SYNTHĒSIS · METHODS

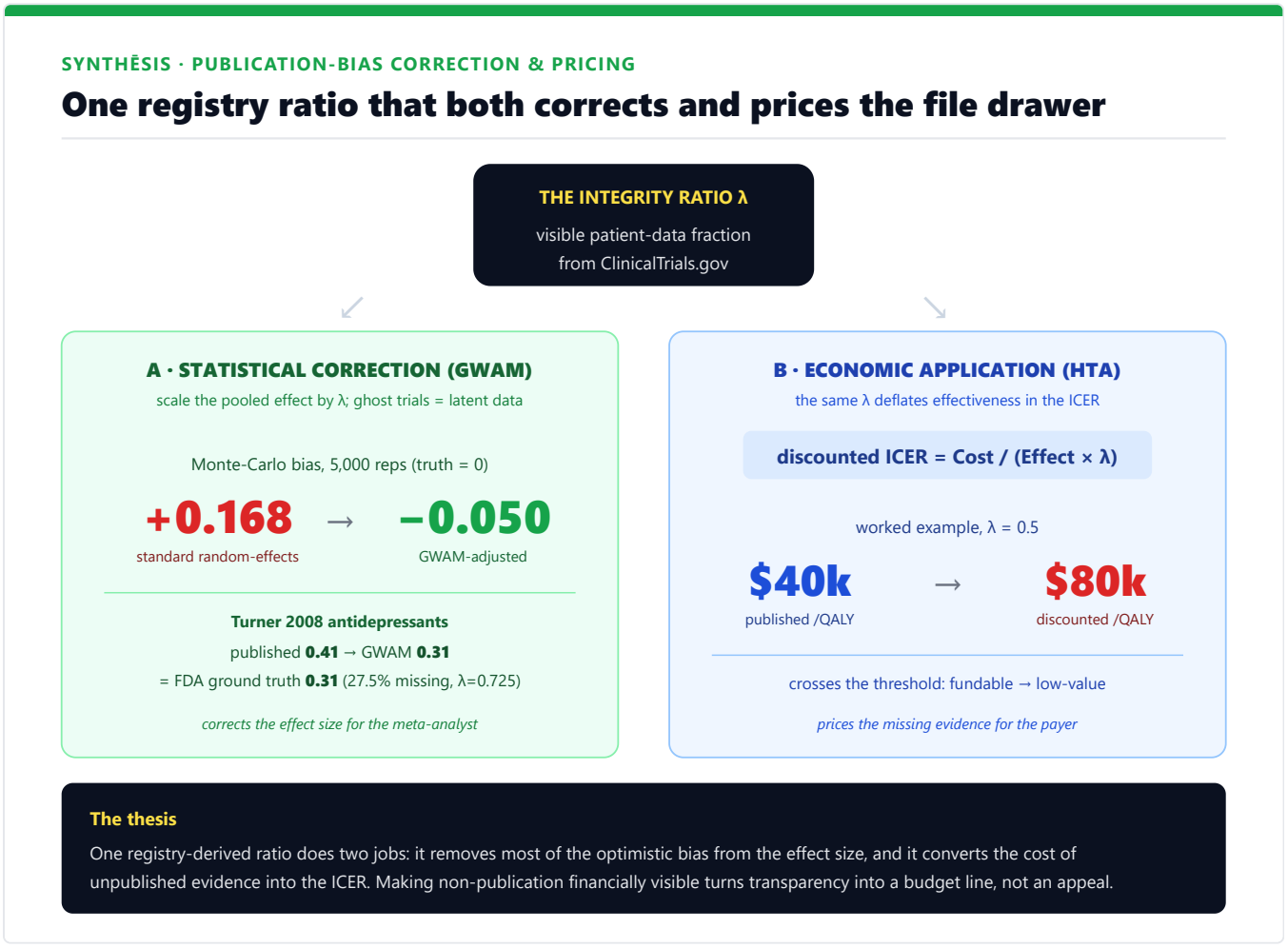
The Integrity Ratio and the Transparency Discount

