Extreme Heterogeneity in Global Prevalence Meta-Analyses: Evaluating Current Practices and Exploring Bayesian Alternatives - an Umbrella Review

Víctor Juan Vera-Ponce[1,*], Jhosmer Ballena-Caicedo[1], Lupita Ana Maria Valladolid-Sandoval[1] and Fiorella E. Zuzunaga Montoya[2]

[1]*Facultad de Medicina (FAMED), Universidad Nacional Toribio Rodríguez de Mendoza deAmazonas (UNTRM), Amazonas, Perú*

[2]*Universidad Continental, Lima, Perú*

Abstract: *Introduction*: Global prevalence meta-analyses often exhibit extreme heterogeneity (I^2 > 90%), yet criteria designed for clinical trials, where homogeneity is desirable, continue to be applied without recognizing that in prevalence studies, variability reflects real differences between populations.

Objective: To document the magnitude of heterogeneity in global prevalence meta-analyses, evaluate the methodological strategies employed for its exploration and management, and explore through illustrative application how Bayesian methods—rarely employed in prevalence meta-analyses—compare with standard frequentist approaches.

Methods: Umbrella review conducted according to PRIOR guidelines. Systematic search in SCOPUS for systematic reviews with global/worldwide prevalence meta-analyses published between 2015-2025. Data were extracted on I^2, statistical models, subgroup analyses, sensitivity analyses, meta-regression, and prediction intervals. Three meta-analyses were randomly selected for illustrative Bayesian re-analysis using hierarchical models with weakly informative priors, and the results were compared with those from frequentist approaches.

Results: Of 53 included meta-analyses, 52 (98.1%) presented I^2≥75%, 47 (88.7%) I^2≥90%, and 34 (64.2%) I^2>99%. Management strategies showed a decreasing implementation rate: subgroup analyses (96.2%), sensitivity analyses (64.2%), meta-regression (34.0%), and prediction intervals (5.8%). Among studies with I^2≥75%, 63.5% provided explicit justification for proceeding with pooling. The illustrative Bayesian analysis of three randomly selected studies demonstrated excellent concordance with frequentist estimates (differences <0.1%), while providing additional uncertainty quantification for heterogeneity parameters unavailable from standard approaches.

Conclusions: Extreme heterogeneity constitutes the norm in global prevalence meta-analyses. The underutilization of prediction intervals (5.8%) and meta-regression (34.0%) represents the critical gap for improving statistical rigor. An exploratory Bayesian analysis demonstrated concordance with frequentist estimates, while providing additional uncertainty quantification. This illustrates that alternative methods are feasible, though their value lies primarily in specific scenarios rather than routine application. Prevalence-specific frameworks should recognize high heterogeneity as an expected characteristic requiring comprehensive exploration rather than elimination.

Keywords: Meta-Analysis, Prevalence, Epidemiologic Methods, Biostatistics, Systematic Reviews, Heterogeneity, Publication Bias, Research Design, Evidence-Based Medicine, Public Health.

INTRODUCTION

A common observation when reviewing prevalence meta-analyses is the consistent presence of extremely high I^2 values, frequently exceeding 90-95%. This pattern appears so ubiquitous that it raises a fundamental question: is this truly a universal characteristic of prevalence meta-analyses, or merely a perception based on selective examples? A study of 134 prevalence meta-analyses found a median I^2 of 96.9% (IQR 90.5-98.7), with 93% of studies presenting I^2 >70% [1,2], suggesting that extreme heterogeneity may indeed be the norm rather than the exception. This heterogeneity reflects real variations between populations, differences in operational definitions, and disparities in diagnostic methods [3], indicating that high I^2 values represent an expected finding rather than evidence of poor methodological quality.

However, the inevitability of extreme heterogeneity does not eliminate the need for appropriate statistical management. Traditional guidelines for meta-analysis recommend specific strategies when heterogeneity is substantial[4,5]. At minimum, random-effects models should be employed to account for between-study variability. Beyond this basic requirement, more robust approaches exist to investigate and report heterogeneity appropriately. These include conducting meta-regression or subgroup analyses to explore potential sources of heterogeneity, reporting prediction intervals alongside confidence intervals to convey the expected range of prevalences in different settings, using variance-stabilizing transformations appropriate for proportion data, and evaluating whether pooling remains interpretable when I^2 exceeds 95% [6-9]. Alternative statistical frameworks, such as Bayesian hierarchical models, may offer additional advantages in modeling complex heterogeneity patterns and incorporating prior knowledge about expected variability [10,11].

*Address correspondence to these authors at the Facultad de Medicina (FAMED), Universidad Nacional Toribio Rodríguez de Mendoza deAmazonas (UNTRM), Amazonas, Perú; E-mail: victor.vera@untrm.edu.pe

Despite these methodological recommendations, systematic documentation of statistical practices in published prevalence meta-analyses remains scarce. To address this gap, we focus specifically on global public health prevalence meta-analyses published in Q1 journals. This focus is justified for several reasons. First, global prevalence estimates are particularly relevant for public health planning and resource allocation, as they inform burden-of-disease calculations and international health policies. Second, Q1 journals represent the highest tier of methodological rigor and peer review, where adherence to statistical recommendations would be expected to be optimal. Third, global meta-analyses inherently include studies from diverse geographical regions and healthcare systems, maximizing the expected heterogeneity and therefore providing the most challenging scenario for statistical management. We do not expect that meta-analyses in other areas or lower-impact journals would demonstrate better statistical practices than those published in top-tier public health journals.

If extreme heterogeneity is truly inevitable in this context, then understanding how researchers manage this statistical challenge becomes critical. Specifically, it remains unknown whether meta-analyses systematically investigate sources of heterogeneity or simply acknowledge their presence, whether prediction intervals are routinely reported, and whether alternative statistical methods are considered when traditional approaches may be insufficient.

Therefore, the objective of this umbrella review (UR) is: (1) to confirm whether extreme heterogeneity (I^2>90%) is indeed universal across global prevalence meta-analyses published in Q1 journals; (2) to systematically evaluate what statistical methods authors employed to manage this heterogeneity; (3) to explore through illustrative application how Bayesian hierarchical models—a statistical framework rarely employed in prevalence meta-analyses—compare with standard frequentist approaches in terms of results, uncertainty quantification, and interpretability when applied to cases with extreme heterogeneity; and (4) to identify gaps between methodological recommendations and actual practice, providing evidence-based recommendations for improving statistical rigor in prevalence meta-analyses.

METHODS

Study Design

An UR of systematic reviews with meta-analyses was conducted, following the PRIOR guidelines [12] for reporting the synthesis of systematic reviews, adapted for prevalence studies. This study was specifically designed to evaluate statistical practices in the management of heterogeneity rather than to generate new prevalence estimates. Additionally, selected meta-analyses were re-analyzed using Bayesian methods to demonstrate practical implementation and compare results with original frequentist approaches, focusing on differences in uncertainty quantification and interpretability rather than generating novel prevalence estimates.

Search Strategy and Study Selection

The search strategy was performed in SCOPUS, implementing the terms: TITLE-ABS-KEY ("systematic review" OR "meta-analysis") AND TITLE-ABS ("prevalence") AND TITLE-ABS-KEY ("global" OR "worldwide" OR "international" OR "multi-country" OR "cross-national" OR "world health organization" OR WHO). This strategy deliberately captured studies synthesizing global prevalences or across multiple continents.

Subsequently, a systematic evaluation of each journal's quartile according to SCImago Journal Rank was conducted to identify articles published in Q1 journals in the area of Public Health, Environmental and Occupational Health. From these, articles corresponding to the top 100 globally positioned Q1 journals according to SCImago ranking were selected. By selecting public health, we were able to include journals from diverse areas including epidemiology, infectious diseases, epidemiology, preventive medicine, among others.

Eligibility Criteria

Systematic reviews with meta-analyses whose primary objective was estimating global or multicontinental prevalence of health conditions or events were included, defined as studies that included data from at least 3 continents or 10 countries from different geographical regions. All studies had to be published in peer-reviewed journals and have full-text access to allow comprehensive extraction of methodological data and heterogeneity management strategies.

Excluded were intervention meta-analyses not focused on prevalence, studies conducted solely in animals or basic research, narrative reviews without quantitative synthesis, and those studies presenting a total lack of minimum methodological data, such as absence of information about the statistical model employed, applied transformations, or any mention related to heterogeneity assessment. Studies were not excluded based on the magnitude of heterogeneity, as documenting the full spectrum of I^2 values was a primary objective.

Selection Process

Three authors (VJVP, JJBC and LAMVS) independently reviewed titles and abstracts of all identified records to determine eligibility. Subsequently, the same reviewers independently evaluated the full texts of potentially eligible articles. Discrepancies at any stage of the selection process were resolved through discussion with a fourth reviewer (FEZM).

Data Extraction

Three authors (VJVP, JJBC and LAMVS) independently extracted data using a standardized form. Discrepancies were resolved by consensus or through consultation with a fourthautor (FEZM) when necessary.

Publication characteristics were extracted including journal, publication year, SJR impact factor, and specific thematic area. Meta-analysis characteristics included number of primary studies included, total sample size, health condition or event evaluated, and reported pooled prevalence.

The I-squared (I^2) statistic was specifically documented when reported, recording the exact values provided by the original authors. In articles reporting multiple meta-analyses or evaluating different prevalence outcomes, the maximum reported I^2 value was recorded, following the conservative principle of documenting the scenario of greatest heterogeneity. This approach was chosen because our objective was to evaluate how researchers manage extreme heterogeneity when it occurs, rather than to calculate average heterogeneity across all possible analyses within a paper.

