Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Validity of self-reported tooth count compared with clinical examination in adults: a systematic review and meta-analysis

Carlos Andrés Moreno Alarcón
Universidad de los Andes
Andrés Osvaldo Celis Sersen
Universidad de los Andes
Zabdiel Andrés Faundez Maureira
Universidad de los Andes
Rocío Maturana Zaviezo
Universidad de los Andes
Nicolás Andrés Portalier Duble
Universidad de los Andes
Camilo Andrés Montero Olivares
Universidad de los Andes
Jorge Andrés Celis Dooner
Universidad de los Andes
Duniel Ricardo Ortuño Borroto
dortuno@uandes.cl
Universidad de los Andes

Systematic Review

Keywords: self-reported tooth count, tooth loss, validity, agreement, epidemiological surveillance

Posted Date: September 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10806706/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Tooth loss is a critical marker of oral disease burden and impacts masticatory efficiency, speech, self-esteem and psychosocial well-being. Given the substantial implications of tooth loss, accurate assessment of dental status is essential for public health surveillance, planning, and the design of targeted interventions. While clinical examination is the reference method for tooth counting, it is resource-intensive. Self-reported measures offer a cost-effective alternative, but their validity is debated. This systematic review and meta-analysis aimed to assess the validity of self-reported tooth count in adults.

Methods

A comprehensive literature search was conducted in MEDLINE, EMBASE, Scopus, and Web of Science up to April 2026. We included studies that compared self-reported tooth count with clinical examinations in adult populations. A random-effects meta-analysis was performed to calculate the pooled mean difference between self-reported and clinically examined tooth count.

Results

Sixteen studies met the inclusion criteria. Association and agreement measures between self-reported and clinically examined tooth count varied across studies, with Pearson's r ranging from 0.77 to 0.96, Spearman's ρ from 0.47 to 0.93, Cohen's κ from 0.56 to 0.63, ICC from 0.78 to 0.96, and CCC from 0.40 to 0.95. Seven studies were included in meta-analysis. The pooled mean difference was −0.036 teeth (95% CI: −0.263 to 0.191; p = 0.712). However, substantial heterogeneity was observed across studies ($\tau^2 = 0.042$; $I^2 = 84.92\%$), indicating considerable variability in the direction and magnitude of the discrepancy across populations.

Conclusion

The pooled mean difference between self-reported and clinically examined tooth count was small and non-significant; however, the heterogeneity observed across studies limits the strength of this overall estimate and indicates that the accuracy of self-report varies across populations. These findings suggest that self-reported tooth count may hold potential as a screening tool for large-scale epidemiological surveillance.

PROSPERO Register number: CRD42024507785

BACKGROUND

Oral health is recognized as a vital component of overall health, influencing social interactions and general quality of life (1, 2). Tooth loss, including both partial and complete edentulism, remains prevalent globally, particularly among older adults, despite significant advances in preventive and restorative dentistry (3, 4). Tooth loss is often regarded as the “final marker” of oral disease, reflecting the cumulative impact of untreated dental caries, periodontal disease, and limited access to adequate dental care (5, 6). This condition not only impairs essential functions, such as chewing and speech, but also affects self-esteem and psychosocial well-being (2, 7).

Furthermore, multiple studies have documented associations between tooth loss and systemic conditions, including cardiovascular disease, diabetes, and increased all-cause mortality (8, 9). The loss of natural dentition impacts masticatory efficiency, which is critical for consuming a nutritionally balanced diet. Individuals with fewer teeth often face difficulties in consuming fibre-rich foods, fruits, and vegetables, which are essential for preventing chronic metabolic conditions and supporting overall health (10, 11). Additionally, poor diet quality linked to tooth loss has been implicated in the development of frailty and functional limitations, particularly in older adults (12).

Given the substantial implications of tooth loss on both oral and systemic health, accurate assessment of dental status is essential for public health surveillance, health planning, and the design of targeted interventions (13). While clinical dental examination is considered the gold standard for determining the number of remaining teeth, it is logistically challenging and resource-intensive, particularly for large-scale epidemiological studies (14). In response, self-reported measures of oral health have been increasingly utilized as cost-effective alternatives, allowing for broader population coverage in epidemiological surveys (15, 16).

The validity of self-reported oral health measures, however, remains a topic of debate. Research on the diagnostic accuracy of self-reported data, particularly on the number of teeth and periodontal health status, has yielded mixed results. For instance, some studies

report high levels of agreement between self-reported and clinically examined tooth counts, indicating that self-reported data can serve as a reliable proxy in large-scale surveys (14, 15). Notably, Pitiphat et al. found a significant correlation between self-reported and clinical tooth counts in adult populations, suggesting its potential applicability in epidemiological research (14). Conversely, other studies highlight substantial variability influenced by factors such as age, socioeconomic status, cognitive ability, and awareness of oral health, which may affect the reliability of self-reported measures (13, 16).

Currently, to our knowledge, no systematic review with meta-analysis has synthesized the evidence comparing self-reported and clinically examined tooth counts to generate a consensus on their applicability in large-scale epidemiological studies. This study aimed to critically assess the validity of self-reported number of teeth compared to clinical examination, focusing on the adult population. In this review, validity refers to the extent to which self-reported tooth counts correspond to clinically examined tooth counts.

METHODS

A protocol for this study was registered in the International Prospective Register of Systematic Reviews, registration number: CRD42024507785. This systematic review evaluates the validity of self-reported tooth count compared to clinical examination, adhering to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (17).

Literature Search

A comprehensive literature search was conducted across multiple databases through the Ovid platform in MEDLINE, EMBASE, Scopus, and Web of Science, covering studies published in English or Spanish from database inception up to April 2026. Additionally, reference lists of all eligible articles were reviewed to identify studies not captured in the initial database search. The search strategy combined free-text terms with Boolean operators AND and OR to ensure sensitivity and specificity. The detailed search structure is presented in Supplementary Table 1.

Study Selection

Three independent reviewers (C.M.A., R.M.Z. and Z.F.M.) conducted the study selection using the Rayyan.ai screening tool (18). Studies were included if they assessed self-reported tooth count compared with clinical dental examination and the accuracy of self-report data in estimating the number of natural teeth in adults. In addition, studies must be published in English or Spanish.

Exclusion criteria were studies that have participants with cognitive impairment or other conditions that precluded reliable completion of the self-reported tooth-count assessment. Studies that focused solely on other dental conditions without reference to tooth count, those utilizing therapies unrelated to self-reported assessments (e.g., clinical examinations only), those that do not include comparison between the self-reported assessment and the clinical assessment, case reports, journal or conference abstracts (structured or unstructured), reviews, letters to the editor, and book chapters. In cases where studies addressed general oral health disorders, data specific to self-reported tooth count validity were extracted. Any discrepancies regarding the inclusion or exclusion of studies were resolved through discussion or consultation with a fourth reviewer (A.C.S.).

