Drifting in the Network: Catching Baseline Shifts Before They Wreck Your Network Meta-Analysis

Saswata Paul Choudhury, MSc, Kalpesh Chatterjee, MSc, Sekhar Kumar Dutta, MSc, Abhirup Dutta Majumdar, MSc.

PharmaQuant Insights Private Limited, Kolkata, India.

MSR78

PharmaQuant

Focused Insights, Better outcomes

Background

- Network meta-analysis (NMA) allows for the simultaneous comparison of multiple treatments by combining direct and indirect evidence from a network of clinical trials. This method yields robust estimates of relative treatment effects when head-to-head trials are unavailable.
- The transitivity assumption is a core prerequisite for valid indirect comparisons in NMAs. Transitivity requires that studies be comparable in their distributions of key effect modifiers, such as baseline patient characteristics (BCxs), so that indirect estimates are not biased. If trials compare A versus B and B versus C, transitivity implies that the trial populations are sufficiently similar that participants could, in principle, have been randomized to any of the evaluated treatments.
- Substantial differences in BCxs across studies can lead to violation of the transitivity assumption and result in distributional drift.
- We propose a clustering framework to identify and quantify distributional drift in BCxs, thereby improving the internal validity and interpretability of NMA results.

Objectives

- Identify and group studies with comparable distributions of BCxs to support stratified analyses or targeted sensitivity checks that help preserve transitivity.
- Quantify the degree and drivers of distributional drift in BCxs to evaluate its potential impact on NMA results.

Methods

- A clustering analysis was performed to assess similarity across studies and detect early signals of non-comparability in BCxs. This approach generated clusters of studies that were more homogeneous in their BCx distributions while highlighting those that differed substantially. Such heterogeneity between clusters may indicate potential bias in indirect comparisons within the NMA.
- A simulated dataset comprising 15 studies was generated. Each study reported eight BCxs.
- K-means clustering was performed using the Hartigan and Wong algorithm.[4] The choice of number of clusters (k) was guided by the elbow method. Each cluster is represented by a centroid that approximates optimized, lower-entropy marginal distributions of the BCxs **(Figure 1)**.
- BCxs were ranked according to their contribution to variability across studies. Those contributing most to observed distributional drift were identified as primary drivers of heterogeneity between clusters.
- **Jensen-Shannon Divergence (JSD)** was used to quantify the magnitude of marginal distribution shifts for each BCx across clusters. JSD measures the dissimilarity of BCx distributions between clusters. Lower JSD values indicate greater similarity in BCx distributions across studies, while higher values reflect greater distributional drift and potential threats to transitivity.
- **Relative importance analysis (RIA)** was performed to determine the primary drivers of heterogeneity between clusters. While JSD quantifies the extent of drift, RIA identifies its key contributors by ranking BCxs according to their influence on cluster separation. Higher RIA values indicate the BCxs most responsible for distributional drift and heterogeneity between clusters.
- All analyses were conducted using R-V4.5.1 software.

Results

- The within-cluster sum of squares (WCSS) curve demonstrated a clear bend at k = 2, indicating that 2 clusters were optimal for the analysis. This suggested a natural grouping structure among studies such that increasing the number of clusters beyond two did not substantially reduce within-cluster variance. This confirmed the stability and adequacy of the two-cluster solution **(Figure 1)**.
- The K-means clustering model partitioned the 15 studies into Cluster 1 (n = 12) and Cluster 2 (n = 3). This indicated that 12 studies shared broadly similar BCx distributions, whereas the 3 studies in Cluster 2 exhibited distinct BCx patterns (Figure 2).
- **Jensen-Shannon Divergence**: Jensen-Shannon divergence values confirmed that X1 (0.585), X5 (0.422), and X6 (0.362) exhibited the highest levels of distributional drift. This indicated that these variables contributed most strongly to heterogeneity between clusters. The remaining BCxs demonstrated comparatively lower drift, reflecting greater consistency in their distributions across clusters (Figure 3)

