MARTICORENA, F.M., CARVALHO, A., DE OLIVEIRA, L.F., DOLAN, E., GUALANO, B., SWINTON, P. and SAUNDERS, B. 2021. Nonplacebo controls to determine the magnitude of ergogenic interventions: a systematic review and meta-analysis. *Medicine and science in sports and exercise* [online], 53(8), pages 1766-1777. Available from: https://doi.org/10.1249/MSS.0000000000002635

Nonplacebo controls to determine the magnitude of ergogenic interventions: a systematic review and meta-analysis.

MARTICORENA, F.M., CARVALHO, A., DE OLIVEIRA, L.F., DOLAN, E., GUALANO, B., SWINTON, P. and SAUNDERS, B.

2021

Non-placebo controls to determine the magnitude of ergogenic interventions: A systematic review and meta-analysis

Felipe Miguel Marticorena[1], Arthur Carvalho[1], Luana Farias de Oliveira[1], Eimear Dolan[1], Bruno Gualano[1,2], Paul Swinton[3], Bryan Saunders[1,4]

Running title: The importance of non-placebo controls

1 – Applied Physiology and Nutrition Research Group, School of Physical Education and Sport; Rheumatology Division; Faculdade de Medicina FMUSP, Universidade de Sao Paulo, Sao Paulo, SP, BR, University of São Paulo, São Paulo, Brazil.

2 – Food Research Center (FoRC), University of São Paulo, São Paulo, Brazil

3 – School of Health Sciences, Robert Gordon University, Aberdeen, United Kingdom.

4 – Institute of Orthopaedics and Traumatology, Faculty of Medicine FMUSP, University of São Paulo, Brazil.

Correspondence:

Dr Bryan Saunders

Applied Physiology & Nutrition Research Group,

Rheumatology Division, Faculty of Medicine FMUSP,

Av. Dr. Arnaldo, 455 - Cerqueira César - CEP: 01246903

University of São Paulo,

São Paulo, SP, Brazil.

E-mail: drbryansaunders@outlook.com

Phone: +55 11 3061-8789

Fax: +55 11 3813-5921

Abstract

Introduction: Placebos are used as a control treatment that is meant to be indistinguishable from the active intervention. However, where substantive placebo effects may occur, studies that do not include a non-placebo control arm may underestimate the overall effect of the intervention (active plus placebo components). This study aimed to determine the relative magnitude of the placebo effect associated with nutritional supplements (caffeine and extracellular buffers) by meta-analysing data from studies containing both placebo and non-placebo control sessions. **Methods:** Bayesian multilevel meta-analysis models were used to estimate pooled effects and express the placebo effect as a percentage of the overall intervention effect. **Results:** Thirty-four studies were included, with the median pooled effect size ($ES_{0.5}$) indicating a very small ($ES_{0.5}$=0.09 [95%CrI:0.01 to 0.17]) improvement in performance of placebo compared to control. There was no moderating effect of exercise type (capacity or performance), exercise duration or training status. The comparison between active intervention and control indicated a small to medium effect ($ES_{0.5}$=0.37 [95%CrI:0.20 to 0.56]). Expressed in relative terms, the placebo effect was equivalent to 25% [75%CrI:16 to 35%] and 59% [75%CrI:34 to 94%] of the total intervention effect for buffers and caffeine. **Conclusion:** These results demonstrate a very small, but potentially important placebo effect with nutritional supplementation studies. A substantive proportion of supplement effects may be due to placebo effects, with the relative proportion influenced by the magnitude of the overall ergogenic effect. Where feasible, intervention studies should employ non-placebo control-arm comparators to identify the proportion of the effect estimated to come from placebo effects and avoid underestimating the overall benefits that the physiological plus psychobiological aspects associated with an intervention provide in the real world.

Key words: Exercise; nutrition; physiology, belief, expectation, neurobiological.

Introduction

Historically in research, placebos have been used as a control treatment that is indistinguishable from the active intervention but has no active component. This allows the intervention arm to be compared to the non-active placebo treatment which, theoretically, should allow determination of the "true" effect of the intervention, through reducing the potential that expectancy, or belief, about the intervention may artificially inflate the estimated response to that intervention. These placebo effects are induced by administering an inert intervention (e.g., sham drugs, nutrients, or equipment), which can elicit a neurobiological response due to verbal suggestions, explicit expectations, or implicit experience (1). Placebo effects in sport and exercise nutrition are well reported with a systematic review showing small to moderate positive effects on exercise outcomes (d = 0.35 [CI = 0.20 to 0.51]) (2) when participants are administered an inert substance believed to be a beneficial treatment. In the real-world, the actual effect of interventions is likely to comprise both the active physiological component and the placebo effect. This means that studies could underestimate the overall real-world impact of a given intervention if they do not include a non-placebo control arm. Accordingly, studies that comprise three conditions, namely active treatment, placebo, and a no-treatment control, are required to determine the total effect of the intervention (active treatment plus placebo effects), and to estimate the relative contribution of these component parts.