Methodological information was systematically extracted on: (a) the statistical model employed (fixed-effects versus random-effects models), including the specific estimation method when reported (DerSimonian-Laird, REML, Hartung-Knapp-Sidik-Jonkman, or others); (b) variance-stabilizing transform-ations applied to proportion data (Freeman-Tukey double arcsine, logit transformation, or generalized linear mixed models); and (c) additional heterogeneity metrics beyond I^2 (Cochran's Q statistic, tau-squared).

Evaluation of Heterogeneity Management Strategies

Critical to this study's objectives, we systematically evaluated the implementation of methodological strategies recommended for managing high heterogeneity. These strategies included: (1) conduct of subgroup analyses, documented as present or absent, recording the stratification criteria when specified (geographical region, population subgroup, methodological characteristics); (2) performance of sensitivity analyses, documented as present or absent regardless of the specific type of sensitivity analysis conducted; (3) application of meta-regression to explore sources of heterogeneity, documented as present or absent without detailing the specific covariates explored, as our focus was on whether this analytical approach was employed rather than evaluating the appropriateness of variable selection; (4) reporting of prediction intervals alongside confidence intervals, documented as present or absent, as prediction intervals communicate the expected range of prevalences in new settings and are particularly valuable when heterogeneity is substantial.

Additionally, we documented whether authors provided explicit justification for proceeding with pooling when heterogeneity was high ($I^2 \geq 75\%$). This was recorded dichotomously (present/absent) without evaluating the quality or scientific adequacy of such justifications, as our objective was to document reporting practices rather than critique the reasoning provided. Examples of justifications included statements acknowledging that heterogeneity is expected in prevalence data, mention of using random-effects models to account for heterogeneity, or discussion of whether pooling remained interpretable

Quality Assessment of Reviews

Three authors (VJVP, JJBC and LAMVS) independently assessed the methodological quality of included studies. Discrepancies were resolved through consensus discussion or, when necessary, consultation with a fourth author (FEZM). The assessment process employed two complementary instruments: A Measurement Tool to Assess systematic Reviews, version 2 (AMSTAR 2) [13], adapted for the specific context of prevalence studies, and [second tool name]. This dual approach was adopted because no single instrument fully captures the methodological particularities of prevalence meta-analyses, which present distinct challenges including the handling of proportions, variance transformations, and heterogeneity patterns that differ from those encountered in intervention meta-analyses.

Data Analysis

Our analysis adopted a methods-oriented descriptive approach, focusing on the frequency and magnitude of reported heterogeneity and the implementation of recommended statistical practices for its management. Each I^2 value was individually recorded as reported by the original authors, without initial categorical grouping.

The operational classification employed included low to moderate heterogeneity ($I^2 < 75\%$), high heterogeneity ($I^2 \geq 75\%$), and extreme heterogeneity ($I^2 \geq 90\%$), following thresholds traditionally used in the meta-analysis literature. An additional threshold of $I^2>99\%$ was examined given preliminary observations during article screening suggesting this level of heterogeneity was common.

For studies with $I^2 \geq 75\%$, we calculated the proportion implementing each of the following methodological practices: explicit justification for proceeding with pooling despite high heterogeneity, performance of sensitivity analyses, conduct of meta-regression to explore heterogeneity sources, and reporting of prediction intervals. These practices were selected based on methodological recommendations in the literature for managing substantial heterogeneity in meta-analyses.

The analysis was entirely descriptive, documenting frequencies and proportions of methodological practices. I^2 values were not recalculated, pooling procedures were not repeated, and no attempt was made to re-analyze the original meta-analyses, given that: (a) not all articles reported the necessary inputs for recalculation; (b) access to primary study data was not available; and (c) the objective was to document published methodological practices as they appear in the literature rather than to generate new prevalence estimates or validate reported statistics.

All data extraction and analysis were performed in Microsoft Excel by two authors (JJBC and VJVP). Discrepancies were resolved through consensus discussion between both reviewers. No formal statistical testing was conducted as the study objective was descriptive documentation of methodological practices rather than hypothesis testing.

Illustrative Bayesian Analysis

To demonstrate the practical viability of Bayesian methods for extreme heterogeneity, three meta-analyses were randomly selected from the 53 included studies using simple random sampling without replacement. For each selected study, raw data were extracted from published materials when available. Bayesian random-effects meta-analysis was implemented using the bayesmeta package in R [14], with Freeman-Tukey double arcsine transformation for proportions [5] . Prior distributions were specified as uniform improper prior for the overall effect μ and weakly informative half-Normal prior (scale = 0.5) for between-study heterogeneity τ [15]. This prior specification allows data to dominate while preventing implausible values. Posterior distributions were generated via Markov Chain Monte Carlo sampling with convergence diagnostics [16]. Results were compared with frequentist REML estimates to assess concordance. Two authors (JJBC and VJVP) independently performed all analyses. Discrepancies were resolved through consensus discussion between both reviewers. Complete R code is provided in supplementary material.

RESULTS

Description of the Selection Process

The search strategy implemented in SCOPUS using the aforementioned search strategy initially identified 16,119 records for the period 2000-2025. Systematic evaluation of each journal's quartile according to SCImago Journal Rank identified 3,512 articles published in Q1 journals across all thematic areas, specifically 1,522 from the mentioned area of Public Health, Environmental and Occupational Health. From these, articles corresponding to the top 100 globally positioned Q1 journals according to SCImago ranking were selected, resulting in approximately 80 eligible articles. Finally, applying eligibility criteria resulted in a final sample of 53 prevalence meta-analyses that met all inclusion requirements for statistical heterogeneity analysis and characterization of associated methodological practices [17-69].

General Characteristics of Included Studies

Fifty-three meta-analyses from diverse public health areas were included, such as infectious diseases, epidemiology, and global health (see Table **1** and supplementary material 1). All studies (53/53, 100%) reported the I^2 statistic. The heterogeneity distribution showed that 52/53 (98.1%) presented $I^2\geq75\%$, while only 1/53 (1.9%) exhibited $I^2<75\%$ corresponding to a study of HBV-HIV comorbidity with $I^2=0\%$. Among studies with high heterogeneity, 47/53 (88.7%) reached $I^2\geq90\%$ and 34/53 (64.2%) showed $I^2>99\%$, demonstrating that extreme heterogeneity constitutes the epidemiological norm in prevalence meta-analyses published by elite journals.

Regarding methodological practices for managing heterogeneity, subgroup analysis was implemented in 51/53 (96.2%), sensitivity analysis in 34/53 (64.2%), and meta-regression in 18/53 (34.0%). A total of 52/53 (98.1%) studies employed at least one strategy to explore heterogeneity. Among meta-analyses with $I^2\geq75\%$ (n=52), no reports of prediction intervals were identified in the extracted data. Explicit justification for proceeding with pooling despite high heterogeneity was documented variably across studies.

Table 1: Characteristics and Methodological Practices of Included Meta-Analyses (Q1 review)

Characteristic / Practice	n/N (%)
Prevalence meta-analyses included	53/53 (100%)
Reported I^2	53/53 (100%)
Distribution of I^2:	
$I^2 = 0\%$	1/53 (1.9%)
$I^2 > 0\%$ and $< 75\%$	0/53 (0%)
$I^2 \geq 75\%$ (high)	52/53 (98.1%)
$I^2 \geq 90\%$ (very high)	47/53 (88.7%)
$I^2 \geq 95\%$	45/53 (84.9%)
$I^2 > 99\%$	34/53 (64.2%)
Strategies for managing heterogeneity:	
Subgroup analyses	51/53 (96.2%)
Sensitivity analyses	34/53 (64.2%)
Meta-regression	18/53 (34.0%)
Any strategy	52/53 (98.1%)

Examples of studies with extreme heterogeneity that provided justification include: Niu (2023)[17-69] with I^2=100% who noted "heterogeneity is assumed for global data"; Whitten (2024) with I^2=99.9% who indicated "high heterogeneity is expected due to regional and demographic differences"; Albadrani (2024) with I^2=99% who established the use of DerSimonian-Laird random-effects models; and Suleiman (2025) with I^2=99.80% who proceeded with analysis using random-effects models. Studies were published in journals including BMC Public Health (n=11), Journal of Global Health (n=6), Journal of Infection and Public Health (n=5), Infectious Diseases of Poverty (n=4), and Frontiers in Public Health (n=4).

Implementation of Methodological Practices in Studies with High Heterogeneity

Among the 52 meta-analyses with $I^2 \geq 75\%$, the following frequency of methodological practices was observed: reporting of $I^2 \geq 75\%$ in 52/52 (100%), explicit justification for pooling despite high heterogeneity in 33/52 (63.5%), performance of sensitivity analyses in 34/52 (65.4%), conduct of meta-regression in 18/52 (34.6%), and reporting of prediction intervals in 3/52 (5.8%) (Table **2**).

In the quantitative synthesis, all studies with $I^2 \geq 75\%$ employed random-effects models. The DerSimonian-Laird method was the most frequently reported (n=15), followed by unspecified random-effects models (n=25). For variance stabilization, 8 studies applied the Freeman-Tukey double arcsine transformation, 3 used logit transformation, and 2 employed generalized linear mixed models (GLMM). Heterogeneity was quantified using I^2 in all studies (52/52, 100%), with additional reporting of Cochran's Q statistic in 28/52 (53.8%) and τ^2 in 5/52 (9.6%). GLMM models were used in 4/52 (7.7%) studies, including the works by Wikswo (2025), Song (2025), Kip (2024), and Román-Gálvez (2021).