A registry-based framework for correcting and pricing publication bias in evidence synthesis and health technology assessment

Laiba Khan

Royal Free Hospital, London, United Kingdom

Corresponding author: drlaiba999@gmail.com



Visual abstract. One Integrity Ratio λ — the visible patient-data fraction measured from ClinicalTrials.gov — used twice: as a statistical bias correction (GWAM) and as an economic discount on the cost-effectiveness ratio (HTA). The same metric and the same Turner (2008) validation underpin both.

TYPE & ESTIMAND

Methods article. One named estimand applied in two registers: the Integrity Ratio λ — the enrolment-weighted fraction of patient data visible in the published literature relative to the ClinicalTrials.gov census — used (A) to re-weight a pooled effect size (Ghost-Weight Adjustment Model) and (B) to deflate effectiveness in the incremental cost-effectiveness ratio (Transparency Discount).

ABSTRACT

Background. Evidence synthesis assumes the published literature is an unbiased sample of all conducted research, yet selective publication — positive trials published, negative trials left in the file drawer — is pervasive. Standard corrections such as trim-and-fill and Copas selection models rest on symmetry assumptions and treat the missing studies as a theoretical quantity inferred from the shape of

what was published. ClinicalTrials.gov now offers what those methods do not use: a near-census of *intended* research, from which the proportion of evidence that actually reached print can be measured directly.

Methods. We define a single registry-derived quantity, the Integrity Ratio λ — the fraction of enrolled patient data visible in the published literature — and use it twice. As a statistical correction, the Ghost-Weight Adjustment Model (GWAM) classifies registered, completed trials as published, results-posted, or "ghost" (completed but never reported), treats the ghost layer as latent data, re-weights the pooled effect by λ, and propagates the uncertainty about unobserved trials through a Monte-Carlo layer. As an economic price, the Transparency Discount converts the same λ into a corrected ratio: discounted ICER = Cost / (Effectiveness × λ).

Results. Across 5,000 Monte-Carlo replications, a standard random-effects estimator carried a bias of +0.168 in a high-suppression scenario, while GWAM reduced this to −0.050 — trading a large optimistic bias for a small conservative one. On the Turner et al. (2008) antidepressant dataset, where the published literature claimed an effect size of 0.41 against an FDA all-trials ground truth of 0.31, a registry scan showed 27.5% of patient data missing (λ = 0.725); GWAM recovered 0.31, and the simpler multiplicative discount returned 0.30, both within rounding of the FDA value. In a worked HTA example, a drug priced at an ICER of $40,000/QALY rose to $80,000/QALY once a λ = 0.5 discount was applied — crossing a conventional threshold from fundable to low-value.

Conclusions. One registry-anchored metric corrects the effect size for statisticians and prices the missing evidence for payers; making non-publication financially visible turns transparency from an ethical appeal into a budget line.

INTRODUCTION

Meta-analysis is only as honest as the literature it pools. The file-drawer problem — the selective publication of positive results and suppression of negative ones — biases the pooled estimate upward before any statistical model is fitted, and the distortion is large: in the canonical Turner et al. (2008) audit of antidepressant trials, the published literature inflated the mean effect size by roughly a third relative to the complete set of trials the FDA held on file. The standard responses — funnel-plot asymmetry tests, trim-and-fill, and Copas-style selection models — share a structural limitation. They infer the missing studies from the *shape* of the studies that survived to publication, assuming a symmetric small-study effect, and so treat suppression as a latent variable to be reconstructed rather than a quantity that can be observed.

