

The surface under the cumulative ranking curve

Shengping Yang PhD, Gilbert Berdine MD

I am planning a network meta-analysis of four treatment options for asthma. Given the multiple available treatments, I intend to use a network meta-analytic approach and am exploring the use of the Surface Under the Cumulative Ranking Curve (SUCRA) to rank their relative effectiveness. Is SUCRA considered a standard approach in this context?

Network meta-analysis (NMA) is a powerful method for comparing multiple treatments simultaneously, especially when direct head-to-head comparisons are limited. The Surface Under the Cumulative Ranking Curve (SUCRA) is a widely accepted and standard approach for ranking treatments in NMA. It is a metric used to summarize the relative effectiveness or safety of competing interventions.

1. BACKGROUND

In evidence-based medicine, treatment comparisons often involve multiple competing interventions. Traditional pairwise meta-analyses are limited because they only compare two treatments at a time, whereas clinical decision-making typically requires evaluating all available options simultaneously.

To address this, NMA was developed. It integrates both direct evidence (from head-to-head trials) and indirect evidence (e.g., trials comparing A vs. B and B vs. C to infer A vs. C) into a unified framework. This allows for the coherent ranking of all treatments, even when some pairs have never been directly compared.¹

As NMAs became more common, researchers sought intuitive ways to summarize treatment hierarchies. A metric allows comparison of a property between two objects at the same time or the same object at different times. A metric is a numeric scale such that

if two objects have the same value for the metric, the two objects are considered equal in terms of the property being considered. If two objects have different values for the metric, the values can be used to compare, rank, or sort the objects. Simple metrics like “probability of being best” were limited because they ignored the full distribution of rankings (e.g., how often a treatment ranks second or third). The SUCRA method was developed to address this need and was first proposed by Salanti et al.² It has since become a standard for summarizing treatment rankings in NMAs.

2. THE BASIC CONCEPT OF SUCRA

SUCRA is a quantitative metric derived from the cumulative ranking probabilities of treatments in a NMA. Unlike traditional effect-size metrics, such as mean differences or odds ratios, SUCRA captures the entire ranking distribution, providing a broader perspective on a treatment’s performance relative to others, in Bayesian or frequentist NMAs.⁴

Mathematical definition:

For a treatment i in an NMA comparing K competing treatments:

$$SUCRA_i = \frac{1}{K-1} \sum_{r=1}^{K-1} F_i(r),$$

where $F_i(r) = P(\text{Rank} \leq r)$, is the cumulative probability that treatment i ranks at or above position r .³ Following are the steps for calculating SUCRA.

2.1 RANK TREATMENTS

For each treatment, based on the NMA results (e.g., relative risks or odds ratios), derive the probability that the treatment ranks at each possible position.

2.2 CUMULATIVE PROBABILITY

For each treatment, compute the cumulative probability $F_i(r)$, i.e., $\text{Rank} \leq r$. In Bayesian NMAs, these

Corresponding author: Shengping Yang
Contact Information: Shengping.Yang@pbrc.edu
DOI: 10.12746/swjm.v13i55.1445

probabilities can be derived from posterior MCMC samples. In frequentist NMAs, they can be estimated via bootstrapping or analytical approximations. For example, if there are 4 treatments (A, B, C, and D), treatment D might have the following cumulative probabilities based on the data: $F_D(1) = 0.05$, $F_D(2) = 0.25$, and $F_D(3) = 0.5$, while $F_D(4)$ is always 1.

2.3 SUCRA CALCULATION

For the treatment D, $SUCRA_D$ is calculated using the formula above. For instance, in the example $K = 4$:

$$SUCRA_D = \frac{1}{3} \sum_{r=1}^3 F_i(r) = \frac{0.05 + 0.25 + 0.5}{3} = 0.27.$$

Note that for any of the treatments, $SUCRA_i$ represents the normalized area under the cumulative ranking curve for treatment i . A higher $SUCRA_i$ value indicates a better overall ranking of treatment i relative to competing treatments. A $SUCRA_i$ of 100% means treatment i is almost always ranked the best, while a $SUCRA_i$ of 0% means it is almost always ranked the worst. Meanwhile, a $SUCRA_i$ of, for example, 0.8 (80%) means the treatment has an 80% probability of being among the top options (when compared to all other treatments).

