Throwing cold water on muscle growth: A systematic review with meta-analysis of the effects of post-exercise cold water immersion on resistance training-induced hypertrophy

For correspondence:
brad.schoenfeld@lehman.cuny.edu

Alec Piñero[1], Ryan Burke[1], Francesca Augustin[1], Adam E. Mohan[1], Kareen DeJesus[1], Max Sapuppo[1], Max Weisenthal[1], Max Coleman[1], Patroklos Androulakis-Korakakis[1], Jozo Grgic[2], Paul A. Swinton[3], Brad J. Schoenfeld[1]

[1] Department of Exercise Science and Recreation, Applied Muscle Development Laboratory, CUNY Lehman College, Bronx, NY
[2] Institute for Health and Sport, Victoria University, Melbourne, VIC 3011, Australia.
[3] School of Health Sciences, Robert Gordon University, Aberdeen, UK

Please cite as: Piñero, A., Burke, R., Augustin, F., Mohan, A.E., Sapuppo, M., Weisenthal, M., Coleman, M., Androulakis-Korakakis, P., Grgic, J., Swinton, P.A., Schoenfeld, B.J. (2023). Throwing cold water on muscle growth: A systematic review with meta-analysis of the effects of post-exercise cold water immersion on resistance training-induced hypertrophy. *SportRχiv*.

All authors have read and approved this version of the manuscript. This article was last modified May, 2023.

ABSTRACT

The purpose of this paper was to systematically review the literature and perform a meta-analysis of the existing data on the effects of post-exercise cooling coupled with resistance training (RT) on gains in measures of muscle growth. To locate relevant studies for the topic, we comprehensively searched the PubMed/MEDLINE, Scopus, and Web of Science databases. A total of 8 studies met inclusion criteria; all employed cold water immersion (CWI) as the means of cold application. Preliminary analyses conducted on non-controlled effect sizes provided strong evidence of hypertrophic adaptations with RT that were likely to be at least small in magnitude ($SMD_{0.5}$ = 0.36 [95%CrI: 0.10 to 0.61]; *p*(>0) = 0.995, *p*(>0.1) = 0.977). In contrast, non-controlled effect sizes provided some evidence of hypertrophic adaptations with CWI + RT that were likely to be between small and zero in magnitude ($SMD_{0.5}$ = 0.14 [95%CrI: -0.08 to 0.36]; *p*(>0) = 0.906, *p*(>0.1) = 0.68). The primary analysis conducted on comparative effect sizes provided some evidence of greater relative hypertrophic adaptations with RT compared to CWI + RT ($cSMD_{0.5}$ = -0.22 [95%CrI: -0.47 to 0.04]), with differences likely to be greater than zero (*p*(<0) = 0.957) and of at least a small magnitude of effect (*p*(< -0.1) = 0.834). Meta-regression did not indicate a potential moderation effect of training status ($\beta_{\mathrm{Trained:Untrained}_{0.5}}$= -0.10 [95%CrI: -0.65 to 0.43] $p < 0)$=0.653). In conclusion, the current data suggest that the application of CWI immediately following bouts of RT may attenuate hypertrophic changes.

KEYWORDS: recovery strategies; cold application; cooling; cross-sectional area; fat-free mass; lean mass

INTRODUCTION

Practitioners employ a variety of strategies to attenuate the fatigue and discomfort of resistance training (RT) or physical competition, or in an attempt to improve performance-defined measures of recovery. Of these strategies, extreme temperature exposure in various forms (e.g. sauna, cold and hot water immersion, cryotherapy, phase-change material) have been found to reduce the severity of muscle soreness [1], perceived fatigue [2] [3], and time to recovery [4]. Notably, despite the target of cold therapy often cited as being a reduction of acute post-exercise inflammation [5] [6], there is some evidence that cold therapy may not actually reduce biological markers of inflammation [7] nor improve recovery from eccentric exercise-induced muscle damage [8]. Of the various cold therapy strategies, cold water immersion (CWI), generally practiced by immersing the torso and limbs or individual limbs in water of <15°C for 10-20 minutes following an exercise bout [9], has been found to improve recovery for certain types of subsequent athletic or training performance [3] [10] [9], though perhaps not all [3].

Although athletes commonly use CWI to enhance recovery and acutely improve exercise performance [10] [11], the same might not hold true for chronic adaptations to RT. While there seem to be little or no negative effects of post-exercise CWI on endurance training adaptations [12], CWI following RT may result in muted strength-related adaptations, such as absolute strength and muscular power, according to a series of recent narrative and meta-analytical reviews [13] [14] [15] [16] [9].

To date, no meta-analysis has been conducted on the effect of post-exercise cold application on RT-induced muscle hypertrophy. However, there are various mechanistic reasons that suggest CWI may have detrimental effects on longitudinal skeletal muscle accretion. Most notably, CWI has been found to acutely attenuate post-RT mechanistic target of rapamycin complex 1 signaling [17], ribosome biogenesis [18], muscle protein synthesis (MPS) [19], satellite cell activity [20], and increases in circulating testosterone and cytokines [21] –

responses which may, to varying degrees, negatively impact muscular adaptations [22] [23] [24] [25] [26]. The purpose of this paper was to systematically review the literature and perform a meta-analysis of the existing data on the effects of post-exercise cooling coupled with RT on gains in measures of muscle growth.

METHOD

We conducted this review in accordance with the guidelines of the "Preferred Reporting Items for Systematic Reviews and Meta-Analyses" (PRISMA) [27]. The study was preregistered on the Open Science Framework (https://osf.io/gx69b).

Search strategy

To locate relevant studies for the topic, we comprehensively searched the PubMed/MEDLINE, Scopus, and Web of Science databases using the following Boolean search syntax: ("cold water immersion" OR "CWI" OR "cryotherapy" OR "cryo" OR "cryostimulation" OR "cryochamber" OR "ice bath*" OR "ice-bath*" OR "ice water bath" OR "ice-water bath" OR "cold exposure" OR "cold application" OR "cold plunge" OR "cold stress" OR "cold treatment" OR "post-exercise cooling" OR "post exercise cooling" OR "cooling therap*" OR "contrast-water therapy" OR "contrast water therapy") AND ("resistance training" OR "resistance exercise" OR "weight lifting" OR "weightlifting" OR "strength exercise" OR "strength training" OR "strengthening" OR "resistive exercise" OR "resistive training") AND ("muscle hypertrophy" OR "muscular hypertrophy" OR "muscle growth" OR "muscular growth" OR "muscle mass" OR "muscle development" OR "muscular development" OR "muscle volume" OR "lean body mass" OR "fat-free mass" OR "fat free mass" OR "lean mass" OR "muscle fiber" OR "muscle size" OR "muscular size" OR "myofiber" OR "myofibre" OR "muscle fibre" OR "muscle thickness" OR "cross-sectional area" OR "cross sectional area"). In addition, we screened the reference lists of articles retrieved and applicable review papers to uncover any additional studies that might meet inclusion criteria, as per Greenhalgh and Peacock [28].

The search process was carried out separately by 3 researchers (AP, MW, and KD). The initial search consisted of screening all titles and abstracts for studies potentially meeting inclusion/exclusion criteria. For papers deemed potentially relevant, full texts were evaluated and decisions were then made as to whether a given study warranted inclusion. Any disputes the search team could not resolve were settled by a fourth researcher (BJS). The search was finalized on February 24th, 2023.

