

Mental Toughness and Self-Efficacy as Predictors of Ultramarathon Performance: A Systematic Review and Meta-Analysis

Owen R. Thornton

othornt@unc.edu

University of Michigan–Ann Arbor

Systematic Review

Keywords: ultramarathon, trail running, mental toughness, self-efficacy, performance, meta-analysis

Posted Date: August 11th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10646465/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Abstract

Background

Ultramarathon participation has grown over 345% in the last decade, yet the psychological determinants of performance in these events remain quantitatively unsynthesized. Mental toughness (MT) and self-efficacy (SE) are the two constructs most consistently proposed as indicators of ultrarunning success, but primary findings conflict, ranging from null associations in an elite sample to small positive effects in broader field samples. No meta-analysis has pooled these effects.

Methods

Quantitative studies reporting an extractable association between MT or SE and an ultramarathon performance outcome (finish time, placing, performance index, or completion versus non-finish) were identified through electronic searches of openly indexed literature and systematic citation chasing of two prior reviews and all included studies. Effects were converted to a common correlation metric, transformed to Fisher's z , and pooled with a DerSimonian–Laird random-effects model. One study reporting only non-significant omnibus statistics without direction was conservatively imputed as a null effect and examined in sensitivity analysis. The review was not prospectively registered.

Results

Four studies met inclusion criteria (total $N = 940$). The pooled association between MT/SE and ultramarathon performance was small and positive, $r = .20$ (95% CI $.13-.26$, $p < .001$), with negligible heterogeneity ($I^2 = 1\%$; 95% prediction interval $.05-.33$). Sensitivity analysis excluding the imputed null yielded $r = .21$ (95% CI $.14-.27$), and the completion-outcome subgroup yielded $r = .20$ (95% CI $.12-.27$).

Conclusions

MT and SE show a small, consistent positive association with ultramarathon performance and completion — approximately two-thirds the magnitude of the pre-event self-efficacy effect reported across sport generally. This attenuation is consistent with a threshold account in which high MT and SE are near-universal among those who reach an ultramarathon start line, restricting their range and limiting their discriminative value within the competing population.

1. Introduction

Ultrarunning has grown over 345% in the last ten years alone, and with it a research literature attempting to explain what separates finishers from non-finishers across events lasting from 6 hours to multiple days [1]. Among the endurance community it is commonly said that completing an ultramarathon takes

“90% mind and 10% body,” yet the scientific literature testing that claim has developed unevenly: physiological determinants of ultra-endurance performance are comparatively well characterized, while the psychological determinants have only recently attracted quantitative attention [1, 2].

Mental toughness (MT) is the personal capacity to produce consistently high performance toward objective goals despite challenges, stressors, and significant adversities [19]. Self-efficacy (SE) refers to an individual's belief in their capability to execute the behaviors required to produce a specific attainment [9]. Both constructs have been repeatedly nominated as the psychological signature of the successful ultrarunner. Elite ultramarathon athletes score higher on MT than professional athletes from other sports, and MT and SE correlate strongly with one another in ultra samples ($r = .72$) [3]. Our previous narrative review concluded that successful ultramarathoners demonstrate higher levels of MT, SE, and emotional intelligence, coupled with lower mood variability [1].

However, the primary evidence is conflicting in an instructive way. In the most elite sample studied to date – competitors at the Hawaiian Ultra Running Team's Trail 100-mile endurance run – neither MT nor SE significantly predicted placing, finishing time, world-tour ranking, or race completion, despite the sample's exceptionally high absolute MT levels [3]. The authors proposed a threshold account: a criterion level of MT and SE is prerequisite to prepare for and reach the start line of an elite ultramarathon, and beyond that criterion, other factors govern performance [3]. By contrast, in broader field samples, MT correlates positively with objective performance indices ($r = .23$ with the ITRA Performance Index among 307 trail runners) [4], SE distinguishes finishers from withdrawers and doubles the odds of completion ($OR = 2.03$) [5], and finishers score significantly higher than withdrawers on MT in ultra-trail mountain races [6]. Whether these findings reflect one small underlying effect observed with varying precision, or genuinely different effects across contexts, cannot be resolved narratively.