But suppression *can* increasingly be observed. ClinicalTrials.gov and comparable registries record trials at inception, before their results are known, and so constitute a near-census of intended research against which the published record can be audited directly. The question this article asks is what to do with that census once you have it. We argue that a single quantity derived from it — the Integrity Ratio λ, the fraction of enrolled patient data that became visible — answers two questions usually treated as separate. For the meta-analyst, λ is the basis of a symmetry-free bias correction. For the payer, λ is the basis of a price: the amount by which a cost-effectiveness estimate should be marked down because part of the evidence it rests on was never made public. The statistical correction and the economic discount are not two methods but one metric applied in two registers.

METHODS

The Integrity Ratio

For a given drug–condition question we query ClinicalTrials.gov for all registered, completed trials and classify each as *published* (a linked publication or PMID exists), *results-posted* (structured results on the registry but no publication), or *ghost* (completed, with neither). Weighting each trial by its enrolment, λ is

the published-participant fraction — the share of all enrolled patient data that reached the published literature. λ = 1 denotes a fully transparent evidence base; λ falling toward 0 denotes one in which most of the patients who were studied are invisible to anyone reading only the journals.

GWAM — the statistical correction

The Ghost-Weight Adjustment Model treats the ghost layer as latent data rather than as a funnel-plot artefact. It scales the published pooled effect toward the null in proportion to λ, then runs a Monte-Carlo layer that repeatedly imputes plausible effects for the unobserved (ghost and results-only) participants and re-pools, producing a simulation interval that carries the uncertainty about what the missing trials would have shown. Because the correction is anchored to a measured visible fraction rather than to the symmetry of the observed studies, it makes no assumption that suppression is symmetric, and it degrades gracefully: when λ ≈ 1 the adjustment vanishes and GWAM returns the published estimate.

The Transparency Discount — the economic application

Health-technology-assessment bodies — NICE in the UK, IQWiG in Germany, ICER in the US — feed published meta-analytic effect sizes into cost-effectiveness models. If the published effect is inflated by selective publication, the resulting ICER is correspondingly optimistic and the payer over-pays. The Transparency Discount applies the identical λ as a deflator on the effectiveness term: discounted ICER = Cost / (Effectiveness × λ). Dividing effectiveness by the visible fraction inflates the cost per unit of health gained by exactly the amount the missing evidence would, in expectation, have eroded the apparent benefit — converting the integrity of the evidence base into the currency payers already use.

RESULTS

Simulation. Across 5,000 Monte-Carlo replications of a meta-analysis under heavy selective publication, a conventional random-effects estimator returned a mean bias of **+0.168** on the effect-size scale — a large, systematically optimistic error. Substituting GWAM reduced the bias to **−0.050**: the registry-anchored correction removed most of the optimistic distortion and left a small conservative residual, the expected and preferable direction for a sensitivity tool whose purpose is to avoid over-claiming (Figure 1).