Inclusion criteria

We included studies that satisfied the following criteria: (a) had a randomized design (either within- or between-group design) and directly compared cold application + RT vs. RT with a sham or active/passive recovery (both with and without adjuvant dietary interventions) for estimates of changes in lean/muscle mass using a validated measure (dual-energy X-ray absorptiometry [DXA], bioelectrical impedance analysis, magnetic resonance imaging [MRI], computerized tomography [CT], ultrasound [US], muscle biopsy or limb circumference measurement) in healthy adults; (b) involved at least 2 RT sessions per week for a duration of at least 4 weeks[1] (NOTE: in our preregistration we indicated a minimum duration of 6 weeks, but after perusing the data we decided to accept studies with a duration of at least 4 weeks and then subanalyze by study length); (c) published in a peer-reviewed English language journal or on a pre-print server. We excluded studies that utilized participants with co-morbidities that might impair muscle hypertrophy responses (musculoskeletal disease/injury/cardiovascular impairments).

Data extraction

For each included study, 2 researchers (RB and AM) independently extracted and coded the following data: Author name(s), title and year of publication, sample size, participant characteristics (i.e. sex, training status, age), description of the training intervention (i.e.

[1] NOTE: In our preregistration we indicated a minimum duration of 6 weeks, but after perusing the data we decided to accept studies with a duration of at least 4 weeks given evidence of appreciable hypertrophy at this timepoint [67] and then subanalyze data by study length.

duration, volume, load, frequency, proximity to failure, body region), method for hypertrophy assessment (i.e., DXA, MRI, CT, US, biopsy, circumference measurement), and mean pre- and post-study lean/muscle mass values with corresponding standard deviations. In cases where measures of changes in lean/muscle mass were not reported, we attempted to contact the corresponding author(s) to obtain the data. If unattainable, we extracted the data from graphs (when available) via online software (https://automeris.io/WebPlotDigitizer/). To account for the possibility of coder drift, a third researcher (FA) recoded 30% of the studies, which were randomly selected for assessment [29]. Per case agreement was determined by dividing the number of variables coded the same by the total number of variables. Acceptance required a mean agreement of 0.90. Any discrepancies in the extracted data were resolved through discussion and mutual consensus of the coders.

Methodological quality

As noted in the preregistration, we originally planned to use the Downs and Black assessment tool [30] to assess study quality. However, after discussion with the research team, it was determined that the proposed tool was too generic to properly evaluate the complexities of longitudinal RT research. Thus, we developed an alternative tool specifically designed to assess the quality (both in terms of risk of bias as well as transparency of reporting) of longitudinal RT interventions. We named the tool: Standards Method for Assessment of Resistance Training in Longitudinal Designs (SMART-LD).

The SMART-LD tool consists of 20 questions that address the following aspects of a study's methodology: general (items 1-2); participants (items 3-7), training program (items 8-11), outcomes (items 12-16), and statistical analyses (17-20). Each item in the checklist is given 1 point if the criterion is satisfied or 0 points if the criterion is not satisfied. The values of all questions are summed, with the final total used to classify studies as follows: "good quality" (16-20 points); "fair quality" (12-15 points); or "poor quality" (≤ 11 points).

As per Resnick et al. [31], we established content validity of the tool by initially creating a list of items that addressed the primary aspects of repeated measures RT protocols (participants, program, outcomes, statistics). Our team of 5 experienced researchers reached a consensus on the content and wording of the items included in the tool. We then sought input from colleagues who provided additional feedback on areas of relevance and ambiguity. After addressing the input from our colleagues, we sent the tool to 4 independent researchers, all experienced with carrying out longitudinal RT trials, to rate the relevance of the items on a scale of 1 (not very relevant) to 4 (very relevant). The mean relative rating for all items between the 4 raters was 3.56 (89%) and no item was rated "not very relevant", indicating the tool has high relevance for evaluating the quality of longitudinal RT designs. An overview of the items included in the SMART-LD tool and explanation of the grading criteria for each item can be found at: https://osf.io/nhva2/.

We thus employed the SMART-LD tool as the primary quality assessment of studies included in this meta-analysis. Four reviewers (AP, MC, RB, and PAK) independently rated each study; any disputes were resolved by majority consensus. Given our initial intention to employ the Downs and Black checklist as noted in the preregistration, we also carried out a quality assessment of studies using this tool. Three reviewers (AP, MS, and FA) independently rated each study; any disputes were resolved by majority consensus.

Statistical analyses

A Bayesian framework was chosen over a frequentist approach as it can provide more flexible modeling enabling results to be presented intuitively by reporting subjective probabilities [32]. As muscle growth was measured using different methods in the included studies, the primary analysis was conducted using comparative standardized mean difference effect sizes (cSMD) calculated from direct comparisons between CWI + RT and RT. To provide additional context, preliminary analyses were conducted with non-controlled standardized mean difference effect sizes (SMD) to determine whether the CWI + RT and RT programs tended to result in hypertrophic adaptations. Three-level random-effects Bayesian hierarchical models were used to pool effect sizes and model the mean comparative effect, variance within studies, variance between studies, and covariance of multiple outcomes reported in the same study (i.e., multiple outcomes and/or single outcome reported at multiple time points following baseline). Within-study variances were calculated using standard distributional assumptions [33] [34] with adjustment for cross-over designs where required [35]. Within-study variances are dependent on pre-post correlations [33] that are generally not reported. Rather than specify a single correlation value, this was estimated but constrained using an informative prior distribution. Similarly, informative prior distributions were used for the comparative effect sizes based on previous meta-analysis data [36]. Sensitivity analyses were conducted using weakly-informative prior distributions and presented in supplementary file (see supplemental file S1).

Inconsistency in models was described by comparing variances across the three levels. Inferences from all analyses were performed on posterior samples generated using the Hamiltonian Markov Chain Monte Carlo method and through credible intervals (CrI) and calculated probabilities (*p*). Interpretations were based on medians (e.g. $cSMD_{0.5}$), range of values within CrIs, and calculation of probabilities that the magnitude of the pooled mean effect size exceeded qualitative thresholds (i.e. small, medium, and large) specific to strength and conditioning interventions [37] [36]. For non-controlled effect sizes the small, medium, and large thresholds selected were +0.10, +0.35, and +0.70 [37], with values of ±0.10, ±0.30, and ±0.50 used for comparative effect sizes [34]. Meta-regression or subgroup analyses were performed where sufficient data were available including a minimum of 4 data points per category level or 10 data points for continuous variables [38]. Small-study effects (publication bias, etc.) were visually inspected with funnel plots and quantified with a multi-level extension of Egger's regression-intercept test [39]. Analyses were performed using the R wrapper package *brms* interfaced with *Stan* to perform the sampling [40]. Full model details including prior distributions for all meta-analyses are presented in the supplementary file S1, with summary descriptions presented in text.