A quantitative synthesis is therefore the natural next step, and none exists. Prior systematic reviews of ultrarunner psychology are narrative in method [1, 2]. Meta-analyses of MT and of SE in sport pool across all disciplines – the pre-event SE–performance association across sport generally is $r = .31$ [7, 12] – but no meta-analysis has pooled these constructs within ultramarathon running specifically, where the threshold account predicts systematic attenuation of exactly these effects. The present study fills this gap. We conducted a systematic review and meta-analysis of quantitative studies linking MT and SE to ultramarathon performance outcomes, with three aims: to estimate the pooled magnitude of the association; to compare it against the general-sport benchmark; and to examine whether outcome type (completion versus graded performance) or construct (MT versus SE) patterns the effect.

2. Methods

2.1 Design and reporting

This review was conducted and is reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement [18]. The protocol was not

prospectively registered; the absence of registration is disclosed as a limitation in § 4.1.

2.2 Search strategy

The openly indexed literature was searched electronically (PubMed/MEDLINE and Google Scholar–indexed sources, from inception through August 2026) using combinations of: (ultramarathon OR “ultra-marathon” OR “ultra-endurance” OR “ultra-trail” OR “trail running”) AND (“mental toughness” OR “self-efficacy” OR SMTQ OR MTQ OR MTI OR ESSES OR GSES) AND (performance OR finish* OR completion OR DNF OR time OR rank*). This search was supplemented by systematic backward citation chasing of the reference lists of two prior reviews of ultrarunner psychology – Roebuck et al. (51 primary studies) [2] and Thornton et al. [1] – and of every study meeting inclusion criteria. Subscription-only databases (Scopus, Web of Science, SPORTDiscus, PsycINFO) were not searched; the implications of this restriction are addressed in § 4.1.

2.3 Eligibility criteria

Studies were eligible if they (a) sampled adult ultramarathon runners, defined as competitors in footraces exceeding 42.195 km or 6 hours; (b) measured MT or SE with a validated instrument; (c) reported an ultramarathon performance outcome – finish time, placing or rank, a performance index, or completion versus non-finish; and (d) reported an effect size (r , β , d , OR) or statistics convertible to one (means and SDs, F , t , χ^2 , η^2). Studies were excluded if they were purely qualitative, sampled marathon-or-shorter distances or non-running ultra-endurance disciplines, reported no performance outcome, or presented no convertible statistics. Where multiple reports drew on the same sample, the most complete report was retained.

2.4 Study selection and data extraction

Titles and abstracts were screened against the eligibility criteria, followed by full-text assessment of remaining records. From each included study the following were extracted: sample size and composition, event characteristics, predictor and instrument, outcome and its operationalization, and all reported statistics linking predictor to outcome. Screening and extraction were performed by a single reviewer (O.R.T.); this is disclosed as a limitation in § 4.1.

2.5 Effect size computation

The Pearson correlation coefficient served as the common effect metric, coded such that positive values indicate that higher MT or SE is associated with better performance or higher completion likelihood. Between-group comparisons of finishers and non-finishers were converted using standard identities: from η^2 , $r = \sqrt{\eta^2}$ [5]; from an independent-samples t statistic, $r = \sqrt{t^2/(t^2 + df)}$ [6]; signs were assigned from reported group means. One study [3] reported only non-significant omnibus regression and χ^2 statistics without zero-order correlations or direction; following conservative meta-analytic convention for non-significant unreported effects, this study's effect was imputed as $r = 0$ and the influence of this imputation was examined in a pre-specified sensitivity analysis. Correlations were transformed to Fisher's z for pooling, with variance $1/(n - 3)$, and pooled estimates back-transformed to r .

2.6 Statistical analysis

A random-effects model with the DerSimonian–Laird estimator of between-study variance (τ^2) was fitted. Heterogeneity was quantified with Cochran's Q and I^2 , and a 95% prediction interval was computed for the distribution of true effects. Two supplementary analyses were conducted: a sensitivity analysis excluding the imputed null effect, and a subgroup estimate for completion outcomes. Construct (MT versus SE) contrasts are reported descriptively given the small k. Formal publication-bias tests (Egger's regression) are severely underpowered below $k = 10$ and were therefore not conducted; this is addressed as a limitation. Analyses were implemented in Python 3 (NumPy/SciPy), reproducing the DerSimonian–Laird estimator as implemented in the R package metafor.

2.7 Risk of bias

Risk of bias for each included study was appraised qualitatively using criteria drawn from the AXIS tool for cross-sectional studies (sampling frame, response, instrument validity, outcome objectivity, analysis adequacy), applied to the prospective design [5] with attention to cohort-specific concerns. Ratings are summarized in Table 2.

