Methodological Challenges with Conducting Network Meta-Analyses Assessing Long-term Comparative Efficacy in Psoriasis

A Critique of Assumptions Underpinning Recent Indirect Treatment Comparisons

Tim Disher; Steve Peterson; Kiefer Eaton; Sandhya Nair; Reginald Villacorta; Brian Hutton; Chris Cameron; Jeroen P Jansen

CONFLICT OF INTEREST

Tim Disher is an employee of EVERSANA. EVERSANA consults for various pharmaceutical and medical device companies. Development of the content for this presentation was funded by Janssen.

Background

THE DISEASE:

Plaque psoriasis is a chronic and disfiguring dermatological condition associated with autoimmune-mediated inflammation of the skin which can have a significant impact on patients' quality of life.[1]

THE ANALYSES:

Given the lack of head to head trials, network meta-analysis (NMA) is required to assess the relative effectiveness and safety of therapies for psoriasis.[2–6]

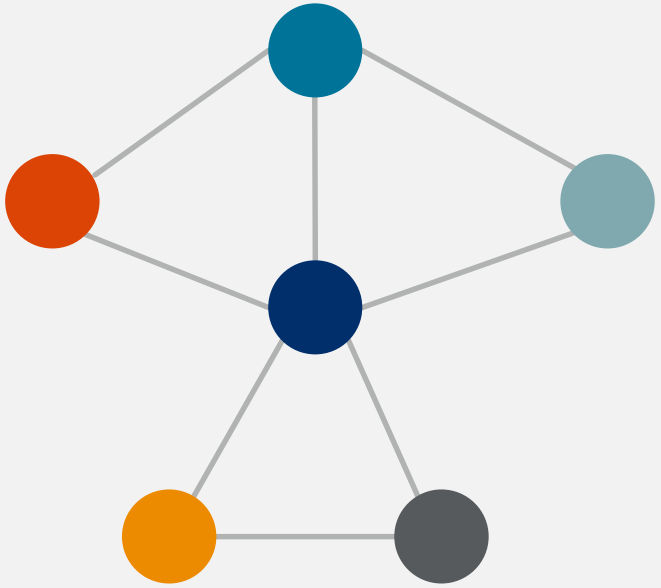
THE ISSUES:

While several past NMAs[7–13] have assessed the short-term comparative efficacy of treatments during the placebo-controlled component of treatment, few have assessed their related long-term benefits.

Challenges in Long-term NMAs

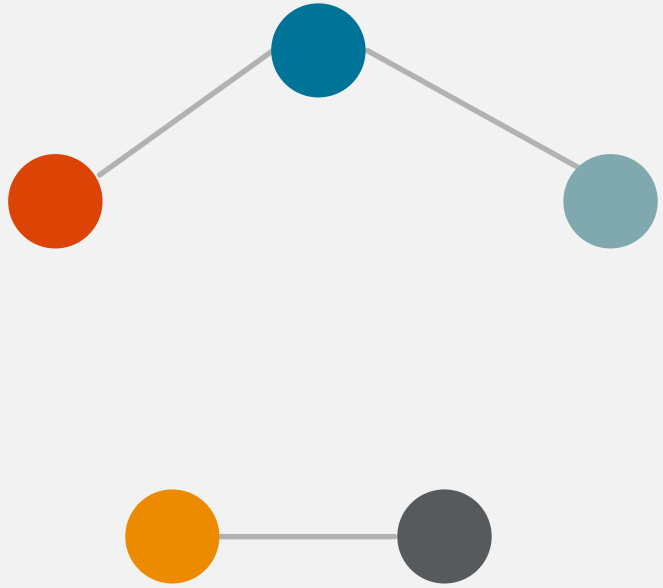
A key reason for the lack of long-term analysis is that trial designs typically include a planned cross-over to active treatment after an initial placebo-controlled period.

A. Placebo-Controlled Period (10-16 weeks)



Studies in the placebo-controlled period form a single connected network via comparisons to placebo.

B. Long-term Period (52 weeks)



At the end of the placebo-controlled period, placebo patients are transitioned to an active comparator resulting in two or more sub-networks.

Objectives

NMA Overview

Provide a brief overview of the rationale and framework for NMA

PsO Long-Term NMAs

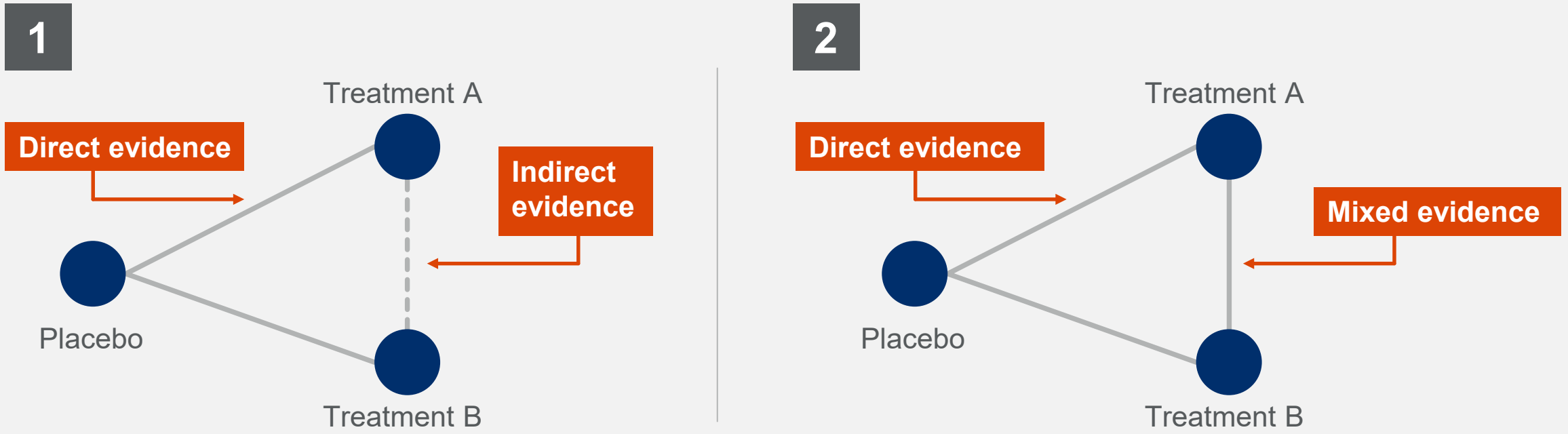
Review applications in PsO and assess appropriateness of assumptions

Alternatives and Next Steps

Conclude with a summary of the difficulties and discuss the potential for future approaches