Caffeine and buffering supplements, such as sodium bicarbonate, are ergogenic supplements considered effective to improve exercise capacity and performance (3). However, placebo effects may occur with some nutritional supplements. Specifically, individuals who believe they have ingested caffeine (4-6) or sodium bicarbonate (7), but have actually ingested placebo, can improve their exercise performance. On the contrary, performance improvements might be blunted when individuals consume caffeine but believe they have received placebo (8). Thus, in double-blind

placebo-controlled designs, there may be a variable placebo or nocebo effect related to the psychosocial context of ingesting a supplement (active or inactive) which can influence the exercise outcome due to expectancy (5). Studies investigating the efficacy of supplements commonly include an active treatment arm and a placebo comparator in a double-blind crossover design, but do not always employ a non-placebo control session. Those studies that do incorporate both a placebo and a non-placebo control session provide an excellent platform to determine the proportion and variability of ergogenic effects that can be attributed to placebo effects. Moreover, knowing the resulting net effect of a given active intervention plus its placebo effect would help inform prescription for athletes.

The aim of this study was to estimate the size of the placebo effects associated with caffeine and buffering supplements, and to determine the proportion of the overall ergogenic effect that is explained by placebo. This was achieved through meta-analysis of aggregate data from studies including both placebo and non-placebo control sessions. A secondary aim was to determine which factors might modify the size of these placebo effects, including exercise, population, and supplement characteristics. Caffeine and buffering supplements were chosen as a model for this study since they are considered effective ergogenic aids, and both have been shown to incur placebo effects.

Methods

Study Eligibility:

The study protocol was designed in accordance with PRISMA guidelines (9) (see Table, Supplemental Digital Content 1) and the inclusion criteria defined according to PICOS criteria (Population, Intervention, Comparator, Outcomes and Study design). Only English-language peer-reviewed, original human studies were included within this review. The ***population*** included healthy human males and females of any age, but studies conducted with diseased-state participants were excluded. Recreationally active and trained individuals and professional athletes were considered for inclusion. The ***intervention*** required a supplementation protocol comprising any dose of caffeine or sodium bicarbonate, sodium citrate, calcium lactate or sodium lactate prior to performing an exercise test. These supplements (*i.e.*, caffeine and extracellular buffers) were selected to test our hypotheses as there is a substantial body of placebo-controlled studies estimating ergogenic and placebo effects. In relation to the ***comparator***, the aims of this study determined that studies necessarily employed both a placebo (inert substance) and a control (no treatment) trial. Studies that reported on ***outcomes*** based on exercise performance or capacity tests (*e.g.* total work done, mean power output) were considered for inclusion. ***Study design*** included any randomised and blinded, crossover or parallel-group design. Studies that employed balanced placebo designs, in which participants were informed (correctly or deceptively) what supplement they had received were not included, as this has been reviewed extensively elsewhere (2). The study was not pre-registered.

Search Strategy and Quality Assessment:

An electronic search of the literature was undertaken by FMM using three databases (MedLine, Embase and SPORTDiscus) to identify relevant articles. Caffeine studies were searched using the term “caffeine” concatenated with “exercise”, “performance”, “physical performance” and

"training". Extracellular buffer studies were searched using the search terms "sodium bicarbonate", "sodium citrate", "calcium lactate", "sodium lactate" and "alkalosis" concatenated with "supplementation", "exercise", "training", "athlete" and "performance". The extracellular buffers search was originally conducted to inform a systematic review and meta-analysis on the use of extracellular buffers on exercise outcomes.

Duplicates were removed before a 2-phase search strategy was performed independently by two reviewers for buffering supplements (LFO and ED) and caffeine (FMM and AC) using Rayyan software for systematic reviews (10). Phase one assessed the eligibility of the title and abstract of every article generated from the search terms against the inclusion/exclusion criteria. Studies with uncertain suitability were included at this stage and a final decision was reached at the next phase, namely phase two in which full articles were retrieved and assessed against the eligibility criteria. Reference lists of all included studies and review articles were screened to ensure all relevant studies were included. Any differences of opinion relating to study eligibility were resolved through discussion. The original searches were conducted in November 2019 and these searches were updated in April 2020, to identify any eligible studies published in the interim. No date limit was applied to the search.