Results

Descriptive data

A total of 8 interventions met inclusion criteria (see supplemental file S2); all employed CWI as the means of cold application (see Table 1). The duration of the studies ranged from 4 to 12 weeks. All studies included young adults (aged 20–26 years). Seven studies included only males [41] [42] [20] [17] [43] [44] [45], and 1 study included both males and females [46]. Four studies employed resistance-trained participants [46] [41] [42] [20] and the others employed untrained participants [45] [44] [43] [17]. Six studies incorporated a parallel group design [45] [44] [43] [41] [20] [17] and the other 2 employed a within-subject crossover design [46] [42]. All the RT sessions were performed 2-3 times per week. Two studies solely focused on training handgrip [40] [41], 1 study solely focused on training the wrist flexors [45], 3 studies trained just the lower body [46] [41] [20] , and the other 2 employed full body training protocols [42] [17]. Only 2 studies reported intensity of effort, with 1 reporting that participants trained to failure [46] and 1 reporting participants stopping shy of failure [42] and none of the other studies reporting proximity to failure. Three studies reported that training was directly supervised [46] [42] [20] while the other 5 did not report whether training was supervised or unsupervised [45] [44] [43] [41] [17]. Three studies exposed only upper limbs to CWI [45] [44] [43], 2 studies exposed only lower limbs to CWI [41] [20], and 3 studies exposed participants to full-body CWI [46] [42] [17]. Three studies applied CWI for 10 minutes [46] [41] [20], 2 studies applied CWI for 15 minutes [42] [17], and 3 applied CWI for 20 minutes [45] [44] [43]. Water temperature was 10°C for 6 of the studies [45] [44] [43] [41] [20] [17], 15°C for 1 of the studies [42], and between 14-15°C for the last [46]. CWI was administered 3 minutes post-RT in 2 studies [45] [43], 5 minutes post-RT in 2 studies [20] [17], and 15 minutes post-RT in 2 studies [41] [42], while 2 studies only specified that CWI followed the last set of exercise [44] and immediately followed each training session [46]. Three studies reported total lean body mass via DXA [41] [42] [17], with only 1 reporting distinct upper body and lower body measurements [17]. For site-specific measures of hypertrophy 2 studies used biopsy of the vastus lateralis to analyze type I and type II muscle fibers [20] [17], 1 study used US of the vastus medialis [46], 1 study measured leg circumference [46], 1 study used MRI of the quadriceps [20], 1 study used US of the forearm [43], 1 study used US for wrist flexors [45], and 2 studies measured forearm circumference [45] [44].

Table 1. Summary of the methods of included studies.

Study	Sample	Design	RT Protocol	CWI Protocol	Hypertrophy Measure	Duration
Fyfe et al. (2019)	8 untrained men	Random assignment to 1 of 2 groups: (1) RT + CWI; (2) RT + PASS	TB protocol performed 3 d/wk consisting of 3–5 sets per exercise at either 20RM or 12RM	TB CWI for 15 minutes at 10ºC beginning 5 minutes post-RT	- DXA: UB, LB, TLM - Biopsy: VL	7 wks
Horgan et al. (2023)	17 trained men	Random order assignment to 1 of 3 conditions with a crossover design: (1) RT + CWI; (2) RT + SS	TB supervised protocol performed 2 d/wk consisting of 2-4 sets per exercise at 66-87% 1RM or body weight with 1 set completed every 4 minutes; core exercises consisted of 3 sets with 15s inter-set rest intervals	TB CWI for 15 minutes at 15ºC beginning 15 minutes post-RT. SS group performed static stretching at 23°C	- DXA: TLM	4 wks
Ohnishi et al. (2004)	16 untrained men	Random assignment to 1 of 2 groups: (1) RT + CWI; (2) RT + PASS	Unilateral handgrip protocol performed 3 d/wk consisting of 3 sets at 8RM	UL CWI for 20 minutes at 10ºC following RT	- CIR: forearm	6 wks
Poppendieck et al. (2021)	11 trained men and women (M = 9; W = 2)	Random order assignment to 1 of 2 conditions with a crossover design: (1) RT + CWI; (2) RT + PASS	LB supervised protocol performed 2 d/wk consisting of 3 sets per exercise at 10RM with 3min inter-set rest intervals	TB CWI for 10 minutes at 14-15ºC immediately following RT	- US: VM - CIR: leg	8 wks
Roberts et al. (2015)	21 trained men	Random assignment to 1 of 2	LB supervised protocol performed 2 d/wk	LL CWI for 10 minutes at 10ºC beginning 5	- biopsy: VL - MRI: QUAD	12 wks

		groups: (1) RT + CWI; (2) RT + ACT	consisting of 3–6 sets per exercise at 8-12RM for machine exercises, 20%+ PTBM for WWL, body weight for CDJ, 50% of lunge load for SESJ, SLJ, and CBJ with 1min inter-set rest intervals and 3min inter-exercise rest intervals	minutes post-RT. ACT group performed 10 minutes low intensity stationary cycling		
Wilson et al. (2021)	13 trained men	Random assignment to 1 of 2 groups: (1) RT + CWI; (2) RT + SI	LB protocol performed 2 d/wk consisting of 2 blocks. - Strength block consisted of 3–4 sets per exercise at 4-6RM except for JS (2 reps @load eliciting peak power) - Power block consisted of 5 sets at: final strength block weight (1/4 squat for 2 reps and JS for 3 reps), body weight (box jumps for 3 reps), same weight as JS (squat jumps for 3 reps), 30-cm box (drop jumps for 3 reps)	LL CWI for 10 minutes at 10ºC beginning 15 minutes post-RT. SI group were told they received extra leucine supplementation	- DXA: TLM	8 wks
Yamane et al. (2006)	16 untrained men	Random assignment to 1 of 2 groups: (1) RT + CWI; (2) RT + PASS	Bilateral handgrip protocol performed 3 d/wk consisting of 3 sets of handgrip ergometer exercise at 8RM (70-80% 1RM) with 2min inter-set rest intervals	UL CWI for 20 minutes at 10ºC beginning 3 minutes post-RT	US: forearm	4 wks
Yamane et al. (2015)	14 untrained men	Random assignment to 1 of 2 groups: (1) RT + CWI; (2) RT + PASS	Unilateral WF protocol performed 3 d/wk consisting of 5 sets of 8 reps of wrist flexion ergometer	UL CWI for 20 minutes at 10ºC beginning 3 minutes post-RT	- US: WF - CIR: forearm	6 wks

	exercise 70-80% 1RM with 2min inter-set rest intervals	

PASS = passive recovery; SS = static stretching; ACT = active recovery; SI = sham intervention; TB = total body; LB = lower body; PTBM = pre-training body mass; WWL = weighted walking lunges; CDJ = countermovement drop jumps; SESJ = slow eccentric squat jumps; SLJ = split lunge lumps; CBJ = countermovement box jumps; JS = jump shrugs; WF = wrist flexor; UL = upper limb; LL = lower limbs; DXA: dual-energy x-ray absorptiometry; UB = upper body; TLM = total lean mass; VL = vastus lateralis; CIR = circumference; US = ultrasound; VM = vastus medialis; MRI = magnetic resonance imaging; QUAD = quadriceps

Meta-analyses

Preliminary analyses conducted on non-controlled effect sizes provided strong evidence of hypertrophic adaptations with RT that were likely to be at least small in magnitude ($SMD_{0.5}$ = 0.36 [95%CrI: 0.10 to 0.61]; *p*(>0) = 0.995, *p*(>0.1) = 0.977); Figure 1). In contrast, non-controlled effect sizes provided some evidence of hypertrophic adaptations with CWI + RT that were likely to between small and zero in magnitude ($SMD_{0.5}$ = 0.14 [95%CrI: -0.08 to 0.36]; *p*(>0) = 0.906, *p*(>0.1) = 0.68); Figure 1). Full model details including information of prior distributions are presented in supplementary file S1 (Table S1).

