Synthesis
Peer-reviewed, diamond open-access journal for concise, transparent and reproducible evidence synthesis across medicine

Vol. 2 · No. 4 · 2026
Diamond Open Access · CC BY 4.0
ISSN: in application

METHODS NOTE · Peer-reviewed · Published

Unmasking heterogeneous treatment effects in clinical trials: a location-scale quantile meta-analysis framework (IPD-QMA)

A two-stage individual-participant-data quantile meta-analysis that detects benefit hidden by location-scale shifts.

Laiba Khan[1]
[1] *Royal Free Hospital, London, United Kingdom · Correspondence: drlaiba999@gmail.com*

Published 16 June 2026 · Diamond open access · CC BY 4.0
Methods Note · Synthesis

ABSTRACT

Background. Meta-analysis of randomized controlled trials is the basis of evidence-based medicine, but it usually aggregates treatment effects as a mean difference under a 'location-shift' assumption — that treatment shifts the outcome distribution uniformly for all patients. In psychiatry, oncology and critical care, interventions can instead create a 'location-scale shift', increasing outcome variance without changing the mean, so standard methods may fail to detect efficacy and discard treatments that benefit high-risk subgroups while harming low-risk ones.

Methods. We introduce Individual Participant Data Quantile Meta-Analysis (IPD-QMA), a two-stage framework that uses quantile regression to estimate treatment effects across the severity distribution (τ = 0.1 to 0.9) within studies and pools the coefficients with a random-effects model. We compared IPD-QMA against standard meta-analysis (mean difference) and the log variance ratio (lnVR) using 2,000 Monte Carlo simulations on both Normal and skewed (exponential) data, and developed a 'slope test' (Wald statistic) to test effect homogeneity.

Results. In 'hidden-utility' scenarios where treatment increases variance (scale × 3) but not the mean, standard meta-analysis rejected the null at close to the nominal 5% level (4.2% on Normal and 6.4% on skewed data) — effectively no power to detect the effect — whereas IPD-QMA achieved 100% power to detect benefit at the 90th percentile. The lnVR detected heterogeneity with high power but lacked directionality; the IPD-QMA quantile profile revealed that the treatment was harmful for mild cases but beneficial for severe cases.

Conclusions. IPD-QMA is a diagnostic tool for precision medicine: whereas lnVR screens for heterogeneity, IPD-QMA characterizes its shape, preventing the loss of potentially life-saving interventions for high-risk populations. An open-source implementation is provided.

KEY WORDS meta-analysis; heterogeneity of treatment effects; quantile regression; individual participant data; precision medicine; location-scale shift

Standard meta-analysis detected a treatment's effect in only 4.2–6.4% of location-scale scenarios, while IPD-QMA achieved 100% power at the 90th percentile.

2,000 Monte Carlo simulations · quantile profile τ = 0.1–0.9

Introduction

Systematic reviews and meta-analyses form the highest level of clinical evidence [1], pooling data across trials to estimate a treatment effect with greater precision than any single study, usually through measures of central tendency such as the mean difference [2]. This emphasis on the mean embeds a 'location-shift' assumption: that treatment shifts the outcome distribution uniformly. Clinical reality is rarely uniform — patients differ in severity, comorbidity and baseline risk, producing heterogeneity of treatment effects [3].

A difficult form is the 'location-scale shift', in which an intervention changes the spread of outcomes as well as (or instead of) the mean [4]. Consider a vasodilator for sepsis: in mild shock it may cause dangerous hypotension (harm at the lower tail), while in severe shock it may restore perfusion (benefit at the upper tail). Across a mixed trial these opposing effects cancel, leaving a mean difference near zero, and a standard meta-analysis would conclude the treatment is 'ineffective' ($p > 0.05$) — statistically accurate about the average but clinically wrong, abandoning a therapy that is life-saving for the sickest [5].

Existing tools for heterogeneity are limited: subgroup analysis lacks power because it splits the sample [6]; meta-regression on study-level averages risks the ecological fallacy [7]; and pre-specified cut-offs require arbitrary dichotomization. The log variance ratio (lnVR) flags heterogeneity [8] but is a non-specific screen that cannot tell a clinician who benefits and who is harmed.