Data Extraction and Variable Categorisation:

Data extraction was conducted by FMM using a standardised and pre-piloted extraction spreadsheet created using Microsoft excel. Extracted information included authors and year of publication, population characteristics (age, gender and training status), supplementation protocol (dose, timing and form of administration), exercise protocol, type (exercise capacity or performance) and duration (<30 s; 30 s – 10 min; >10 min) of exercise, and the exercise outcomes. All extracted data are available in a Table (Supplemental Digital Content 2). To avoid duplication

bias, a solitary outcome measure from each exercise protocol was extracted based upon an *a priori* hierarchy agreed upon by all authors to ensure consistency in data extraction. Data were extracted according to availability, prioritising exercise measures over physiological measures, according to the following hierarchical profile (11):

1. Total work done
2. Mean output throughout the test (*i.e.,* mean power output; mean velocity; mean height)
3. Time to completion (performance test)/time to exhaustion (capacity test)

Following data extraction, an *a posteriori* decision was made to determine the contribution of factors that might modify the placebo response to supplementation. As such, data were categorised according to the following factors for analysis:

i) Exercise protocols were separated by exercise duration [*exercise duration*] according to the approach of Saunders et al. (12), namely 0-0.5 min; 0.5-10 min; >10 min, which were chosen considering energy system contribution to tests that differ in duration.

ii) Exercise protocols were also categorised according to whether they measured exercise capacity or performance [*exercise type*], as these different exercise types have been shown to modify the effects of supplements (13).

iii) Studies were separated according to the sample population recruited [*training status*] since trained athletes may have greater belief in the power of placebos to enhance sporting performance (14) which might differ from non-trained individuals. Trained individuals were considered those engaged in a structured training programme with a training plan relevant to the exercise task employed in the study, while remaining populations that did not fit these criteria (i.e., recreationally active; non-trained; sedentary) were categorised as non-trained.

iv) Supplement characteristics may modify the placebo effects of an intervention (15-17). Supplementation protocols were thus separated according to delivery method [*supplement delivery*], namely whether they provided the placebo and active interventions as capsules or dissolved in solution.

Risk of Bias and Quality Assessment

Risk of bias was assessed using the most recent Cochrane tool for assessing risk of bias in randomized crossover trials (18), and study quality was evaluated using an adapted Downs and Black questionnaire (19) (Supplemental Digital Content 3 – Table). Using these two tools allowed greater certainty and reliability to identify the quality of the included studies. Evaluation of risk of bias and study quality was performed in a blinded fashion by two independent reviewers (FMM and AC), and any disagreements were resolved between these two reviewers via discussion and, when necessary, with the help of a third reviewer (B.S.). We adapted the Downs and Black questionnaire checklist, since some questions were not relevant for the purpose of this review, resulting in a 14-point adapted questionnaire. According to each question, the answers were classified with points from 0 to 1 or 2 that were summed to provide an overall score. The quality score was ranked according to the following intervals: High (12 - 14); Moderate (9 - 11); Low (6 - 8); and Very Low (≤5).

Data Analysis

Extracted data were transformed into pairwise effect sizes (intervention vs placebo, intervention vs control and placebo vs control) by calculating Hedges' g standardized mean difference. Calculation of standardized effect sizes enabled results from exercise tests conducted on different scales to be pooled in the meta-analysis. Standard distributional assumptions were used to calculate effect size standard errors (20). Most previous meta-analyses have been conducted within a