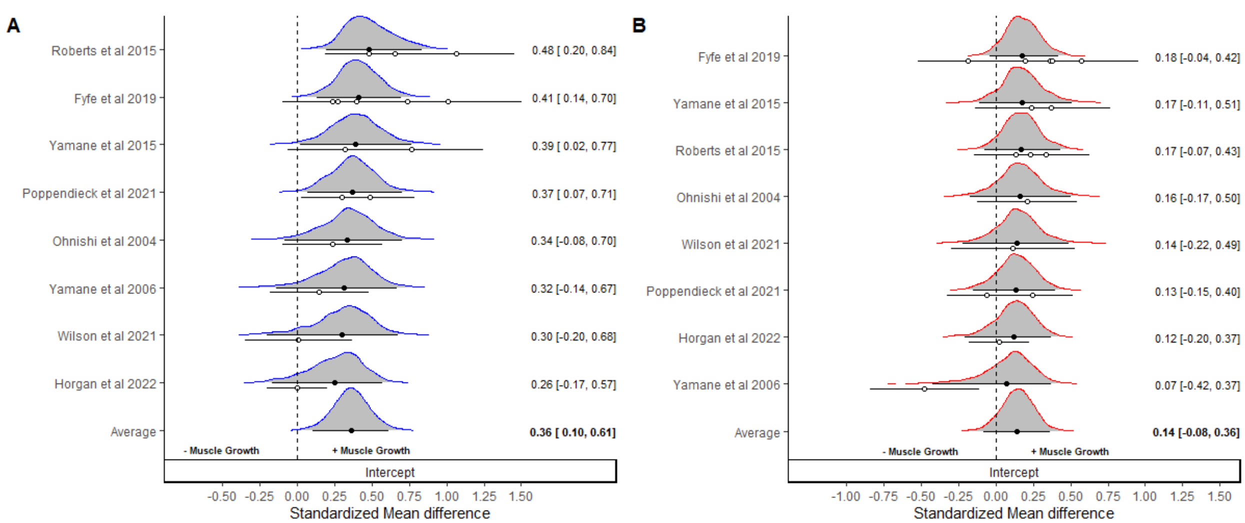

Figure 1. Bayesian forest plots illustrating pooling of non-controlled standardized mean difference effect sizes for resistance training (A) and with cold water immersion with resistance training (B).

Positive values indicate muscle growth and negative values indicate a reduction in muscle following intervention. Distributions represent "shrunken estimates" based on all effects sizes included, the random effects model fitted and borrowed information across studies to reduce uncertainty. Black circles and connected intervals represent the median value and 95% credible intervals for the shrunken estimates. White circles and intervals represent the raw estimates and sampling variance calculated directly from study data. Bottom distributions illustrate uncertainty in the pooled means.

The primary analysis conducted on comparative effect sizes (Figure 2) provided some evidence of greater relative hypertrophic adaptations with RT compared to CWI + RT ($cSMD_{0.5}$ = -0.22 [95%CrI: -0.47 to 0.04]), with differences likely to be greater than zero ($p(<0) = 0.957$) and of at least a small effect ($p(< -0.1) = 0.834$). Full model details including information of prior distributions are presented in supplementary file S1 (Table S2).

Two meta-regressions were conducted to investigate the potential moderation effect of intervention duration (shorter: <8 weeks; longer: ≥8 weeks) and training status. Limited evidence of a moderation effect was obtained for both factors with the CrIs demonstrating high uncertainty (Shorter:Longer 0.5= -0.04 [95%CrI: -0.61 to 0.55], p(<0)=0.570; Trained:Untrained 0.5= -0.10 [95%CrI: -0.65 to 0.43] p(<0) = 0.653). Full model details including information of prior distributions are presented in supplementary file S1 (Table S3).

Egger's regression intercept test produced wide intervals and a visual inspection of the funnel plot (Figure 3) did not identify any small-study-related issues.

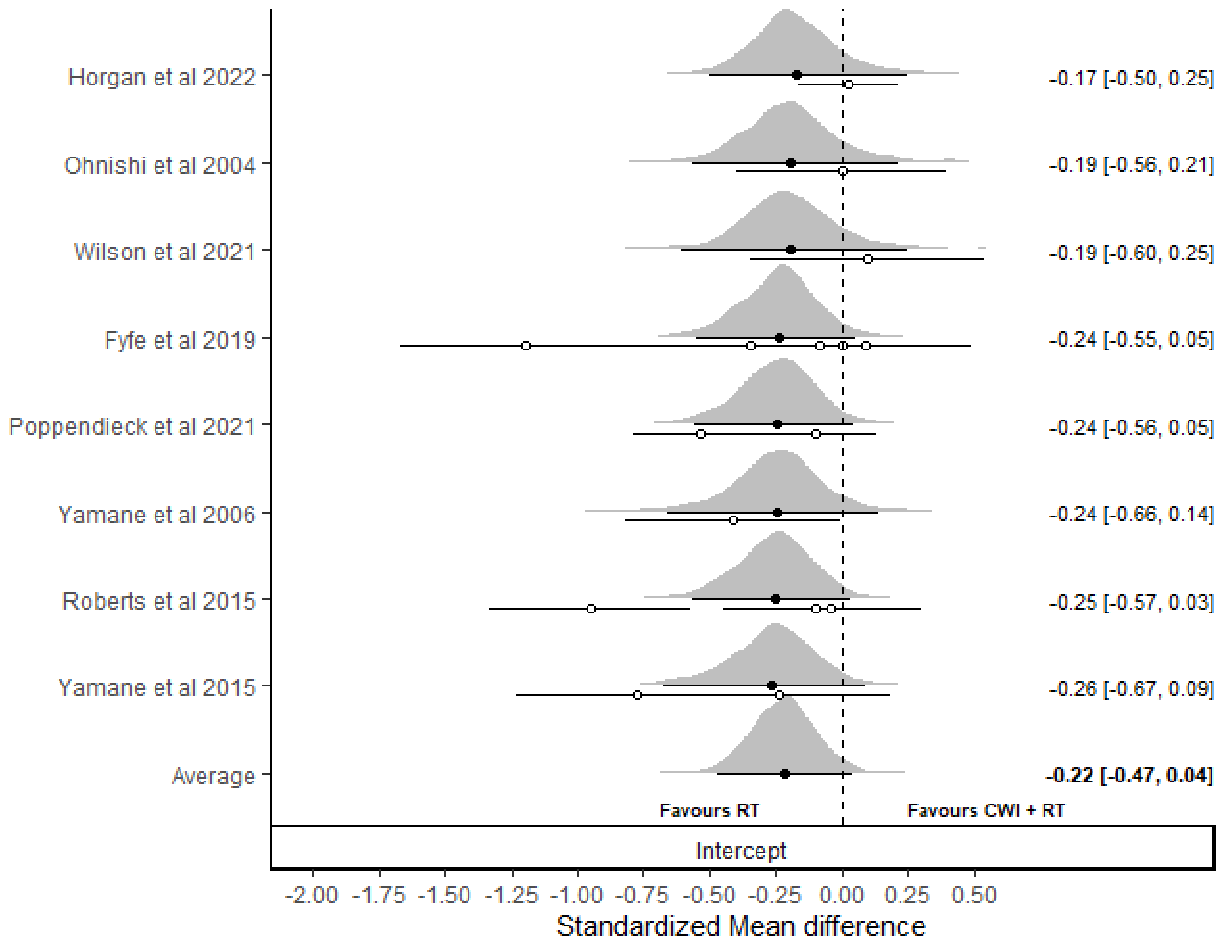

Figure 2: Bayesian forest plot illustrating pooling of comparative standardized mean difference effect sizes directly comparing resistance training (RT) and cold water immersion with resistance training (CWI + RT). Positive values favor cold water immersion and resistance training (CWI + RT) and negative values favor resistance training (RT). Distributions represent "shrunken estimates" based on all effects sizes included, the random effects model fitted and borrowed information across studies to reduce uncertainty. Black circles and connected intervals represent the median value and 95% credible intervals for the shrunken estimates. White circles and intervals represent the raw estimates and sampling variance calculated directly from study data. Bottom distribution illustrates uncertainty in the pooled mean.