Overview of NMA

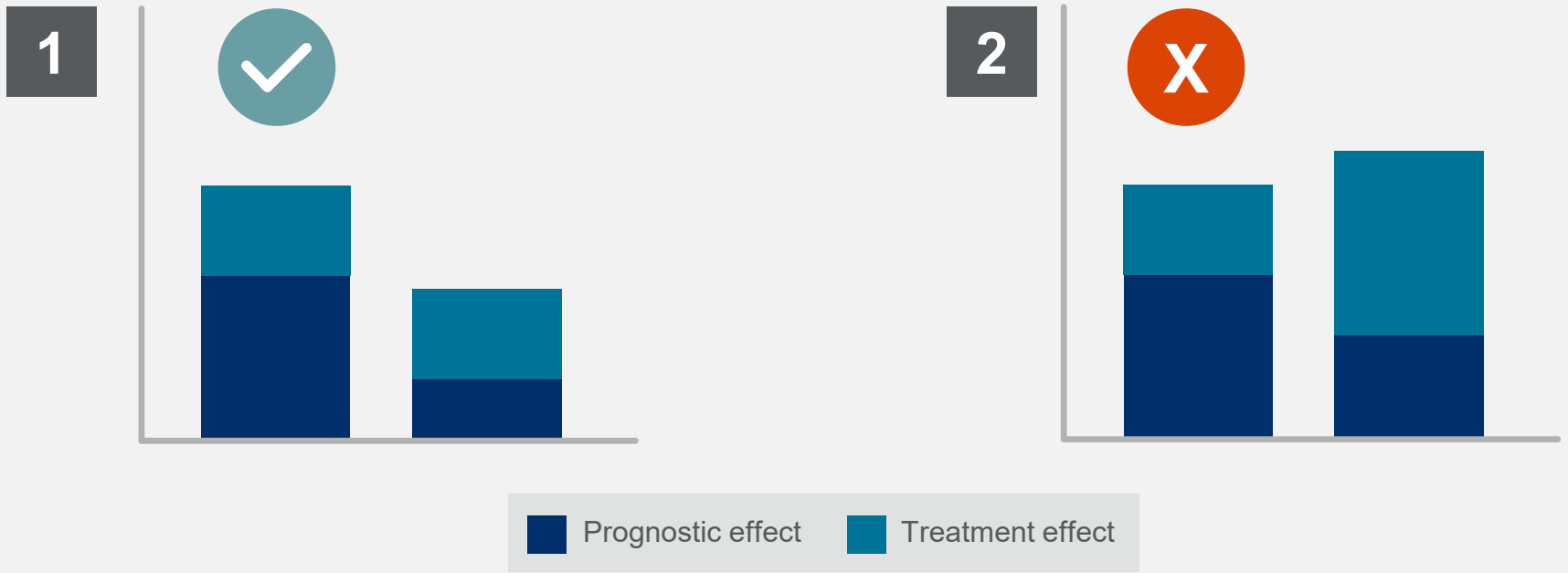
- In the absence of head to head trials (direct evidence), researchers and clinicians sometimes resort to comparisons of absolute outcomes between active arms of interest.
  - These naïve treatment comparisons should be avoided because they eliminate the randomization of RCTs.[14,15]
- NMA allows for comparisons between treatments based on indirect evidence when no direct evidence exists (1), and combines both direct and indirect evidence when both types are present (2)



EVERSANA™

Assumptions of NMA

- Comparisons in NMAs are anchored through a common comparator, which preserves randomization of the included studies
- Relative effects will be unbiased even when trials differ in prognostic variables (1) as long as they are similar with regard to effect modifiers (2).



EVERSANA™

Three Approaches to Analysis of Long-term Networks in Psoriasis

- Planned cross-over from placebo to active treatment in long-term networks for psoriasis result in disconnected networks that don't allow for standard anchored comparisons
- At least three approaches based on available published data have been used to address this in psoriasis

ARMSTRONG[16]

Conducted naïve indirect comparisons based on meta-analysis of individual active treatment arms

DIELS[17]

Assumed that all tumor necrosis factor inhibitors had the same effect, allowing all comparisons to be anchored through them

SAWYER[7]

Anchored long-term comparisons to week 16 placebo results, allowing for anchored comparisons