frequentist framework where parameters such as the pooled effect size are estimated, and the uncertainty expressed with a 95% confidence interval (i.e. the values that would not be rejected by p<0.05 (21)). However, confidence intervals contain no distributional information, such that there is no direct sense by which the parameter values in the middle of the interval are more probable that the ends (16). In contrast, Bayesian frameworks combine prior beliefs regarding the most plausible values with data to provide values that can be directly interpreted as probabilities. Results can therefore be interpreted easily and more applied questions such as the probability that parameters of interest exceed relevant thresholds can be addressed. In the present meta-analysis, the Bayesian framework was implemented through three-level hierarchical models with random effects to account for variation in the mean effect and covariance where multiple outcomes were reported in the same study (22). To investigate potential moderating effects of factors such as [*supplement type*], [*supplement delivery*], [*exercise type*] and [*exercise duration*] and [*training status*], meta-regressions were performed. To investigate the relative proportions of the placebo and intervention effect, models combining both intervention and placebo study effect sizes were conducted in a meta-regression. Inferences from all analyses were performed on posterior samples generated by Hamiltonian Markov Chain Monte Carlo with Bayesian 95% credible intervals (CrIs). Interpretations were based on visual inspection of the posterior sample, the median value ($ES_{0.5}$: 0.5- quantile) and 95% CrIs. Threshold values of .01, 0.2, 0.5 and 0.8 were used to describe pooled effect sizes as very small, small, medium and large, respectively (23). Interpretations of meta-regressions were based on location and spread of posterior distributions of regression coefficients ($\beta_{0.5;\mathrm{Reference:Comparison}}$). Expression of placebo as a proportion was achieved by taking the ratio of the two posterior samples. Analyses were performed using the R wrapper package brms, which interfaced with Stan to perform sampling (24). Weakly informative Student-t prior and half-t priors with 3 degrees of freedom and scale parameter equal to 2.5 were used for intercept and variance parameters (25). Convergence of parameter estimates was obtained for all

models with Gelman-Rubin R-hat values below 1.1 (26). Where outliers ($g \geq 2.0$, 15 out of 168, calculated effect sizes) were present, sensitivity analyses were conducted by performing robust hierarchical models employing a t-distribution for the likelihood. No substantive differences in findings were obtained for any sensitivity analysis conducted.

Results

Study search, characteristics, and quality appraisal

The primary search resulted in 7824 articles for caffeine and 3621 for buffers (Figure 1). Following removals of duplicates (Caffeine: 2193; Buffers: 334), Phase One resulted in the exclusion of 5143 caffeine and 2994 buffer papers. The remaining papers were screened in their entirety for suitability and following removal of studies without a non-placebo control session, thirty-four published studies met the criteria for inclusion in the analyses (Table 1), containing a total of 56 exercise outcomes from 363 participants. These comprised 10 caffeine studies yielding 11 outcomes, and 24 articles for buffering supplements yielding 45 outcomes. All studies were randomised crossover study designs.

Risk of bias and quality assessment

Almost all studies included in the meta-analysis were classified as having "some concerns" according to ROB2, except one which was categorised as "high" risk of bias due to issues in the randomization process (Figure 2). Although all studies were randomised, twenty studies had some concerns in Domain 1 (Randomization process) due to a lack of detail while all studies were classified as having some concerns owing to a lack of a pre- specified analysis plan (as outlined in Domain 5). The Downs and Black quality appraisal showed all studies attained a high score of between 12-14 (Supplemental Digital Content 3 – Table).

Meta-analysis

Absolute placebo effects (Placebo vs. Non-placebo control)

An initial assessment of the placebo effect was obtained by pooling pairwise effect sizes (placebo vs non-placebo control) across studies investigating caffeine and/or buffering supplements. The pooled effect size indicated a very small effect ($ES_{0.5}$ = 0.09 [95%CrI: 0 to 0.17]; $\tau_{0.5}$ = 0.04 [0.00

to 0.14]; ICC: 0.38 [0.10 to 0.69]; Figure 3, Panel A) of placebo compared to non-placebo control (Plot including effect sizes prior to creation of shrunken estimates are included in Supplemental Digital Content 4 – Figure). The probability that the pooled effect size represented at least a small effect was p = 0.005. There was some evidence that the placebo effect was moderated by [*supplement type*] and [*supplement delivery*], with greater effects obtained with buffers ($\beta_{0.5;\text{Caffeine:Buffers}}$ = 0.06 [95%CrI: -0.12 to 0.24]) and with solution ($\beta_{0.5;\text{Capsule:Solution}}$ = 0.10 [95%CrI: -0.07 to 0.28]). However, when the meta-regression was performed with both factors included in the model, the analysis indicated that [*supplement delivery*] was more likely to act as a moderator ($\beta_{0.5;\text{Intercept:Solution}}$ = 0.11 [95%CrI: -0.08 to 0.31; $\beta_{0.5;\text{Intercept:Buffers}}$ = -0.01 [95%CrI: -0.21 to 0.20]). There was no evidence of a moderating effect of [*training status*] ($\beta_{0.5;\text{Trained:Non-trained}}$ = 0.00 [95%CrI: -0.17 to 0.17]), [*exercise type*] ($\beta_{0.5;\text{Performance:Capacity}}$ = 0.00 [95%CrI: -0.21 to 0.20]), or [*exercise duration*] ($\beta_{0.5;<30s:30s-10\text{min}}$ = 0.01 [95%CrI: -0.21 to 0.22]; $\beta_{0.5;<30s:+10\text{min}}$ = -0.01 [95%CrI: -0.26 to 0.25]). Finally, there was some evidence of an effect of year of publication, with greater effects reported in studies published prior to 2000 than after ($\beta_{0.5;<2000:>2000}$ = -0.10 [95%CrI: -0.27 to 0.05]). All coefficients for meta-regressions are presented in Table 2.