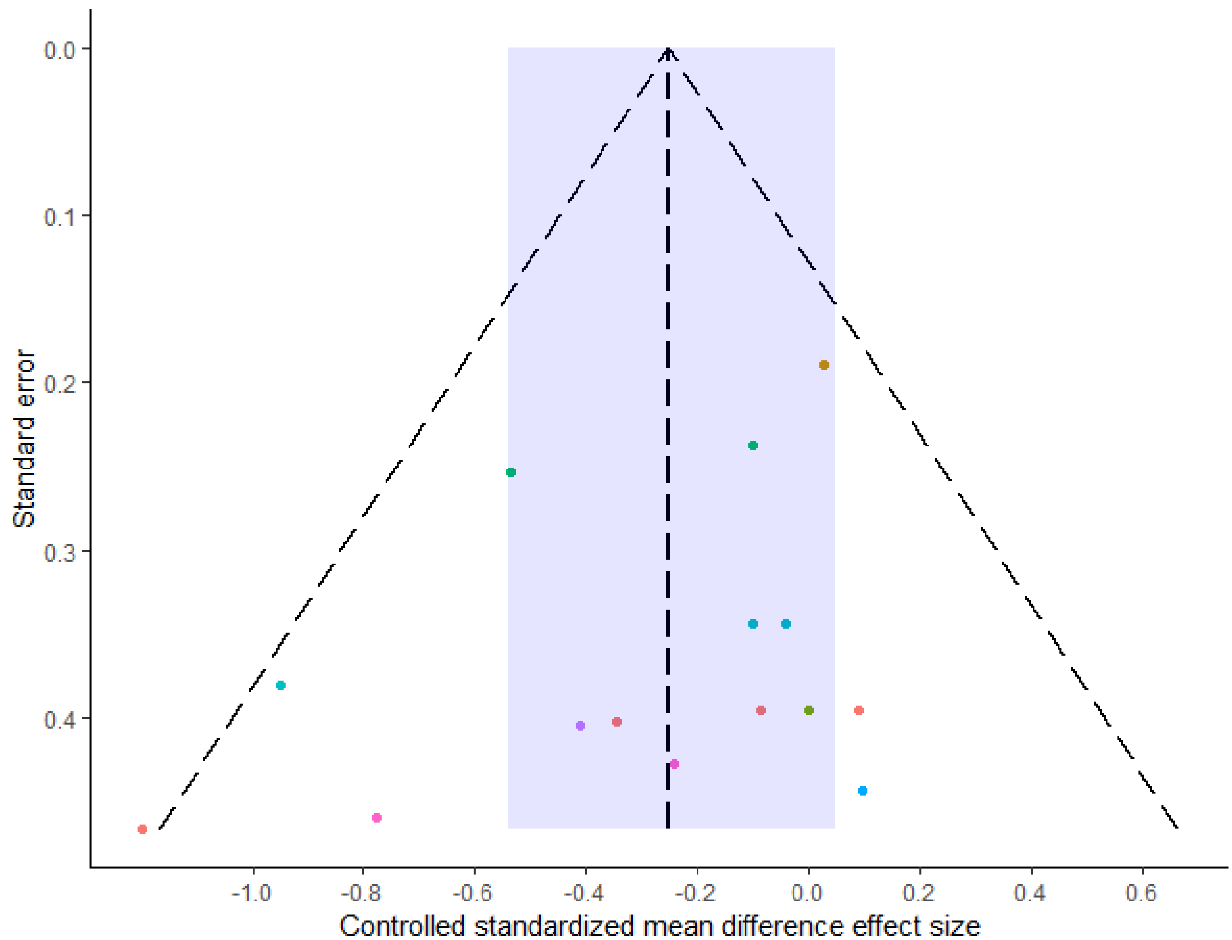

Figure 3: Funnel plot of all comparative effect sizes. Data are colored according to the individual studies. Blue region illustrates the pooled mean estimate and 95% credible interval.

Study quality

A qualitative assessment of the studies via the SMART-LD tool indicated a mean score of 9.8 out of a possible 20 points (range: 4 to 13 points). No studies were deemed to be of good quality, 3 studies were deemed to be of fair quality [42] [17] [20], and 5 studies were deemed to be of poor quality [41] [43] [44] [46] [45].

A qualitative assessment of the studies via the Downs and Black checklist indicated a mean score of 18.4 out of a possible 29 points (range: 12 to 24 points). Two studies were deemed to be of good quality [20] [42], 6 studies were classified as being of fair quality [17] [46] [41] [43] [45] [44], and no studies were found to be of poor quality.

Discussion

This is the first systematic review with meta-analysis to examine the effects of CWI on RT-induced skeletal muscle hypertrophy. Although evidence indicates that CWI does not completely prevent muscular gains, our results suggest it likely attenuates adaptations compared with RT alone. Analyses indicated the strategy is likely to result in at least a small magnitude of reduction in hypertrophy, with the upper credible interval identifying a relatively low probability of a moderate detrimental effect. Subanalysis using meta-regression provided a lack of evidence that training status altered the likely attenuation of muscle hypertrophy with CWI.

Our findings are consistent with acute data showing that CWI blunts the anabolic response to RT. For example, Fuchs et al. [19] found that CWI administered 20 minutes after performance of lower body RT reduced MPS rates for up to 5 hours post-application. The researchers also demonstrated an impaired MPS response to CWI during prolonged RT [19]. In addition, evidence indicates that the post-RT exposure to CWI attenuates activation of transcriptional factors involved in ribosome biogenesis [18] and suppresses satellite cell activity [20], both of which are purported to be important mediators of skeletal muscle hypertrophy [47]. These alterations persisted for up to 48 hours after application, suggesting a prolonged deleterious effect.

While the scope of this analysis did not include mechanistic drivers of skeletal muscle accretion, potential physiological mechanisms behind attenuated hypertrophy outcomes following CWI have been proposed in previous research. One such hypothesis is that CWI alters the acute inflammatory response to RT, which has been implicated in the kinase domain of titan hypertrophy [22]. A reduction in inflammatory responses to RT could also conceivably attenuate reactive oxygen species production [48] and associated activation of the mitogenactivated protein kinase pathway [49], thereby downregulating MPS [50] and potential anabolism. To this point, several human trials have reported blunted inflammatory-related cytokine responses to RT in the minutes and hours following cold exposure including inflammation markers such as interleukin-6 (IL-6), tumor necrosis factor-α and plasma chemokine ligand 2 [1] [21]. On the other hand, Peake et al. [7] found similar acute blunting of pro-inflammatory cytokines between CWI and active recovery groups after a bout of intense RT. Similarly, Ahokas et al. [51] reported little difference in inflammation markers following CWI and thermoneutral water immersion following high-intensity sprinting and jumping. The conflicting data call into question whether post-exercise reductions in acute inflammatory responses induced by CWI play a role in altering muscle development and, if so, to what extent.