Methodological Approaches to Adjustment

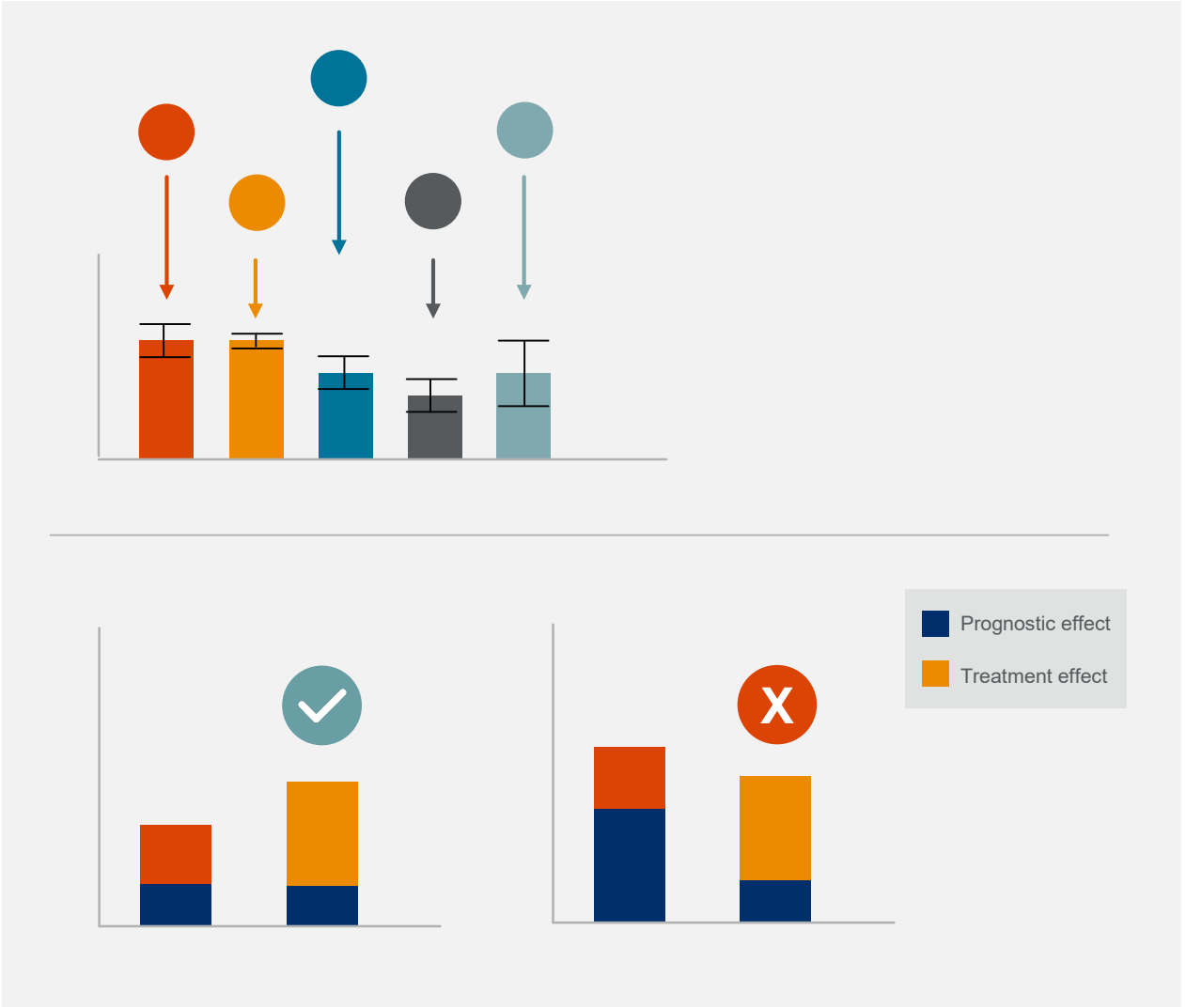
Armstrong et al

Approach: Meta-analysis of active arms of comparators followed by naïve indirect comparison

Strengths: Easy to implement

Limitations:

- Adds assumption of balance on prognostic variables
- Similar to comparing unadjusted observational trials
- Typically considered a fatal flaw in the presence of more reasonable alternatives



Methodological Approaches to Adjustment

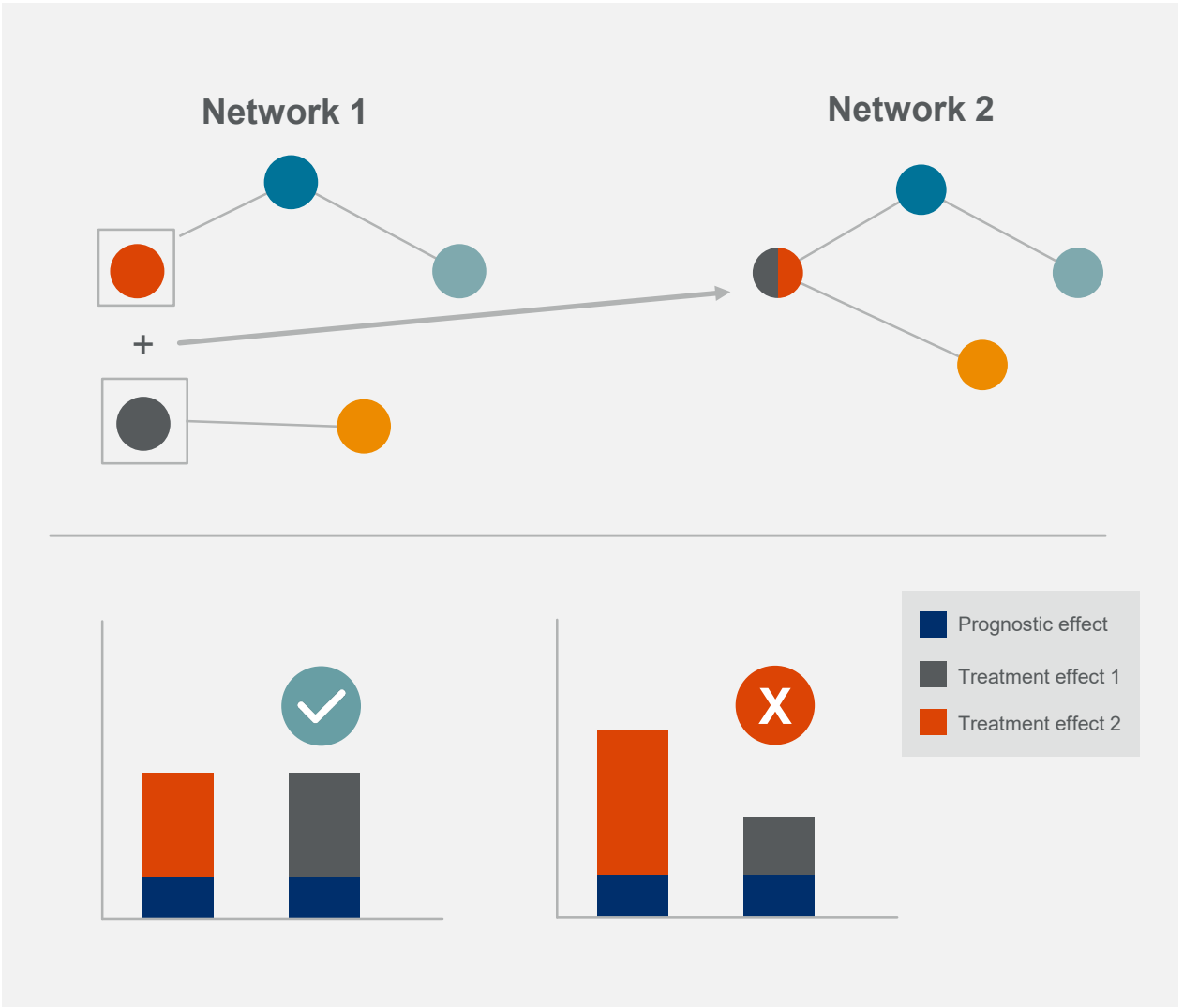
Diels et al

Approach: Treatments in the same class of drugs are combined together, allowing for a connected network

Strengths: Leverages assumptions regarding equivalent effects within mechanisms of action to reduce the analysis to a standard NMA

Limitations:

- Requires the assumption of a class effect in order to be valid
  - Note that validity of results still may apply if there is certain variability across TNF use.