Figure 1 provides a graphical illustration of the example above. Since $F(4)$ is always equal to 1 for any treatment, it is excluded from the SUCRA calculation and is also not shown in the figure. The SUCRA values for treatments A, B, C, and D are 0.83, 0.52, 0.38, and 0.27, respectively. Based on these values, treatment A is considered the most favorable option in this analysis.

3. APPLICATIONS OF SUCRA

SUCRA is a key metric derived from NMA, designed to simplify the interpretation of complex comparative effectiveness data. By converting multidimensional ranking distributions into a single numerical value (ranging from 0 to 1), SUCRA allows users to efficiently prioritize treatment options. Its primary applications include:

3.1 CLINICAL GUIDELINES

SUCRA helps translate NMA results into actionable recommendations by quantifying the likelihood that

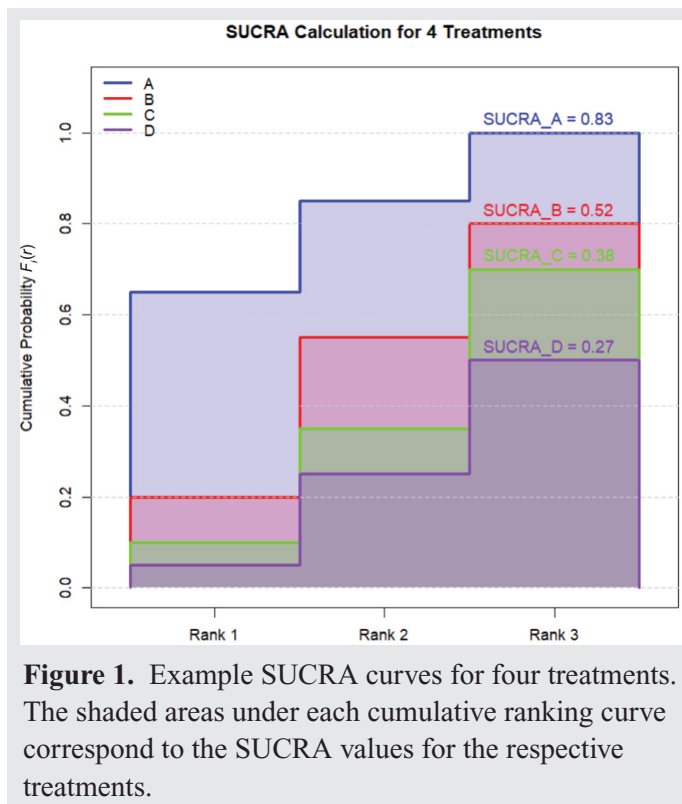


Figure 1. Example SUCRA curves for four treatments. The shaded areas under each cumulative ranking curve correspond to the SUCRA values for the respective treatments.

a treatment outperforms its competitors. For example, a guideline panel may favor Treatment A (SUCRA = 0.90) over Treatment B (SUCRA = 0.60) when efficacy is the primary concern.

3.2 HEALTH TECHNOLOGY ASSESSMENT (HTA)

In HTA, SUCRA complements traditional cost-effectiveness analyses (e.g., cost-per-QALY) by integrating multiple dimensions:

- Relative efficacy: Treatments are ranked based on outcomes such as response rates or risk reduction.
- Safety: Lower-ranked options may indicate less favorable safety profiles.
- Cost-effectiveness: Treatments with high SUCRA values may justify higher pricing or reimbursement priority.

3.3 PUBLIC HEALTH EVIDENCE SYNTHESIS

SUCRA is particularly valuable in public health settings, such as during a pandemic, where direct

head-to-head comparisons are limited. It helps identify and prioritize promising interventions (e.g., vaccines or preventive measures) using both direct and indirect evidence.

3.4 DRUG FORMULARY DECISIONS

SUCRA also aids formulary committees in selecting therapies for inclusion by emphasizing relative effectiveness across treatments, rather than absolute outcomes alone.

SUCRA values are commonly presented alongside visual tools to enhance interpretability.⁵ For example, rankograms illustrate the probability distribution of each treatment's rank (e.g., "Treatment A has a 70% probability of being ranked 1st"); league tables show pairwise comparisons (e.g., odds ratios) with SUCRA rankings for context; and forest plots may include SUCRA values with confidence intervals, presenting both rankings and associated uncertainty.