Data Extraction and Analysis

Data were systematically extracted from the included studies following a qualitative assessment based on the criteria. One reviewer (C.M.A.) performed the data extraction, with confirmation by other reviewers (A.C.S., D.O.B. and R.M.Z.). Extracted data included title, author, journal, year, country, study design, study objectives, sample size, population characteristics, methods of self-reported tooth count assessment, self-reported questions, self-report criteria, type of examiner, clinical count criteria, primary findings, validity, association, and agreement measures, and any identified biases or limitations. The primary outcome was the paired mean difference in the number of teeth between self-reported and clinically examined counts, calculated as self-reported minus clinically examined tooth count.

Risk of Bias Assessment

Two reviewers (C.M.A. and R.M.Z.) independently assessed the risk of bias using an adaptation of the Newcastle-Ottawa Scale (NOS) for cross-sectional studies (19). Discrepancies were resolved through discussion and consensus.

Quality Assessment

To evaluate the methodological quality of the included studies, the Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies from the National Heart, Lung, and Blood Institute (NHLBI) was applied (20). Two reviewers (C.M.A. and R.M.Z.) independently assessed each study's quality, resolving disagreements through consensus.

Statistical Analysis

Statistical analyses and graphical representations were performed using STATA version 18.0 (StataCorp). A narrative synthesis was elaborated to summarize the characteristics of the included studies. A random-effects meta-analysis was conducted to estimate the pooled paired mean difference in tooth count between self-reported and clinically examined tooth count and its corresponding standard error. For studies reporting only standard deviations for self-reported and clinically examined tooth count, the standard deviation of the paired difference was calculated using the correlation coefficient reported by the study (Supplementary Fig. 1). When the SD of the paired difference was not reported but a 95% confidence interval was available, it was derived from the confidence interval and sample size. When the underlying total sample data were available, the mean of each group, their difference, and the standard deviation of this difference were calculated directly from the raw data. Between-study variance (τ^2) was estimated using restricted maximum likelihood (REML), and the confidence interval of the pooled effect was calculated using the Hartung-Knapp-Sidik-Jonkman (HKSJ) adjustment. As sensitivity analyses, (1) the pooled estimate was recalculated using the conventional DerSimonian-Laird estimator with a Wald-type confidence interval, and (2) a leave-one-out analysis was performed to assess the influence of individual studies on the pooled estimate and heterogeneity.

Certainty of evidence assessment

The certainty of evidence was evaluated using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework (21), with two reviewers (C.M.A. and R.M.Z.) independently assessing risk of bias, inconsistency, indirectness, imprecision, and publication bias. The overall certainty was categorized as high, moderate, low, or very low, with disagreements resolved by consensus.

RESULTS

The database search yielded a total of 2,051 records. After removing 110 duplicates, 1,941 records were screened by title and abstract. From this initial filter, 20 articles were identified for full-text review; none were excluded for inaccessibility. Following full-text review, 16 articles met the inclusion criteria for this review. The selection process is detailed in the PRISMA flowchart (Fig. 1).

Study Characteristics

The included studies represented a wide geographical diversity, encompassing 11 countries across Asia, Europe, the Americas, and the Middle East. The total sample size across all studies exceeded 11,000 participants. Regarding sex composition, 93.7% of the studies (15 out of 16) included both men and women, with the exception of Nascimento et al. (31), which focused exclusively on older women.

The statistical tests used to measure association and/or agreement between self-reported and clinical examination tooth count were as follows: Pearson's correlation (7 studies), Spearman's rank correlation (7 studies), Cohen's κ (2 studies), Intraclass Correlation Coefficient (ICC) (5 studies), and Concordance Correlation Coefficient (CCC) (2 studies). Two studies did not apply a test for association or agreement. A summary of the included studies is presented in Table 2.

Table 2
Summary of included articles

Author	Title	Country	Sample size	Sex	Age	Association/agreement test
Ueno et al. (22)	Validity of self-reported tooth counts and masticatory status study of a Japanese adult population	Japan	2356	Women and men	40–74	Pearson Spearman ICC
Buhlin et al. (23)	Validity and limitations of self-reported periodontal health	Sweden	129	Women and men	20–84	Did not perform
Axelsson et al. (24)	Comparison of oral health data from self-administered questionnaire and clinical examination.	Iceland	318	Women and men	18, 35–44, ≥ 65	Cohen's κ
Hakeem et al. (25)	Validity of self-reported number of teeth among Saudi older adults: A public health perspective	Saudi Arabia	337	Women and men	≥ 60	Spearman Pearson
Margozzini et al. (26)	Validity of the self-reported number of teeth in Chilean adults	Chile	101	Women and men	15–35, 36–60, ≥ 60	Spearman ICC
Sekundo et al. (27)	Patients' self-reported measures of oral health—A validation study on basis of oral health questions used in a large multi-country survey for populations aged 50+	Germany	186	Women and men	≥ 50, ≥ 18	CCC
Matsui et al. (28)	Validity of self-reported number of teeth and oral health variables	Japan	1501	Women and men	≥ 35	Spearman
Ueno et al. (29)	Validity of the self-reported number of natural teeth in Japanese adults	Japan	1152	Women and men	40–56	Pearson ICC
Shimazaki et al. (30)	Validity of the self-reported number of teeth in independent older people in Japan	Japan	2621	Women and men	75, 80	Pearson Cohen's κ
Nascimento et al. (31)	Validity of self-reported oral conditions among Brazilian older women: Do socio-economic factors matter?	Brazil	72	Women	≥ 60	CCC
Høvik et al. (32)	The validity of self-reported number of teeth and edentulousness among Norwegian older adults, the HUNT Study	Norway	586	Women and men	≥ 70	Spearman Pearson
Similä et al. (33)	Validity of self-reported number of teeth in middle-aged Finnish adults: the Northern Finland Birth Cohort Study 1966	Finland	1669	Women and men	46	Spearman Pearson ICC
Pitiphat et al. (14)	Validation of Self-reported Oral Health Measures	USA	58	Women and men	≥ 18	Spearman
Gilbert et al. (34)	Validity of Self-reported Tooth Counts During a Telephone Screening Interview	USA	506	Women and men	≥ 48	ICC

Author	Title	Country	Sample size	Sex	Age	Association/agreement test
Douglass et al. (35)	The Validity of Self-reported Oral Health Status in the Elderly	USA	48	Women and men	≥ 70	Pearson
Blizniuk et al. (36)	Association between self-reported and clinical oral health status in Belarusian adults	Belarus	281	Women and men	18–60	Did not perform

Main Findings

The difference between these means varied from − 1.4 to 0.6 teeth. However, the percentage of perfect agreement (participants with an exact match between self-reported and clinical examination tooth counts) varied considerably, ranging from 28.5% in Belarus to 73.4% among the older adult group in Iceland (Table 3). Association and agreement estimates varied across studies, with Pearson's r ranging from 0.77 to 0.96, Spearman's ρ from 0.47 to 0.93, Cohen's κ from 0.56 to 0.63, ICC from 0.78 to 0.96, and CCC from 0.40 to 0.95. Of the 16 included articles, 12 provided results in terms of the mean tooth counts for each group.

