisons (MAICs)² can overcome these issues by making *targeted* comparisons of outcomes observed for the new treatment and those observed in the treatment arms of the comparators of interest. Thus, the units of analysis in these targeted comparisons are outcome measures like event rates rather than effect estimates like hazard ratios as in MTCs. This poses an important challenge, however; outcomes observed in treatment arms from different studies are not necessarily comparable. These not only reflect the effects of the treatments received but are also impacted by the profiles of the populations and possibly design

EXAMPLE OF AN EVIDENCE NETWORK based on four studies providing indirect data about the relative effectiveness of treatments A and B

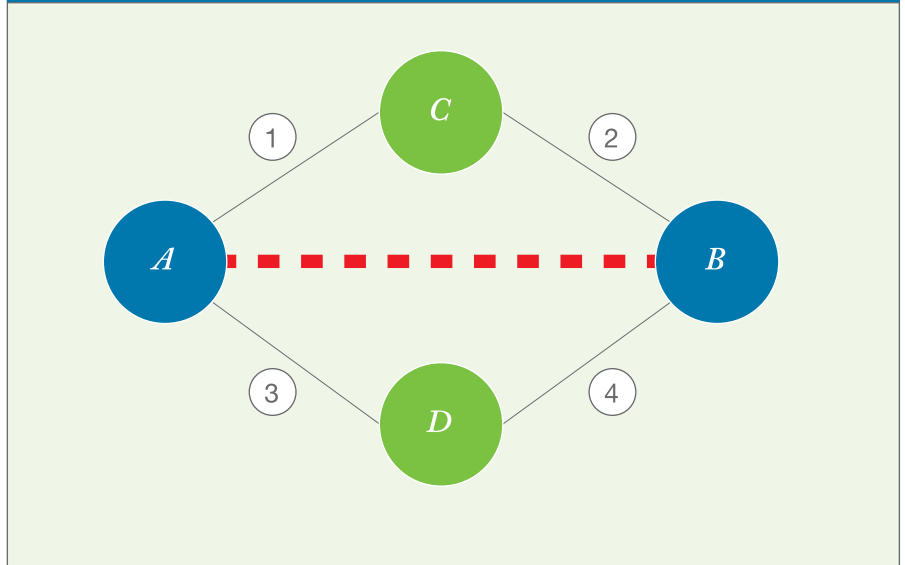


figure 1

STCs AND MAICs CAN BE APPLIED, OR AT LEAST CONSIDERED AND ASSESSED FOR FEASIBILITY, IN SITUATIONS WHERE STANDARD TECHNIQUES HAVE SIGNIFICANT LIMITATIONS OR CANNOT BE APPLIED AT ALL.

features of the studies. STCs and MAICs are designed to deal with these issues and produce reliable comparisons by making analytical adjustments to balance the populations being compared. Unlike MTCs which rely only on published data, these novel methods require patient-level data on at least one of the treatments to be able to adjust for differences in populations.

WHEN SHOULD STCs OR MAICs BE CONSIDERED?

STCs and MAICs can be applied, or at least considered and assessed for feasibility, in situations where standard techniques have significant limitations or cannot be applied at all. Three specific scenarios are described below.

HETEROGENEITY

Figure 1 illustrates a simple evidence network (i.e., representation of the studies and treatments involved in the MTC) to evaluate a comparison of treatments A and B. The network includes four studies, identified

by lines connecting the treatments compared in each of these. For instance, trial ① compared treatment A to C, and trial ④ compared treatment B to D. Thus, the indirect comparison of A and B (represented by the dashed red line) is informed by the relative effects of these treatments to their effects compared with common comparators C and D. Suppose, however, trials ② and ③ have similar populations and design, and differ significantly from the other two studies.

Such variation causes *heterogeneity* in the results being pooled and compared, which is dealt with in MTCs by adding parameters that account for *excess variability* in results. It is assumed, however, that differences between trials only cause random fluctuation, so that the indirect comparison derived from the MTC effectively averages over differences in populations, design features, measurement techniques, etc., across studies. This can be problematic, however, when there is significant

EXAMPLE OF AN INCOMPLETE OR DISCONNECTED EVIDENCE NETWORK in which treatments A and B cannot be linked through common comparators

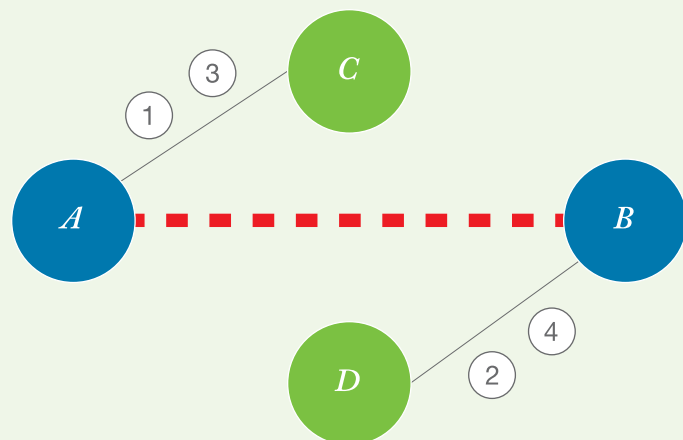


figure 2

heterogeneity, and specific differences that may distort results are noted. Published data are often too limited to allow a closer examination and adjustment for such factors in MTCs.