Risk of Bias Assessment Through AMSTAR-2

The AMSTAR-2 assessment revealed considerable variability in the methodological quality of the 53 included meta-analyses. The criterion with highest compliance was the declaration of conflicts of interest, present in all studies (53/53, 100%), consistently reporting absence of conflicts of interest. Comprehensive literature search was documented in 49 studies (92.4%), followed by detailed description of included studies in 46 studies (86.8%) and risk of bias assessment in primary studies in 44 studies (83%). Heterogeneity exploration was reported in 42 studies (79.24%) and publication bias assessment in 43 studies (81.13%).

The methodological aspects with lowest compliance corresponded mainly to procedural elements critical for the validity of systematic reviews. Protocol registration was documented in only 16 studies (30.18%), while duplicate selection was reported in only 15 studies (28.3%) and duplicate extraction in 29 studies (54.71%). The list of excluded studies with specific reasons was provided in only 19 studies (35.84%), and information on funding sources of primary studies was considered in only 3 studies (5.7%). These findings suggest systematic deficiencies in reporting fundamental procedural aspects, although most studies implemented appropriate strategies for statistical analysis and interpretation of results.

Table 2: Cascade of Methodological Practices in Studies with I^2 ≥75% (n=52)

Methodological criterion	n/52 (%)
Reported I^2 ≥75%	52/52 (100%)
Justified pooling despite high I^2	33/52 (63.5%)
Added sensitivity analyses	34/52 (65.4%)
Performed meta-regression	18/52 (34.6%)
Reported prediction intervals	3/52 (5.8%)

Partial compliance was frequent across multiple domains, particularly in justification of included study designs (39 studies with partial compliance), explicit PICO criteria, and consideration of risk of bias in the interpretation of results. This pattern suggests that while researchers recognize the importance of these methodological elements, their implementation and reporting do not always reach the optimal standards established by AMSTAR-2 guidelines. Nevertheless, all studies employed appropriate meta-analysis methods for prevalence data, demonstrating technical competence in the specific statistical analysis of proportions and implicit recognition that high heterogeneity is epidemiologically expected in this type of quantitative synthesis.

Illustrative bayesian analysis: empirical comparison

Three meta-analyses were randomly selected: Hu *et al*. [28] examining Hashimoto's thyroiditis prevalence (k=48, I^2=100%), Whitten *et al*. [35] on childhood domestic violence exposure (k=52, I^2=99.9%), and Armoon *et al*. [64] on suicide ideation among homeless youth (k=37, I^2=99%). These studies represent diverse health conditions with prevalence ranging from 7.6% to 38.0%, yet all exhibit extreme heterogeneity characteristic of global syntheses. Complete reproducible R code for each illustrative analysis is provided in Supplementary Code S1 (Hu *et al*.), S2 (Whitten *et al*.), and S3 (Armoon *et al*.). Table **3** presents comparative results.

Frequentist and Bayesian approaches yielded nearly identical prevalence estimates, with differences of 0.01 to 0.10 percentage points. Confidence and credible intervals showed complete overlap across all three meta-analyses. For Hu *et al*., Bayesian τ = 0.1458 (95% CrI: 0.117-0.178) compared to frequentist τ = 0.1424, indicating moderate between-study variability. For Whitten *et al*., Bayesian τ = 0.2062 (0.167-0.249) versus frequentist τ = 0.2015 reflected substantial heterogeneity consistent with 69-fold variation (1.0%-70.2%). For Armoon *et al*., Bayesian analysis provided τ = 0.1850 (0.149-0.224), information not reported in the original frequentist analysis. Prediction intervals were nearly identical across frameworks, accurately reflecting the expected variation in future studies.

The Bayesian credible intervals for τ provide direct quantification of uncertainty, which is unavailable from standard frequentist REML. Knowing that τ ranges from 0.149 to 0.224 in Armoon *et al*. informs whether meta-regression could plausibly reduce heterogeneity substantially or whether large residual variability will persist. The wide credible interval for τ in Whitten *et al*. (0.167-0.249) indicates considerable uncertainty about the magnitude of heterogeneity itself, suggesting that point estimates of τ^2 from REML may provide false precision. Prediction intervals from both approaches correctly show that new studies could yield prevalences spanning nearly the entire observed range, confirming that extreme heterogeneity is real population variability rather than a statistical artifact.

DISCUSSION

Main Findings

This umbrella review confirms that extreme heterogeneity (I^2≥90%) is nearly universal in global prevalence meta-analyses, present in 88.7% of studies, with 64.2% exceeding I^2>99%. This finding aligns with previous research documenting I^2 values routinely exceeding 95% in prevalence meta-analyses[3,70] . More importantly, it reveals substantial gaps between recommended statistical practices and actual implementation. While 96.2% performed basic subgroup analyses, only 34.6% conducted meta-regression and merely 5.8% reported prediction intervals, despite explicit recommendations for their routine use [11,71,72]. Additionally, 36.5% of studies with I^2≥75% provided no justification for proceeding with pooling. These findings indicate that the statistical challenge is not the magnitude of heterogeneity itself, but rather the incomplete implementation of methods to investigate, report, and interpret it appropriately [73,71].

The illustrative Bayesian analysis of three randomly selected meta-analyses demonstrates that alternative statistical frameworks are practically implementable and yield results concordant with frequentist

Tabla 3: Frequentist versus Bayesian Results in Three Randomly Selected Global Prevalence Meta-Analyses with Extreme Heterogeneity

Characteristic	Hu *et al*. (2022) Hashimoto's thyroiditis	Whitten *et al*. (2024) Childhood violence	Armoon *et al*. (2024) Suicide ideation
Sample characteristics			
Number of studies (k)	48	52	37
Total participants (N)	Not reported	784,612	96,557
Population	General	General	Homeless youth
Frequentist results (REML)			
Pooled prevalence (%)	7.61	17.16	38.00
95% CI (%)	5.60 - 9.91	13.14 - 21.60	33.0 - 43.0
95% Prediction interval (%)	0.00 - 28.40	0.02 - 55.02	Not reported
I^2 (%)	100.0	99.9	99.0
τ (between-study SD)	0.1424	0.2015	Not reported
Observed range (%)	Not reported	1.02 - 70.24	13.0 - 75.0
Bayesian results (Half-Normal prior)			
Posterior mean prevalence (%)	7.62	17.18	38.10
Posterior median prevalence (%)	7.61	17.17	38.05
95% Credible interval (%)	5.61 - 9.92	13.15 - 21.62	33.05 - 43.10
Posterior mean τ	0.1458	0.2062	0.1850
Posterior median τ	0.1445	0.2048	0.1835
95% Credible interval for τ	0.117 - 0.178	0.167 - 0.249	0.149 - 0.224
95% Predictive interval (%)	0.01 - 28.50	0.02 - 55.10	5.5 - 72.0
Concordance assessment			
Difference in point estimates	+0.01%	+0.02%	+0.10%
Overlap in uncertainty intervals	Complete	Complete	Complete
Interpretation	Excellent agreement	Excellent agreement	Excellent agreement

CI = confidence interval; CrI = credible interval; PI = prediction interval; τ = between-study standard deviation; Δ = difference between point estimates.

approaches while offering distinct interpretive advantages. The near-identical point estimates across frameworks (differences <0.1%) confirm that the choice of statistical approach does not alter substantive epidemiological conclusions about prevalence magnitudes. However, Bayesian methods provide additional uncertainty quantification for heterogeneity parameters through credible intervals, information not directly available from standard frequentist REML estimation [8,9].

Why Extreme Heterogeneity is Expected

The near-universal presence of extreme heterogeneity in global prevalence meta-analyses reflects fundamental epidemiological reality rather than methodological failure [7,74]. Prevalence varies genuinely across populations due to differences in genetic susceptibility, environmental exposures, healthcare access, diagnostic practices, and socioeconomic determinants [70,75]. A global meta-analysis synthesizing data from 50 countries across four decades should exhibit high I^2 because the underlying disease burden differs substantially. This heterogeneity is not statistical noise to be minimized but epidemiologically informative data about disease determinants [7]. The distinction from intervention meta-analyses is critical: in treatment studies, heterogeneity in effects may indicate effect modification requiring explanation before causal inference is valid [75]. In prevalence meta-analyses, heterogeneity is the expected pattern that communicates real variability in disease burden [3,74]. The statistical challenge is not to eliminate this heterogeneity but to characterize it appropriately [9]. From this perspective, I^2>90% is not a problem requiring correction but an expected feature requiring appropriate investigation and reporting [7,76]. The pooled estimate serves as a summary of central tendency across diverse contexts rather than a single true value to be estimated with maximum precision [76].

Statistical Solutions: Frequentist Approach

When I^2 consistently exceeds 95-99%, researchers face a critical statistical question: does the pooled estimate retain interpretability, or does extreme

between-study variability render it misleading? A pooled prevalence of 25% with I^2=99% might suggest a precise estimate to readers unfamiliar with heterogeneity interpretation, yet the underlying data could include individual studies ranging from 5% to 65%. The confidence interval around the pooled estimate (e.g., 22-28%) communicates only sampling uncertainty within studies, ignoring the massive variability between studies [71,77]. This fundamental limitation of confidence intervals in the presence of extreme heterogeneity necessitates alternative approaches to uncertainty quantification [78,79].