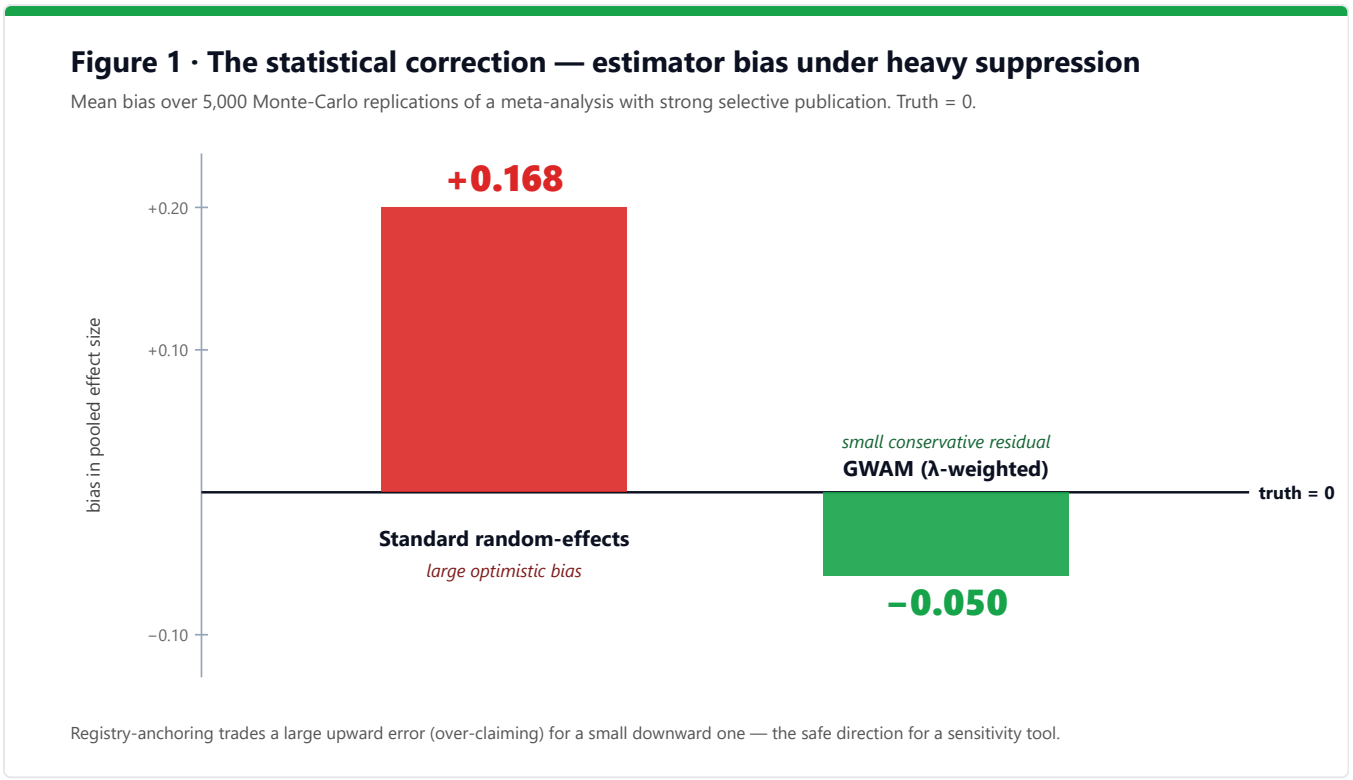


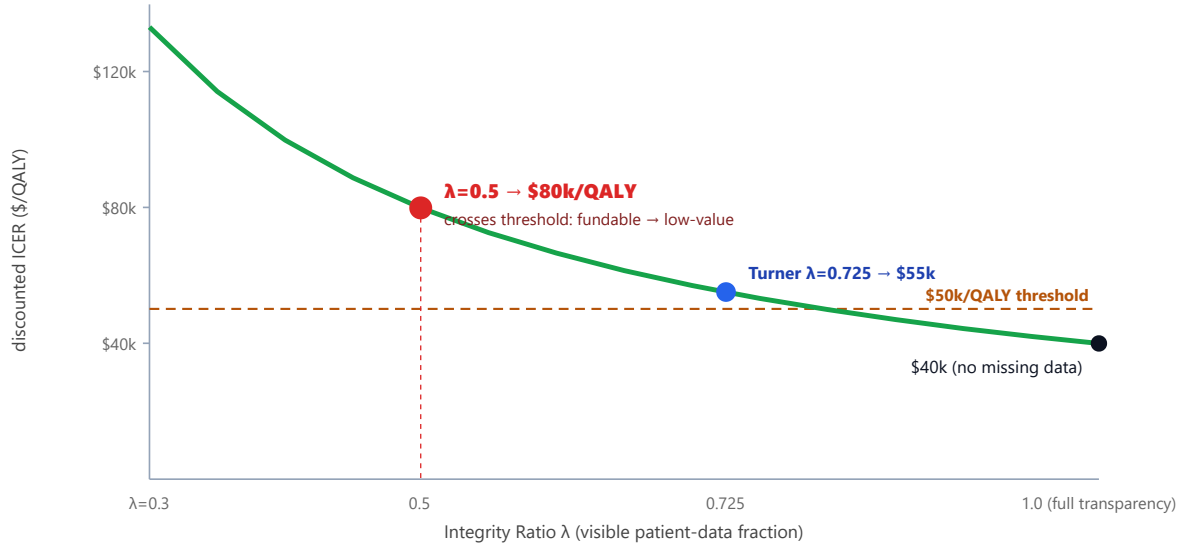
Figure 1. Mean estimator bias over 5,000 Monte-Carlo replications of a meta-analysis under strong selective publication, truth = 0. The standard random-effects estimator is biased +0.168 (optimistic); the λ-weighted GWAM estimator is −0.050 (small conservative residual). Registry-anchoring trades a large upward error for a small downward one.