Substantial variability was observed across studies in both the self-report and clinical examination methodologies. Regarding self-report, studies differed in questionnaire administration mode, question wording (e.g., "natural teeth" vs. "teeth," with or without clarification on prostheses), and instructions provided to participants (e.g., use of a mirror, removal of removable prostheses). Regarding clinical examination, heterogeneity was observed in criteria for counting a "tooth" (e.g., inclusion/exclusion of root remnants, crowned teeth) and examiner type (calibrated dentists vs. non-dental professionals, with variable reporting of inter-examiner reliability) (Supplementary Table 2).

Table 3
Main results summary

Author	Self-report mean (SR)	Clinical examination mean (CE)	Difference (SR-CE)	Pearson	Spearman	Cohen's κ	ICC	CCC	Exact Agreement (%)
Ueno et al. (2018)	23.78	23.68	0.098	0.845	0.771	-	0.843	-	-
Buhlin et al.	26.60	26.20	0.40	-	-	-	-	-	41.10%
Axelsson et al.	-	-	-	-	-	18: 0.56 35–44: 0.6 ≥ 65: 0.63	-	-	73.33% 65.90% 73.40%
Hakeem et al.	18.80	18.40	0.40	0.96	0.91	-	-	-	33.80%
Margozzini et al.	20.00	20.00	0.00	-	0.93	-	0.96	-	-
Sekundo et al.	-	-	-	-	-	-	-	≥ 50: 0.95	-
Matsui et al.	25.10	26.50	-1.40	-	0.69	-	-	-	-
Ueno et al. (2010)	25.52	25.84	-0.32	0.8	-	-	0,78	-	47.50%
Shimazaki et al.	-	-	-	0.96	-	0.559	-	-	57.50%
Nascimento et al.	Maxillary: 8.60 Mandible: 6.60	Maxillary: 7.10 Mandible: 9.10	Maxillary: 1.50 Mandible: -2.50	-	-	-	-	Maxillary: 0.4 Mandible: 0.5	-
Høvik et al.	20.80	21.00	-0.20	0.91	0.86	-	-	-	35.00%
Similä et al.	27.48	27.46	0.02	Men: 0,81 Women: 0.77	Men: 0.78 Women: 0.82	-	Men: 0.8 Women: 0.76	-	-
Pitiphat et al.	25.00	25.70	-0.70	-	0.47	-	-	-	-
Gilbert et al.	21.10	21.20	-0.10	-	-	-	0.87	-	23.71%
Douglass et al.	12.44	12.79	-0.35	0.97	-	-	-	-	-
Blizniuk et al.	27.60	27.00	0.60	-	-	-	-	-	-

Risk of Bias and Quality Assessment

Of the total articles, 11 were classified as "Good" and 5 as "Fair" in terms of methodological quality based on the NHLBI tool (Table 4). Common weaknesses across the included studies were the lack of justification for sample size or statistical power, limited reporting of examiner calibration or reliability, and incomplete adjustment for potential confounders. In general, most studies had clear research questions, defined populations, uniform inclusion and exclusion criteria, and well-defined exposure and outcome measures.

Table 4
Summary of quality assessment

Author	Score	Rating	Critical Points
Ueno et al. (2018)	13/14	Good	Sample size was not justified.
Buhlin et al.	13/14	Good	Sample size was not justified.
Axelsson et al.	13/14	Good	Sample size was not justified.
Hakeem et al.	10/14	Fair	Convenience sampling, participation rate not reported, sample size not justified, clinical exam without report of calibration/reliability between examiners.
Margozzini et al.	12/14	Good	Low participation rate (50% of those contacted), limited adjustment for confounding variables.
Sekundo et al.	12/14	Good	Convenience sampling at a single center, multiple exposure levels not evaluated (partial).
Matsui et al.	10/14	Fair	Participation rate not clearly reported, sample size not justified, inter-examiner reliability not calculated, limited adjustment for confounding variables.
Ueno et al. (2010)	9/14	Fair	Very low participation rate (5.3%), sample size not justified, inter-examiner reliability not calculated, multiple examiners without formal calibration.
Shimazaki et al.	9/14	Fair	Low participation rate (16.5% at age 75, 13.2% at age 80), sample size not justified, no examiner training/calibration.
Nascimento et al.	13/14	Good	Sample size was not justified.
Hovik et al.	13/14	Good	Sample size was not justified.
Similä et al.	13/14	Good	Sample size was not justified.
Pitiphat et al.	12/14	Good	Moderate participation rate (62%), limited adjustment for confounding variables.
Gilbert et al.	12/14	Good	Moderate participation rate (48.5% at examination phase), statistical power not reported.
Douglass et al.	13/14	Good	Sample size was not justified.
Blizniuk et al.	10/14	Fair	Convenience sampling among dental patients, sample size not justified, single examiner without report of intra-examiner reliability, non-representative population (patients only).

The risk of bias analysis, using the Newcastle-Ottawa Scale adaptation, yielded the following results: low risk, 7 articles; moderate risk, 8 articles; and high risk, 1 article (Table 5). The most recurrent sources of bias were the representativeness of the sample (not using random or representative sampling), an unjustified sample size, and a lack of adjustment for confounders.

Table 5
Summary of risk of bias assessment.