3. Results

3.1 Study selection

Electronic searching and citation chasing identified eleven records, of which one was a duplicate, leaving ten unique records screened at title and abstract (Fig. 1). Three were excluded at screening: one measured stress–recovery states rather than MT or SE, one sampled Olympic-distance triathletes rather than ultramarathon runners [11], and one was not a primary ultramarathon study. Of seven reports sought in full text, one could not be retrieved – a prospective analysis of 189 competitors at the 2014 Western States Endurance Run reporting a positive MT–performance association after controlling for demographic and training covariates [10]; it is classified as awaiting assessment and its potential influence is considered in § 4.1. Six reports were assessed for eligibility, of which two were excluded because they reported no MT/SE–performance association: a runner-versus-control comparison of resilience, personality, and emotion regulation with no performance outcome [21], and a descriptive survey of ultramarathoners' psychological attributes [22]. Four studies met all criteria and contributed effects to the meta-analysis.

3.2 Characteristics of included studies

Table 1 summarizes the four included studies (total $N = 940$). Samples ranged from elite 100-mile trail competitors to broad national trail-running populations; outcomes comprised graded performance indices (rank, time, ITRA index) and binary completion. MT was measured with versions of the Sports Mental Toughness Questionnaire in all MT studies; SE was measured with the Endurance Sport Self-

Efficacy Scale and a theory-of-planned-behavior scale. Every pooled effect derives from a statistic reported in the original article, via the conversions specified in § 2.5.

Table 1
Characteristics of included studies and effect size derivations.

Study	N	Sample / event	Predictor (instrument)	Outcome	Effect r (derivation)
Brace et al. (2020) [3]	56	HURT100, 100-mi trail (elite)	MT (SMTQ); SE (ESSES)	UTWT rank; placing; time; completion	.00 imputed (NS omnibus: $F(2,53) = 0.74-1.74$; $\chi^2=0.56$; direction unreported)
Corrion et al. (2018) [5]	221	140-km ultra-trail (France)	SE (TPB, 3-item)	Completion vs. DNF	.16 (from $F(1,219) = 9.17$, $\eta^2=.025$; finishers higher; OR = 2.03)
Méndez-Alonso et al. (2021) [6]	356	Ultra-trail mountain races (Spain)	MT	Completion vs. withdrawal	.22 (from $t(354) = 4.25$; finishers higher)
Gameiro et al. (2023) [4]	307	Portuguese trail runners (ITRA)	MT (SMTQ)	ITRA Performance Index	.23 (reported Pearson r)

3.3 Risk of bias

All four studies were rated at moderate risk of bias overall (Table 2). Recurring concerns were self-selected convenience samples, reliance on self-report instruments administered at a single pre-race time point, and, in one case [3], low internal consistency for two SMTQ subscales. Recurring strengths were objective, event-verified performance outcomes in all four studies – an uncommon asset in sport-psychology field research.

Table 2
Risk-of-bias summary (AXIS-informed qualitative appraisal).

Study	Principal strengths	Principal concerns	Overall
Brace et al. (2020)	Objective, event-verified outcomes; a priori power analysis; validated instruments; pre-race measurement	41% response rate of invited field; low subscale alphas (Confidence .52, Constancy .42); n = 33 for the time analysis	Moderate
Corrion et al. (2018)	Prospective design; outcome recorded independently of self-report; adequate N; validated TPB items ($\alpha = .75$)	Single race (first edition); 12% female; brief 3-item SE measure; self-selection of respondents	Moderate
Méndez-Alonso et al. (2021)	Largest N; pre- and post-race measurement; objective completion outcome	Effect derived from t statistic rather than reported r; convenience recruitment; instrument translation details limited in report	Moderate
Gameiro et al. (2023)	Objective normative outcome (ITRA index); validated Portuguese SMTQ and CD-RISC-10; SEM measurement model tested	Cross-sectional; online convenience sample; performance index aggregates 36 months of results temporally distant from MT measurement	Moderate

3.4 Pooled analysis

Across the four studies, the random-effects pooled association between MT/SE and ultramarathon performance was $r = .20$ (95% CI .13 to .26; $z = 6.02$, $p < .001$). Heterogeneity was negligible ($Q(3) = 3.03$, $p = .39$; $I^2 = 1\%$; $\tau^2 < .001$), and the 95% prediction interval (.05 to .33) excluded zero, indicating that the true effect in a comparable future study would be expected to remain positive. Figure 2 presents the forest plot.

