METHODS & TOOLS · OPEN MATERIALS

Meta-analytic entropy: information-theoretic heterogeneity beyond I²

An information-theoretic complement to I² across 403 Cochrane reviews

Mahmood Ahmad · Tahir Heart Institute, Rabwah · ORCID 0000-0001-9107-3704

ABSTRACT — E156 (7 sentences, 155 words; estimand: Normalized Entropy Index, 0–1)

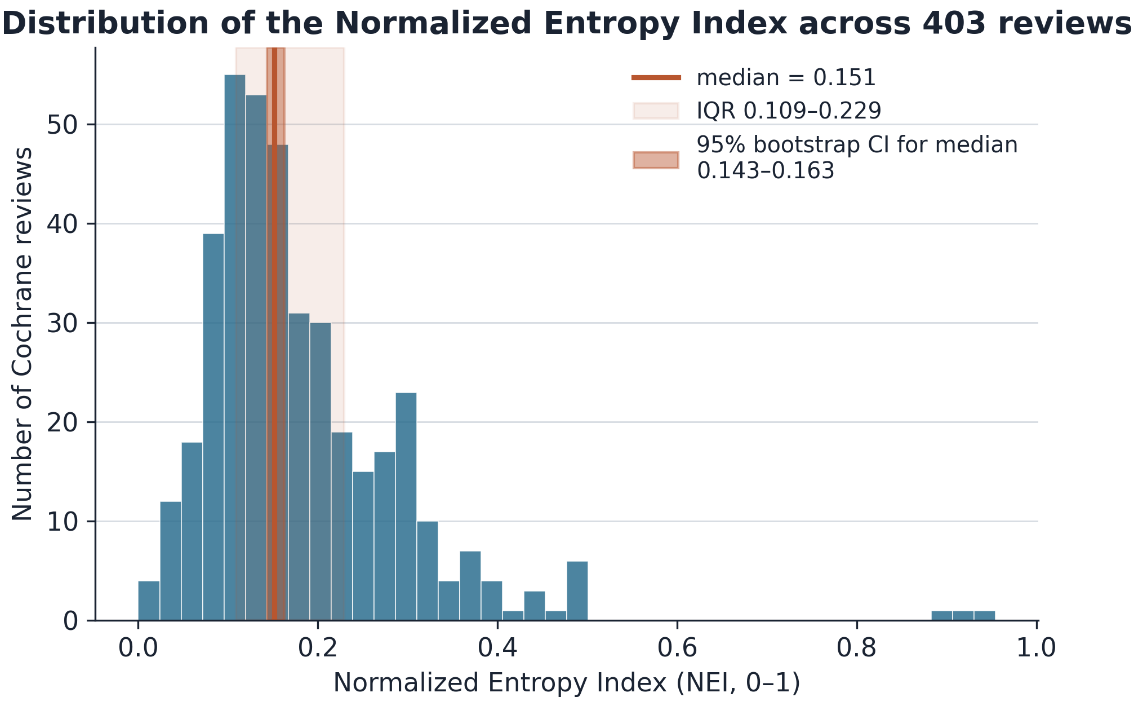
Can information-theoretic metrics detect heterogeneity features that I-squared is blind to, namely multimodality, skewness, and distributional shape? We computed Shannon entropy, Kullback-Leibler divergence, effect-precision mutual information, Fisher information, and a normalized entropy index for 403 Cochrane reviews from Pairwise70. Estimation used Monte Carlo sampling with 10000 draws per review; the index compared mixture entropy against theoretical bounds, while Silverman kernel density counted modes on precision-weighted distributions. The median index was 0.151 (IQR 0.109-0.229; 95% bootstrap CI 0.143-0.163), and 45% of reviews were multimodal despite I-squared below 60%. Because such estimates grow noisier below ten studies, we stratified by k; multimodality rose from 33% (k<10) to 62% (k>=10), a real signal, not a small-sample artifact. Effect-precision mutual information was positive in every computable review and significant by permutation in 22%, capturing dependence consistent with, but not proving, small-study mechanisms. The Normalized Entropy Index (0-1) thus captures distributional shape that complements, rather than replaces, variance-based heterogeneity measures.