Methodological Discussion

Sawyer et al

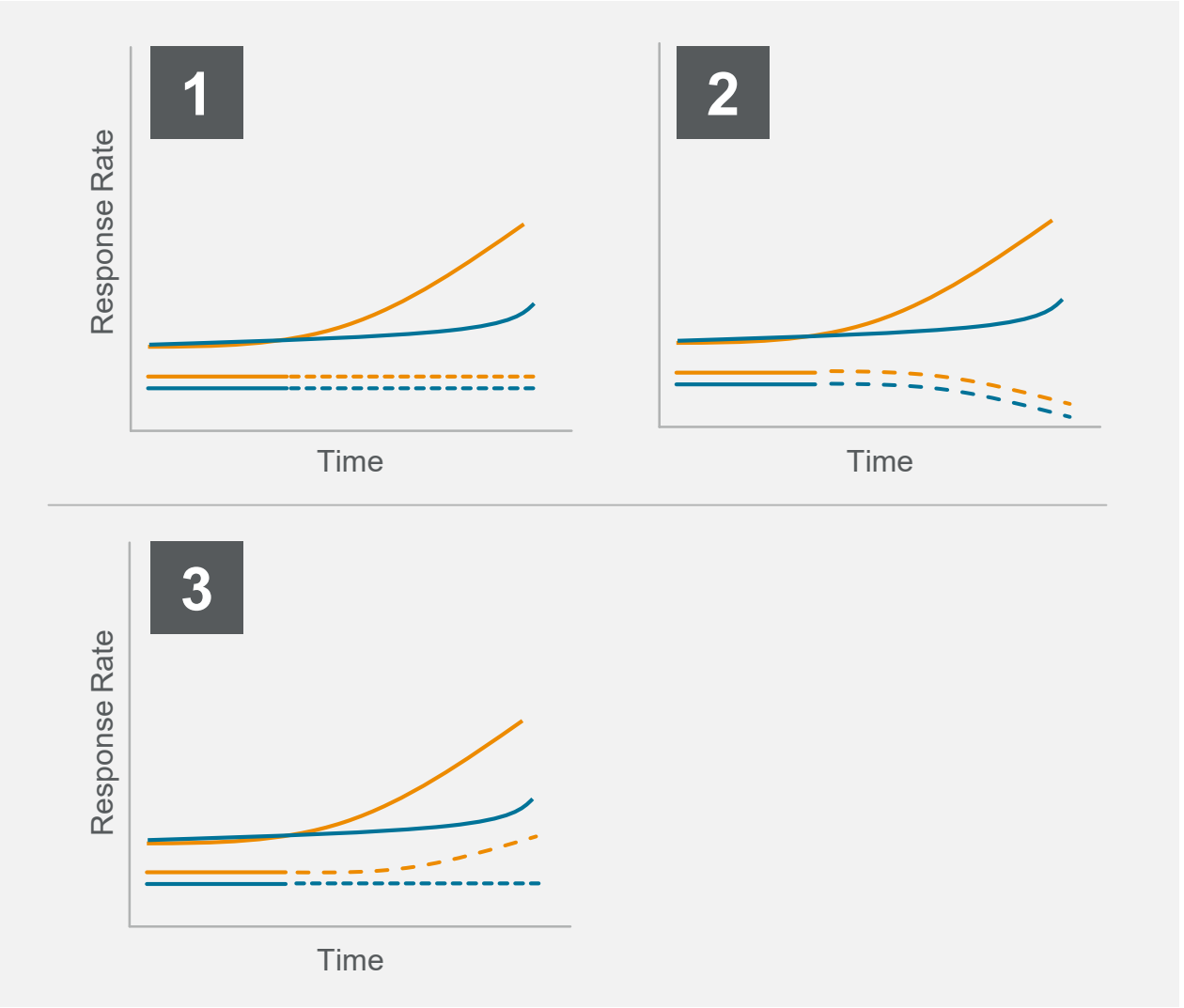
Approach: Anchor long term comparisons to last observed placebo response

Strengths:

- Maintains randomization
- All comparisons are valid if placebo would have stayed constant (1)
- ITCs are valid if placebo response would change in the same way (2)

Limitations:

- If changes in best-supportive care or other trial characteristics could lead to different placebo trends over time, comparisons will be biased (3)
- Need to be careful to avoid double counting, over-precision