Intervention effects (Active intervention vs. Placebo and vs. Non-placebo control)

The intervention effect was initially investigated by pooling the pairwise non-placebo control effect sizes across studies investigating both caffeine and buffering supplements. The pooled effect size indicated a small to medium effect ($ES_{0.5}$ = 0.37 [95%CrI: 0.20 to 0.56]; $\tau_{0.5}$ = 0.25 [0.04 to 0.48]; ICC: 0.33 [0 to 0.69]); Figure 3, Panel B) compared to non-placebo control (Plot including effect sizes prior to creation of shrunken estimates are included in Supplemental Digital Content 4 – Figure). The probability that the pooled effect size represented at least a small effect was p = 0.976, with the probability of at least a medium or large effect equal to p = 0.076 and p <0.001,

respectively. There was strong evidence of a differential effect between supplements ($\beta_{0.5;\text{Caffeine:Buffers}}$ = 0.33 [95%CrI: -0.03 to 0.68]), with buffering supplements ($ES_{0.5;\text{Buffers}}$ = 0.46 [95%CrI: 0.27 to 0.68]) producing substantially larger effects than caffeine ($ES_{0.5;\text{Caffeine}}$ = 0.13 [95%CrI: -0.14 to 0.44]). The median effect of intervention was reduced when compared to placebo (as opposed to versus non-placebo control), although remained between small and medium ($ES_{0.5}$ = 0.30 [0.13 to 0.48]). All coefficients for meta-regressions are presented in Table 2.

Proportion of total effect due to placebo effects

Given this difference in ergogenic effect, but somewhat consistent placebo effect, further analyses were completed with each supplement separately to estimate the percentage of each supplement effect that could be explained by placebo. Separating the effect size estimates according to supplement resulted in positive skews between supplementation type. Although larger absolute placebo effects were shown for buffers, the percentage of the effect that could be explained by placebo was estimated to be 25% [75%CrI: 16 to 35%]. For the smaller supplement effect of caffeine, the percentage of the effect that could be explained by placebo was estimated to be 59% [75%CrI: 34 to 94%]. Plots to visualise the relative placebo and intervention effects for each study are contained within supplemental figures (Supplemental Digital Content 5 for sodium bicarbonate and Supplemental Digital Content 6 for caffeine). A final sensitivity analysis was completed by removing the caffeine mouth rinse data and completing for ingested caffeine effects alone. The estimated effect for caffeine increased slightly ($ES_{0.5}$caffeine = 0.27 [95%CrI: -0.15 to 0.83]) and, as a result, the estimated percentage of the effect that could be explained by placebo reduced to 49% [75%CrI: 30 to 77%]. Visual inspection of funnel plots from both placebo and intervention effect sizes identified no clear signs of asymmetry and therefore of small study effects such as publication bias (Figure 4).

Discussion

The results of this meta-analysis of studies involving caffeine and buffering supplements showed that there is a very small, but significant placebo effect on exercise performance, which means that the real-world impact of these supplements is likely to be somewhat larger than commonly indicated by the scientific literature. There was evidence that the magnitude of the placebo effect may be influenced by supplement form with greater effects obtained when the placebo was presented as a solution compared with a capsule. The size of the placebo effect was very consistent across the ergogenic agents investigated, and although the absolute magnitude of the reported effect was very small, it was substantive when expressed relative to the overall ergogenic effect of the supplements. This was particularly apparent for the lower ergogenic effect of caffeine. These data confirm the notion that the use of a non-placebo control arm is influential in describing the total effect of a given ergogenic aid which, in the real world, is the net addition of the active intervention and its placebo effects.