It also is possible that CWI negatively affects anabolism via reductions in post-exercise blood flow to the musculature, a potential impact that should not be overlooked as blood flow and nutrient-dependent skeletal muscle proteolysis and MPS regulatory effects of insulin are well established [52] [53]. In brief, studies have found that CWI may reduce blood flow to musculature exposed to RT compared to both thermoneutral water [54] and whole-body cryotherapy [55] interventions. While the post-RT anabolic window for muscle growth associated with MPS may not be as narrow as once thought [56], an acute reduction in post-RT nutrient delivery, and therefore an extended period of muscle protein catabolism and diminished maximal MPS capacity via CWI-induced blood flow impairment, is theoretically plausible. Whether these acute outcomes directly or indirectly affect long-term hypertrophic adaptations remains undetermined.

Taken in concert, these findings may contribute to an acute mechanistic understanding of the attenuated skeletal muscle hypertrophy discussed in this analysis. However, it should be noted that findings related to acute responses to RT should not be extrapolated to chronic mechanisms of hypertrophic adaptations out of hand (e.g. [57]).

Regarding the quality of the included studies, we employed 2 separate assessments to determine their risk of bias and transparency of reporting. The frequently used Downs and Black checklist indicated that the included studies were generally of good to fair quality. Alternatively, the SMART-LD tool, created to specifically assess the quality of longitudinal RT research, indicated studies were generally of fair to poor quality. In combination, these results suggest that higher quality studies are needed to draw stronger inferences as to the effects of CWI on muscle development. Moreover, the discrepancies in results between the 2 tools indicate that commonly employed quality assessment methods may not be sufficient for evaluating longitudinal RT interventions.

Limitations

There are several limitations that must be acknowledged when attempting to draw evidence-based conclusions from our analysis. First, the majority of studies included in this analysis lasted between 4-8 weeks, with only one intervention exceeding 8 weeks [20]. Although results showed evidence that the combination of CWI and RT modestly impairs measures of muscular hypertrophy, it is not clear if the comparison between groups would have varied over longer time frames.

Second, the heterogeneity of measurement methods between studies can be considered a limitation to the validity of the findings, as the ability to draw inferences regarding the efficacy of a RT protocol is largely predicated on the assessment tools used [58]. The studies included in this meta-analysis employed a wide array of measurements including biopsy, DXA, circumference, US, and MRI. However, direct imaging modalities (i.e. MRI, CT, and US) have been shown to be more accurate for assessing hypertrophic adaptations compared to indirect modalities (i.e. DXA) [59] [60]. Moreover, biopsy shows high coefficients of variation (~13%) for assessing fiber CSA [61], indicating questionable reliability for this modality. Thus, future studies investigating the effects of CWI on muscular adaptations should seek to employ direct imaging methods either alone or in combination with other modalities. On the other hand, the fact that a variety of modalities indicate CWI impairs hypertrophy, even those less sensitive to detecting subtle changes in muscle mass, would seem to strengthen the confidence in our conclusions.

Third, the RT protocols varied greatly between studies, including total weekly training volume, frequency, and proximity to failure. Several of these studies included only exercises targeting smaller muscle groups during single-joint movements, which may not accurately reflect the training programs of most athletes. Only 3 studies [45] [43] [44] included exercises involving multi-joint movements commonly seen in athletic settings. Furthermore, only 1 study [17] used whole-body RT, which may be essential for determining if CWI has localized or systemic effects on RT adaptations.

Fourth, most of the included studies did not attempt to compile nutritional intake across the respective study periods. It is well-documented that both total daily energy and protein consumption influences the hypertrophic response to RT [62] [63]. Although the randomized designs would seemingly help to account for nutritional discrepancies between groups, the relatively small sample sizes of studies could have unduly influenced the responses in the respective groups. Thus, future studies should seek to account for dietary intake to ensure this variable does not confound results.

Fifth, all the included studies administered CWI therapy following every RT session using a similar approach (i.e. 10-20 minutes, <15 minutes following training, 10-15^{o}C), and there may be alternative CWI approaches to consider. Realistically, CWI therapy may be applied only intermittently throughout a certain period of time (e.g. a week or month), and does not necessarily have to be applied immediately following each training session. It therefore is possible that alternative approaches to CWI application might yield different results. Future research should look to establish consistent and ecologically valid standards as to the timing and frequency of CWI application to enhance the generalizability of findings.

Sixth, our results are specific to the use of CWI as a recovery strategy. We therefore cannot necessarily extrapolate findings to other cold application strategies such as cryotherapy, which warrant further investigation as to their chronic effects on muscular adaptations.

Finally, the pooled subject population consisted primarily of young men; only 1 of the 8 studies [46] involved female participants and no studies involved adolescents or older adults. Thus, our findings cannot necessarily be generalized to other populations. Given the influences that recovery, muscular sensitivity, and endocrine factors can have on RT adaptations [64] [65] [66], future research should investigate the impact of cooling strategies on muscular hypertrophy across populations.

Conclusion

The current data suggest that the application of CWI immediately following bouts of RT may modestly attenuate gains in muscle hypertrophy. When considering the practical implications of these findings, it is important to note that the results of this analysis apply solely to CWI application within 15 minutes of exercise cessation, which may not accurately reflect ecologically valid scenarios where CWI is employed several hours post-RT and/or implemented periodically rather than exclusively on RT days. It is unknown as to whether, or the degree to which, intermittent use of CWI or more time between RT sessions and CWI application may influence gains in muscle mass. Thus, individuals seeking to maximize muscle hypertrophy should avoid using CWI immediately following bouts of RT and further consider the frequency and timing of application. In addition, the current results suggest that RT in combination with CWI may still induce gains in muscle mass, but to a lesser degree compared to RT alone. These findings may have practical implications for athletes looking to limit RT-induced gains in muscle mass (e.g. distance runners). Further research is needed to understand the effects of different frequencies and timing strategies of CWI on RT-induced muscular adaptations, especially in resistance trained individuals and endurance athletes.

Contributions

AP conceived of the idea for the study. BJS and PAS designed the methodology for the study. MW, KD and AP carried out the search; RB, AEM, and FA coded the data. MC, RB, PAK, and AP performed the SMART-LD quality assessment. MS, FA, and AP performed the Downs and Black quality assessment. PAS carried out the statistical modeling and analyses. All authors contributed to the writing and critical editing of the manuscript. All authors approved the final manuscript.

Conflict of interest

The authors report no competing interests with the content of this manuscript.

Funding information

No funding was received for this manuscript.

Data and Supplementary Material Accessibility

Data and supplementary material are available on the Open Science Framework project page: https://osf.io/gx69b

References

1. Crystal NJ, Townson DH, Cook SB, LaRoche DP. Effect of cryotherapy on muscle recovery and inflammation following a bout of damaging exercise. Eur J Appl Physiol. 2013 October 01;113(10):2577-86.

2. Wang Y, Li S, Zhang Y, Chen Y, Yan F, Han L, et al. Heat and cold therapy reduce pain in patients with delayed onset muscle soreness: A systematic review and meta-analysis of 32 randomized controlled trials. Phys Ther Sport. 2021 March 01;48:177-87.