REPORTED QUANTITIES — 403 Cochrane reviews (Pairwise70)

Quantity	Value	Uncertainty / stratum
Median Normalized Entropy Index	**0.151**	IQR 0.109–0.229; 95% boot CI 0.143–0.163
Multimodal (≥2 modes), overall	60.3%	liberal local-maxima detector
Multimodal & I^2<60%	**44.7%**	k<10: 33% · k≥10: 62%
…prominence-filtered bound	3.0%	valley ≥10% of peak
MI(effect;precision) > 0	307/307	binned MI ≥0 by construction
MI significant (perm. p<0.05)	**22%**	66/307 computable

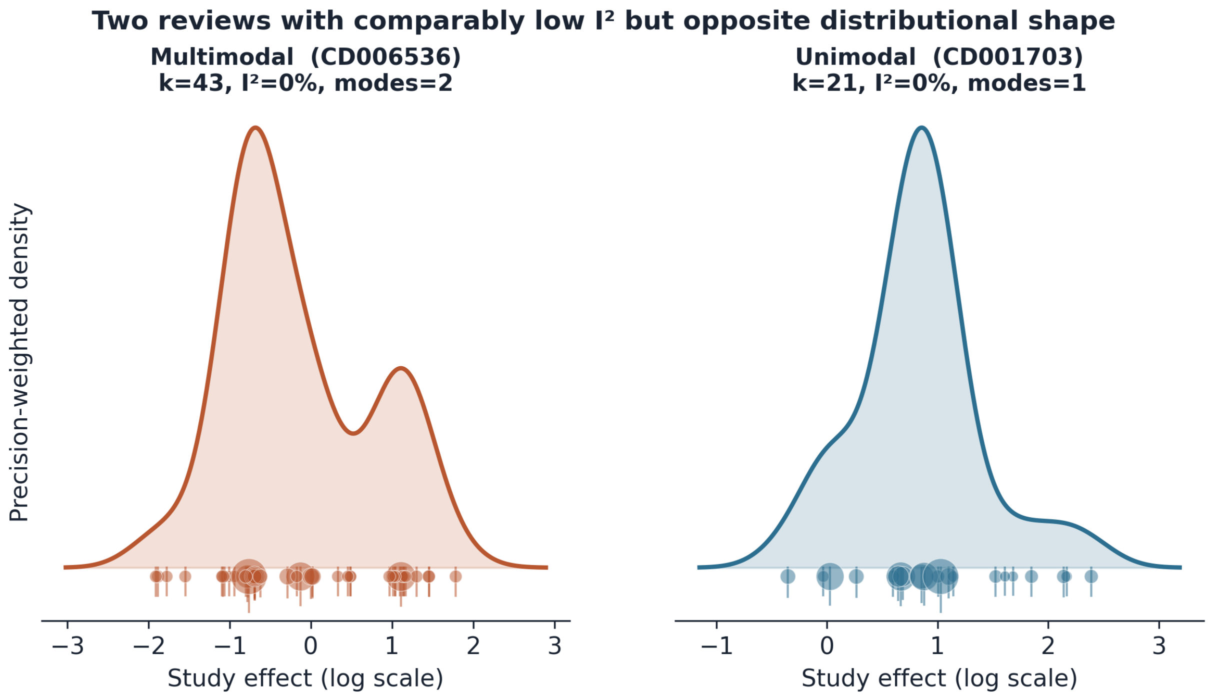
Verification: the published MetaEntropy engine was re-run over the source Pairwise70 corpus; all 403 per-review NEI and mode-count values reproduce bit-exactly. Summary statistics were independently re-computed by a separate Python agent and by an independent Node.js re-implementation (different RNG), and confirmed by Codex — all agreeing to four decimals. Bootstrap CI: 10,000 resamples, seed 2024.

FIGURES



NEI concentrated near the low end: most pooled effect distributions are information-poor relative to their entropy bounds.