STCs and MAICs can deal with this type of heterogeneity by focusing the comparison of the studies that are deemed more closely comparable—② and ③ in this example. Outcomes observed for treatment B in study ② are compared with outcomes for C in study ③. It is possible that the profiles of the populations of these studies may differ, even if only due to chance and requires adjustment to obtain an unbiased comparison. The way this is handled in each approach is further described below.

INCOMPLETE EVIDENCE NETWORK

STCs and MAICs would also be useful in situations where the evidence network is incomplete or disconnected. That is, the treatments to be compared cannot be linked through common comparators. This is illustrated in

Figure 2, where two trials comparing A to C and two trials comparing B to D make up the evidence network. Since the comparators in the trials of A and B are different, it is impossible to obtain an indirect comparison of these treatments with an MTC. Approaches like STC or MAIC may be the only way to achieve an indirect comparison in these situations, since this would be obtained from a targeted comparison of the specific arms of interest in the trials of A and B. This may be done by selecting two specific trials that are most compatible, as in the example of the previous section, or by using data from all four of the trials, and pooling data as appropriate to serve as the basis of analyses.

MULTI-STEP COMPARISON

STCs and MAICs may also be useful in situations where the treatments of interest can only be linked through multiple intermediate comparisons. This is illustrated in *Figure 3*. In this evidence network, trials of A and B do not have a common comparator,

STCs AND MAICs WOULD ALSO BE USEFUL IN SITUATIONS WHERE THE EVIDENCE NETWORK IS INCOMPLETE OR DISCONNECTED. THAT IS, THE TREATMENTS TO BE COMPARED CANNOT BE LINKED THROUGH COMMON COMPARATORS.

and must rely on trials that compared their respective comparators to make the link. That is, A is linked to B through a comparison of C to E and F and D (i.e., A vs. C, C vs. E, E vs. B, and A vs. F, F vs. D and D vs. B). The reliability of MTCs in this situation may be compromised as heterogeneity may impact comparisons at intermediate steps and distort the main comparison of interest. The problem is amplified as the number of steps involved to link treatments increases (e.g., to link A to D in *Figure 3*). The targeted comparisons involved in STCs and MAICs bypass the issue by targeting the analyses on specific arms of interest, as long as the trials of treatment A and B can be considered sufficiently compatible for a targeted comparison.

WHEN ARE STC AND MAIC FEASIBLE?

The first consideration in the assessment of the feasibility of these novel methods lies in the availability of patient-level data

on at least one of the treatments being compared. This should generally be possible when these analyses are initiated by a manufacturer. One or more trials of the manufacturer's product (the index trial(s)) would then serve as the basis of the STC and MAIC and would be used to adjust for differences in populations of comparators' trials. In most situations data on the comparator treatments will only be available from publications. This is not a limiting factor, as long as information on the profile of the population and outcomes of interest are reported with adequate detail.

In addition to the availability of patient-level data, the feasibility of these novel techniques depends on the availability of one or more compatible studies for comparators of interest. Compatibility is determined based on the similarity of the populations and the designs of the trials. It is not necessary for the populations to be identical, since the methods are designed

to balance differences. This can only be done, however, when there is sufficient overlap in the profiles of the two samples. For example, a difference of 20% in the proportion of male patients in the two trials is not problematic, but the comparison may be unreliable if one study was based on male patients and the other on female patients. Similarly, the duration of the trials and timing of measurements should be similar but not necessarily identical, and likewise for other design features such as admissibility criteria, concomitant medications, treatment protocols, etc.

Finally, reliable application of STCs and MAICs requires that all determinants of the outcomes of interest that may confound the comparison are available in both the index trial data and reported in the publication(s) for the comparator(s) (which will be in summary form, such as means and percentages). The results are subject to residual confounding in cases where determinants are available in one but not both sources.