4. DISCUSSIONS

While the utility of SUCRA is well recognized, several important considerations, both strengths and limitations, should guide its application and interpretation.

SUCRA offers several notable advantages:

- **Simplification of Complex Data:**

SUCRA distills probabilistic rankings from network meta-analyses into a single, interpretable metric. For example, a SUCRA value of 90% for Treatment A indicates a high likelihood of its ranking among the top interventions, making comparisons across multiple treatments more straightforward for clinicians and decision-makers.

- **Visual Intuitiveness:**

When presented alongside rankograms, SUCRA enhances the interpretability of uncertainty in treatment rankings. This combination, graphical visualization through rankograms and numerical summary via SUCRA, supports transparent and accessible evidence synthesis.

- **Utility in Data-Limited Contexts:**

SUCRA is particularly useful when head-to-head comparisons are lacking. In evaluating novel therapies or emerging interventions, especially in areas with limited direct evidence, SUCRA can help prioritize treatments for further investigation or real-world evaluation.

Despite its strengths, SUCRA also has several limitations that warrant caution:

- **Oversimplification of Uncertainty:**

By collapsing complex, multidimensional uncertainty into a single value, SUCRA may obscure important nuances. A treatment with a high SUCRA might still have wide credible intervals for its effect size, indicating significant uncertainty in its actual efficacy.

- **Model Dependency:**

SUCRA values are highly dependent on the statistical model underlying the NMA. Assumptions related to consistency, transitivity, and the choice of priors (in Bayesian frameworks) can influence the calculated rankings. Violations of these assumptions may lead to misleading SUCRA values.

- **Insensitivity to Effect Size Magnitude:**

SUCRA reflects the relative rank of treatments, not the magnitude of their effect sizes. Treatments with nearly identical efficacy may end up with noticeably different SUCRA scores if their rank probabilities differ slightly. This can lead to overinterpretation of minor differences.

5. CONCLUSION

SUCRA is a valuable tool in the analysis and interpretation of network meta-analyses. By summarizing a treatment's ranking probabilities into a single, easy-to-interpret value, SUCRA supports more informed decision-making in contexts where multiple interventions are compared.

However, as with any statistical summary, SUCRA is not without limitations. It should be used in conjunction

with other outputs – such as effect sizes, credible intervals, GRADE assessments, and clinical considerations – to ensure comprehensive and nuanced interpretation. When applied appropriately, SUCRA can be a powerful aid in the synthesis of complex evidence landscapes.

Article citation: Yang S, Berdine G. The surface under the cumulative ranking curve. *The Southwest Journal of Medicine* 2025;13(55):35–38

From: Department of Biostatistics (SY), Pennington Biomedical Research Center, Baton Rouge, LA; Department of Internal Medicine (GB), Texas Tech University Health Sciences Center, Lubbock, Texas

Submitted: 4/14/2025

Accepted: 4/16/2025

Conflicts of interest: none

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

REFERENCES

1. Yang S, Berdine G. Network meta-analysis. *The Southwest Respiratory and Critical Care Chronicles*. 2022;10(45):79–82. <https://doi.org/10.12746/swrccc.v10i45.1107>
2. Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol*. 2011 Feb;64(2):163–71. doi: 10.1016/j.jclinepi.2010.03.016. Epub 2010 Aug 5. PMID: 20688472.
3. Mbuagbaw L, Rochweg B, Jaeschke R, Heels-Andsell D, Alhazzani W, Thabane L, Guyatt GH. Approaches to interpreting and choosing the best treatments in network meta-analyses. *Syst Rev*. 2017 Apr 12;6(1):79. doi: 10.1186/s13643-017-0473-z. PMID: 28403893; PMCID: PMC5389085.
4. Rücker G, Schwarzer G. Ranking treatments in frequentist network meta-analysis works without resampling methods. *BMC Med Res Methodol*. 2015 Jul 31;15:58. doi: 10.1186/s12874-015-0060-8. PMID: 26227148; PMCID: PMC4521472.
5. Chaimani A, Higgins JP, Mavridis D, Spyridonos P, Salanti G. Graphical tools for network meta-analysis in STATA. *PLoS One*. 2013 Oct 3;8(10):e76654. doi: 10.1371/journal.pone.0076654. PMID: 24098547; PMCID: PMC3789683.