The current novel data showed that double-blind placebo-controlled studies with ergogenic supplements such as caffeine and buffers result in very small ($ES_{0.5}$=0.09 [95%CrI:0.01 to 0.17]), but real, placebo effects on exercise outcomes. There was some evidence that the absolute magnitude of the placebo effect (placebo trial vs. non-placebo control trial) was different between studies investigating buffering supplements and caffeine, which confirms previous research showing that the size of placebo effects differs between ergogenic supplements (2); however, the form in which the supplement was provided appears more important. Not all placebos and placebo effects are created equally, and different characteristics such as colour, taste and administration form may interfere with the placebo effects attributed to an intervention (15-17). Placebo effects here were greater when supplements were provided in a solution as opposed to in capsules. To the authors' knowledge, these are the first data to suggest that placebo effects in sport may depend on

the form in which the supplements are administered. The reasons for these disparate effects are unclear but may relate to their taste since buffering and caffeine supplements dissolved in solution have distinct and strong flavours. Their respective placebo solutions are generally taste-matched to render indistinguishable any differences in flavour between the two solutions; however, the property of taste itself may inadvertently lead to ergogenic effects due to interaction with sweet and bitter taste receptors found in the mouth (27). The current results showed that placebo effects did not differ according to exercise capacity or performance tests, or between different exercise durations. Participants were separated according to training status since we speculated that differences in personality traits and placebo expectations between trained and non-trained individuals might lead to differential placebo effects. However, the magnitude of the placebo effect did not differ between trained and non-trained individuals. Had more detailed information relating to specific personality traits been collected, this might have rendered different and clearer insights. There was also some evidence that studies published prior to the year 2000 had larger placebo effects, this perhaps being indicative of improved blinding methods in more recent studies.

The size of the absolute placebo effects shown here are smaller than the small to moderate placebo effects on exercise outcomes reported previously (2). It is important to highlight, however, that the previous systematic review examined the magnitude of the placebo effect elicited by direct experimental manipulation wherein individuals were specifically led to believe that they were provided with the active treatment (2). All studies included herein consisted of double-blind randomised crossover studies in which individuals unknowingly received an active or placebo intervention, meaning any prior expectation of the participants will have been due to their own individual beliefs and given these double-blinded conditions, it seems likely that mixed expectations (positive, negative or neutral) will have been present (as per Saunders et al. (5)). An important limitation here is that we cannot make generalisations regarding what participants

believed they were ingesting during placebo or active trials. This is important since the psychosocial context of ingesting a substance, as well as verbal cues, expectancy, positivity, beliefs and preconditioning may elicit neurobiological effects, activating specific brain pathways that can affect subsequent exercise performance (1, 28, 29). We have previously shown that knowingly taking an active supplement can lead to further exercise improvements above the mean of the intervention, while knowingly taking a placebo can impair performance (5). Interestingly, openly provided placebos can improve exercise performance (30), although the beneficial effects seems related to the psychosocial context in which the placebo is provided (31). It is likely that when individuals believe they are taking a placebo in a double-blind context, the connotations are more negative since they know that they could have been taking the active supplement. Similarly, an individual who believes they are taking a substance that can improve their performance will likely have more positive reinforcement. Overall, we showed very small placebo effects associated with double-blind crossover supplementation studies, but manipulating an individual's expectation of the supplement may increase or decrease this effect, causing further placebo (*i.e.*, beneficial) or even nocebo (*i.e.*, detrimental) effects (2).

The percentage of the overall effect of the intervention (active intervention vs. non-placebo control) that can be attributed to placebo effects was high and was estimated to be higher for caffeine than for the extracellular buffers (59% vs. 25%). Although the absolute magnitude of the placebo effect with caffeine was slightly lower than for buffers, the larger contribution to the total effect was due to the substantially lower overall effect of caffeine supplementation in the studies included in the current meta-analysis. Removing caffeine mouth rinse studies increased the overall effect of caffeine and subsequently reduced the point estimate of the percentage explained by placebo effects from 59% to 49%. Effect sizes estimated for caffeine research on aerobic exercise (0.22-0.61) are generally higher than those for anaerobic exercise (0.16-0.20) (32). Since placebo

responses were consistent across different exercise modes and duration, it could be speculated that the proportion of the overall response attributed to placebo effects may be lower for endurance vs. anaerobic exercise with caffeine ingestion. This would be due to caffeine's greater overall effect on endurance performance (32), while placebo effects would accordingly likely comprise a very substantial part of caffeine's ergogenic effect on short-duration anaerobic exercise since the overall effects are smaller. This also means that these meta-analytical estimates (32), which were compared to placebo, are likely to be even higher had they been compared to a non-placebo control.