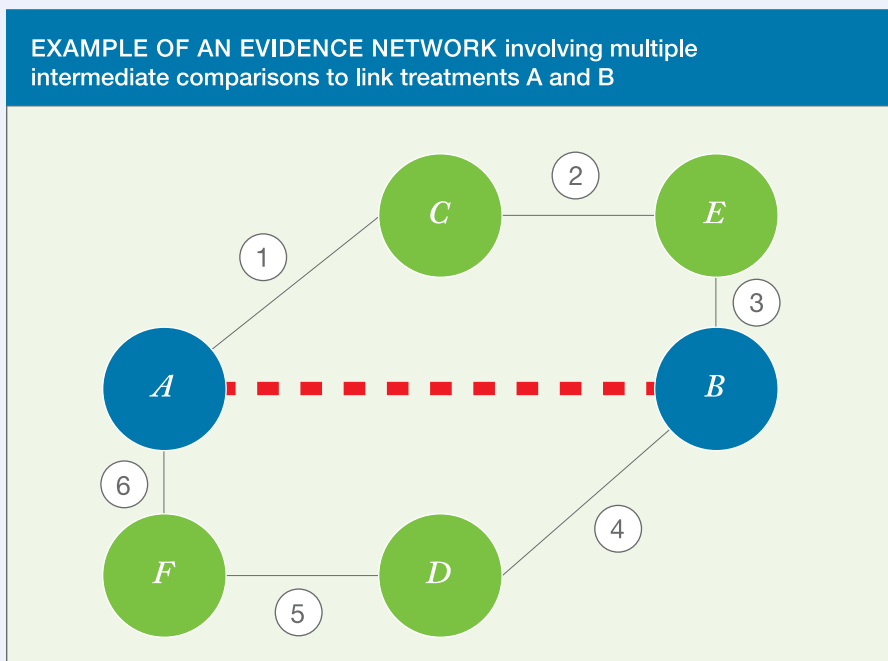


figure 3

CONCEPTUAL REPRESENTATION of how comparisons are derived with adjustment for differences in population profiles with STC and MAIC