Author	Score	Interpretation	Critical points
Ueno et al. (2010)	7/9	Low risk	Sample size: sample size not justified; Exposure assessment: acceptable exposure assessment instrument, but not a gold standard.
Buhlin et al.	6/9	Moderate risk	Sample size: sample size not justified; Exposure assessment: acceptable exposure assessment instrument, but not a gold standard; Adjustment for confounders: partial adjustment for confounders.
Axelsson et al.	4/9	Moderate risk	Adjustment for confounders: inadequate adjustment for confounders; Sample size: sample size not justified; Exposure assessment: acceptable exposure assessment instrument, but not a gold standard; Confounder assessment: confounders not measured using an acceptable method.
Hakeem et al.	6/9	Moderate risk	Sample representativeness: non-representative sample or sampling method not described; Sample size: sample size not justified; Exposure assessment: acceptable exposure assessment instrument, but not a gold standard.
Margozzini et al.	3/9	High risk	Adjustment for confounders: inadequate adjustment for confounders; Sample representativeness: non-representative sample or sampling method not described; Sample size: sample size not justified; Exposure assessment: acceptable exposure assessment instrument, but not a gold standard; Confounder assessment: confounders not measured using an acceptable method.
Sekundo et al.	7/9	Low risk	Exposure assessment: acceptable exposure assessment instrument, but not a gold standard; Adjustment for confounders: partial adjustment for confounders.
Matsui et al.	4/9	Moderate risk	Sample representativeness: non-representative sample or sampling method not described; Sample size: sample size not justified; Exposure assessment: acceptable exposure assessment instrument, but not a gold standard; Outcome assessment: acceptable outcome assessment instrument, but not a gold standard; Adjustment for confounders: partial adjustment for confounders.
Ueno et al. (2018)	7/9	Low risk	Sample size: sample size not justified; Exposure assessment: acceptable exposure assessment instrument, but not a gold standard.
Shimazaki et al.	7/9	Low risk	Sample size: sample size not justified; Exposure assessment: acceptable exposure assessment instrument, but not a gold standard.
Nascimento et al.	6/9	Moderate risk	Sample size: sample size not justified; Exposure assessment: acceptable exposure assessment instrument, but not a gold standard; Outcome assessment: acceptable outcome assessment instrument, but not a gold standard.
Hovik et al.	7/9	Low risk	Sample size: sample size not justified; Exposure assessment: acceptable exposure assessment instrument, but not a gold standard.
Similä et al.	7/9	Low risk	Sample size: sample size not justified; Exposure assessment: acceptable exposure assessment instrument, but not a gold standard.
Pitiphat et al.	7/9	Low risk	Sample size: sample size not justified; Exposure assessment: acceptable exposure assessment instrument, but not a gold standard.
Gilbert et al.	6/9	Moderate risk	Sample representativeness: non-representative sample or sampling method not described; Sample size: sample size not justified; Exposure assessment: acceptable exposure assessment instrument, but not a gold standard.
Douglass et al.	5/9	Moderate risk	Sample size: sample size not justified; Exposure assessment: acceptable exposure assessment instrument, but not a gold standard; Outcome assessment: acceptable outcome assessment instrument, but not a gold standard; Adjustment for confounders: partial adjustment for confounders.
Blizniuk et al.	4/9	Moderate risk	Adjustment for confounders: inadequate adjustment for confounders; Sample size: sample size not justified; Exposure assessment: acceptable exposure assessment instrument, but not a gold standard; Confounder assessment: confounders not measured using an acceptable method.

Meta-Analysis

For the meta-analysis, 7 of the 16 studies provided sufficient data for quantitative synthesis, the remaining 9 studies were included in the qualitative synthesis but could not be quantitatively pooled because the information required to calculate the paired mean difference and its variance was unavailable. This analysis used the paired mean difference between self-reported and clinically

examined tooth counts and the corresponding standard deviation, obtained from reported or calculated variance estimates from these 7 articles (Supplementary Table 3).

The pooled effect using a random effects model was a mean difference of -0.036 teeth (95% CI: -0.263 to 0.191). The pooled estimate was slightly negative, indicating a numerically lower self-reported than the clinically examined tooth count; however, the confidence interval included zero. The analysis revealed substantial heterogeneity across studies: I^2 = 84.92% (Fig. 2).

As a sensitivity analysis, the pooled mean difference was recalculated using the conventional DerSimonian-Laird random-effects estimator with a Wald-type (z-based) confidence interval. This yielded a pooled estimate of − 0.03 (95% CI: −0.18 to 0.11; z = − 0.46, p = 0.65) (Fig. 3)

A leave-one-out sensitivity analysis was performed to assess the influence of individual studies on the pooled estimate. The recalculated pooled mean differences ranged from − 0.09 (95% CI: -0.27 to 0.09) when Study 6 was omitted, to 0.02 (95% CI: −0.20 to 0.24) when Study 2 was omitted (Fig. 4). All leave-one-out estimates remained close to zero, with confidence intervals that included the null value.

Certainty of evidence

Using the GRADE framework, the certainty of evidence for the primary outcome was rated as moderate, downgraded for serious inconsistency and risk of bias. Imprecision was considered in relation to the clinical relevance of the confidence interval, although no established minimal clinically important difference exists for the number of teeth. Publication bias could not be reliably assessed because fewer than 10 studies were included in the meta-analysis.

DISCUSSION

This systematic review and meta-analysis evaluated the validity of self-reported compared to clinically examined tooth count, incorporating 16 observational studies with heterogeneous sample sizes (48 to 2,621 participants) from 11 countries. The pooled estimate was close to zero, with a 95% CI that included zero; however, this null finding must be interpreted alongside the substantial heterogeneity observed across studies (I^2 = 84.92%), which limits the extent to which a single pooled estimate can characterize agreement across diverse populations. Taken together, these findings suggest that self-reported tooth count may hold potential as a proxy for dental status at the population level, particularly in large-scale epidemiological studies where clinical examination is not feasible; however, this potential should be interpreted cautiously given the variability in accuracy observed across settings (22, 25, 26, 29).

Despite these favorable results, several factors could directly affect the difference between self-report and clinical examination. One of the most consistent findings across studies is that oral health complexity appears to be an important potential source of reporting error. Generally, patients have difficulty distinguishing between natural teeth and fixed prostheses (single or multiple). In the study by Ueno et al., the presence of at least one multiple fixed prosthesis was significantly associated with greater inaccuracy in self-report (29). Furthermore, methodological analysis reveals a lack of clarity on whether root remnants should be counted as "teeth," as seen in the Norwegian study where their clinical inclusion increased underreporting (32). Contrary to what might be assumed, a higher number of teeth in the mouth is not associated with better self-report accuracy; in fact, some studies suggest it is easier to count accurately when several teeth have been lost (30, 32).

Another factor that could influence self-report is the way questions are phrased and instructions are given to participants. For example, using the phrase "natural teeth" may lead participants to erroneously exclude teeth with fixed prostheses, even though they constitute a functional unit (26, 28). Conversely, providing specific instructions, such as using a mirror, using fingers to count, or indicating the removal of removable prostheses, could significantly improve the validity of self-reported results (26, 35). This technical ambiguity, combined with recall bias (particularly when substantial time elapsed between the clinical examination and self-report) and potential social desirability bias related to perceptions of tooth loss, may contribute to the high heterogeneity detected in the meta-analysis (25, 26).