Our searches revealed a paucity of studies that incorporated non-placebo controls (11 caffeine and 24 buffer studies from over 200 original articles for each). Although the absolute placebo effects shown here were very small, they could, potentially, lead to substantial differences upon repeat exposure, such as during a training plus supplementation study. Considering the current data, where feasible we recommend that double-blind placebo-controlled studies should include a non-intervention control session to quantify the placebo effect associated with the treatment under examination, and calculate the entire effect associated with the active intervention plus placebo effects. This recommendation is in line with the consensus statement on placebo effects in sports and exercise to adopt methods that aim to quantify placebo effects that could explain some of the interindividual variability seen in response to interventions (33). However, it is acknowledged that this may not always be feasible as the inclusion of an additional non-placebo control may substantially add to the complexity and cost of designs, in particular those involving chronic supplementation strategies (e.g., creatine or beta-alanine) which would require an entire extra non-placebo control group. Participants should be questioned as to what intervention they believed they had taken to account for any effects of expectation on performance (5), and this is of particular importance in studies where a three-arm trial is not possible.

Caffeine and buffering supplementation studies were chosen here due to the well-known ergogenic and placebo effects associated with their ingestion (4, 5, 7), but these only served as a model, and results could be extrapolated to other effective nutritional supplements (3) or any alternative intervention or interaction that might elicit placebo effects (1). For example, the largest placebo effects of nutritional ergogenic aids on exercise performance were shown when individuals believed they were receiving banned performance enhancers like anabolic steroids (2). Placebo effects could also differ depending on the general and recognised efficacy of the supplement under investigation (*e.g.*, greater effects with known effective supplements) or with supplements with clear and obvious side-effects (*e.g.*, easier to determine when ingesting the placebo). It is unclear if placebo effects would be different for supplements which are considered less or ineffective, but the contribution of placebo effects to their overall effects would likely be greater than those shown here and may even represent their entire effect, although this is somewhat speculative. Thus, the proportion of the placebo contribution to the overall effect will likely depend on a multitude of factors including the exercise protocol, belief, expectation, preconditioning and the intervention itself. Also, studies have shown that certain personality traits might influence the placebo response (25, 26), including supplement use and beliefs. Further work should elucidate how much each of these factors can contribute to the placebo effect associated with double-blind placebo-controlled research designs.

The current meta-analysis showed that there is a real placebo effect associated with double-blind crossover studies involving caffeine and extracellular buffers, and these effects are greater when supplements are provided in solution compared to capsules. Although it does not change the statistically significant effect of these ergogenic aids (i.e. their effects vs. placebo remain significant), these placebo effects modify the overall size attributed to the intervention and accounts for a large proportion of their total effects. Coaches and practitioners should be aware

that actual ergogenic effects with these supplements are likely greater than those shown in placebo-controlled studies because in the real world, the net effect of the physiological effects of active interventions and their placebo effects are additive. Practitioners should also consider taking advantage of these placebo effects inherent to effective interventions by instilling positive belief in them to optimize training adaptations and performance outcomes. Nutritional supplementation studies should strive to employ a non-placebo control arm where possible as a comparator to measure the overall effect of the intervention (physiological and psychobiological) and identify the proportion of this effect that can be estimated to come from placebo effects. This approach will allow determination of the net additive effect of a given active intervention and its placebo effects, allowing practitioners and coaches to base their prescription on the totality of evidence for a real-world scenario.

Acknowledgements

No specific funding was received for writing this review. Felipe Marticorena (2019/20614-0), Eimear Dolan (2019/05616 and 2019/26899-6) and Bryan Saunders (2016/50438-0) are financially supported by Fundação de Amparo à Pesquisa do Estado de São Paulo. Bryan Saunders has received a grant from Faculdade de Medicina da Universidade de São Paulo (2020.1.362.5.2). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

Conflicts of Interest

None to declare. The results of the study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation. The present study does not constitute endorsement by ACSM

Author Contributions Statement

BS is responsible for the conception of the work. FMM, LFO and ED performed the searches. FMM, AC, LFO and ED performed the screening and FMM and LFO performed the data extraction. PS performed the meta-analysis. FMM and BS are responsible for the initial writing of the manuscript. PS and BG helped interpret the data and revised the manuscript. All authors approved the final version of the manuscript.