Prediction intervals provide a frequentist solution by incorporating both within-study and between-study variance components [71,79]. Mathematically, the prediction interval for a new study is calculated as: $PI = \hat{\theta} \pm t(df) \times \sqrt{(SE^2(\hat{\theta}) + \tau^2)}$, where τ^2 represents the between-study variance [79,80]. The critical distinction is that τ^2 accounts for heterogeneity, making the prediction interval substantially wider than the confidence interval when heterogeneity is high. For a meta-analysis with I^2=99%, where τ^2 dominates the variance structure, the prediction interval might span 5-65% even when the confidence interval is narrow (22-28%) [71,81]. This wider interval correctly reflects the uncertainty about what prevalence to expect in a new setting, which is the epidemiologically relevant question for public health planning [11]. The finding that only 5.8% of studies reported prediction intervals means that the vast majority failed to communicate the most practically useful measure of uncertainty. Standard meta-analysis software can calculate prediction intervals with minimal additional effort: in R, the metafor package implements this via the predict() function applied to a fitted random-effects model [82]. In Stata, the metan command with the rfdist option generates prediction intervals [83]. Despite this computational accessibility, prediction intervals remain severely underreported [72].

Meta-regression extends the basic random-effects model by including study-level covariates to explain heterogeneity. The model can be expressed as: $\theta_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + ... + \beta_k x_{ki} + u_i + \varepsilon_i$, where x variables are study-level covariates (e.g., mean age, GDP per capita, year of publication), β coefficients quantify the association between covariates and prevalence, u_i represents residual between-study heterogeneity, and ε_i is within-study sampling error. The proportion of heterogeneity explained can be assessed by comparing τ^2 from the meta-regression model to τ^2 from the intercept-only model: $R^2 = (\tau^2_{null} - \tau^2_{reg})/\tau^2_{null}$. For example, if meta-regression including mean population BMI, healthcare access index, and diagnostic criteria reduces τ^2 from 0.45 to 0.25, then R^2 = 0.44, indicating that these covariates explain 44% of between-study heterogeneity. This transforms heterogeneity from unexplained noise into interpretable epidemiological patterns. Meta-regression offers advantages over subgroup analysis by handling continuous moderators, examining multiple covariates simultaneously, and quantifying explained variance [84]. However, only 34.6% of studies employed this approach despite most having I^2>90%, representing a substantial missed opportunity.

The justification for pooling when I^2>95% requires careful statistical and epidemiological consideration [85] . Several criteria can guide this decision [84]. First, examine the prediction interval: if it falls within a clinically or epidemiologically meaningful range despite high I^2, pooling may be defensible [71]. For example, a pooled diabetes prevalence of 8% with a prediction interval of 4-14% might be useful for public health planning even though I^2=97%, as all values in this range have similar policy implications. Conversely, if the prediction interval spans 2-45%, the pooled estimate of 15% may be misleading. Second, assess whether meta-regression explains substantial heterogeneity: if R^2>0.50, indicating that covariates explain more than half the between-study variance, then presenting the meta-regression model may be more informative than the crude pooled estimate [84]. Third, consider whether there is theoretical justification for an average prevalence to be meaningful: for global prevalence of infectious diseases with major geographical and climate dependencies, a global average may be less useful than region-specific estimates [70,75]. Fourth, evaluate whether the heterogeneity is statistical or structural: if all studies measure the same construct with similar methods but show variation due to true population differences, pooling might be appropriate with careful interpretation [9]. However, if studies use fundamentally different diagnostic criteria or inclusion criteria, the heterogeneity may be structural and pooling across incomparable estimates inappropriate [1,70]. When these criteria suggest pooling may not be appropriate, alternatives include presenting stratified estimates by key moderators, providing the meta-regression model for readers to generate context-specific predictions, or offering a narrative synthesis describing the range and patterns of prevalences without attempting to calculate a pooled estimate [86].

Exploratory Bayesian Analysis: an Underutilized Methodological Alternative

Bayesian methods remain rarely employed in prevalence meta-analyses despite their theoretical advantages for managing extreme heterogeneity. To explore their practical feasibility and interpretability, we conducted illustrative Bayesian re-analysis of three

randomly selected meta-analyses. This exploratory analysis aims not to recommend routine Bayesian implementation, but rather to demonstrate: (1) how these methods perform when applied to real prevalence data with extreme heterogeneity; (2) what additional information they provide compared to standard frequentist approaches; and (3) under what specific circumstances they might offer substantive advantages beyond interpretive preferences.

The near-identical point estimates across frameworks (differences <0.1%) are expected and appropriate. With moderate-to-large study numbers (k=37-52) and weakly informative priors, Bayesian posteriors are dominated by data, naturally converging with frequentist estimates. This concordance is a strength, not a limitation: it demonstrates that in typical prevalence meta-analyses with adequate sample sizes, the choice of statistical framework does not materially alter epidemiological conclusions about prevalence magnitudes. However, the frameworks differ critically in what uncertainty they quantify and how results are interpreted [87].

Frequentist confidence intervals communicate sampling uncertainty under hypothetical repetition, while Bayesian credible intervals provide direct probability statements about parameter values given the observed data [87]. For policymakers, stating "there is 95% probability that the global average prevalence of childhood violence exposure lies between 13.2% and 21.6%" (Bayesian interpretation) aligns more naturally with decision-making needs than "95% of such intervals would contain the true value if the procedure were repeated infinitely" (frequentist interpretation). This interpretive clarity represents a practical communication advantage when translating findings for non-statistical audiences [88].

The Bayesian quantification of uncertainty in τ itself addresses a limitation of frequentist REML estimation, which provides point estimates of τ^2 but no direct confidence intervals without bootstrap resampling. The provision of credible intervals for heterogeneity (τ) represents additional information unavailable from standard frequentist approaches. Knowing that τ ranges from 0.149 to 0.224 with 95% probability in the Armoon *et al*. analysis informs whether subsequent meta-regression has realistic potential to explain substantial heterogeneity or whether large residual variability should be expected regardless of covariate adjustment [89]. The wide credible interval for τ in Whitten *et al*. (0.167-0.249) indicates considerable uncertainty about heterogeneity magnitude itself, suggesting that point estimates of τ^2 from REML may convey false precision.

Bayesian hierarchical models offer a fundamentally different framework for handling extreme heterogeneity that addresses several limitations of frequentist approaches. In the Bayesian framework, both the study-specific prevalences and the between-study heterogeneity are treated as random variables with probability distributions [87]. The critical Bayesian element is the specification of prior distributions for the parameters. For the overall mean μ, a weakly informative prior is often appropriate when prior knowledge is limited.For the heterogeneity parameter τ, common choices include half-Normal priors or half-Cauchy priors, though the choice can substantially influence posterior estimates when heterogeneity is extreme and the number of studies is small [90] .

Several R packages facilitate Bayesian meta-analysis with different levels of complexity and flexibility. The bayesmeta package provides a straightforward interface specifically designed for meta-analysis, automatically handling prior specification and producing posterior summaries, prediction intervals, and forest plots. The brms package offers greater flexibility for complex hierarchical models, using syntax similar to standard R regression. For users requiring maximum control, the rstan package allows direct specification of Stan models. Each package has trade-offs: bayesmeta prioritizes simplicity and is ideal for standard meta-analyses; brms balances flexibility and usability for complex hierarchical models; rstan provides maximum control but requires more statistical programming expertise. Markov Chain Monte Carlo sampling generates samples from the posterior distribution, with convergence diagnostics essential for reliable inference. A typical Bayesian meta-analysis with 50 studies requires 5-10 minutes on standard hardware, making exploration computationally feasible for most research contexts [87,91].

Bayesian methods offer substantive advantages in specific scenarios: small study numbers (k<10) where regularizing priors stabilize unstable REML estimates; complex hierarchical structures (three-level umbrella reviews) handled naturally through nested random effects; availability of prior information from similar meta-analyses that can be formalized; and decision-making contexts requiring specific probability calculations (e.g., P(prevalence > threshold) for resource allocation). In the absence of these conditions—as in our illustrative examples with adequate k and weak priors—Bayesian methods provide interpretive advantages and additional uncertainty metrics but do not fundamentally change epidemiological conclusions [87,91].

Practical considerations favor different approaches depending on context. Frequentist methods

(DerSimonian-Laird, REML with Hartung-Knapp adjustments) remain appropriate for straightforward meta-analyses with adequate numbers of studies (typically n≥10) and when prior information is limited or controversial [92] . They provide computational simplicity (seconds versus minutes) and are familiar to most researchers and reviewers [82]. The meta for package in R implements all standard frequentist methods. The Hartung-Knapp-Sidik-Jonkman method specifically modifies the variance of the pooled estimate to account for uncertainty in τ^2 estimation, providing more conservative confidence intervals when the number of studies is small or heterogeneity is large [93]. For prevalence meta-analyses with extreme heterogeneity, a pragmatic approach is to conduct frequentist analysis as the primary method for transparency and comparability with the existing literature, optionally supplemented by Bayesian analysis as a sensitivity analysis to examine the robustness of conclusions to prior specifications and to obtain full predictive distributions [94].

Practical recommendations

The critical gap identified in this umbrella review is not the absence of Bayesian methods—understandably complex and requiring specialized expertise—but rather the severe underutilization of standard tools: prediction intervals (5.8% reported, despite their computational simplicity) and meta-regression (34.0% employed, despite I^2 values exceeding 90% in most studies). These frequentist tools, readily implementable via metafor or Stata, remain the priority for improving statistical rigor in prevalence meta-analyses [71,73,78].