Figure 1: Cluster Selection Elbow Plot

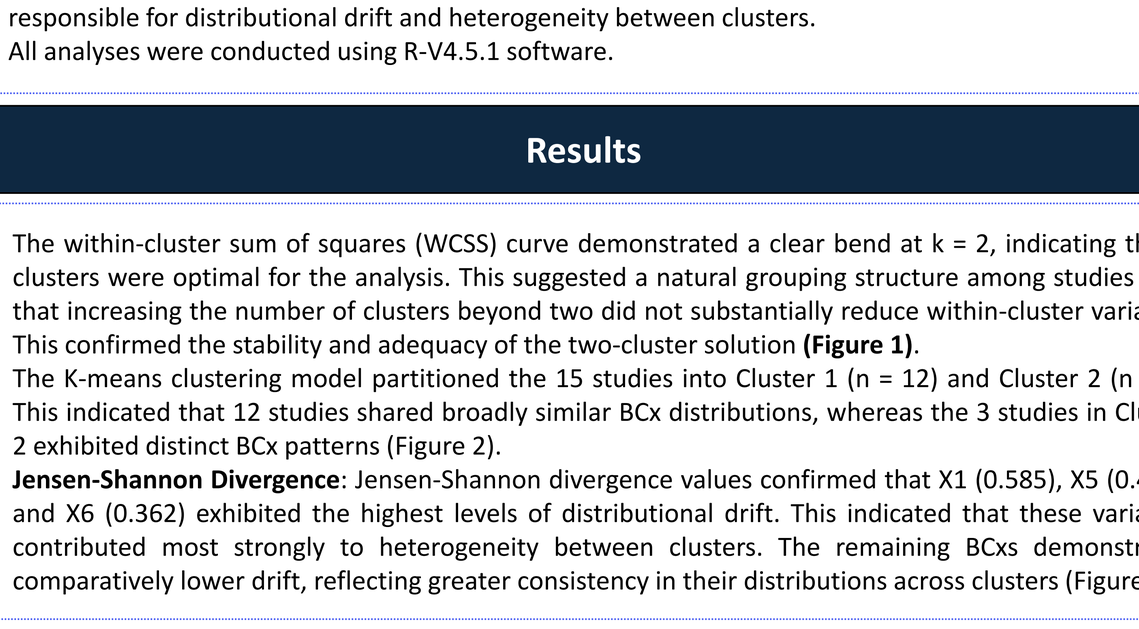


Figure 2: Study-level Cluster Map with Centroids

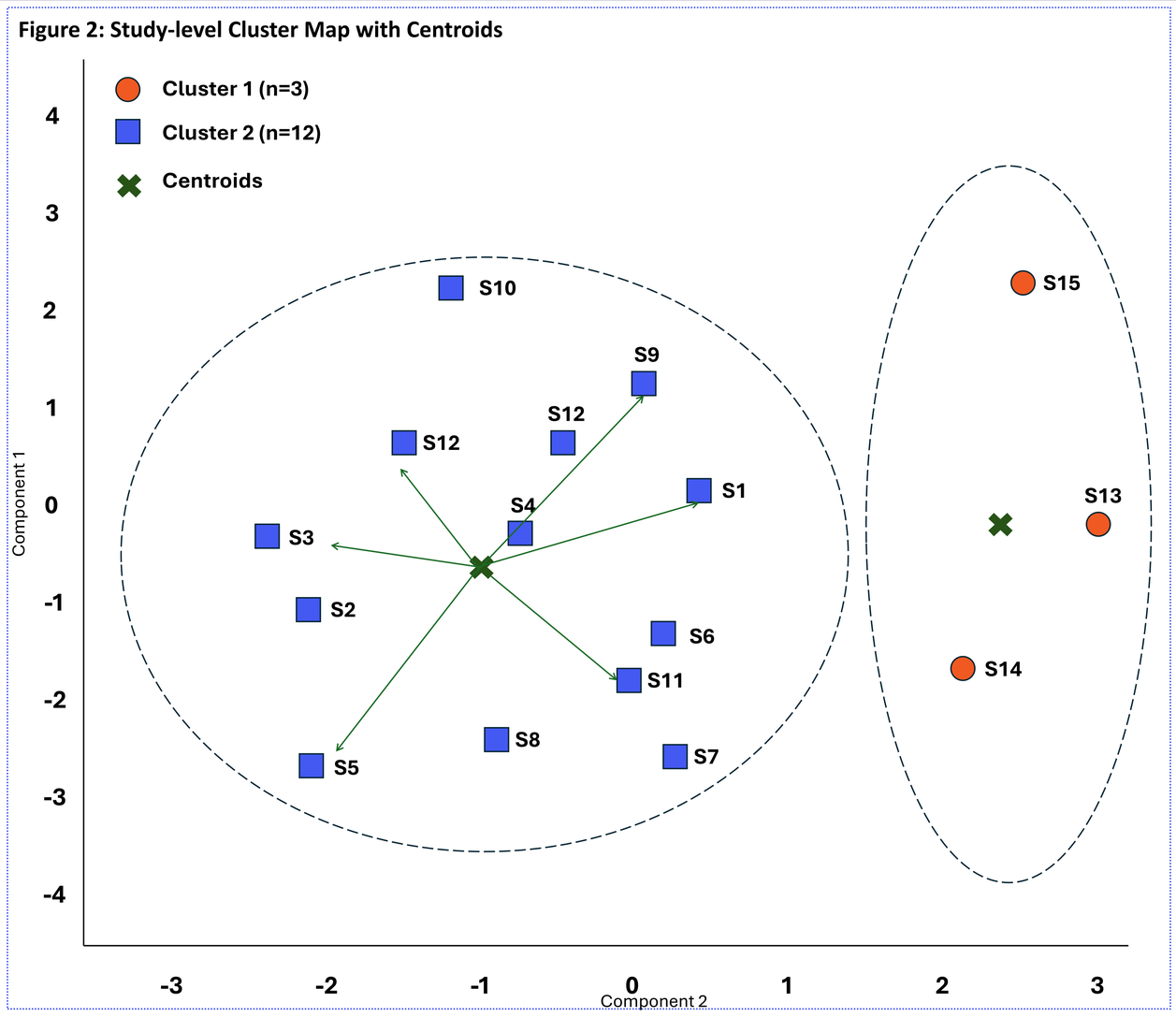


Figure 3: Jensen-Shannon Divergence by Baseline Characteristic

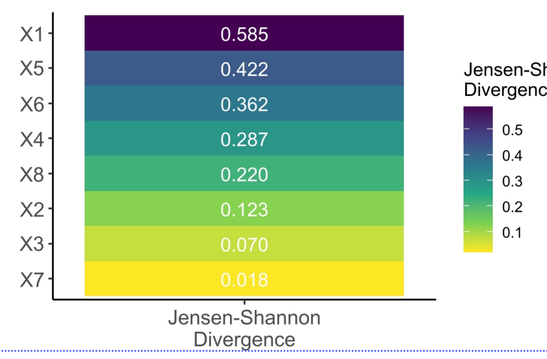
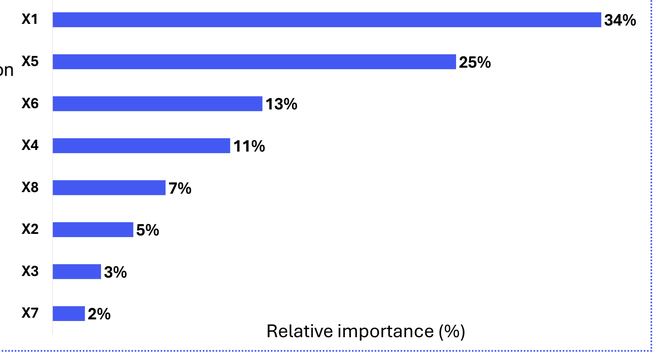
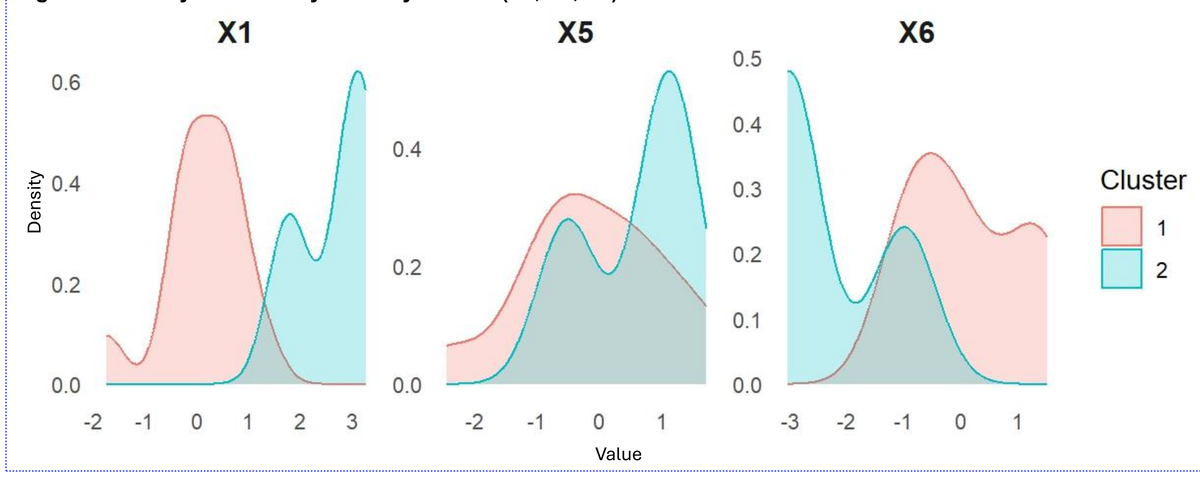


Figure 4: Relative Importance of Baseline Characteristics



- RIA revealed that X1 (34%), X5 (25%), and X6 (13%) contributed most to heterogeneity between clusters, highlighting their influence on study comparability **(Figure 4)**.
- X1, X5, and X6 showed clear distributional differences between Cluster 1 and Cluster 2. X1 showed the greatest shift in both peak and shape. X6 showed a moderate change with some overlap between clusters. X5 showed the smallest difference, though a smaller but still visible difference **(Figure 5)**.