Socioeconomic position (by education and/or income) also influences reporting accuracy, although it shows no clear trend or association, as its impact may vary according to the country's context. In Brazil, greater discrepancies were found in women with basic

education and low income, who tended to overestimate upper teeth and underestimate lower teeth (31). In Finland, conversely, agreement was higher among men with basic education (33).

Psychological and cultural factors surrounding self-perception of oral health may also influence the differences in self-report, although evidence for these mechanisms remains limited (25). Although not an exclusion criterion, Cognitive impairment may also compromise the reliability of self-reported tooth count, particularly in older populations. This factor should therefore be considered when interpreting self-reported tooth count in populations with a high prevalence of cognitive impairment (32). This information suggests that data obtained in populations with high prevalences of cognitive impairment should consider this factor both methodologically and in the interpretation of results.

Limitations

The present review has several limitations. The high variability in the statistical tests used by the included studies to evaluate association and/or agreement, as well as differences in how results were presented, limited the inclusion of only 7 of 16 articles in the meta-analysis. Additionally, the high heterogeneity ($I^2 = 84.92\%$) could not be fully explained by the available covariates. The primary studies differed in how the self-report question was formulated, the clinical criterion for a "tooth," the type of examiner, and the population (older adults, institutionalized, mixed-age), which limits the ability to obtain a truly homogeneous combined estimate. Likewise, the absence of test-retest evaluations of self-report in the primary studies prevents differentiation between random error and systematic bias associated with the collection method.

In addition to the limitations arising from the primary studies, several methodological choices inherent to this review may have influenced its conclusions. The meta-analysis was restricted to the mean difference between self-reported and clinically examined tooth counts; association and agreement measures were synthesized narratively rather than quantitatively, so the overall association between self-report and clinical examination could not be expressed as a single combined estimate. Despite the very high heterogeneity observed ($I^2 = 84.92\%$), no formal meta-regression or subgroup analysis by region, age group, or self-report question wording was conducted to explore its sources beyond the leave-one-out sensitivity analysis, because only seven studies contributed to the quantitative synthesis, meta-regression and formal subgroup analyses were not considered sufficiently powered or reliable. Finally, the search was restricted to studies published in English or Spanish and to indexed databases, without a dedicated search of gray literature or trial registries, which may have excluded relevant studies and introduced language or publication bias.

CONCLUSION

Self-reported tooth count may provide an acceptable population-level estimate, but high association coefficients do not necessarily imply accurate individual-level agreement. These findings are especially relevant for epidemiological surveillance and public health planning, contexts where feasibility and sample coverage are often prioritized.

In conclusion, the pooled estimate was close to zero, suggesting little average systematic discrepancy between self-reported and clinically examined tooth count. However, the substantial between-study heterogeneity limits the generalizability of this pooled estimate.

It is a priority to unify the formulation of questions, explicitly define what is meant by "tooth" both in self-report and clinical examination, and establish standardized instructions, given their potential impact on the accuracy of self-report. Furthermore, future studies should control or adjust for potential sources of bias and confounding, including socioeconomic status, oral health complexity (use of removable prostheses, fixed prostheses, caries, fillings), and potential social desirability biases linked to the stigma of tooth loss.

Abbreviations

CCC
Concordance Correlation Coefficient
CE
Clinical examination
CI
Confidence interval
GRADE

Grading of Recommendations Assessment, Development and Evaluation
HKSJ
Hartung-Knapp-Sidik-Jonkman
ICC
Intraclass Correlation Coefficient
NHLBI
National Heart, Lung, and Blood Institute
NOS
Newcastle-Ottawa Scale
PRISMA
Preferred Reporting Items for Systematic Reviews and Meta-Analyses
REML
Restricted Maximum Likelihood
SD
Standard Deviation
SR
Self-report

Declarations

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable.

Availability of data and materials: All data generated or analyzed during this study are included in this article and its supplementary information.

Competing interests: The authors declare that they have no competing interests.

Funding: This study did not receive any specific funding.

Authors' contributions: R.M.Z.: Data curation, Investigation, Writing-original draft. Z.F.M.: Data curation, Investigation, A.C.S.: Conceptualization, Writing-review & editing, Methodology, Supervision, Validation, C.M.A.: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing-original draft, Visualization, N.P.D.: Data curation, Investigation, C.M.O.: Data curation, Investigation, J.C.D.: Conceptualization, Writing-original draft, D.O.B.: Conceptualization, Writing-review & editing, Methodology, Supervision, final validation of the manuscript.

Acknowledgements: Not applicable.

References

1. Emami E, de Souza RF, Kabawat M, Feine JS. The impact of edentulism on oral and general health. Int J Dent. 2013;2013:498305. 10.1155/2013/498305 . PubMed PMID: 23737789; PubMed Central PMCID: PMC3664508.
2. Griffin SO, Jones JA, Brunson D, Griffin PM, Bailey WD. Burden of Oral Disease Among Older Adults and Implications for Public Health Priorities. Am J Public Health. 2012;102(3):411–8. 10.2105/AJPH.2011.300362 . PubMed PMID: 22390504; PubMed Central PMCID: PMC3487659.
3. Shin HS. The number of teeth is associated with diet quality in Korean adult population. Arch Oral Biol. 2020;118:104882. 10.1016/j.archoralbio.2020.104882 . PubMed PMID: 32835987.
4. Al-Rafee MA. The epidemiology of edentulism and the associated factors: A literature Review. J Fam Med Prim Care. 2020;9(4):1841–3. 10.4103/jfmpc. .jfmpc_1181_19 PubMed PMID: 32670928; PubMed Central PMCID: PMC7346915.
5. Gennai S, Izzetti R, Pioli MC, Music L, Graziani F. Impact of rehabilitation versus edentulism on systemic health and quality of life in patients affected by periodontitis: A systematic review and meta-analysis. J Clin Periodontol. 2022;49(Suppl 24):328–58. 10.1111/jcpe.13526 . PubMed PMID: 34761419.