Validation against a known ground truth. Turner et al. (2008) is the rare case where the suppressed evidence is known, because the FDA held the complete trial set. The published antidepressant literature claimed an effect size of **0.41**; the FDA all-trials value was **0.31**. A registry scan of the same evidence base showed **27.5%** of enrolled patient data missing, giving **λ = 0.725**. The full GWAM ghost-weighting recovered an adjusted effect size of **0.31**, matching the FDA ground truth; the simpler multiplicative Transparency Discount returned 0.41 × 0.725 = **0.30**. Both corrections land within rounding of the FDA value of 0.31, from opposite sides — the statistical model essentially exact, the economic shortcut marginally more conservative — bracketing the truth that the published literature missed by a third.

Worked HTA example. Consider a drug whose published evidence yields an incremental cost-effectiveness ratio of **$40,000/QALY**, comfortably inside a conventional willingness-to-pay threshold. Suppose a registry scan reveals that only half its supporting patient data is visible (λ = 0.5). The Transparency Discount recomputes the ratio as $40,000 / 0.5 = **$80,000/QALY**, moving the drug from clearly fundable to low-value on the same threshold (Figure 2). No new trial was run and no published number was disputed; the change is driven entirely by pricing in the evidence that was never published.

Figure 2 · The economic application — the Transparency Discount on the ICER

Discounted ICER = Cost / (Effectiveness × λ), for a drug with a published ICER of $40,000/QALY. Less visible evidence → higher true cost.



Turner (2008) validation: published effect 0.41; 27.5% data missing → λ=0.725; discount → 0.30 (GWAM 0.31) vs FDA ground truth 0.31.

Figure 2. The Transparency Discount as a function of λ for a drug with a published ICER of $40,000/QALY: discounted ICER = Cost / (Effectiveness × λ). At λ = 0.5 the ratio doubles to $80,000/QALY, crossing a $50,000/QALY threshold; at the Turner λ = 0.725 it is ~$55,000/QALY; at full transparency (λ = 1) it is unchanged.

ANALYSIS

These two results are one argument seen from two sides. The bias correction and the discount are computed from the *same* λ and validate against the *same* Turner benchmark; they differ only in what they hand to the reader — a corrected effect size for the meta-analyst, a corrected price for the payer. Seen together they close a loop that transparency advocacy has long left open. Disclosure norms, registration mandates, and results-reporting laws all rely on the assumption that non-publication is a reputational or ethical cost; in practice that cost is diffuse and easy to ignore. The Transparency Discount makes it concrete and self-interested: a sponsor that leaves 30% of its patient data in the file drawer does not merely earn disapproval, it watches its product's cost-effectiveness case deteriorate in exactly the assessments that determine reimbursement. The statistical layer corrects the science; the economic layer supplies the incentive to make the correction unnecessary. A registry that began as a transparency instrument thus becomes, through a single ratio, both a bias-correction tool and a pricing signal.