The sensitivity analysis excluding the imputed null effect [3] ($k = 3$, $N = 884$) produced a nearly identical pooled estimate, $r = .21$ (95% CI .14 to .27; $Q(2) = 0.80$, $p = .67$; $I^2 = 0\%$), indicating that the conservative imputation does not drive the result. The completion-outcome subgroup — SE [5] and MT [6] predicting finishing versus withdrawal ($k = 2$, $N = 577$) — pooled at $r = .20$ (95% CI .12 to .27), closely matching the overall estimate.

3.5 Construct and outcome patterns

At $k = 4$, construct contrasts are descriptive only. The three non-imputed effects were closely clustered ($r = .16$ to .23) across both constructs and both outcome types, and the single null arose in the only exclusively elite sample [3] — the pattern the threshold account predicts, though the present k cannot test elitism as a moderator inferentially.

4. Discussion

This meta-analysis provides the first pooled estimate of the association between mental toughness, self-efficacy, and ultramarathon performance: small, positive, and strikingly consistent ($r = .20$; $I^2 = 1\%$). Three features of this result deserve emphasis.

First, the effect is roughly two-thirds the magnitude of the pre-event self-efficacy–performance association across sport generally ($r = .31$) [7], and half the elite-athlete subgroup estimate from that literature ($r = .40$) [7]. This attenuation is precisely what the threshold account predicts [3]. Ultramarathon populations are self-selected for extreme MT and SE — elite ultrarunners out-score professional athletes from other sports on MT [3], and ultra-trail samples out-score both other athletes and sedentary populations on MT and resilience [6]. Range restriction in the predictor mechanically attenuates observable correlations; substantively, once the criterion level of MT and SE required to train for and start an ultramarathon is met, residual variation in these traits buys comparatively little additional performance. The pooled estimate quantifies, for the first time, how much it still buys: a small but reliable amount.

Second, the null result in Brace et al. [3] is not an anomaly to be explained away but a data point the model accommodates. The homogeneity of effects ($I^2 = 1\%$) indicates the elite-sample null and the field-sample positives are statistically compatible with a single small underlying effect estimated with different precision — although the descriptive gradient across sample elitism remains suggestive and constitutes the single most important question for future primary research: does the MT/SE–performance association shrink monotonically as samples become more elite?

Third, the practical implication is calibrated rather than dismissive. A correlation of .20 for completion-type outcomes corresponds to a meaningful shift in finish probability across the SE distribution — consistent with the doubling of completion odds per SE unit reported by Corrion et al. [5] — and MT is trainable, with intervention meta-analytic support across sport [13]. Psychological skills training targeting race-specific self-efficacy and MT is therefore a defensible, evidence-consistent component of ultramarathon preparation, particularly for completion-focused (rather than podium-focused) athletes. What the present results counsel against is the stronger community claim that the mind dominates ultra performance outright: among those who make the start line, the measured psychological contribution is real but modest.

4.1 Limitations

It is important to note the constraints on these conclusions. The evidence base is small ($k = 4$), and although observed heterogeneity was negligible, heterogeneity estimates at this k are imprecise. The search was restricted to openly indexed literature supplemented by citation chasing; subscription-only databases were not searched, and one known eligible-appearing study [10] remains awaiting assessment — notably, its reported direction (positive) matches the pooled estimate, making directional bias from its absence unlikely, though its inclusion would sharpen precision. The protocol was not prospectively registered, and screening and extraction were single-reviewer; both depart from best

practice and temper confidence in the review process. Predictors (MT, SE) and outcomes (graded performance, completion) were pooled across conceptually distinct pairings, although the near-identical subgroup and overall estimates mitigate this concern empirically. One effect was conservatively imputed as null from non-significant omnibus statistics; the sensitivity analysis indicates the pooled estimate is robust to this choice. All included designs are observational, precluding causal inference; self-report instruments carry shared-method concerns. Finally, publication bias could not be formally assessed at this k, and a literature this small is plausibly vulnerable to it.

4.2 Future research

Proposed questions include: Does the MT/SE–performance association attenuate monotonically with sample elitism, as the threshold account predicts? Do prospective designs measuring SE race-specifically (rather than dispositionally) recover larger completion effects? And can within-person designs — tracking MT across a season of races [6] — separate trait selection from race-induced training of these constructs? Primary studies would serve future syntheses greatly by routinely reporting zero-order correlations between psychological predictors and all performance outcomes, significant or not.