Figure 1. Distribution of the Normalized Entropy Index across 403 reviews. Red line = median 0.151; light band = IQR 0.109–0.229; dark band = 95% bootstrap CI for the median (0.143–0.163). Most pooled effect distributions are information-poor relative to their bounds.



Silverman precision-weighted KDE. Marker size ∝ study precision ($1/SE^2$). I^2 is near-zero in both, yet entropy/mode-counting separates them.

Figure 2. Two reviews with I^2 = 0% but opposite shape. Left (CD006536, k=43): genuinely bimodal precision-weighted effect distribution. Right (CD001703, k=21): unimodal. I^2 calls both perfectly homogeneous; entropy and mode-counting separate them.

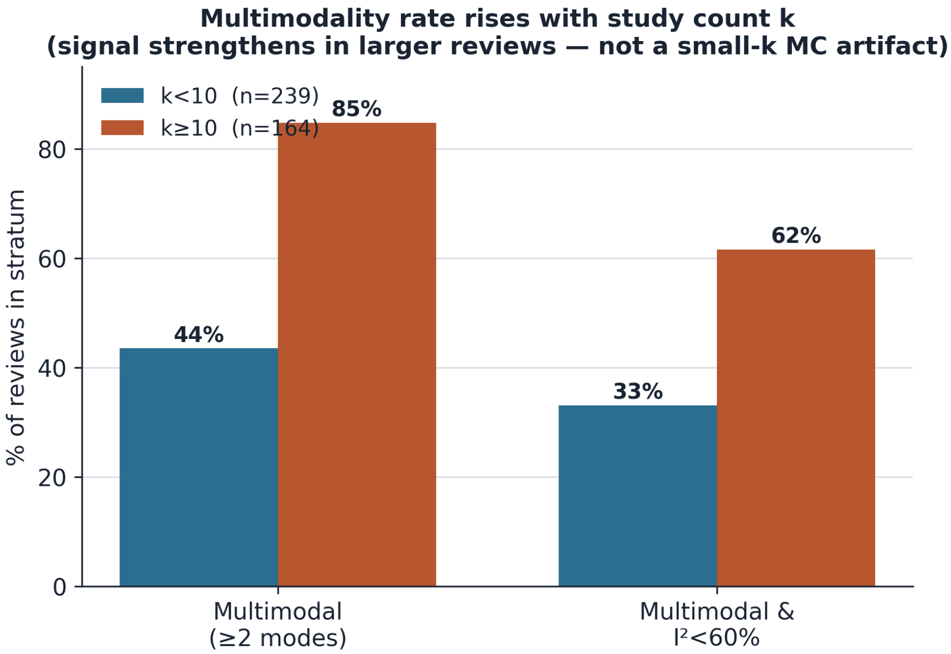
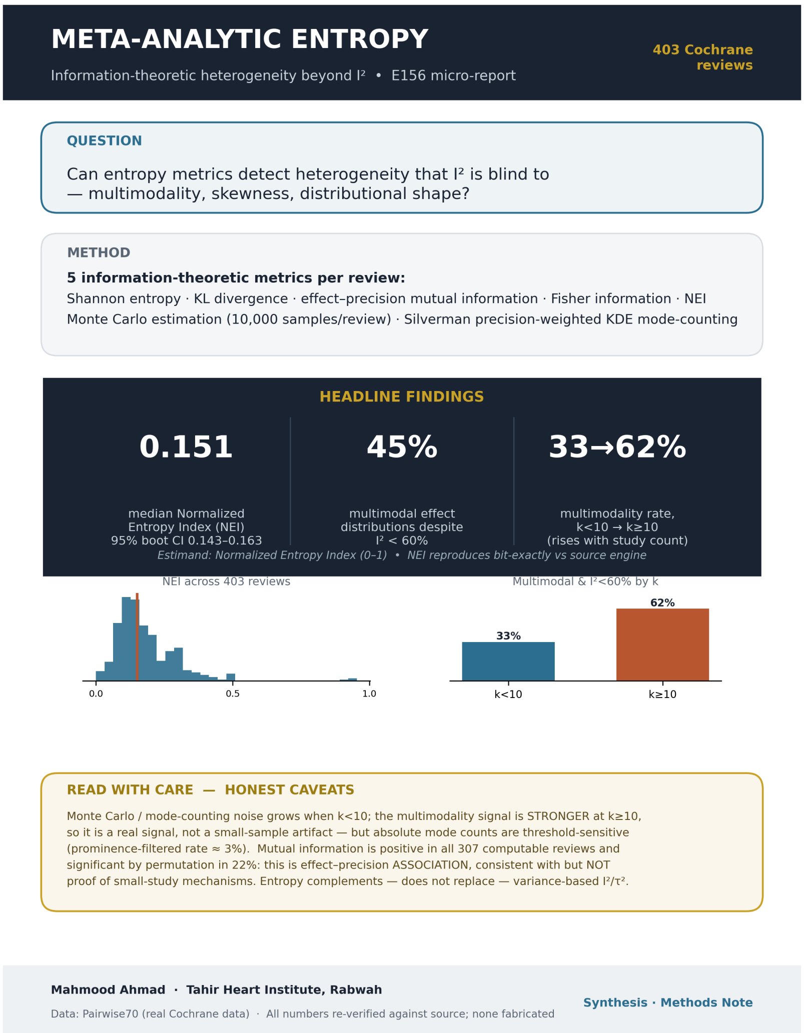