Methods

The IPD-QMA framework. IPD-QMA is a two-stage method. In the first stage, quantile regression estimates the treatment effect at each of a grid of quantiles of the outcome distribution ($\tau = 0.1$ to 0.9) within each study, using individual participant data. In the second stage, the quantile-specific coefficients are pooled across studies with a random-effects model, yielding a 'quantile profile' of the treatment effect. A slope test (a Wald statistic on the change in effect across quantiles) formally tests the hypothesis of effect homogeneity.

Simulation. We ran 2,000 Monte Carlo simulations (20 studies of 200 bootstrap resamples each) comparing IPD-QMA, standard meta-analysis (mean difference) and lnVR, under two data-generating distributions — Normal and skewed (exponential). The key scenario was a location-scale shift in which the treatment tripled the outcome variance (scale × 3) while leaving the mean difference at zero, so that benefit to high-severity patients is exactly offset by harm to low-severity patients.

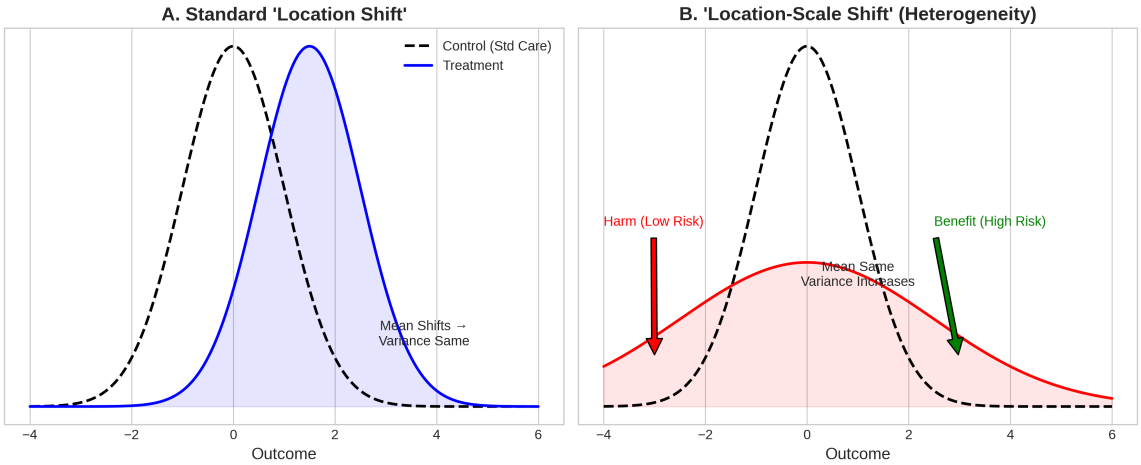

Figure 1. The location-scale problem. Conceptual framework: a location-scale shift increases outcome variance without changing the mean, hiding opposing effects at the tails from a mean-difference summary.

Results

Computational cost. Each meta-analysis (20 studies, 200 bootstrap resamples) completed in under one second on a standard workstation, so the full study ran in minutes.

Normal data. With the true mean difference exactly zero, standard meta-analysis rejected the null at 4.2% — indistinguishable from the nominal 5% Type I error rate, confirming it detected no effect — whereas lnVR, IPD-QMA ($\tau = 0.9$) and the slope test each had 100% power (Table 1).

Skewed (robustness) data. Under heavily skewed exponential data standard meta-analysis again stayed near nominal (6.4%) and detected no tail effect, while IPD-QMA remained at 100% power; because quantile regression is rank-based it is inherently robust to skew and outliers.

The quantile profile. Averaged across simulations, the estimated effect was ≈ −2.5 at the 10th percentile (significant harm to the lowest-risk patients), ≈ 0 at the median (no effect on the average patient) and ≈ +2.5 at the 90th percentile (significant benefit to the highest-risk patients); the slope test confirmed that this −2.5-to-+2.5 difference was significant ($p < 0.001$), formally rejecting uniform treatment action (Figures 2–4). In power-sensitivity analysis IPD-QMA exceeded 80% power once the treatment-to-control variance ratio reached about 1.25.