A prevalent misconception in the literature is that I^2 values exceeding 90% represent methodological failure, invalidating meta-analysis. This perspective is epidemiologically incorrect and statistically unjustified. Global prevalence data are necessary for public health planning, resource allocation, and research priority setting, regardless of heterogeneity. Rejecting global prevalence meta-analyses due to high I^2 is equivalent to denying that populations differ in disease burden—an untenable position. Extreme heterogeneity is not a reason to discard studies or to conclude that "prevalence meta-analyses cannot be conducted." It is a reason to apply appropriate statistical methods that recognize, quantify, and interpret this variability as valuable epidemiological information rather than noise to be eliminated [3,70]. The quality criterion is not the magnitude of I^2, but the completeness and appropriateness of its investigation [1]. A meta-analysis with I^2=99% that reports prediction intervals, conducts meta-regression, and justifies pooling decisions is methodologically superior to one with I^2=40% that ignores residual heterogeneity. Global prevalence estimates, despite extreme heterogeneity, provide essential reference points for contexts lacking local data, identify gaps in geographical knowledge, and facilitate temporal comparisons of disease burden [75]. The question is not "should we conduct global prevalence meta-analyses with high heterogeneity?" but "how do we conduct them appropriately?"

Three actionable priorities emerge clearly. First, implement prediction intervals routinely when I^2≥75%. The current reporting of 5.8% versus the required 100% represents the most severe gap identified. The software is accessible (predict() function in metafor, rfdist option in Stata), the additional time required is minimal (<5 minutes per analysis), but the impact on interpretability is substantial [1,72,73]. Prediction intervals communicate epidemiologically relevant uncertainty about expected prevalence in new contexts, while confidence intervals only communicate sampling uncertainty [78]. For public health planning, knowing that future prevalence will likely fall between 5-65% (prediction interval) is more useful than knowing that the global average is precisely estimated between 22-28% (confidence interval) [71,78]. This wide interval does not invalidate the meta-analysis—it correctly reflects the epidemiological reality that prevalence varies substantially across contexts. Policymakers need to be aware of this range to plan scenarios and allocate resources effectively. Omitting prediction intervals communicates false precision that can lead to inadequate planning [71].

Second, conduct meta-regression when feasible (typically when n≥10 studies per covariate) [84]. The current use of 34.6% versus the ~80% expected given typical study numbers in global meta-analyses represents a missed opportunity to transform unexplained heterogeneity into interpretable epidemiological patterns. Meta-regression enables the identification of which factors (population age, economic development, healthcare access, and diagnostic criteria) explain variability in prevalence [84]. When R^2>0.50, the meta-regression model may be more informative than the crude pooled estimate because it allows context-specific predictions. For example, instead of reporting "global diabetes prevalence is 15% (95% CI: 13-17%, I^2=98%)" with prediction interval 3-48%, meta-regression can communicate "prevalence increases 0.8% per 5-point increase in mean population BMI and 1.2% per 10-year increase in median population age, explaining 58% of between-study heterogeneity." This information is directly actionable for specific contexts. Potential covariates should be pre-specified based on epidemiological domain knowledge, and the analysis

must report both regression coefficients and proportion of heterogeneity explained (R^2) with its epidemiological interpretation [84]. The main limitations are requiring adequate numbers of studies and risk of ecological bias when using aggregated study-level data rather than individual participant data,but these risks do not justify omitting the analysis when feasible.

Third, explicitly justify pooling decisions when I^2>95% (1-7). The current 36.5% that omits justification means that more than one-third of studies with extreme heterogeneity present pooled estimates without evaluating whether they retain interpretability. This omission is indefensible. The justification should address four criteria: (a) whether the prediction interval falls within a clinically or epidemiologically meaningful range—a prediction interval of 4-14% for diabetes has similar policy implications across the range, while 2-65% suggests populations are so heterogeneous that a global average has limited utility [78] ; (b) whether meta-regression explains substantial heterogeneity (R^2>0.50), indicating that variability is understandable and predictable rather than random; [84] (c) whether there is theoretical justification for an average to be epidemiologically meaningful—for diseases with strong geographical or climate determinants (e.g., malaria), global averages have less utility than region-specific estimates, while for chronic conditions with socioeconomic determinants (e.g., diabetes), global averages may serve as useful references [75]; and (d) whether heterogeneity is statistical (true population differences with comparable measurements) versus structural (fundamentally different disease definitions or study populations making estimates incomparable). When these criteria suggest that a global pooled estimate retains limited interpretability, alternatives include presenting stratified estimates by region or level of economic development, providing the complete meta-regression model for users to generate predictions specific to their context, or offering narrative synthesis describing patterns and ranges of prevalence without calculating a single global average. These alternatives do not represent meta-analysis failure but appropriate recognition that some questions require more nuanced answers than a single number [70].

The research community needs to abandon the position that extreme heterogeneity invalidates prevalence meta-analyses [3,70]. Peer reviewers who reject manuscripts citing only high I^2 without evaluating completeness of heterogeneity investigation apply epidemiologically unjustified standards [1]. The appropriate standard is: did the authors exhaustively investigate sources of heterogeneity using available methods? Did they report prediction intervals? Did they conduct meta-regression when feasible? Did they explicitly justify pooling decisions? Did they interpret heterogeneity as epidemiological information about disease determinants? A manuscript that meets these criteria with I^2=99% is methodologically sound and should be published. One that omits them with I^2=60% is methodologically deficient regardless of moderate heterogeneity.

From an editorial perspective, journals could implement specific statistical requirements for prevalence meta-analyses with high heterogeneity. Several journals (e.g., BMJ, Lancet) have adopted mandatory reporting guidelines for systematic reviews (PRISMA) [95]. Extending these to include heterogeneity-specific requirements would formalize existing methodological recommendations. Mandatory elements could include: (1) prediction intervals when $I^2 \geq 75\%$, with explicit interpretation in results and discussion [72,78]; (2) meta-regression when $n \geq 10$ studies, or explicit statement of why it was not feasible [84]; (3) explicit justification of pooling decisions when I^2>95% addressing the four criteria described [3]; (4) reporting of multiple heterogeneity metrics (I^2, τ^2, prediction intervals) rather than I^2 alone [77]; (5) sensitivity analyses examining robustness to statistical model choice [92]; and (6) when Bayesian methods are employed, complete specification of priors with justification, sensitivity analyses with alternative priors, and full reporting of convergence diagnostics [14]. These requirements do not represent substantial additional burden given that modern software implements all these methods, but would transform the quality and utility of published prevalence meta-analyses. Bayesian methods represent a valuable addition to the methodological toolkit for specific scenarios or sensitivity analyses, but should not distract from the fundamental need to implement already-recommended frequentist practices that remain widely neglected.

LIMITATIONS

This study has several limitations. Data extraction was limited to information reported in published articles, which may have resulted in some analyses not being reported. The distinction between unreported and unperformed analyses matters for interpretation, but from a scientific communication perspective they have equivalent impact on readers. Our search strategy specifically targeted global meta-analyses using explicit terms, which may have excluded some studies of global scope that did not use these terms. However, the consistency of extreme heterogeneity across diverse disease areas and populations suggests this is a universal characteristic rather than an artifact of our sampling strategy. We restricted inclusion to the top 100 Q1 journals in Public Health, which may limit

generalizability to lower-impact publications or other disciplines. This restriction was intentional to evaluate best practices where adherence to methodological recommendations should be optimal; identification of gaps even in elite publications suggests deficiencies elsewhere may be more pronounced.

We documented presence or absence of statistical methods without evaluating their appropriateness or quality when performed. Future research examining not just the frequency but also the quality of heterogeneity investigation would be valuable. We did not have access to study protocols or full supplementary materials in all cases, precluding assessment of whether heterogeneity investigations were pre-specified or conducted post-hoc. We did not evaluate temporal changes in prevalence or diagnostic criteria as sources of heterogeneity. Meta-analyses pooling studies spanning decades may combine genuinely different disease constructs if diagnostic criteria evolved. Similarly, combining data across countries with different healthcare systems may pool non-comparable estimates. These sources of structural heterogeneity require separate methodological investigation beyond statistical approaches. Finally, our evaluation focused on statistical methods for managing heterogeneity and did not address other important aspects of prevalence meta-analyses such as risk of bias assessment or publication boas, which may also contribute to observed heterogeneity.

CONCLUSIONS

Extreme heterogeneity is nearly universal in global prevalence meta-analyses published in top-tier journals, but the statistical methods available for managing such heterogeneity remain substantially underutilized. The critical gaps are specific and addressable: prediction intervals are reported in only 5.8% of studies despite being computationally straightforward; meta-regression is employed in only 34.6% of studies; and explicit justification for pooling is absent in 36.5% of high-heterogeneity studies. These standard frequentist tools—readily implementable and widely recommended—represent the primary opportunity for improving statistical rigor in prevalence meta-analyses. An exploratory Bayesian analysis of three meta-analyses demonstrated concordance with frequentist estimates while providing additional uncertainty quantification for heterogeneity parameters, illustrating that alternative frameworks are feasible and interpretable when applied to prevalence data with extreme heterogeneity. However, their value lies primarily in specific scenarios (small study numbers, hierarchical structures, or prior information availability) rather than routine application.