6. Kassebaum NJ, Bernabé E, Dahiya M, Bhandari B, Murray CJL, Marcenes W. Global Burden of Severe Tooth Loss: A Systematic Review and Meta-analysis. J Dent Res. 2014;93(7 Suppl):S20–8. 10.1177/0022034514537828 . PubMed PMID: 24947899; PubMed Central PMCID: PMC4293725.
7. Yu YH, Cheung WS, Steffensen B, Miller DR. Number of teeth is associated with all-cause and disease-specific mortality. BMC Oral Health. 2021;21(1):568. 10.1186/s12903-021-01934-0 . PubMed PMID: 34749715; PubMed Central PMCID: PMC8574051.
8. Zhu Y, Hollis JH. Associations between the number of natural teeth and metabolic syndrome in adults. J Clin Periodontol. 2015;42(2):113–20. 10.1111/jcpe.12361 . PubMed PMID: 25581485.
9. Nowjack-Raymer RE, Sheiham A. Numbers of natural teeth, diet, and nutritional status in US adults. J Dent Res. 2007;86(12):1171–5. doi:10.1177/154405910708601206 PubMed PMID: 18037650.
10. Chalittikul W, Kassim S, Sabbah W. The association between number of teeth and physical function limitation among older adults in the USA. Gerodontology. 2020;37(4):389–94. 10.1111/ger.12489 . PubMed PMID: 32710496.
11. Bortoluzzi MC, Traebert J, Lasta R, Da Rosa TN, Capella DL, Presta AA. Tooth loss, chewing ability and quality of life. Contemp Clin Dent. 2012;3(4):393–7. 10.4103/0976-237X. .107424 PubMed PMID: 23633796; PubMed Central PMCID: PMC3636836.
12. Ramos RQ, Bastos JL, Peres MA. Diagnostic validity of self-reported oral health outcomes in population surveys: literature review. Rev Bras Epidemiol Braz J Epidemiol. 2013;16(3):716–28. 10.1590/s1415-790x2013000300015 . PubMed PMID: 24896284.
13. Blicher B, Joshipura K, Eke P. Validation of self-reported periodontal disease: a systematic review. J Dent Res. 2005;84(10):881–90. doi:10.1177/154405910508401003 PubMed PMID: 16183785.
14. Pitiphat W, Garcia RI, Douglass CW, Joshipura KJ. Validation of self-reported oral health measures. J Public Health Dent. 2002;62(2):122–8. 10.1111/j.1752-7325.2002.tb. 03432.x PubMed PMID: 11989207.
15. Axelsson G, Helgadóttir S. Community Dent Oral Epidemiol. 1995;23(6):365–8. 10.1111/j.1600-0528.1995.tb. 00264.x PubMed PMID: 8681520. Comparison of oral health data from self-administered questionnaire and clinical examination.
16. Gilbert GH, Chavers LS, Shelton BJ. Comparison of two methods of estimating 48-month tooth loss incidence. J Public Health Dent. 2002;62(3):163–9. 10.1111/j.1752-7325.2002.tb. 03438.x PubMed PMID: 12180044.
17. PRISMA statement [Internet]. [cited 2026 Jul 20]. PRISMA statement. Available from: https://www.prisma-statement.org
18. AI-Powered Systematic Review & Literature Review Platform [Internet]. [cited 2026 Jul 6]. Available from: https://www.rayyan.ai/
19. Risk of Bias Evaluation of Cross-Sectional Studies. Adaptation of the Newcastle-Ottawa Scale - Carra - Journal of Periodontal Research - Wiley Online Library [Internet]. [cited 2026 Jun 29]. Available from: https://onlinelibrary.wiley.com/doi/10.1111/jre.13405
20. Study. Quality Assessment Tools | NHLBI, NIH [Internet]. [cited 2026 Jun 29]. Available from: https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools
21. Brignardello-Petersen R, Guyatt GH. Assessing the certainty of the evidence in systematic reviews: importance, process, and use. Am J Epidemiol. 2025;194(6):1681–6. 10.1093/aje/kwae. 332 PubMed PMID: 39218429; PubMed Central PMCID: PMC12133264.
22. Ueno M, Shimazu T, Sawada N, Tsugane S, Kawaguchi Y. Validity of self-reported tooth counts and masticatory status study of a Japanese adult population. J Oral Rehabil. 2018;45(5):393–8. 10.1111/joor.12615 . PubMed PMID: 29420835.
23. Buhlin K, Gustafsson A, Andersson K, Håkansson J, Klinge B. Validity and limitations of self-reported periodontal health. Community Dent Oral Epidemiol. 2002;30(6):431–7. 10.1034/j.1600-0528.2002. 00014.x PubMed PMID: 12453114.
24. Axelsson G, Helgadóttir S. Community Dent Oral Epidemiol. 1995;23(6):365–8. 10.1111/j.1600-0528.1995.tb. 00264.x PubMed PMID: 8681520. Comparison of oral health data from self-administered questionnaire and clinical examination.
25. Hakeem FF, AlQobaly L, Mubarak AM, Almuzaini SA. Validity of self-reported number of teeth among Saudi older adults: A public health perspective. J Taibah Univ Med Sci. 2024;19(1):122–9. 10.1016/j.jtumed.2023.10.001 . PubMed PMID: 37964862; PubMed Central PMCID: PMC10641271.
26. Margozzini P, Berríos R, Cantarutti C, Veliz C, Ortuno D. Validity of the self-reported number of teeth in Chilean adults. BMC Oral Health. 2019;19(1):99. 10.1186/s12903-019-0794-5 . PubMed PMID: 31164110; PubMed Central PMCID: PMC6549282.
27. Sekundo C, Stock C, Jürges H, Listl S. Patients' self-reported measures of oral health-A validation study on basis of oral health questions used in a large multi-country survey for populations aged 50. Gerodontology. 2019;36(2):171–9. 10.1111/ger.12398 . PubMed PMID: 30821850.
28. Matsui D, Yamamoto T, Nishigaki M, Miyatani F, Watanabe I, Koyama T, et al. Validity of self-reported number of teeth and oral health variables. BMC Oral Health. 2016;17(1):17. 10.1186/s12903-016-0248-2 . PubMed PMID: 27422287; PubMed Central

PMCID: PMC4947278.