3. Xiao F, Kabachkova AV, Jiao L, Zhao H, Kapilevich LV. Effects of cold water immersion after exercise on fatigue recovery and exercise performance--meta analysis. Front Physiol. 2023 January 20;14:1006512.

4. Roberts LA, Nosaka K, Coombes JS, Peake JM. Cold water immersion enhances recovery of submaximal muscle function after resistance exercise. Am J Physiol Regul Integr Comp Physiol. 2014 October 15;307(8):R998-R1008.

5. McPhee JS, Lightfoot AP. Post-exercise recovery regimes: Blowing hot and cold. J Physiol. 2017 February 01;595(3):627-8.

6. Thorpe RT. Post-exercise recovery: Cooling and heating, a periodized approach. Front Sports Act Living. 2021 September 01;3:707503.

7. Peake JM, Roberts LA, Figueiredo VC, Egner I, Krog S, Aas SN, et al. The effects of cold water immersion and active recovery on inflammation and cell stress responses in human skeletal muscle after resistance exercise. J Physiol. 2017 February 01;595(3):695-711.

8. Tseng C, Lee J, Tsai Y, Lee S, Kao C, Liu T, et al. Topical cooling (icing) delays recovery from eccentric exercise-induced muscle damage. J Strength Cond Res. 2013 May 01;27(5):1354-61.

9. Broatch JR, Petersen A, Bishop DJ. The influence of post-exercise cold-water immersion on adaptive responses to exercise: A review of the literature. Sports Med. 2018 June 01;48(6):1369-87.

10. Versey NG, Halson SL, Dawson BT. Water immersion recovery for athletes: Effect on exercise performance and practical recommendations. Sports Med. 2013 November 01;43(11):1101-30.

11. Pointon M, Duffield R. Cold water immersion recovery after simulated collision sport exercise. Med Sci Sports Exerc. 2012 February 01;44(2):206-16.

12. Ihsan M, Abbiss CR, Allan R. Adaptations to post-exercise cold water immersion: Friend, foe, or futile? Front Sports Act Living. 2021 July 16;3:714148.

13. Chaillou T, Treigyte V, Mosely S, Brazaitis M, Venckunas T, Cheng AJ. Functional impact of post-exercise cooling and heating on recovery and training adaptations: Application to resistance, endurance, and sprint exercise. Sports Med Open. 2022 March 07;8(1):37-022.

14. Malta ES, Dutra YM, Broatch JR, Bishop DJ, Zagatto AM. The effects of regular cold-water immersion use on training-induced changes in strength and endurance performance: A systematic review with meta-analysis. Sports Med. 2021 January 01;51(1):161-74.

15. Petersen AC, Fyfe JJ. Post-exercise cold water immersion effects on physiological adaptations to resistance training and the underlying mechanisms in skeletal muscle: A narrative review. Front Sports Act Living. 2021 April 08;3:660291.

16. Grgic J. Effects of post-exercise cold-water immersion on resistance training-induced gains in muscular strength: A meta-analysis. Eur J Sport Sci. 2023 March 01;23(3):372-80.

17. Fyfe JJ, Broatch JR, Trewin AJ, Hanson ED, Argus CK, Garnham AP, et al. Cold water immersion attenuates anabolic signaling and skeletal muscle fiber hypertrophy, but not strength gain, following whole-body resistance training. J Appl Physiol (1985). 2019 November 01;127(5):1403-18.

18. Figueiredo VC, Roberts LA, Markworth JF, Barnett MP, Coombes JS, Raastad T, et al. Impact of resistance exercise on ribosome biogenesis is acutely regulated by post-exercise recovery strategies. Physiol Rep. 2016 February 01;4(2):10.14814/phy2.12670.

19. Fuchs CJ, Kouw IWK, Churchward-Venne TA, Smeets JSJ, Senden JM, Lichtenbelt WDvM, et al. Postexercise cooling impairs muscle protein synthesis rates in recreational athletes. J Physiol. 2020 February 01;598(4):755-72.

20. Roberts LA, Raastad T, Markworth JF, Figueiredo VC, Egner IM, Shield A, et al. Post-exercise cold water immersion attenuates acute anabolic signalling and long-term adaptations in muscle to strength training. J Physiol. 2015 September 15;593(18):4285-301.

21. Earp JE, Hatfield DL, Sherman A, Lee EC, Kraemer WJ. Cold-water immersion blunts and delays increases in circulating testosterone and cytokines post-resistance exercise. Eur J Appl Physiol. 2019 August 01;119(8):1901-7.

22. Koh TJ, Pizza FX. Do inflammatory cells influence skeletal muscle hypertrophy? Front Biosci (Elite Ed). 2009 June 01;1:60-71.

23. Ogasawara R, Suginohara T. Rapamycin-insensitive mechanistic target of rapamycin regulates basal and resistance exercise-induced muscle protein synthesis. FASEB J. 2018 May 14:fj201701422R.

24. Kraemer WJ, Ratamess NA. Hormonal responses and adaptations to resistance exercise and training. Sports Med. 2005;35(4):339-61.

25. Figueiredo VC, McCarthy JJ. Regulation of ribosome biogenesis in skeletal muscle hypertrophy. Physiology (Bethesda). 2019 January 01;34(1):30-42.

26. Crewther B, Keogh J, Cronin J, Cook C. Possible stimuli for strength and power adaptation: Acute hormonal responses. Sports Med. 2006;36(3):215-38.

27. 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 March 29;372:n71.

28. Greenhalgh T, Peacock R. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: Audit of primary sources. BMJ. 2005 November 05;331(7524):1064-5.

29. Cooper H, Hedges L, Valentine J. The handbook of research synthesis and meta-analysis. . 2nd ed. New York: Russell Sage Foundation; 2009.

30. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health. 1998 June 01;52(6):377-84.

31. Resnick B, Zimmerman SI, Orwig D, Furstenberg AL, Magaziner J. Outcome expectations for exercise scale: Utility and psychometrics. J Gerontol B Psychol Sci Soc Sci. 2000 November 01;55(6):S352-6.

32. Kruschke JK, Liddell TM. The bayesian new statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a bayesian perspective. Psychon Bull Rev. 2018 February 01;25(1):178-206.

33. Morris B. Estimating effect sizes from pretest-posttest-control group designs. Organizational Research Methods. 2008;11(2):364-86.

34. Morris SB, DeShon RP. Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. Psychol Methods. 2002 March 01;7(1):105-25.

35. Madeyski L, Kitchenham B. In: Effect sizes and their variance for AB/BA crossover design studies. Proceedings of the 40th international conference on software engineering; Gothenburg, Sweden. New York, NY: Association for Computing Machinery; May 27 2018. p. 420.

36. Swinton PA, Murphy A. Comparative effect size distributions in strength and conditioning and implications for future research: A meta-analysis. SportRxiv. 2022:DOI: 10.51224/SRXIV.202.

37. Swinton PA, Burgess K, Hall A, Greig L, Psyllas J, Aspe R, et al. Interpreting magnitude of change in strength and conditioning: Effect size selection, threshold values and bayesian updating. J Sports Sci. 2022 September 01;40(18):2047-54.

38. Fu R, Gartlehner G, Grant M, Shamliyan T, Sedrakyan A, Wilt TJ, et al. Conducting quantitative synthesis when comparing medical interventions: AHRQ and the effective health care program. J Clin Epidemiol. 2011 November 01;64(11):1187-97.