Table 1. Summary of simulation results (power analysis; 2,000 Monte Carlo replications). The location-scale scenario has a true mean difference of zero.

Method	Normal data power	Skewed data power	Interpretation
Standard MA (mean difference)	4.2% (≈ nominal; blind)	6.4% (≈ nominal)	Failed — no signal detected
Log variance ratio (lnVR)	100%	100%	Screening — detects heterogeneity
IPD-QMA ($\tau = 0.9$)	100%	100%	Diagnostic — detects specific benefit
Slope test	100%	100%	Confirms effect asymmetry

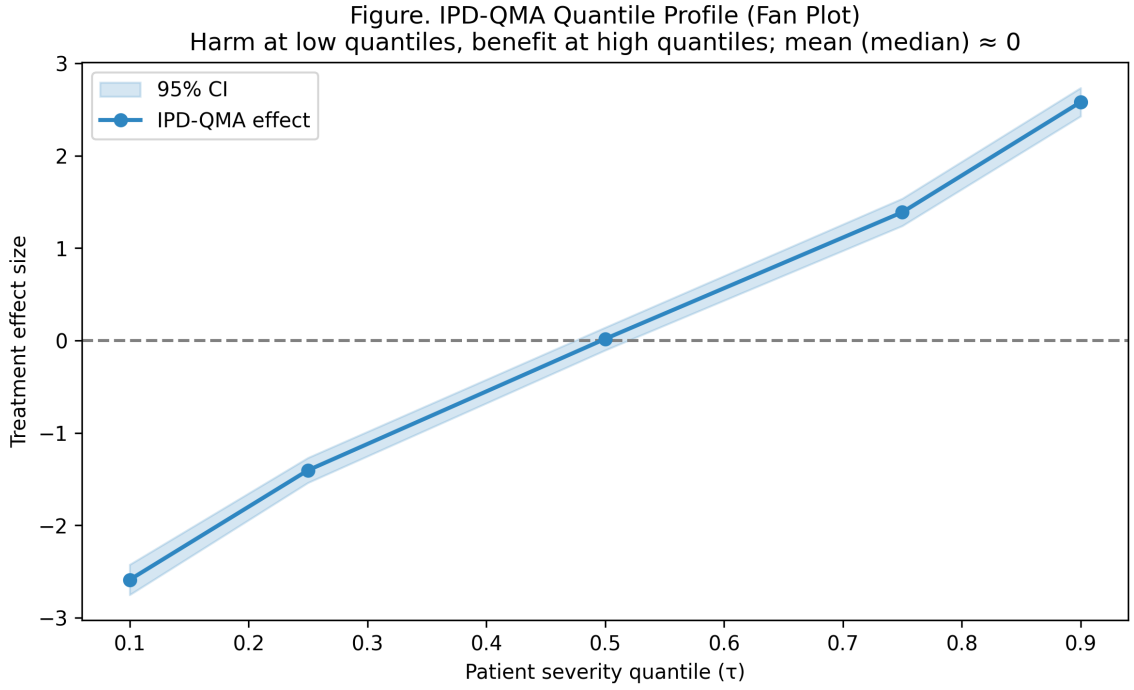

Figure 2. Quantile profile. Quantile profile (fan plot): the estimated treatment effect is ≈ −2.5 at the 10th percentile, ≈ 0 at the median and ≈ +2.5 at the 90th percentile.