Figure 5: Density Plots of Key BCxs by Cluster (X1, X5, X6)



Conclusion

- Studies with similar BCxs were identified and grouped, quantifying the key drivers of heterogeneity using Jensen-Shannon divergence and relative importance analysis. The most comparable subsets of BCxs were then identified.
- This novel method identified outlier trials with mismatched baseline profiles, enabling focused sensitivity analyses. Conducting the NMA with and without these clusters facilitates an assessment of the stability treatment effect stability and rankings. #
- This method also serves as a valuable pre-processing step by identifying non-comparable studies, thereby supporting more robust and reliable evidence synthesis.

References

1. Cope, S., Zhang, J., Saletan, S. *et al.* A process for assessing the feasibility of a network meta-analysis: a case study of everolimus in combination with hormonal therapy versus chemotherapy for advanced breast cancer. *BMC Med* **12**, 93 (2014).
2. Cooper NJ, Sutton AJ, Morris D, Ades AE, Welton NJ. Addressing between-study heterogeneity and inconsistency in mixed treatment comparisons: Application to stroke prevention treatments in individuals with non-rheumatic atrial fibrillation. Stat Med. 2009 Jun 30;28(14):1861-81. doi: 10.1002/sim.3594. PMID: 19399825.
3. Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. Med Decis Making. 2013 Jul;33(5):607-17. doi: 10.1177/0272989X12458724. Epub 2012 Oct 26. PMID: 23104435; PMCID: PMC3704203.
4. Hartigan JA, Wong MA. Algorithm AS 136: A k-means clustering algorithm. Journal of the royal statistical society. series c (applied statistics). 1979 Jan 1;28(1):100-8.

Presented at ISPOR Europe 2025; Glasgow, Scotland, UK; November 9-12, 2025