39. Fernandez-Castilla B, Declercq L, Jamshidi L, Beretvas S, Onghena P, Van den Noorthgate W. Detecting selection bias in meta-analyses with multiple outcomes: A simulation study . J Exp Educ. 2021;89(1):125-44.

40. Bürkner P. Brms: An R package for bayesian multilevel models using stan. J Stat Softw. 2017;80(1):1-28.

41. Wilson LJ, Dimitriou L, Hills FA, Gondek MB, van Wyk A, Turek V, et al. Cold water immersion offers no functional or perceptual benefit compared to a sham intervention during a resistance training program. J Strength Cond Res. 2021 October 01;35(10):2720-7.

42. Horgan BG, Halson SL, Drinkwater EJ, West NP, Tee N, Alcock RD, et al. No effect of repeated post-resistance exercise cold or hot water immersion on in-season body composition and performance responses in academy rugby players: A randomised controlled cross-over design. Eur J Appl Physiol. 2023 February 01;123(2):351-9.

43. Yamane M, Teruya H, Nakano M, Ogai R, Ohnishi N, Kosaka M. Post-exercise leg and forearm flexor muscle cooling in humans attenuates endurance and resistance training effects on muscle performance and on circulatory adaptation. Eur J Appl Physiol. 2006 March 01;96(5):572-80.

44. Ohnishi N, Yamane M, Uchiyama N, Shirasawa S, Kosaka M, Shiono H. Adaptive changes in muscular performance and circulation by resistance training with regular cold application. J Thermal Biol. 2004;29:839-43.

45. Yamane M, Ohnishi N, Matsumoto T. Does regular post-exercise cold application attenuate trained muscle adaptation? Int J Sports Med. 2015 July 01;36(8):647-53.

46. Poppendieck W, Wegmann M, Hecksteden A, Darup A, Schimpchen J, Skorski S, et al. Does cold-water immersion after strength training attenuate training adaptation? Int J Sports Physiol Perform. 2021 February 01;16(2):304-10.

47. Brook MS, Wilkinson DJ, Smith K, Atherton PJ. It's not just about protein turnover: The role of ribosomal biogenesis and satellite cells in the regulation of skeletal muscle hypertrophy. Eur J Sport Sci. 2019 August 01;19(7):952-63.

48. Uchiyama S, Tsukamoto H, Yoshimura S, Tamaki T. Relationship between oxidative stress in muscle tissue and weight-lifting-induced muscle damage. Pflugers Arch. Invalid date Invalid date;452(1):109-16. Available from: http://www.hubmed.org/display.cgi?uids=16402246.

49. Roux PP, Blenis J. ERK and p38 MAPK-activated protein kinases: A family of protein kinases with diverse biological functions. Microbiol Mol Biol Rev. 2004 June 01;68(2):320-44.

50. Takarada Y, Nakamura Y, Aruga S, Onda T, Miyazaki S, Ishii N. Rapid increase in plasma growth hormone after low-intensity resistance exercise with vascular occlusion. J Appl Physiol. Invalid date Invalid date;88(1):61-5. Available from: http://www.hubmed.org/display.cgi?uids=10642363.

51. Ahokas EK, Kyröläinen H, Mero AA, Walker S, Hanstock HG, Ihalainen JK. Water immersion methods do not alter muscle damage and inflammation biomarkers after high-intensity sprinting and jumping exercise. Eur J Appl Physiol. 2020 December 01;120(12):2625-34.

52. Timmerman KL, Lee JL, Dreyer HC, Dhanani S, Glynn EL, Fry CS, et al. Insulin stimulates human skeletal muscle protein synthesis via an indirect mechanism involving endothelial-dependent vasodilation and mammalian target of rapamycin complex 1 signaling. J Clin Endocrinol Metab. 2010 August 01;95(8):3848-57.

53. Fujita S, Rasmussen BB, Cadenas JG, Grady JJ, Volpi E. Effect of insulin on human skeletal muscle protein synthesis is modulated by insulin-induced changes in muscle blood flow and amino acid availability. Am J Physiol Endocrinol Metab. 2006 October 01;291(4):E745-54.

54. Gregson W, Black MA, Jones H, Milson J, Morton J, Dawson B, et al. Influence of cold water immersion on limb and cutaneous blood flow at rest. Am J Sports Med. 2011 June 01;39(6):1316-23.

55. Mawhinney C, Jones H, Low DA, Green DJ, Howatson G, Gregson W. Influence of cold-water immersion on limb blood flow after resistance exercise. Eur J Sport Sci. 2017 June 01;17(5):519-29.

56. Schoenfeld BJ, Aragon AA. Is there a postworkout anabolic window of opportunity for nutrient consumption? clearing up controversies. J Orthop Sports Phys Ther. 2018 December 01;48(12):911-4.

57. Schoenfeld BJ. Postexercise hypertrophic adaptations: A reexamination of the hormone hypothesis and its applicability to resistance training program design. J Strength Cond Res. 2013 June 01;27(6):1720-30.

58. Haun CT, Vann CG, Roberts BM, Vigotsky AD, Schoenfeld BJ, Roberts MD. A critical evaluation of the biological construct skeletal muscle hypertrophy: Size matters but so does the measurement. Front Physiol. 2019 March 12;10:247.

59. Delmonico MJ, Kostek MC, Johns J, Hurley BF, Conway JM. Can dual energy X-ray absorptiometry provide a valid assessment of changes in thigh muscle mass with strength training in older adults? Eur J Clin Nutr. 2008 December 01;62(12):1372-8.

60. Levine JA, Abboud L, Barry M, Reed JE, Sheedy PF, Jensen MD. Measuring leg muscle and fat mass in humans: Comparison of CT and dual-energy X-ray absorptiometry. J Appl Physiol. 2000 February 01;88(2):452-6.

61. Horwath O, Envall H, Röja J, Emanuelsson EB, Sanz G, Ekblom B, et al. Variability in vastus lateralis fiber type distribution, fiber size, and myonuclear content along and between the legs. J Appl Physiol (1985). 2021 July 01;131(1):158-73.

62. Aragon AA, Schoenfeld BJ. Magnitude and composition of the energy surplus for maximizing muscle hypertrophy: Implications for bodybuilding and physique athletes. Strength & Conditioning Journal. 2020;42(5):79-86.

63. Morton RW, Murphy KT, McKellar SR, Schoenfeld BJ, Henselmans M, Helms E, et al. A systematic review, meta-analysis and meta-regression of the effect of protein supplementation on resistance training-induced gains in muscle mass and strength in healthy adults. Br J Sports Med. 2017 July 11.

64. Baláš J, Kodejška J, Krupková D, Giles D. Males benefit more from cold water immersion during repeated handgrip contractions than females despite similar oxygen kinetics. J Physiol Sci. 2020 March 05;70(1):13-020.

65. Castellani JW, Young AJ. Human physiological responses to cold exposure: Acute responses and acclimatization to prolonged exposure. Auton Neurosci. 2016 April 01;196:63-74.

66. Petrofsky JS, Laymon M. Heat transfer to deep tissue: The effect of body fat and heating modality. J Med Eng Technol. 2009;33(5):337-48.

67. DeFreitas JM, Beck TW, Stock MS, Dillon MA, Kasishke PR. An examination of the time course of training-induced skeletal muscle hypertrophy. Eur J Appl Physiol. 2011 November 01;111(11):2785-90.