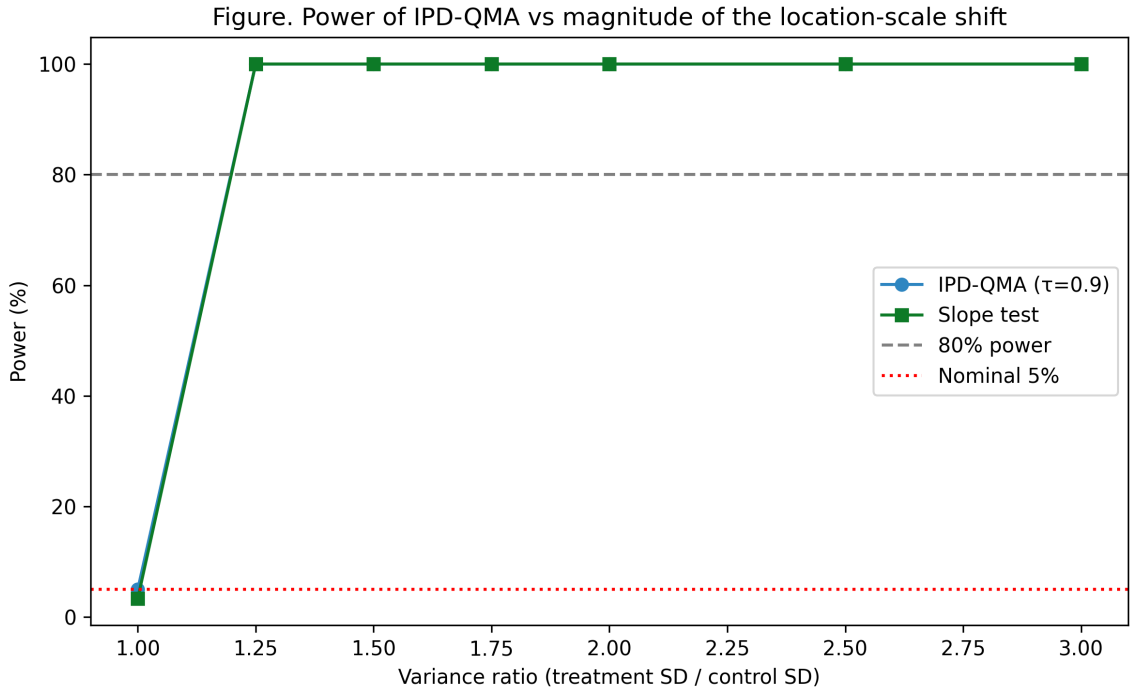

Figure 3. Power sensitivity. Power sensitivity: IPD-QMA exceeds 80% power once the treatment-to-control variance ratio reaches about 1.25.