5. Conclusion

Pooling the available quantitative evidence, mental toughness and self-efficacy show a small, positive, and homogeneous association with ultramarathon performance and completion ($r = .20$), reliably weaker than the same constructs' association with performance in sport generally. Conclusively, the psychology of the ultramarathoner appears to operate chiefly as a gate rather than a gradient: extraordinary MT and SE get an athlete to the start line, and beyond that threshold they confer a real but modest additional advantage. Quantifying that threshold — and what governs performance beyond it — is the field's next task.

Declarations

Conflicts of interest:

The author declares no conflicts of interest.

Ethics

Not applicable; this study is a synthesis of previously published, ethically approved research and involved no new data collection from human participants.

Funding:

The author received no specific funding for this work.

Data availability:

All effect sizes, conversion formulas, and analysis code are available from the author on request; all included studies are open access.

References

1. Thornton OR, Ly S, Colón I, Cole H, Li W (2023) The psychological indicators of success in ultrarunning – a review of the current psychological predictors in ultrarunning. *Ann Med Health Sci Res.* ;13(7)
2. Roebuck GS, Fitzgerald PB, Urquhart DM, Ng SK, Cicuttini FM, Fitzgibbon BM (2018) The psychology of ultra-marathon runners: a systematic review. *Psychol Sport Exerc* 37:43–58
3. Brace AW, George K, Lovell GP (2020) Mental toughness and self-efficacy of elite ultra-marathon runners. *PLoS ONE* 15(11):e0241284
4. Gameiro N, Rodrigues F, Antunes R, Matos R, Amaro N, Jacinto M, Monteiro D (2023) Mental toughness and resilience in trail runner's performance. *Percept Mot Skills* 130(3):1202–1220
5. Corrion K, Morales V, Bergamaschi A, Massiera B, Morin JB, d'Arripe-Longueville F (2018) Psychosocial factors as predictors of dropout in ultra-trailers. *PLoS ONE* 13(11):e0206498
6. Méndez-Alonso D, Prieto-Saborit JA, Bahamonde JR, Jiménez-Arberás E (2021) Influence of psychological factors on the success of the ultra-trail runner. *Int J Environ Res Public Health* 18(5):2704
7. Lochbaum M, Sisneros C, Cooper S, Terry PC (2023) Pre-event self-efficacy and sports performance: a systematic review with meta-analysis. *Sports* 11(11):222
8. Sheard M, Golby J, van Wersch A (2009) Progress toward construct validation of the Sports Mental Toughness Questionnaire (SMTQ). *Eur J Psychol Assess* 25(3):186–193
9. Bandura A (1997) *Self-Efficacy: The Exercise of Control*. Freeman, New York
10. Christensen DA, Brewer BW, Hutchinson JC (2018) Psychological predictors of performance in a 161 km ultramarathon run. *Int J Sport Psychol* 49:74–90
11. Anstiss PA, Meijen C, Madigan DJ, Marcora SM (2018) Development and initial validation of the Endurance Sport Self-Efficacy Scale (ESSES). *Psychol Sport Exerc* 38:176–183
12. Moritz SE, Feltz DL, Fahrbach KR, Mack DE (2000) The relation of self-efficacy measures to sport performance: a meta-analytic review. *Res Q Exerc Sport* 71:280–294
13. Stamatis A, Grandjean P, Morgan G, Padgett RN, Cowden R, Koutakis P (2020) Developing and training mental toughness in sport: a systematic review and meta-analysis. *BMJ Open Sport Exerc Med* 6(1):e000747
14. Zeiger JS, Zeiger RS (2018) Mental toughness latent profiles in endurance athletes. *PLoS ONE* 13(2):e0193071