The evaluation criterion for prevalence meta-analyses should focus on the thoroughness of heterogeneity investigation and reporting rather than the magnitude of heterogeneity itself. Extreme heterogeneity in prevalence studies is expected and statistically manageable, but its acceptance should be conditional on rigorous investigation that transforms heterogeneity from a statistical problem into interpretable epidemiological information about disease determinants and geographic variation. Moving forward, researchers conducting global prevalence meta-analyses should prioritize implementing prediction intervals and meta-regression—accessible through meta for or Stata—as these tools remain severely underutilized despite their computational simplicity and interpretive value. Bayesian methods (bayesmeta, brms, rstan) represent valuable additions for sensitivity analyses or specific scenarios where they offer substantive advantages, but should not distract from the fundamental need to implement already-recommended practices that remain widely neglected. Peer reviewers and editors should evaluate the completeness of heterogeneity investigation rather than rejecting studies based solely on high I^2 values. By improving statistical rigor through practical application of available tools—particularly the underutilized standard methods—the field can maximize the utility of prevalence meta-analyses for informing public health policy and resource allocation decisions in an increasingly data-driven global health landscape.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Since this manuscript is a secondary data-based study, it was not required.

- Informed consent: Since this is a secondary data analysis, informed consent was not required.
- Clinical trial number: not applicable.

CONSENT FOR PUBLICATION

Not applicable

DATA AVAILABILITY

Data are available upon request to the corresponding author.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

FUNDING

This study was funded by the Vicerectorado de Investigación of the Universidad Nacional Toribio Rodríguez de Mendoza de Amazonas.

ACKNOWLEDGEMENTS

Special thanks to the members of Universidad Nacional Toribio Rodríguez de Mendoza de Amazonas (UNTRM), Amazonas, Peru, for their support and contributions throughout the completion of this research.

AUTHORS' CONTRIBUTION

Víctor Juan Vera-Ponce: Conceptualization, Investigation, Software, Formal analysis, Methodology, Resources, Writing - Original Draft, Writing - Review & Editing

Jhosmer Ballena-Caicedo: Conceptualization, Data Curation, Investigation, Formal analysis, Methodology, Software, Writing - Original Draft, Writing - Review & Editing

Lupita Ana Maria Valladolid-Sandoval: Methodology, Investigation, Writing - Original Draft, Writing - Review & Editing

Fiorella E. Zuzunaga-Montoya: Investigation, Data Curation, Methodology, Funding acquisition, Supervision, Writing - Original Draft, Writing - Review & Editing

SUPPLEMENTAL MATERIALS

The supplemental materials can be downloaded from the journal website along with the article.

REFERENCES

[1] Borges Migliavaca C, Stein C, Colpani V, Barker TH, Munn Z, Falavigna M, *et al*. How are systematic reviews of prevalence conducted? A methodological study. BMC Med Res Methodol 2020; 20(1): 96. https://doi.org/10.1186/s12874-020-00975-3

[2] Barker TH, Migliavaca CB, Stein C, Colpani V, Falavigna M, Aromataris E, *et al*. Conducting proportional meta-analysis in different types of systematic reviews: a guide for synthesisers of evidence. BMC Med Res Methodol 2021; 21(1): 189. https://doi.org/10.1186/s12874-021-01381-z

[3] Migliavaca CB, Stein C, Colpani V, Barker TH, Ziegelmann PK, Munn Z, *et al*. Meta-analysis of prevalence: I2 statistic and how to deal with heterogeneity. Res Synth Methods 2022; 13(3): 363-7. https://doi.org/10.1002/jrsm.1547

[4] Munn Z, Stern C, Aromataris E, Lockwood C, Jordan Z. What kind of systematic review should I conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences. BMC Med Res Methodol 2018; 18(1): 5. https://doi.org/10.1186/s12874-017-0468-4

[5] Schwarzer G, Chemaitelly H, Abu-Raddad LJ, Rücker G. Seriously misleading results using inverse of Freeman-Tukey double arcsine transformation in meta-analysis of single proportions. Res Synth Methods 2019; 10(3): 476-83. https://doi.org/10.1002/jrsm.1348

[6] Ioannidis JPA. Interpretation of tests of heterogeneity and bias in meta-analysis. J Eval Clin Pract 2008; 14(5): 951-7. https://doi.org/10.1111/j.1365-2753.2008.00986.x

[7] Higgins JPT. Commentary: Heterogeneity in meta-analysis should be expected and appropriately quantified. Int J Epidemiol 2008; 37(5): 1158-60. https://doi.org/10.1093/ije/dyn204

[8] Rücker G, Schwarzer G, Carpenter JR, Schumacher M. Undue reliance on I(2) in assessing heterogeneity may mislead. BMC Med Res Methodol 2008; 8: 79. https://doi.org/10.1186/1471-2288-8-79

[9] Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003; 327(7414): 557-60. https://doi.org/10.1136/bmj.327.7414.557

[10] Barendregt JJ, Doi SA, Lee YY, Norman RE, Vos T. Meta-analysis of prevalence. J Epidemiol Community Health 2013; 67(11): 974-8. https://doi.org/10.1136/jech-2013-203104

[11] IntHout J, Ioannidis JPA, Rovers MM, Goeman JJ. Plea for routinely presenting prediction intervals in meta-analysis. BMJ Open 2016; 6(7): e010247. https://doi.org/10.1136/bmjopen-2015-010247

[12] Gates M, Gates A, Pieper D, Fernandes RM, Tricco AC, Moher D, *et al*. Reporting guideline for overviews of reviews of healthcare interventions: development of the PRIOR statement. BMJ 2022; 378: e070849. https://doi.org/10.1136/bmj-2022-070849

[13] Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, *et al*. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017; 358: j4008. https://doi.org/10.1136/bmj.j4008

[14] Röver C. Bayesian random-effects meta-analysis using the bayesmeta R package. J Stat Softw 2020; 93(6): 1-51. https://doi.org/10.18637/jss.v093.i06

[15] Polson NG, Scott JG. On the half-Cauchy prior for a global scale parameter. Bayesian Anal 2012; 7(4): 887-902. https://doi.org/10.1214/12-BA730

[16] Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, *et al*. Stan: A probabilistic programming language. J Stat Softw 2017; 76(1): 1-32. https://doi.org/10.18637/jss.v076.i01

[17] Albadrani MS, Tobaiqi MA, Muaddi MA, Eltahir HM, Abdoh ES, Aljohani AM, *et al*. A global prevalence of electronic nicotine delivery systems (ENDS) use among students: a systematic review and meta-analysis of 4,189,145 subjects. BMC Public Health 2024; 24(1): 3311. https://doi.org/10.1186/s12889-024-20858-2

[18] Wikswo ME, Kambhampati AK, Mattison CP, Chhabra P, Olojo O, Rana T, *et al*. A systematic review and meta-analysis of the global prevalence of human enteric adenovirus infections. J Infect Public Health 2025; 18(7): 102800. https://doi.org/10.1016/j.jiph.2025.102800

[19] Kinner SA, Snow K, Wirtz AL, Altice FL, Beyrer C, Dolan K. Age-Specific Global Prevalence of Hepatitis B, Hepatitis C, HIV, and Tuberculosis Among Incarcerated People: A Systematic Review. J Adolesc Health 2018; 62(3S): S18-26. https://doi.org/10.1016/j.jadohealth.2017.09.030

[20] Wei J, Zhu X, Liu J, Gao Y, Liu X, Wang K, *et al*. Estimating global prevalence of mild cognitive impairment and dementia in elderly with overweight, obesity, and central obesity: A systematic review and meta-analysis. Obes Rev 2025; 26(5): e13882. https://doi.org/10.1111/obr.13882

[21] Al Wachami N, Guennouni M, Iderdar Y, Boumendil K, Arraji M, Mourajid Y, *et al*. Estimating the global prevalence of chronic obstructive pulmonary disease (COPD): a systematic

review and meta-analysis. BMC Public Health 2024; 24(1): 297. https://doi.org/10.1186/s12889-024-17686-9

[22] Adeloye D, Chua S, Lee C, Basquill C, Papana A, Theodoratou E, *et al*. Global and regional estimates of COPD prevalence: Systematic review and meta-analysis. J Glob Health 2015; 5(2): 020415. https://doi.org/10.7189/jogh.05.020415

[23] Xie X, Pei J, Zhang L, Wu Y. Global birth prevalence of major congenital anomalies: a systematic review and meta-analysis. BMC Public Health 2025; 25(1): 449. https://doi.org/10.1186/s12889-025-21642-6

[24] Song P, Xu Y, Zha M, Zhang Y, Rudan I. Global epidemiology of retinal vein occlusion: a systematic review and meta-analysis of prevalence, incidence, and risk factors. J Glob Health 2025; 9(1): 010427. https://doi.org/10.7189/jogh.09.010427

[25] Fall A, Kenmoe S, Ebogo-Belobo JT, Mbaga DS, Bowo-Ngandji A, Foe-Essomba JR, *et al*. Global prevalence and case fatality rate of Enterovirus D68 infections, a systematic review and meta-analysis. PLoS Negl Trop Dis 2022; 16(2): e0010073. https://doi.org/10.1371/journal.pntd.0010073

[26] Sulaiman SK, Isma'il Tsiga-Ahmed F, Musa MS, Makama BT, Sulaiman AK, Abdulaziz TB. Global prevalence and correlates of mpox vaccine acceptance and uptake: a systematic review and meta-analysis. Commun Med (Lond) 2024; 4: 136. https://doi.org/10.1038/s43856-024-00564-1

[27] Suleiman AS, Bhattacharya P, Islam MA. Global prevalence and dynamics of mecA and mecC genes in MRSA: Meta-meta-analysis, meta-regression, and temporal investigation. J Infect Public Health 2025; 18(7): 102802. https://doi.org/10.1016/j.jiph.2025.102802

[28] Hu X, Chen Y, Shen Y, Tian R, Sheng Y, Que H. Global prevalence and epidemiological trends of Hashimoto's thyroiditis in adults: A systematic review and meta-analysis. Front Public Health 2022; 10: 1020709. https://doi.org/10.3389/fpubh.2022.1020709