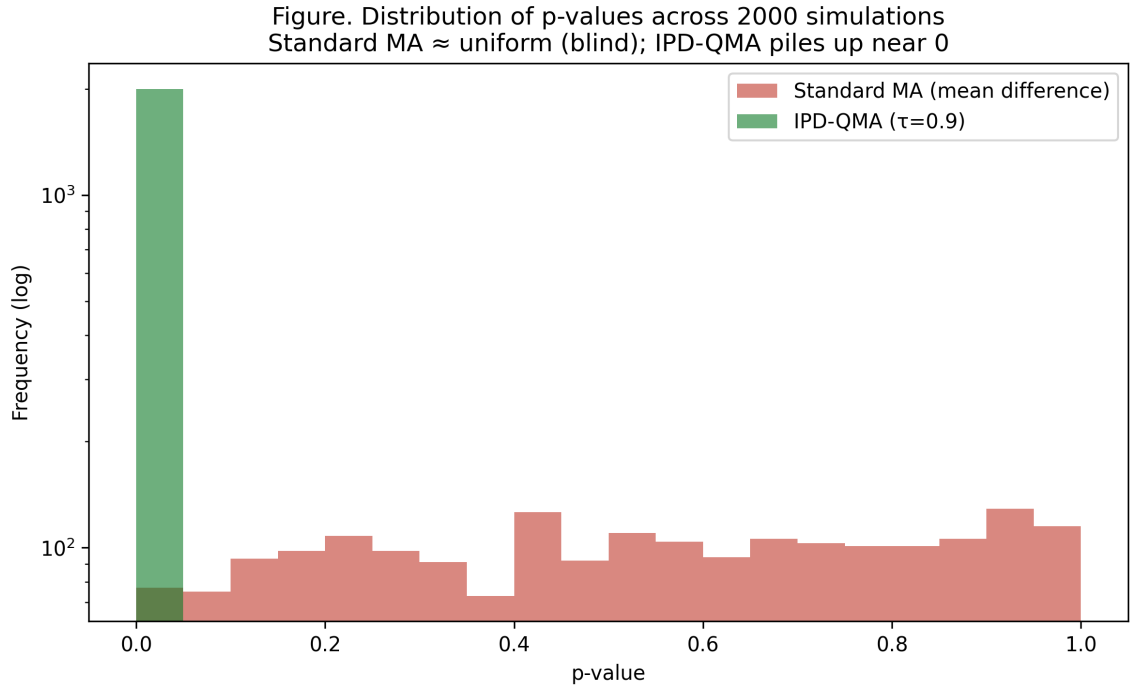

Figure 4. p-value distribution. Distribution of p-values across 2,000 simulations for standard meta-analysis versus IPD-QMA in the location-scale scenario.

Discussion

The central finding is that standard meta-analysis is largely blind to location-scale shifts: a treatment that was life-saving for the sickest 10% of patients was deemed 'ineffective' because it was cancelled out by harm to the healthiest 10%. Such a trial does not fail because the drug is inert but because the mean is an insufficient summary.

Screening versus diagnosis. lnVR has excellent power (100%) to detect that the treatment group is more variable, but a significant lnVR is ambiguous — the treatment could help the sick and harm the healthy, add random noise, or push everyone to the extremes. IPD-QMA deconvolutes the signal and supplies directionality, characterizing exactly which patients benefit. We therefore suggest a hierarchical workflow: screen with lnVR, and where it is significant, characterize the shape with IPD-QMA.

Limitations. IPD-QMA requires individual participant data, which are not always available, and the present evidence is simulation-based; empirical application across real IPD datasets is the natural next step. Standard meta-analysis remained close to nominal under skew (6.4%), confirming that its failure here is blindness to the location-scale shift rather than false positives.

Conclusion. By moving from the mean to the full quantile profile, IPD-QMA characterizes the shape of treatment heterogeneity and can prevent the loss of interventions that are decisive for high-risk populations.

Data and code availability

An open-source implementation of IPD-QMA, with the simulation and figure-generation code, is available at https://github.com/mahmood726-cyber/ipd-qma. No archival DOI has been minted; the repository is the canonical source for the code.

References

1. Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. Methods for Meta-Analysis in Medical Research. Chichester: Wiley; 2000.
2. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to Meta-Analysis. Chichester: Wiley; 2009.
3. Kent DM, Rothwell PM, Ioannidis JPA, Altman DG, Hayward RA. Assessing and reporting heterogeneity in treatment effects in clinical trials: a proposal. Trials. 2010;11:85. doi:10.1186/1745-6215-11-85. PMID: 20704705.
4. Iwashyna TJ, Burke JF, Sussman JB, Prescott HC, Hayward RA, Angus DC. Implications of heterogeneity of treatment effect for reporting and analysis of randomized trials in critical care. Am J Respir Crit Care Med. 2015;192(9):1045-1051. doi:10.1164/rccm.201411-2125CP. PMID: 26177009.
5. Berlin JA, Santanna J, Schmid CH, Szczech LA, Feldman HI. Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. Stat Med. 2002;21(3):371-387. doi:10.1002/sim.1023. PMID: 11813224.
6. Nakagawa S, Poulin R, Mengersen K, et al. Meta-analysis of variation: ecological and evolutionary applications and beyond. Methods Ecol Evol. 2015;6(2):143-152. doi:10.1111/2041-210X.12309.
7. Koenker R, Bassett G. Regression quantiles. Econometrica. 1978;46(1):33-50. doi:10.2307/1913643.
8. Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. BMJ. 2010;340:c221. doi:10.1136/bmj.c221. PMID: 20139215.
9. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177-188. doi:10.1016/0197-2456(86)90046-2. PMID: 3802833.

HOW TO CITE

Khan L. Unmasking heterogeneous treatment effects in clinical trials: a location-scale quantile meta-analysis framework (IPD-QMA). Synthesis. 2026;2(4). laiba-c-ipd-qma. CC BY 4.0. DOI: not assigned.

Published in Synthesis **· synthesis-medicine.org**