15. Rochat N, Hauw D, Antonini Philippe R, von Crettaz F, Seifert L (2017) Comparison of vitality states of finishers and withdrawers in trail running. *PLoS ONE* 12(3):e0173667
16. Hoffman MD, Ong JC, Wang G (2010) Historical analysis of participation in 161 km ultramarathons in North America. *Int J Hist Sport* 27:1877–1891
17. Connor KM, Davidson JRT (2003) Development of a new resilience scale: the Connor–Davidson Resilience Scale (CD-RISC). *Depress Anxiety* 18(2):76–82
18. Page MJ, McKenzie JE, Bossuyt PM et al (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 372:n71
19. Gucciardi DF, Hanton S, Gordon S, Mallett CJ, Temby P (2015) The concept of mental toughness: tests of dimensionality, nomological network, and traitness. *J Pers* 83(1):26–44
20. McCormick A, Meijen C, Marcora S (2015) Psychological determinants of whole-body endurance performance. *Sports Med* 45(7):997–1015
21. Roebuck GS, Urquhart DM, Che X, Knox L, Fitzgerald PB, Cicuttini FM, Lee S, Segrave R, Fitzgibbon BM (2020) Psychological characteristics associated with ultra-marathon running: an exploratory self-report and psychophysiological study. *Aust J Psychol* 72(3):235–247
22. Buck K, Spittler J, Reed A, Khodaei M (2018) Psychological attributes of ultramarathoners. *Wilderness Environ Med* 29:66–71

Figures

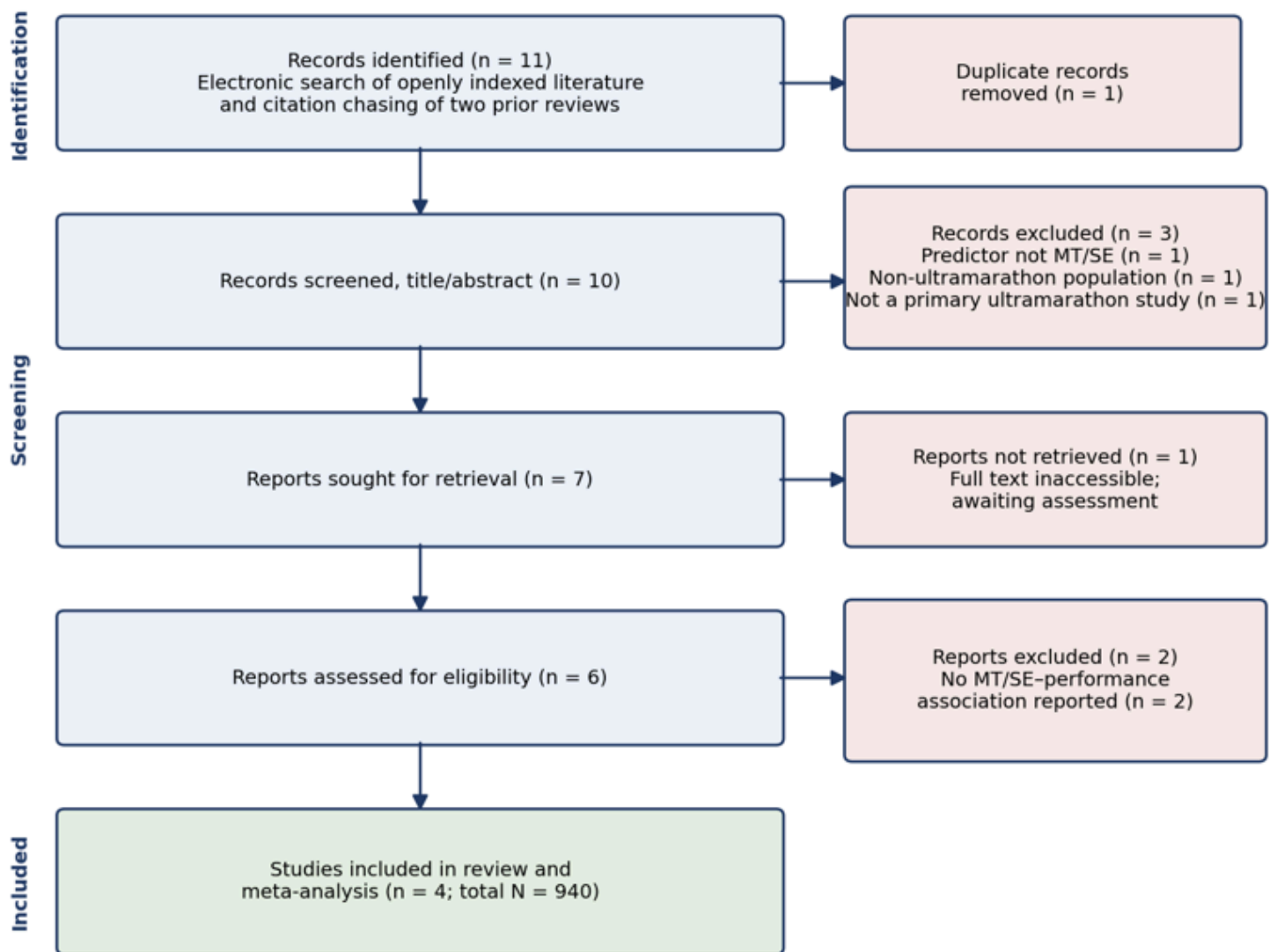


Figure 1

PRISMA 2020 flow diagram of study identification, screening, and inclusion.