29. Ueno M, Zaitsu T, Shinada K, Ohara S, Kawaguchi Y. Validity of the self-reported number of natural teeth in Japanese adults. J Investig Clin Dent. 2010;1(2):79–84. 10.1111/j.2041-1626. 2010.00016.x PubMed PMID: 25427261.
30. Shimazaki Y, Saito M, Nonoyama T, Inamoto Y. Validity of the self-reported number of teeth in independent older people in Japan. BMC Geriatr. 2024;24(1):900. 10.1186/s12877-024-05512-1 . PubMed PMID: 39482622; PubMed Central PMCID: PMC11526520.
31. Nascimento GG, Machado FW, Cascaes AM, Silva AE, Boscato N, Demarco FF. Validity of self-reported oral conditions among Brazilian older women: Do socio-economic factors matter? Int J Dent Hyg. 2024;22(3):521–9. 10.1111/idh.12680 . PubMed PMID: 37183553.
32. Høvik H, Kolberg M, Gjøra L, Nymoen LC, Skudutyte-Rysstad R, Hove LH, et al. The validity of self-reported number of teeth and edentulousness among Norwegian older adults, the HUNT Study. BMC Oral Health. 2022;22(1):82. 10.1186/s12903-022-02116-2 . PubMed PMID: 35313882; PubMed Central PMCID: PMC8935783.
33. Similä T, Nieminen P, Virtanen JI. Validity of self-reported number of teeth in middle-aged Finnish adults: the Northern Finland Birth Cohort Study 1966. BMC Oral Health. 2018;18(1):210. 10.1186/s12903-018-0666-4 . PubMed PMID: 30538003; PubMed Central PMCID: PMC6290517.
34. Gilbert GH, Duncan RP, Kulley AM. Validity of self-reported tooth counts during a telephone screening interview. J Public Health Dent. 1997;57(3):176–80. 10.1111/j.1752-7325.1997.tb. 02970.x PubMed PMID: 9383757.
35. Douglass CW, Berlin J, Tennstedt S. The validity of self-reported oral health status in the elderly. J Public Health Dent. 1991;51(4):220–2. 10.1111/j.1752-7325.1991.tb. 02218.x PubMed PMID: 1941773.
36. Blizniuk A, Ueno M, Zaitsu T, Kawaguchi Y. Association between self-reported and clinical oral health status in Belarusian adults. J Investig Clin Dent. 2017;8(2). 10.1111/jicd.12206 . PubMed PMID: 26762479.
37. Schober P, Boer C, Schwarte LA. Correlation Coefficients: Appropriate Use and Interpretation. Anesth Analg. 2018;126(5):1763–8. 10. 1213/ANE.0000000000002864 PubMed PMID: 29481436.
38. Akoglu H. User's guide to correlation coefficients. Turk J Emerg Med. 2018;18(3):91–3. 10.1016/j.tjem. 2018.08.001 PubMed PMID: 30191186; PubMed Central PMCID: PMC6107969.