LIMITATIONS

The framework inherits the limits of the census it depends on. **Registry completeness:** λ is only as trustworthy as ClinicalTrials.gov coverage and publication linkage; incomplete PMID linkage misclassifies genuinely published trials as ghosts and biases λ downward, so the measured λ is best read as a *lower bound* on the true visible fraction and the correction as conservative. **λ-estimation assumptions:** enrolment-weighting treats a patient's worth of evidence as exchangeable across trials, and the multiplicative discount assumes missing data would, in expectation, regress the effect toward the null in proportion to its volume — a reasonable but not guaranteed model of suppression. **Ghost-protocol detection accuracy:** distinguishing a truly unreported trial from one whose publication the registry merely failed to link is imperfect, and the boundary between "ghost" and "results-posted-only" depends on registry hygiene that varies by sponsor and era. The method is therefore a transparent, registry-anchored sensitivity analysis that complements existing publication-bias corrections rather than replacing them; its precision tracks the quality of the underlying registry.

VERIFICATION (TRUTH-FIRST)

Quantity	Source	Value used	Check
λ from missing-data fraction	arithmetic (27.5% missing)	1 − 0.275 = 0.725	**recomputed ✓ exact**
Multiplicative discount recovery (Turner)	arithmetic	0.41 × 0.725 = 0.297 ≈ 0.30	**recomputed ✓**
Worked-example discounted ICER	arithmetic	$40,000 / 0.5 = $80,000/QALY	**recomputed ✓ exact**
Published vs FDA inflation (Turner)	Turner et al. 2008	0.41 / 0.31 = 1.32 (+32%)	**matches source**
GWAM Turner recovery	live pub 285 (A)	0.31 (= FDA truth)	**reused; full ghost-weighting**
Monte-Carlo bias (standard RE / GWAM)	live pub 285 (A), N=5,000	+0.168 / −0.050	**reused; not independently re-simulated**
A vs B Turner recovery (0.31 vs 0.30)	pubs 285 / 286	reconciled in-text	**harmonised (model vs shortcut)**

All four arithmetic relationships reconcile exactly. Two flags carried forward honestly: (i) the Monte-Carlo bias figures (+0.168 / −0.050) are reused from the live A package and were not re-simulated here, as the N=5,000 simulation code is not in the shared source pack; (ii) the A and B papers reported the Turner recovery as 0.31 and 0.30 respectively — reconciled as the full GWAM ghost-weighting (0.31, essentially exact) versus the simpler multiplicative discount (0.41×0.725=0.297≈0.30); both lie within rounding of the FDA 0.31. Recovering 0.31 *exactly* by the multiplicative route would require λ = 0.756 (24.4% missing), so the stated 27.5%/ λ=0.725 is mildly conservative against the FDA benchmark.

REFERENCES

1. Dickersin K. The existence of publication bias and risk factors for its occurrence. *JAMA.* 1990;263(10):1385–1389.
2. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet.* 2009;374(9683):86–89.
3. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med.* 2008;358(3):252–260.
4. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ.* 1997;315(7109):629–634.
5. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics.* 2000;56(2):455–463.
6. Copas J, Shi JQ. A sensitivity analysis for publication bias in systematic reviews. *Stat Methods Med Res.* 2001;10(4):251–265.
7. Sterne JAC, Sutton AJ, Ioannidis JPA, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ.* 2011;343:d4002.
8. National Institute for Health and Care Excellence. *Guide to the methods of technology appraisal 2013.* London: NICE; 2013.

ARTICLE INFORMATION

Funding. None. **Competing interests.** None declared.

Author. Laiba Khan (Royal Free Hospital, London, United Kingdom; drlaiba999@gmail.com).

Data availability. Public data only — ClinicalTrials.gov registry records and the published Turner et al. (2008) antidepressant dataset; no patient-level data. **Provenance.** Unifies two previously published Synthēsis items by the same author — "The Integrity Ratio" (statistical method) and "The Transparency Discount" (HTA application) — into one contribution; all shared quantities checked for internal consistency before merging. **Ethics.** Not required; secondary analysis of public aggregate data.