MT/SE and Ultramarathon Performance: Random-Effects Meta-Analysis

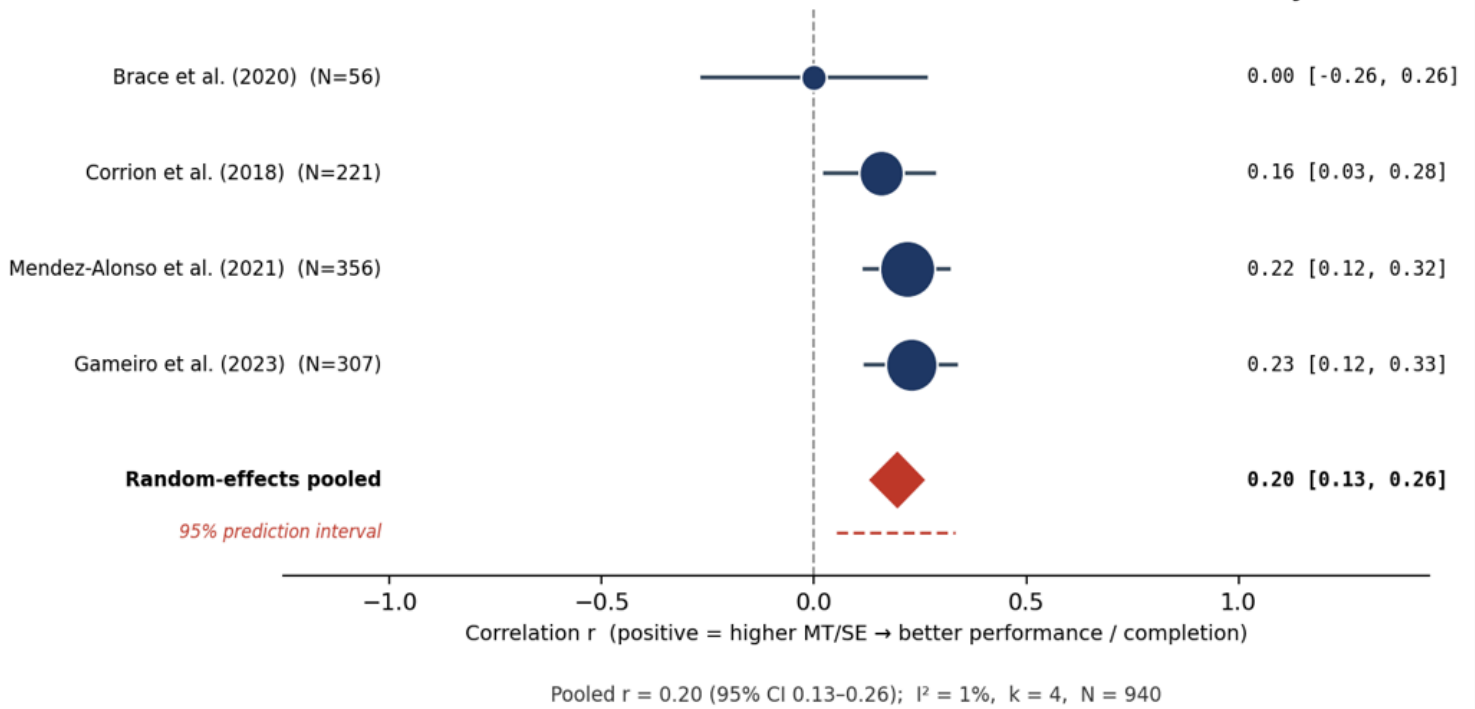


Figure 2

Forest plot of the random-effects meta-analysis. Marker size is proportional to random-effects weight; the diamond represents the pooled estimate and the dashed whisker the 95% prediction interval.