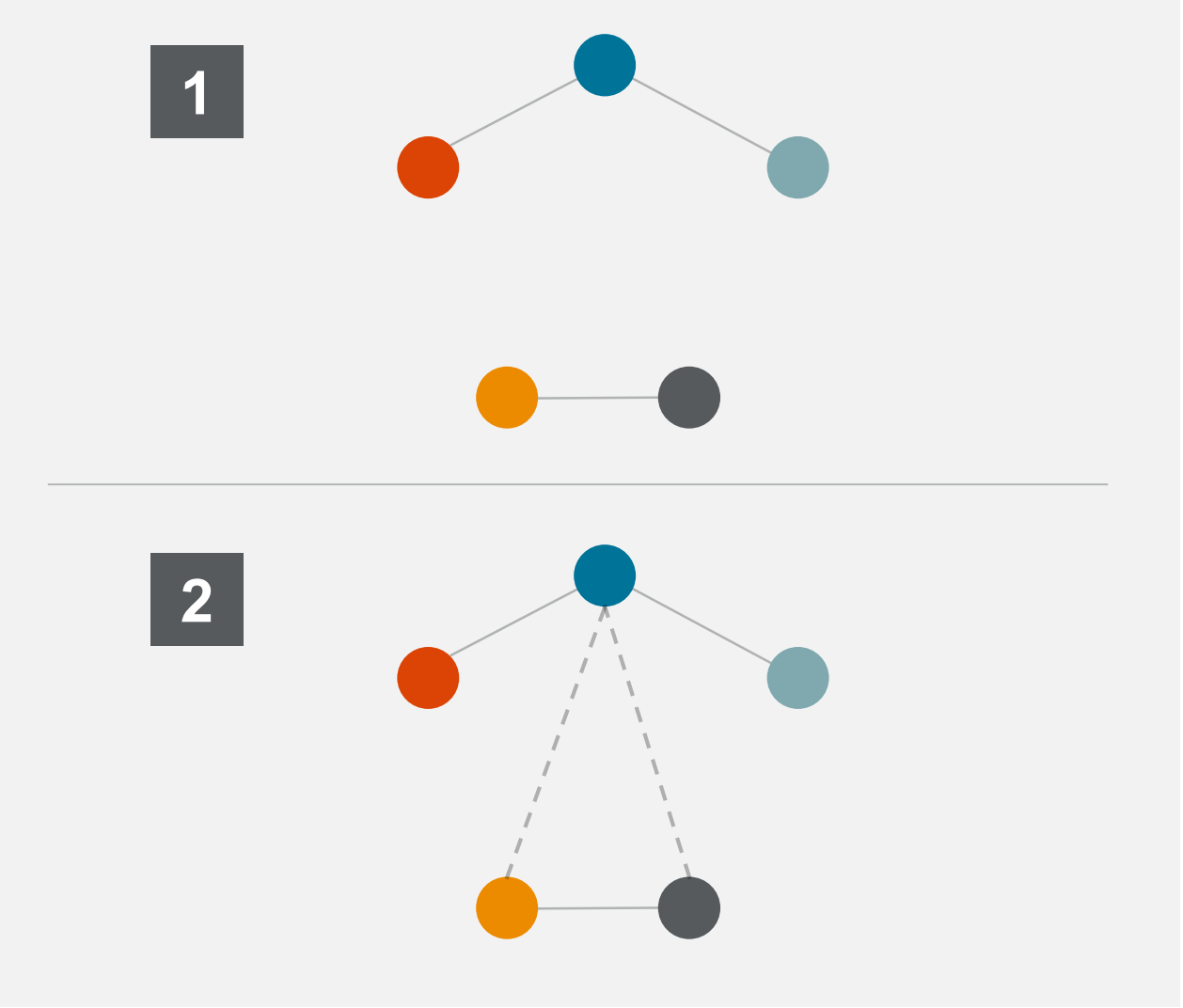
Alternative Approaches

Random Baseline Effects

Approach: Place a model on the baseline treatment response, and impute that baseline into each comparison

Strengths: If the baseline model is correct, inferences will be similar to those from a connected network

Limitations: Allows for information to be shared across trials, breaking randomization and biasing results if the model is incorrect



EVERSANA™

Alternative Approaches

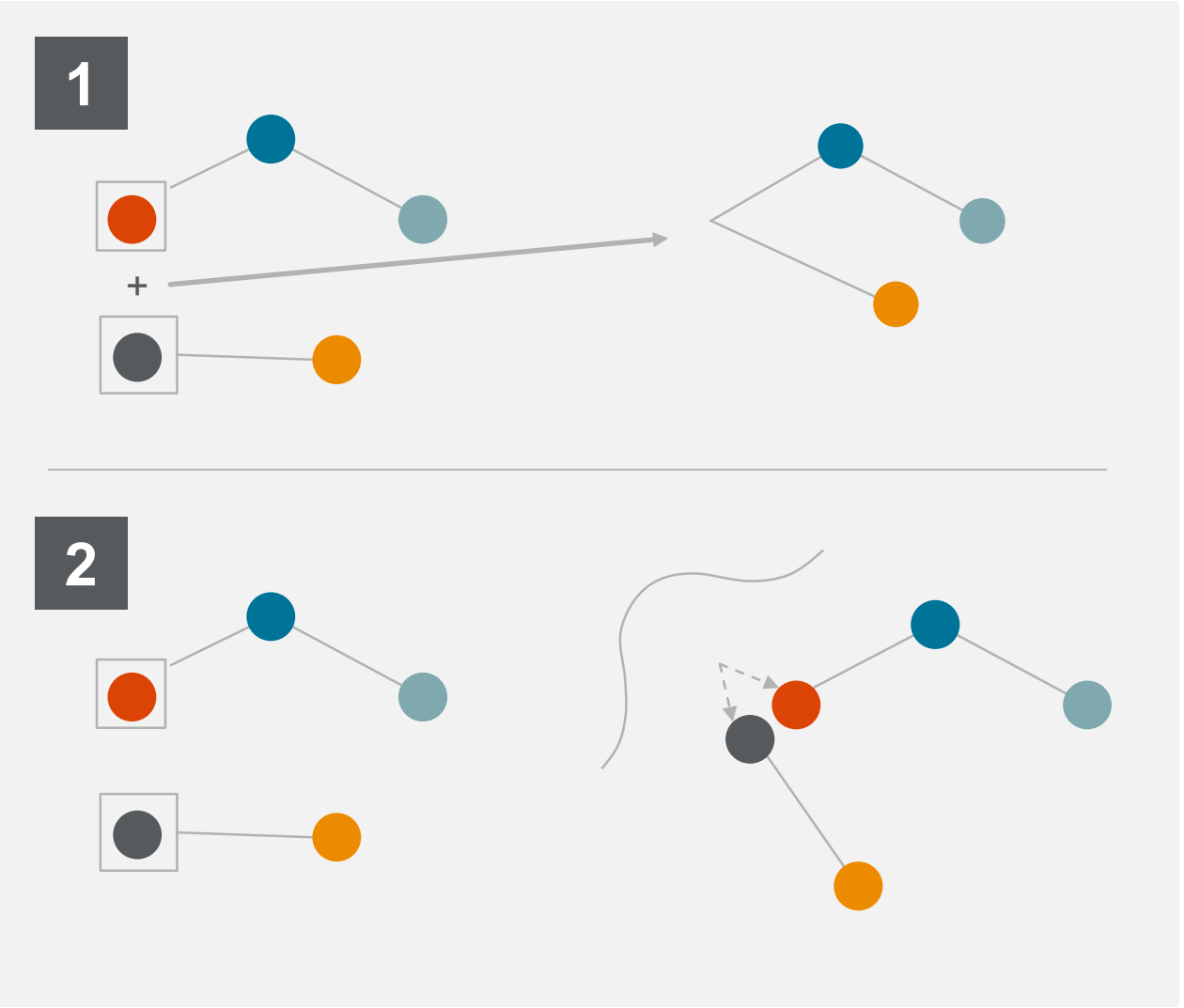
Effects within classes as exchangeable

Approach: Treatments in the same class of drugs are considered exchangeable

Strengths:

- Incorporates additional uncertainty
- Allows for estimates between drugs in the same class
- Provides assessment of feasibility of class assumption with enough data