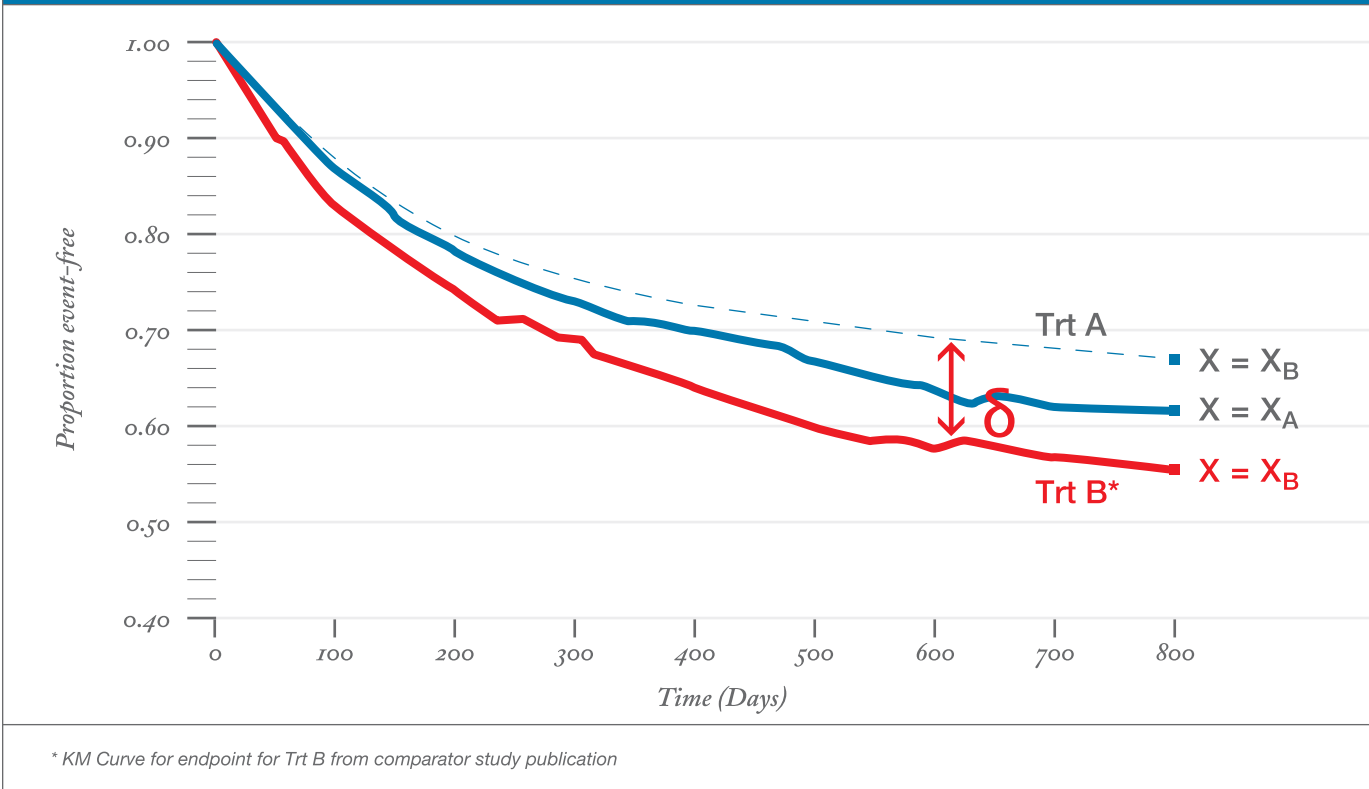


figure 4

HOW DO STCs AND MAICs WORK?

STCs and MAICs are very similar conceptually. *Figure 4* shows a representation of how balanced comparisons are derived in STCs and MAICs. In this illustration, the outcome of interest is a time-to-event endpoint. The solid blue line represents the time-to-event distribution from the index trial of treatment A, while the red solid line represents the distribution for the comparator B obtained from a published report or manuscript. A comparison of these lines is biased by the fact that the profile of the population represented in the blue line (denoted by X_A) may differ, even if only by chance to the profile in the red line (X_B). Thus, to adjust for potential imbalances, these methods

aim to generate an adjusted time-to-event curve that reflects what outcomes may have been with treatment A in a population that matches the profile for treatment B. This is represented by the dashed blue line, which can now be compared directly with the observed outcomes for treatment B (i.e., red line) to measure the relative effectiveness of A and B (denoted by δ).

STCs and MAICs differ in the way they generate the adjusted outcomes for treatment A (dashed blue line). STC accomplishes this by creating a predictive equation for each outcome being compared. The equations are then used to predict outcomes that would have been observed for treatment A in patients with characteristics matching those in X_B . That is, the adjusted line is

produced by setting predictors to their corresponding values in X_B .

MAICs deal with the adjustment by reweighting patients in the index trial so that the weighted average values of determinants of outcomes in the index trial (i.e., X_A) match X_B . These weights are derived from a propensity-score-type analysis using the index trial data, predicting membership into the index vs. comparator's trial. An individual weight is then predicted for each patient in the index trial, and applied in Kaplan-Meier analyses (for example) to generate the adjusted curve.

The methods can be applied following the same process with all types of outcomes (e.g., continuous or dichotomous measures). Furthermore, both approaches produce an estimate



of the relative effectiveness along with measures of uncertainty, like standard errors or confidence intervals.


WHEN TO CHOOSE STCs VS. MAICs?

STCs and MAICs are conceptually very similar and use the same data to accomplish the goal of adjustment for potential confounding. It is, therefore, reasonable to expect that the two methods would yield very similar results. (This is, indeed, what we have observed in actual analyses.) Differences between STCs and MAICs lie in potential efficiencies associated with each approach.

STCs involve generating predictive equations for each of the outcomes of interest. The identification of predictors is added insight and the equations themselves may have utility in other applications. For instance, the equations can be

integrated into a simple disease model to serve as the basis of a trial simulation tool allowing the evaluation of designs for future studies (e.g., to test different population profiles). STCs can be more efficient than MAICs in situations where comparisons with multiple comparators are to be made for a given set of outcomes. Equations for the outcomes would be derived once from the index trial and applied with data from each comparator treatment's study. With MAIC, a separate set of weights would be required for each comparator treatment's study population. By the same token, MAICs would offer efficiencies in situations where there is a single comparator but many outcomes to be compared. A single set of weights would be required to balance the two populations, and could be applied in analyses for each outcome.

SUMMARY

STCs and MAICs are robust and reliable methods to derive indirect comparisons between treatments. These novel methods can produce comparative evidence in situations where standard techniques are inadequate, but can also be complementary to NMA or MTC, providing a more targeted assessment of the relative effectiveness of the treatments. Whereas the MTC may provide an averaged effect estimate, by using the index trial as the basis of the analysis, the STC or MAIC reflects the relative effectiveness that might have been observed if the comparator had been included as an additional arm in the index trial. 

For more information, please contact Jack.Ishak@evidera.com.

References

¹ Caro JJ, Ishak KJ. No Head-to-Head Trial? Simulate the Missing Arms. *Pharmacoeconomics*. 2010 Oct 1; 28(10):957–967.

² Signorovitch JE, Sikirica V, Erder MH, et al. Matching-Adjusted Indirect Comparisons: A New Tool for Timely Comparative Effectiveness Research. *Value Health*. 2012 Sep–Oct; 15(6):940–947.