Figures

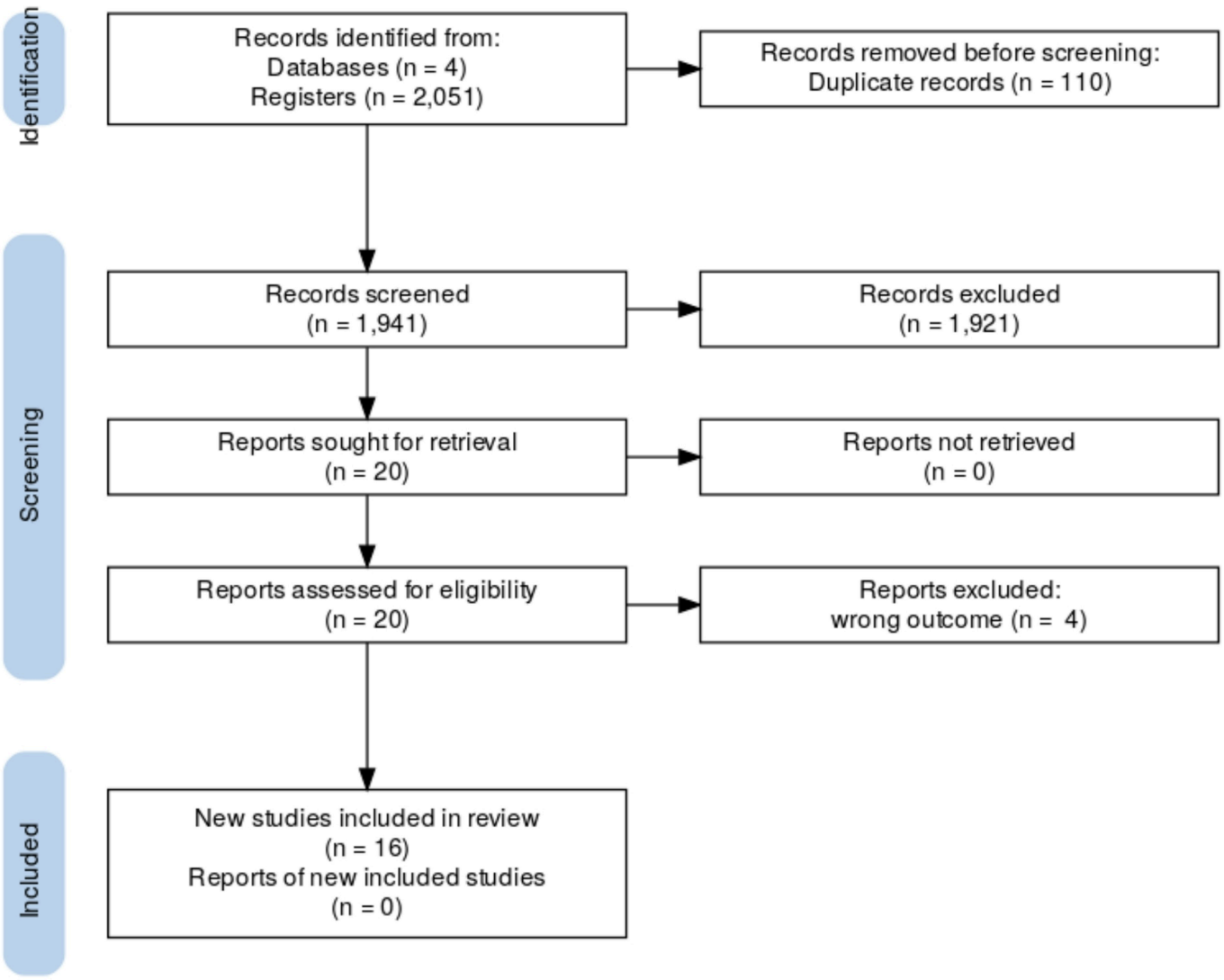


Figure 1

PRISMA Flowchart

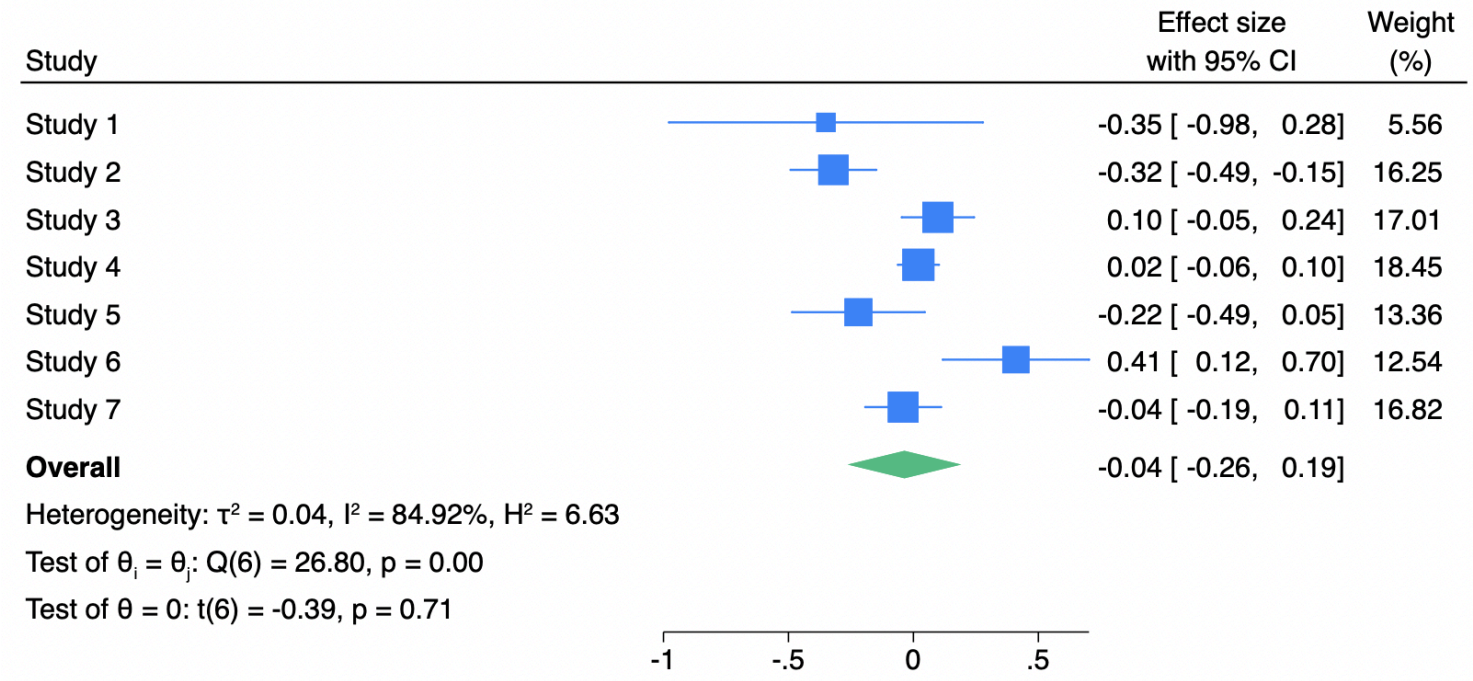


Figure 2

Forest Plot of random effects model, REML/HKSJ meta-analysis

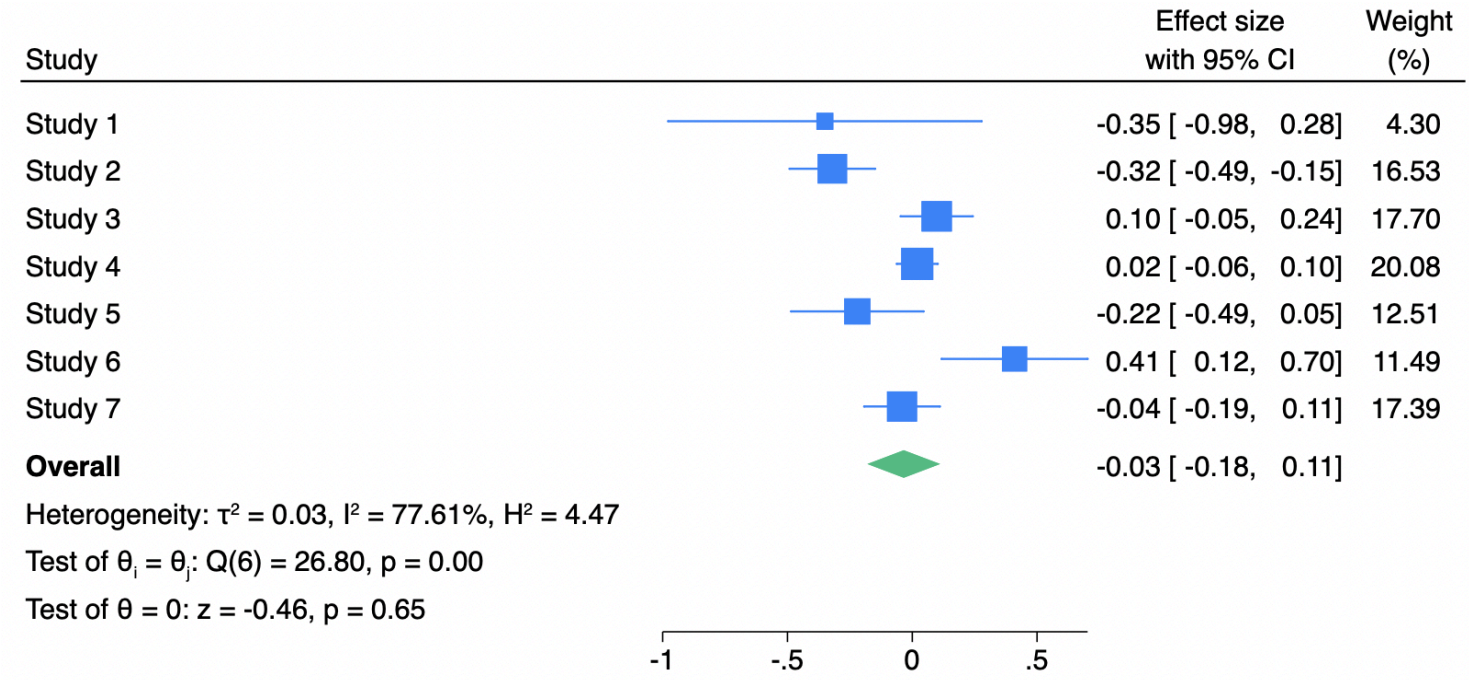


Figure 3

Forest Plot of random effects model, DerSimonian-Laird meta-analysis

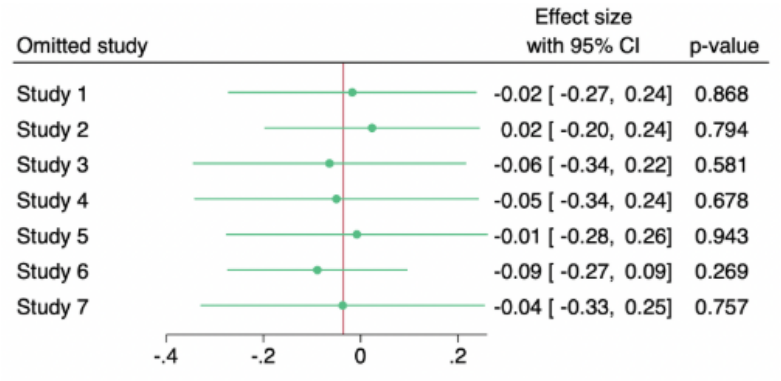


Figure 4

Leave-one-out sensitivity analysis for the meta-analysis

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementary.docx