Limitations: Requires the assumption of a class effect in order to be valid

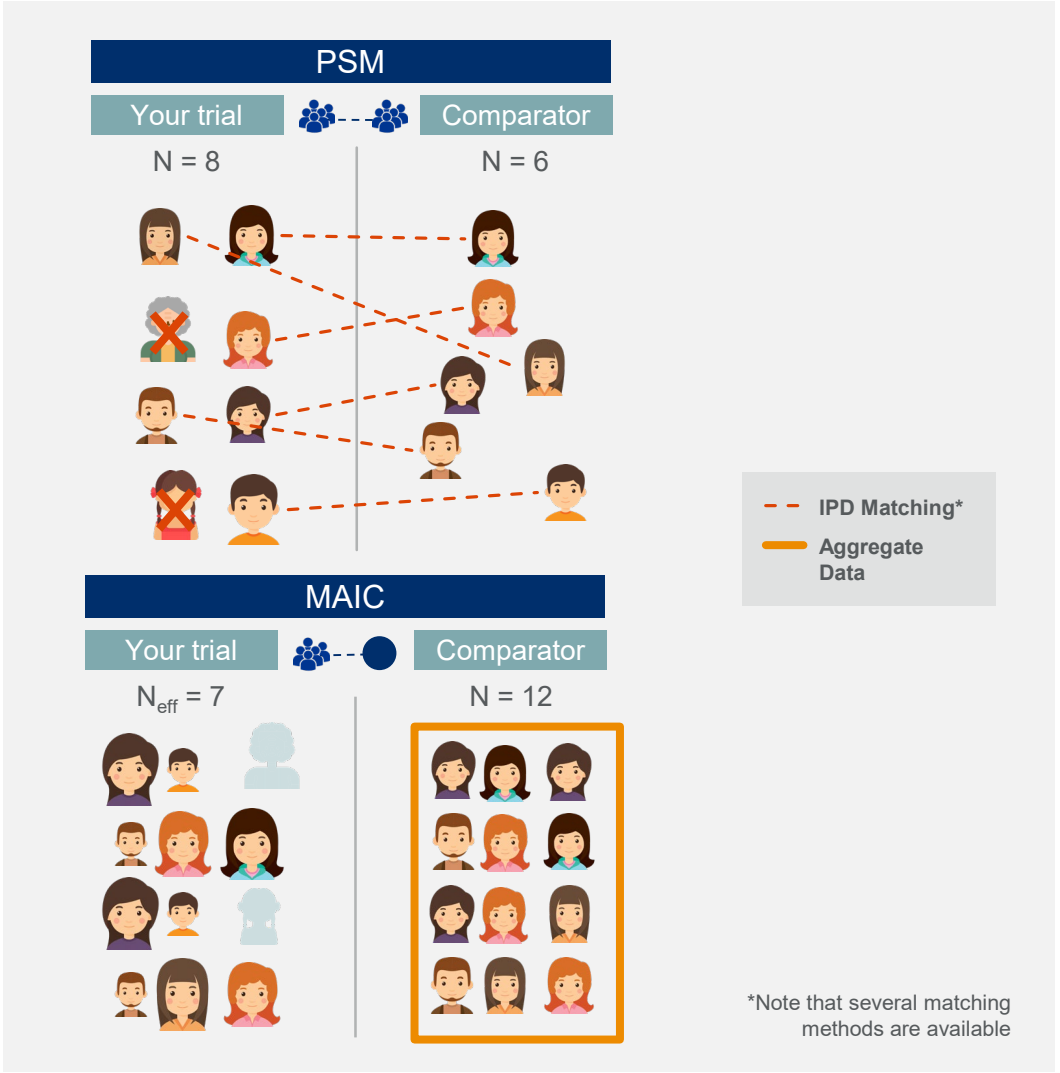


EVERSANA™

Alternative Approaches Involving Individual Participant Data

Propensity score methods, regression models, MAIC

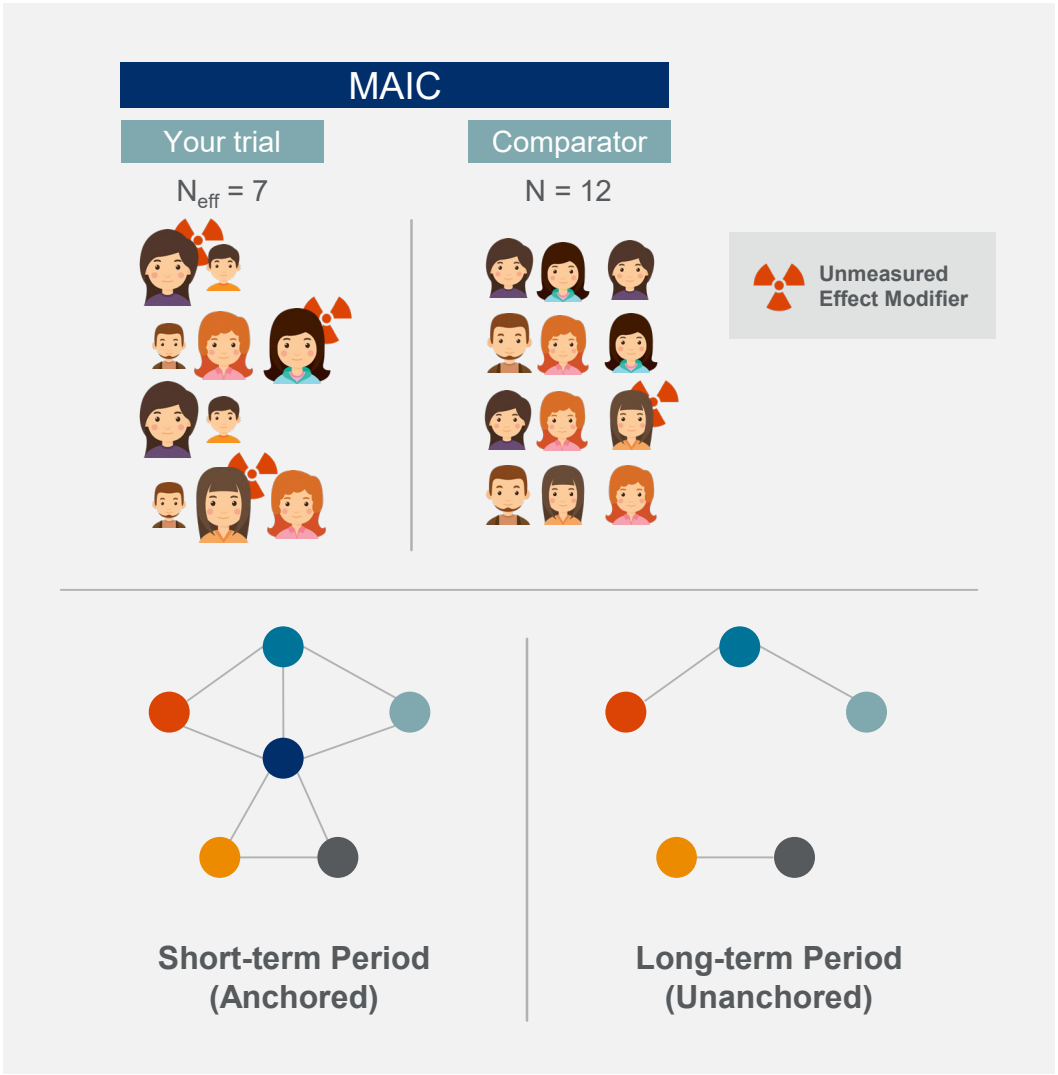
- Use of individual participant data (IPD) allows for several additional analysis methods which may be combined with NMAs[18]
- **Propensity score methods** are appealing with their ability to provide “RCT-like” comparisons
  - A similar approach using regression models can also be used
- If only aggregate data are available for the comparator, **matching adjusted indirect comparisons** (MAIC) can be used
  - Aligns patient populations in terms of eligibility criteria and weighting of IPD to match the comparator population characteristics
  - Simulated treatment comparisons (STC) offer a regression based alternative