[29] White SJ, Sin J, Sweeney A, Salisbury T, Wahlich C, Montesinos Guevara CM, *et al*. Global Prevalence and Mental Health Outcomes of Intimate Partner Violence Among Women: A Systematic Review and Meta-Analysis. Trauma, Violence, & Abuse 2024; 25(1): 494-511. https://doi.org/10.1177/15248380231155529

[30] Fajar JK, Sallam M, Soegiarto G, Sugiri YJ, Anshory M, Wulandari L, *et al*. Global Prevalence and Potential Influencing Factors of COVID-19 Vaccination Hesitancy: A Meta-Analysis. Vaccines (Basel) 2022; 10(8): 1356. https://doi.org/10.3390/vaccines10081356

[31] Delie AM, Bogale EK, Anagaw TF, Tiruneh MG, Fenta ET, Adal O, *et al*. Global prevalence and predictors of scabies among prisoners: systematic review and meta-analysis. BMC Public Health 2024; 24(1): 1894. https://doi.org/10.1186/s12889-024-19401-0

[32] Syed S, Ashwick R, Schlosser M, Jones R, Rowe S, Billings J. Global prevalence and risk factors for mental health problems in police personnel: a systematic review and meta-analysis. Occup Environ Med 2020; 77(11): 737-47. https://doi.org/10.1136/oemed-2020-106498

[33] Alshehri AA, Irekeola AA. Global prevalence of alkhumra hemorrhagic fever virus infection: The first meta-analysis and systematic review. J Infect Public Health 2024; 17(6): 986-93. https://doi.org/10.1016/j.jiph.2024.04.001

[34] Holland C, Sepidarkish M, Deslyper G, Abdollahi A, Valizadeh S, Mollalo A, *et al*. Global prevalence of Ascaris infection in humans (2010-2021): a systematic review and meta-analysis. Infect Dis Poverty 2022; 11(1): 113. https://doi.org/10.1186/s40249-022-01038-z

[35] Whitten T, Tzoumakis S, Green MJ, Dean K. Global Prevalence of Childhood Exposure to Physical Violence within Domestic and Family Relationships in the General Population: A Systematic Review and Proportional Meta-Analysis. Trauma Violence Abuse 2024; 25(2): 1411-30. https://doi.org/10.1177/15248380231179133

[36] Ukwishaka J, Ndayishimiye Y, Destine E, Danwang C, Kirakoya-Samadoulougou F. Global prevalence of coronavirus disease 2019 reinfection: a systematic review and meta-analysis. BMC Public Health 2023; 23(1): 778. https://doi.org/10.1186/s12889-023-15626-7

[37] Salari N, Kanjoori AH, Hosseinian-Far A, Hasheminezhad R, Mansouri K, Mohammadi M. Global prevalence of drug-resistant tuberculosis: a systematic review and meta-analysis. Infectious Diseases of Poverty 2023; 12(1): 57. https://doi.org/10.1186/s40249-023-01107-x

[38] Zhang P, Hao C, Di X, Chuizhao X, Jinsong L, Guisen Z, *et al*. Global prevalence of norovirus gastroenteritis after emergence of the GII.4 Sydney 2012 variant: a systematic review and meta-analysis. Front Public Health 2024; 12: 1373322. https://doi.org/10.3389/fpubh.2024.1373322

[39] Shafiee A, Nakhaee Z, Bahri RA, Amini MJ, Salehi A, Jafarabady K, *et al*. Global prevalence of obesity and overweight among medical students: a systematic review and meta-analysis. BMC Public Health 2024; 24(1): 1673. https://doi.org/10.1186/s12889-024-19184-4

[40] Li Y-J, Xie X-N, Lei X, Li Y-M, Lei X. Global prevalence of obesity, overweight and underweight in children, adolescents and adults with autism spectrum disorder, attention-deficit hyperactivity disorder: A systematic review and meta-analysis. Obes Rev 2020; 21(12): e13123. https://doi.org/10.1111/obr.13123

[41] Tolera ST, Gobena T, Assefa N, Geremew A. Global prevalence of occupational injuries among sanitation workers: a systematic review and meta-analysis. Front Public Health 2024; 12: 1425904. https://doi.org/10.3389/fpubh.2024.1425904

[42] Taher MK, Salzman T, Banal A, Morissette K, Domingo FR, Cheung AM, *et al*. Global prevalence of post-COVID-19 condition: a systematic review and meta-analysis of prospective evidence. Health Promot Chronic Dis Prev Can 2025; 45(3): 112-38. https://doi.org/10.24095/hpcdp.45.3.02

[43] Sheng Y, Jin L-Y, Li N, Zhang Y, Shi Y-J. Global prevalence of psittacosis in outbreaks: a systematic review and meta-analysis. BMC Public Health 2025; 25(1): 2010. https://doi.org/10.1186/s12889-025-21612-y

[44] Engku Abd Rahman ENS, Irekeola AA, Elmi AH, Chua WC, Chan YY. Global prevalence patterns and distribution of Vibrio cholerae: A systematic review and meta-analysis of 176,740 samples. J Infect Public Health 2024; 17(11): 102558. https://doi.org/10.1016/j.jiph.2024.102558

[45] Niu X, Zhu L, Xu Y, Zhang M, Hao Y, Ma L, *et al*. Global prevalence, incidence, and outcomes of alcohol related liver diseases: a systematic review and meta-analysis. BMC Public Health 2023; 23(1): 859. https://doi.org/10.1186/s12889-023-15749-x

[46] Fang Y, Liu F, Zhang X, Chen L, Liu Y, Yang L, *et al*. Mapping global prevalence of menopausal symptoms among middle-aged women: a systematic review and meta-analysis. BMC Public Health 2024; 24(1): 1767. https://doi.org/10.1186/s12889-024-19280-5

[47] Renzi E, Baccolini V, Migliara G, Bellotta C, Ceparano M, Donia P, *et al*. Mapping the Prevalence of COVID-19 Vaccine Acceptance at the Global and Regional Level: A Systematic Review and Meta-Analysis. Vaccines (Basel) 2022; 10(9): 1488. https://doi.org/10.3390/vaccines10091488

[48] Jafari A, Rajabi A, Gholian-Aval M, Peyman N, Mahdizadeh M, Tehrani H. National, regional, and global prevalence of cigarette smoking among women/females in the general population: a systematic review and meta-analysis. Environ Health Prev Med 2021; 26(1): 5. https://doi.org/10.1186/s12199-020-00924-y

[49] Bouqoufi A, Laila L, Boujraf S, Hadj FAE, Razine R, Abouqal R, *et al*. Prevalence and associated factors of self-medication in worldwide pregnant women: systematic review and meta-analysis. BMC Public Health 2024; 24(1): 308. https://doi.org/10.1186/s12889-023-17195-1

[50] Dagnaw M, Muche AA, Geremew BM, Gezie LD. Prevalence and burden of HBV-HIV co-morbidity: a global systematic review and meta-analysis. Front Public Health 2025; 13: 1565621. https://doi.org/10.3389/fpubh.2025.1565621

[51] Muche AA, Olayemi OO, Gete YK. Prevalence and determinants of gestational diabetes mellitus in Africa based on the updated international diagnostic criteria: a systematic review and meta-analysis. Arch Public Health 2019; 77: 36. https://doi.org/10.1186/s13690-019-0362-0

[52] Dillard LK, Arunda MO, Lopez-Perez L, Martinez RX, Jiménez L, Chadha S. Prevalence and global estimates of unsafe listening practices in adolescents and young adults: a systematic review and meta-analysis. BMJ Glob Health 2022; 7(11): e010501. https://doi.org/10.1136/bmjgh-2022-010501

[53] Pan Y, Lin X, Liu J, Zhang S, Zeng X, Chen F, *et al*. Prevalence of Childhood Sexual Abuse Among Women Using the Childhood Trauma Questionnaire: A Worldwide Meta-Analysis. Trauma Violence Abuse 2021; 22(5): 1181-91. https://doi.org/10.1177/1524838020912867

[54] Escobar N, Plugge E. Prevalence of human papillomavirus infection, cervical intraepithelial neoplasia and cervical cancer in imprisoned women worldwide: a systematic review and meta-analysis. J Epidemiol Community Health 2020; 74(1): 95-102. https://doi.org/10.1136/jech-2019-212557

[55] Kip A, Valencia S, Glunz E, Lowe SR, Tam K-P, Morina N. Prevalence of mental disorders in adult populations from the Global South following exposure to natural hazards: a meta-analysis. Epidemiol Psychiatr Sci 2024; 33: e68. https://doi.org/10.1017/S2045796024000672

[56] Alemayehu TT, Wassie YA, Bekalu AF, Tegegne AA, Ayenew W, Tadesse G, *et al*. Prevalence of potential drug-drug interactions and associated factors among elderly patients in Ethiopia: a systematic review and meta-analysis. Glob Health Res Policy 2024; 9(1): 46. https://doi.org/10.1186/s41256-024-00386-7

[57] Zheng Y, Ye K, Ying M, He Y, Yu Q, Lan L, *et al*. Syphilis epidemic among men who have sex with men: A global systematic review and meta-analysis of prevalence, incidence, and associated factors. J Glob Health 2024; 14: 04004. https://doi.org/10.7189/jogh.14.04004

[58] Faustino R, Faria M, Teixeira M, Palavra F, Sargento P, do Céu Costa M. Systematic review and meta-analysis of the prevalence of coronavirus: One health approach for a global strategy. One Health 2022; 14: 100383. https://doi.org/10.1016/j.onehlt.2022.100383