Figure 3. Multimodality rate by study count k. The multimodal-and-I^2<60% rate rises from 33% (k<10) to 62% (k≥10): the signal strengthens in larger reviews — not a small-sample artifact.



Visual abstract. Question → five information-theoretic metrics over 403 reviews → headline median NEI 0.151 and 45% multimodal-despite-low-I^2 → honest caveats. All values computed.

HONEST CAVEATS

Monte-Carlo entropy estimation and KDE mode-counting grow noisier below ten studies; stratifying by k shows the multimodality signal is stronger at $k \geq 10$, evidence it is real rather than a small-k artifact. The absolute rate is sensitive to the mode-definition: the liberal local-maxima detector flags 45%, a prominence-filtered count only 3.0%. The binned mutual-information estimator is non-negative by construction, so "positive in all reviews" is not itself evidence; a permutation test flags 22% as individually significant. Mutual information captures effect–precision dependence, consistent with — but not proof of — a small-study mechanism. Entropy complements, and does not replace, variance-based heterogeneity (I^2, τ^2).

REFERENCES

1. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327:557–560.
2. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21:1539–1558.
3. Shannon CE. A mathematical theory of communication. Bell Syst Tech J. 1948;27:379–423.
4. Cover TM, Thomas JA. Elements of Information Theory. 2nd ed. Wiley; 2006.
5. Silverman BW. Using kernel density estimates to investigate multimodality. J R Stat Soc B. 1981;43:97–99.
6. IntHout J, Ioannidis JPA, Rovers MM, Goeman JJ. Plea for routinely presenting prediction intervals in meta-analysis. BMJ Open. 2016;6:e010247.
7. Riley RD, Higgins JPT, Deeks JJ. Interpretation of random effects meta-analyses. BMJ. 2011;342:d549.
8. Sterne JAC, Sutton AJ, Ioannidis JPA, et al. Recommendations for examining and interpreting funnel plot asymmetry. BMJ. 2011;343:d4002.
9. Efron B, Tibshirani RJ. An Introduction to the Bootstrap. Chapman & Hall; 1993.
10. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to Meta-Analysis. 2nd ed. Wiley; 2021.
11. Higgins JPT, Thomas J, Chandler J, et al., eds. Cochrane Handbook for Systematic Reviews of Interventions v6.5. Cochrane; 2024.

Competing interests & editorial independence. The author is a member of the Synthesis editorial board and had no role in the editorial handling or decision on this manuscript, managed by an independent editor. Data & code: the MetaEntropy pipeline, the re-run/verification scripts, and figure code are openly available; analysis used only publicly available aggregate Cochrane records (Pairwise70). Estimand: Normalized Entropy Index (NEI, 0–1).