Alternative Approaches Involving Individual Participant Data

Limitations

- While IPD methods can allow for adjustment for more variables, they do not guarantee equivalence to estimates from randomized trials
  - Unanchored analysis still require adjustment for all prognostic and effect modifying variables
- Incorporating MAIC results into complete networks is more complicated than it seems (eg, risk for double counting)

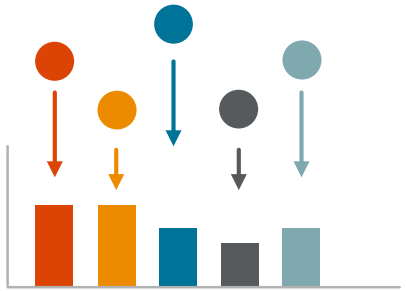


Summary of Recent Long-term NMAs

Recent analyses evaluating the long-term benefits of treatment for plaque psoriasis are associated with a variety of methodological limitations and **should be interpreted cautiously**:

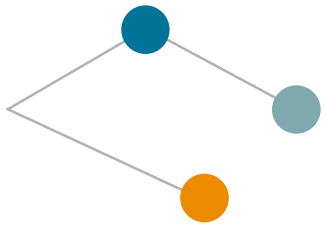
Armstrong et al.

- Conducted naïve treatment comparisons between active arms, breaking randomization
- Easy to implement approach, but is typically considered a **fatal flaw** in the presence of more reliable alternatives



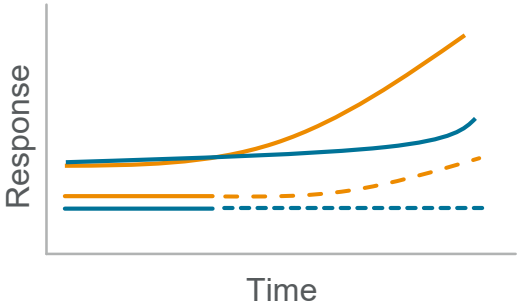
Diels et al.

- Assumed a single class treatment effect for all TNFα inhibitors
- Leverages benefits of an NMA, but **relies on assumption** of class effect within TNFα inhibitors.



Sawyer et al.

- Presumed stability of placebo responses within trials until the end of the maintenance period
- Maintains randomization and **can be valid** if placebo response does not differ between treatments over time



EVERSANA™

Summary and Recommendations

- Some approaches may be more vulnerable to bias than others
  - Notably, the assumptions required for **naïve comparisons** to be valid **are rarely likely to hold**
- Methods that leverage more information (ie, Diels and Sawyer approaches) may continue to be valuable as even new trials allow for linked long-term networks
  - How does well-known between-trial heterogeneity in the induction period translate to sparse networks where baseline risk adjustment is not possible?
- Future studies should consider assumptions underpinning NMAs and assess how alternative NMA methods can be leveraged to yield more rigorous long-term indirect treatment comparisons

References

1. Bhosle, M. J., Kulkarni, A., Feldman, S. R. & Balkrishnan, R. Quality of life in patients with psoriasis. Health Qual. Life Outcomes 4, 35 (2006).
2. Caldwell D, Ades A & Higgins J. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 331, 897–900 (2005).
3. Lu, G. & Ades, A. E. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat. Med.* 23, 3105–3124 (2004).
4. Catala-Lopez F, Tobias A, Cameron C, Moher D & Hutton B. Network meta-analysis for comparing treatment effects of multiple interventions: an introduction. *Rheumatol. Int.* 34, 1489–1496 (2014).
5. Ioannidis J. Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. *CMAJ* 181, 488–493 (2009).
6. Salanti G. Indirect and mixed-treatment comparison, network, or multiple treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Res. Synth. Methods* 3, (2012).
7. Sawyer, L. M. et al. Assessing the relative efficacy of interleukin-17 and interleukin-23 targeted treatments for moderate-to-severe plaque psoriasis: A systematic review and network meta-analysis of PASI response. PloS One 14, e0220868 (2019).
8. Bai, F. et al. Short-Term Efficacy and Safety of IL-17, IL-12/23, and IL-23 Inhibitors Brodalumab, Secukinumab, Ixekizumab, Ustekinumab, Guselkumab, Tildrakizumab, and Risankizumab for the Treatment of Moderate to Severe Plaque Psoriasis: A Systematic Review and Network Meta-Analysis of Randomized Controlled Trials. J. Immunol. Res. 2019, 2546161 (2019).
9. Warren, R. B. et al. Differential Drug Survival of Biologic Therapies for the Treatment of Psoriasis: A Prospective Observational Cohort Study from the British Association of Dermatologists Biologic Interventions Register (BADBIR). J. Invest. Dermatol. 135, 2632–2640 (2015).
10. Sbidian, E. et al. Systemic pharmacological treatments for chronic plaque psoriasis: a network meta-analysis. Cochrane Database Syst. Rev. 12, CD011535 (2017).
11. Xu, G. et al. Comparative efficacy and safety of thirteen biologic therapies for patients with moderate or severe psoriasis: A network meta-analysis. J. Pharmacol. Sci. 139, 289–303 (2019).
12. Jabbar-Lopez, Z. K. et al. Quantitative Evaluation of Biologic Therapy Options for Psoriasis: A Systematic Review and Network Meta-Analysis. J. Invest. Dermatol. 137, 1646–1654 (2017).
13. Signorovitch, J. E. et al. Comparative efficacy of biological treatments for moderate-to-severe psoriasis: a network meta-analysis adjusting for cross-trial differences in reference arm response. Br. J. Dermatol. 172, 504–512 (2015).
14. Song, F. et al. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. BMJ 338, b1147 (2009).
15. Jansen, J. P., Crawford, B., Bergman, G. & Stam, W. Bayesian meta-analysis of multiple treatment comparisons: an introduction to mixed treatment comparisons. Value Health J. Int. Soc. Pharmacoeconomics Outcomes Res. 11, 956–964 (2008).
16. Armstrong, April W., et al. "Comparison of biologics and oral treatments for plaque psoriasis: a meta-analysis." JAMA dermatology 156.3 (2020): 258-269.
17. Diels J *et al.* Novel evidence synthesis methods to assess comparative efficacy in 'disconnected' networks of evidence: a case study assessing comparative efficacy of guselkumab versus interleukin-17 inhibitors for maintenance treatment of moderate-to-severe psoriasis. *TBD* **TBD**, (2017).
18. Phillippo, D., Ades, T., Dias, S., Palmer, S., Abrams, K. R., & Welton, N. (2016). NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submissions to NICE.