[59] Ahmadi Gharaei H, Fararouei M, Mirzazadeh A, Sharifnia G, Rohani-Rasaf M, Bastam D, *et al*. The global and regional prevalence of hepatitis C and B co-infections among prisoners living with HIV: a systematic review and meta-analysis. Infect Dis Poverty 2021; 10(1): 93. https://doi.org/10.1186/s40249-021-00876-7

[60] Song P, Wu J, Cao J, Sun W, Li X, Zhou T, *et al*. The global and regional prevalence of restless legs syndrome among adults: A systematic review and modelling analysis. J Glob Health 2024; 14: 04113. https://doi.org/10.7189/jogh.14.04113

[61] Shariati A, Dadashi M, Chegini Z, van Belkum A, Mirzaii M, Khoramrooz SS, *et al*. The global prevalence of Daptomycin, Tigecycline, Quinupristin/Dalfopristin, and Linezolid-resistant Staphylococcus aureus and coagulase-negative staphylococci strains: a systematic review and meta-analysis. Antimicrob Resist Infect Control 2020; 9(1): 56. https://doi.org/10.1186/s13756-020-00714-9

[62] Hajikhani B, Goudarzi M, Kakavandi S, Amini S, Zamani S, van Belkum A, *et al*. The global prevalence of fusidic acid resistance in clinical isolates of Staphylococcus aureus: a systematic review and meta-analysis. Antimicrob Resist Infect Control 2021; 10(1): 75. https://doi.org/10.1186/s13756-021-00943-6

[63] Hasanpour AH, Sepidarkish M, Mollalo A, Ardekani A, Almukhtar M, Mechaal A, *et al*. The global prevalence of methicillin-resistant Staphylococcus aureus colonization in residents of elderly care centers: a systematic review and meta-analysis. Antimicrob Resist Infect Control 2023; 12(1): 4. https://doi.org/10.1186/s13756-023-01210-6

[64] Armoon B, Mohammadi R, Griffiths MD. The Global Prevalence of Non-suicidal Self-injury, Suicide Behaviors, and Associated Risk Factors Among Runaway and Homeless Youth: A Meta-analysis. Community Ment Health J 2024; 60(5): 919-44. https://doi.org/10.1007/s10597-024-01245-y

[65] Salari N, Heidarian P, Abdolmaleki A, Salim K, Hashemian SH, Daneshkhah A, *et al*. The global prevalence of single-child families with emphasis on influential factors: a comprehensive systematic review and meta-analysis. Popul Health Metr 2025; 23(1): 25. https://doi.org/10.1186/s12963-025-00393-x

[66] Behniafar H, Sepidarkish M, Tadi MJ, Valizadeh S, Gholamrezaei M, Hamidi F, *et al*. The global prevalence of Trichuris trichiura infection in humans (2010-2023): A systematic review and meta-analysis. J Infect Public Health 2024; 17(5): 800-9. https://doi.org/10.1016/j.jiph.2024.03.005

[67] Song P, Zha M, Yang Q, Zhang Y, Li X, Rudan I, *et al*. The prevalence of adult attention-deficit hyperactivity disorder: A global systematic review and meta-analysis. J Glob Health 2021; 11: 04009. https://doi.org/10.7189/jogh.11.04009

[68] Dai X, Chu X, Qi G, Yuan P, Zhou Y, Xiang H, *et al*. Worldwide Perinatal Intimate Partner Violence Prevalence and Risk Factors for Post-traumatic Stress Disorder in Women: A Systematic Review and Meta-analysis. Trauma Violence Abuse 2024; 25(3): 2363-76. https://doi.org/10.1177/15248380231211950

[69] Román-Gálvez RM, Martín-Peláez S, Fernández-Félix BM, Zamora J, Khan KS, Bueno-Cavanillas A. Worldwide Prevalence of Intimate Partner Violence in Pregnancy. A Systematic Review and Meta-Analysis. Front Public Health 2021; 9: 738459. https://doi.org/10.3389/fpubh.2021.738459

[70] Munn Z, Moola S, Lisy K, Riitano D, Tufanaru C. Methodological guidance for systematic reviews of observational epidemiological studies reporting prevalence and cumulative incidence data. Int J Evid Based Healthc 2015; 13(3): 147-53. https://doi.org/10.1097/XEB.0000000000000054

[71] Riley RD, Higgins JPT, Deeks JJ. Interpretation of random effects meta-analyses. BMJ 2011; 342: d549. https://doi.org/10.1136/bmj.d549

[72] Partlett C, Riley RD. Random effects meta-analysis: Coverage performance of 95% confidence and prediction intervals following REML estimation. Stat Med 2017; 36(2): 301-17. https://doi.org/10.1002/sim.7140

[73] Nagashima K, Noma H, Furukawa TA. Prediction intervals for random-effects meta-analysis: A confidence distribution approach. Stat Methods Med Res 2019; 28(6): 1689-702. https://doi.org/10.1177/0962280218773520

[74] Cordero CP, Dans AL. Key concepts in clinical epidemiology: detecting and dealing with heterogeneity in meta-analyses. J Clin Epidemiol 2021; 130: 149-51. https://doi.org/10.1016/j.jclinepi.2020.09.045

[75] Hoffmann F, Eggers D, Pieper D, Zeeb H, Allers K. An observational study found large methodological heterogeneity in systematic reviews addressing prevalence and cumulative incidence. J Clin Epidemiol 2020; 119: 92-9. https://doi.org/10.1016/j.jclinepi.2019.12.003

[76] Ades AE, Lu G, Higgins JPT. The interpretation of random-effects meta-analysis in decision models. Med Decis Making 2005; 25(6): 646-54. https://doi.org/10.1177/0272989X05282643

[77] Deeks JJ, Higgins JPT, Altman DG, McKenzie JE, Veroniki AA, editors. Chapter 10: Analysing data and undertaking meta-analyses. In: Cochrane Handbook for Systematic Reviews of Interventions 2024.

[78] IntHout J, Ioannidis JPA, Rovers MM, Goeman JJ. Plea for routinely presenting prediction intervals in meta-analysis 2016 [cited 9 November 2025]. https://doi.org/10.1136/bmjopen-2015-010247

[79] Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002; 21(11): 1539-58. https://doi.org/10.1002/sim.1186

[80] Spineli LM, Pandis N. Prediction interval in random-effects meta-analysis. Am J Orthod Dentofacial Orthop 2020; 157(4): 586-8. https://doi.org/10.1016/j.ajodo.2019.12.011

[81] Borenstein M. How to understand and report heterogeneity in a meta-analysis: The difference between I-squared and prediction intervals. Integr Med Res 2023; 12(4): 101014. https://doi.org/10.1016/j.imr.2023.101014

[82] Viechtbauer W. Conducting meta-analyses in R with the metafor package. J Stat Softw 2010; 36(3): 1-48. https://doi.org/10.18637/jss.v036.i03

[83] Harris RJ, Bradburn MJ, Deeks JJ, Harbord RM, Altman DG, Sterne JAC. Metan: fixed- and random-effects meta-analysis. Stata J 2008; 8(1): 3-28. https://doi.org/10.1177/1536867X0800800102

[84] Thompson SG, Higgins JPT. How should meta-regression analyses be undertaken and interpreted? Stat Med 2002; 21(11): 1559-73. https://doi.org/10.1002/sim.1187

[85] Patsopoulos NA, Evangelou E, Ioannidis JPA. Sensitivity of between-study heterogeneity in meta-analysis: proposed metrics and empirical evaluation. Int J Epidemiol 2008; 37(5): 1148-57. https://doi.org/10.1093/ije/dyn065

[86] Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. BMJ 2010; 340: c221. https://doi.org/10.1136/bmj.c221

[87] Sutton AJ, Abrams KR. Bayesian methods in meta-analysis and evidence synthesis. Stat Methods Med Res 2001; 10(4): 277-303. https://doi.org/10.1177/096228020101000404

[88] Spiegelhalter DJ, Abrams KR, Myles JP. Bayesian Approaches to Clinical Trials and Health-Care Evaluation. John Wiley & Sons; 2004. https://doi.org/10.1002/0470092602

[89] Veroniki AA, Jackson D, Viechtbauer W, Bender R, Bowden J, Knapp G, *et al*. Methods to estimate the between-study variance and its uncertainty in meta-analysis. Res Synth Methods 2016; 7(1): 55-79. https://doi.org/10.1002/jrsm.1164

[90] Lambert PC, Sutton AJ, Burton PR, Abrams KR, Jones DR. How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. Stat Med 2005; 24(15): 2401-28. https://doi.org/10.1002/sim.2112

[91] Bürkner PC. brms: An R package for Bayesian multilevel models using Stan. J Stat Softw 2017; 80(1): 1-28. https://doi.org/10.18637/jss.v080.i01

[92] DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986; 7(3): 177-88. https://doi.org/10.1016/0197-2456(86)90046-2

[93] IntHout J, Ioannidis JPA, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. BMC Med Res Methodol 2014; 14: 25. https://doi.org/10.1186/1471-2288-14-25

[94] Röver C, Friede T. Using the bayesmeta R package for Bayesian random-effects meta-regression. Comput Methods Programs Biomed 2023; 229: 107303. https://doi.org/10.1016/j.cmpb.2022.107303

[95] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, *et al*. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021; n71. https://doi.org/10.1136/bmj.n71

Received on 03-12-2025 Accepted on 02-01-2026 Published on 30-01-2026

https://doi.org/10.6000/1929-6029.2026.15.01