THANK YOU

Tim Disher
tim.disher@eversana.com

EVERSANA™

Meta-analyses are not infallible

NMA is vulnerable similar limitations of pairwise meta-analysis, where there are cautionary tales regarding the validity of conclusions even from comparisons with many trials[1]

- Intravenous magnesium vs placebo for patients with acute myocardial infarction
  - Meta-analysis of 15 smaller trials lead to the conclusion of a large protective effect
  - ISIS-4 “mega trial” released showing small harm

Even in robust networks with closed loops, modeling choices can lead to different conclusions

- In treatments a network meta-analysis of , Cochrane review using unadjusted approach found small amounts of heterogeneity[2], but baseline risk adjustment leads to change in rank and overall conclusions[3]

1. Dias, S., Sutton, A. J., Welton, N. J., & Ades, A. E. (2012). *NICE technical support document 3: Heterogeneity: Subgroups, meta-regression, bias, and bias-adjustment*. Retrieved from http://www.nicedsu.org.uk

2. Sbidian, E., Chaimani, A., Garcia-Doval, I., Do, G., Hua, C., Mazaud, C., ... & Chosidow, O. (2017). Systemic pharmacological treatments for chronic plaque psoriasis: a network meta-analysis. Cochrane Database of Systematic Reviews, (12).

3. Cameron, C., Hutton, B., Druchok, C., McElligott, S., Nair, S., Schubert, A., ... & Villacorta, R. (2018). Importance of assessing and adjusting for cross-study heterogeneity in network meta-analysis: a case study of psoriasis. Journal of comparative effectiveness research, 7(11), 1037-1051.

Meta-analyses are not infallible

Vibration of Effects/Garden of forking paths

Vibration of effects

- Different methodological choices create a potential “multiverse” of analysis/ data combinations
- ITC of nalmefene vs naltrexone with 60 studies analyzed in 9216 different ways
  - Analysis decisions lead to statistically significant differences in opposite directions

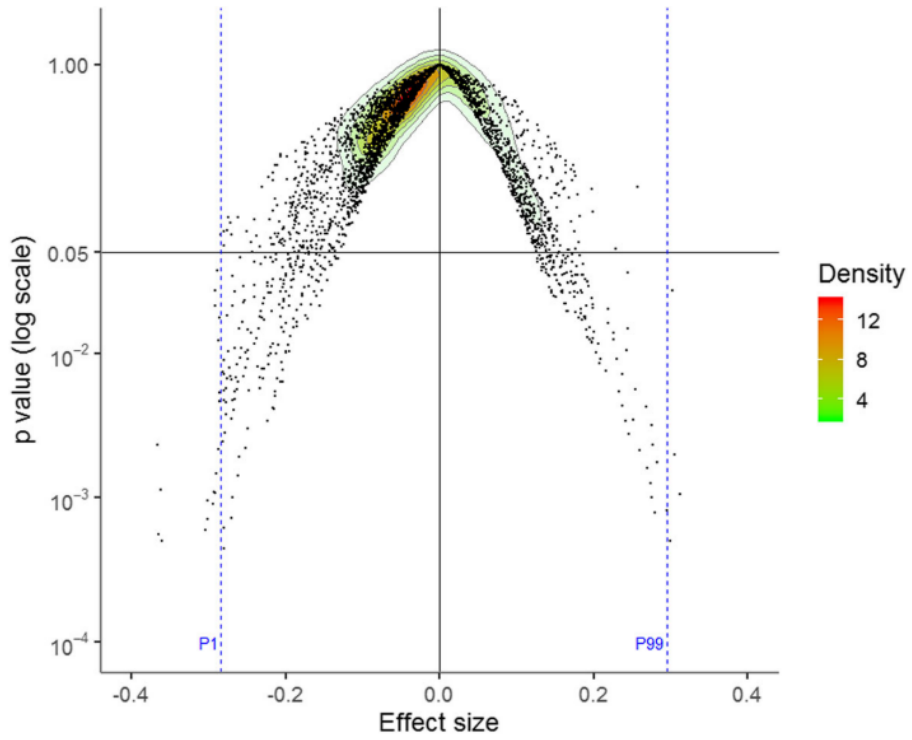


Fig. 2 Vibration of effects for the indirect comparison of nalmefene to naltrexone. A negative effect size favours nalmefene, whereas a positive effect size favours naltrexone. The points represent the meta-analyses. The colours represent the densities

Palpacuer, C., Hammas, K., Duprez, R., Laviolle, B., Ioannidis, J. P. A., & Naudet, F. (2019). Vibration of effects from diverse inclusion/exclusion criteria and analytical choices: 9216 different ways to perform an indirect comparison meta-analysis. *BMC Medicine*, *17*(1), 1–13. https://doi.org/10.1186/s12916-019-1409-3

Meta-analyses are not infallible

The effect of multiplicity + filtering without true prior information

- Conducting 100s of tests creates a multiplicity issue which is compounded by filtering them (picking out best ranking, highest SUCRA, etc).
  - Form of selective reporting that exaggerates differences between the best and worst treatments
- Different modeling choices can be more or less susceptible to this issue, with standard NMA models being the most vulnerable
- Strict null-hypothesis significance testing can be misleading in NMA conclusions
- The effect is less exaggerated in more densely connected networks

In this paper, we highlighted the possible implications of the multiplicity issue in NMA. We presented the problems associated with multiplicity using both theoretical arguments as well as simulations. We showed that the model commonly used for NMA may give exaggerated estimates of the effects of the treatment identified to be the best in the network. It may also be associated with a high probability of (falsely) showing differences between treatments, when actually there is none.

The problems stem from the fact that the NMA model simultaneously estimates tens (or even hundreds, depending on the network size) of relative treatment effects. By chance alone, some of these estimates may be very large. Any

Efthimiou, O., & White, I. R. (2019). The dark side of the force: multiplicity issues in network meta-analysis and how to address them. *Research Synthesis Methods*, (October 2018), 1–18. https://doi.org/10.1002/jrsm.1377