References

1. Davis AJ, Hettinga F, Beedie C. You don't need to administer a placebo to elicit a placebo effect: Social factors trigger neurobiological pathways to enhance sports performance. *Eur J Sport Sci*. 2020;20:302-12.
2. Hurst P, Schipof-Godart L, Szabo A et al. The placebo and nocebo effect on sports performance: A systematic review. *Eur J Sport Sci*. 2020;20:279-92.
3. Maughan RJ, Burke LM, Dvorak J et al. IOC consensus statement: dietary supplements and the high-performance athlete. *Br J Sports Med*. 2018;52(7):439-55.
4. Beedie CJ, Stuart EM, Coleman DA, Foad AJ. Placebo effects of caffeine on cycling performance. *Med Sci Sports Exerc*. 2006;38(12):2159-64.
5. Saunders B, de Oliveira LF, da Silva RP et al. Placebo in sports nutrition: a proof-of-principle study involving caffeine supplementation. *Scand J Med Sci Sports*. 2017;27(11):1240-7.
6. Foad AJ, Beedie CJ, Coleman DA. Pharmacological and psychological effects of caffeine ingestion in 40-km cycling performance. *Med Sci Sports Exerc*. 2008;40(1):158-65.
7. McClung M, Collins D. "Because I know it will!": placebo effects of an ergogenic aid on athletic performance. *J Sport Exerc Psychol*. 2007;29(3):382-94.
8. Hurst P, Schipof-Godart L, Hettinga F, Roelands B, Beedie C. Improved 1000-m Running Performance and Pacing Strategy With Caffeine and Placebo: A Balanced Placebo Design Study. *Int J Sports Physiol Perform*. 2019:1-6.
9. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. *Open Med*. 2009;3(3):e123-30.
10. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. *Systematic reviews*. 2016;5(1):210.
11. Saunders B, Elliott-Sale K, Artioli GG et al. β-alanine supplementation to improve exercise capacity and performance: a systematic review and meta-analysis. *Br J Sports Med*. 2017;51(8):658-69.
12. Saunders B, Elliott-Sale K, Artioli GG et al. Beta-alanine supplementation to improve exercise capacity and performance: a systematic review and meta-analysis. *Br J Sports Med*. 2017;51(8):658-69.
13. Hobson RM, Saunders B, Ball G, Harris R, Sale C. Effects Of beta-alanine Supplementation On Exercise Performance: A Meta-analysis. *Med Sci Sport Exer*. 2012;44:446-.

14. Berdi M, Koteles F, Hevesi K, Bardos G, Szabo A. Elite athletes' attitudes towards the use of placebo-induced performance enhancement in sports. *Eur J Sport Sci*. 2015;15(4):315-21.
15. Kong J, Spaeth R, Cook A et al. Are all placebo effects equal? Placebo pills, sham acupuncture, cue conditioning and their association. *PloS one*. 2013;8(7):e67485.
16. de Craen AJ, Tijssen JG, de Gans J, Kleijnen J. Placebo effect in the acute treatment of migraine: subcutaneous placebos are better than oral placebos. *J Neurol*. 2000;247(3):183-8.
17. de Craen AJ, Roos PJ, de Vries AL, Kleijnen J. Effect of colour of drugs: systematic review of perceived effect of drugs and of their effectiveness. *Bmj*. 1996;313(7072):1624-6.
18. Sterne JAC, Savovic J, Page MJ et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *Bmj*. 2019;366:l4898.
19. 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;52(6):377-84.
20. Morris SB, DeShon RP. Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychol Methods*. 2002;7(1):105-25.
21. Cox DR. *Principles of Statistical Inference*. Cambridge: Cambridge University Press; 2006.
22. Fernandez-Castilla B, Jamshidi L, Declercq L, Beretvas SN, Onghena P, Van den Noortgate W. The application of meta-analytic (multi-level) models with multiple random effects: A systematic review. *Behav Res Methods*. 2020;52:2031-52.
23. Sawilowsky SS. New Effect Size Rules of Thumb. *J Mod Appl Stat Method*. 2009;8(2):597-9.
24. Bürkner P-C. brms: An R Package for Bayesian Multilevel Models Using Stan. *J Stat Softw*. 2017;80(1):28.
25. Gelman A. Prior distributions for variance parameters in hierarchical models. Bayesian analysis. *Bayesian Analysis*. 2006;1(3):515-34.
26. Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian Data Analysis*. Taylor & Francis; 2014.
27. Best R, McDonald K, Hurst P, Pickering C. Can taste be ergogenic? *Eur J Nutr*. 2020. DOI: 10.1007/s00394-020-02274-5

28. Lundh L-G. Placebo, belief, and health. A cognitive–emotion model. *Scand J Psychol*. 1987;28(2):128-43.
29. Price DD, Finniss DG, Benedetti F. A comprehensive review of the placebo effect: recent advances and current thought. *Annu Rev Psychol*. 2008;59:565-90.
30. Saunders B, Saito T, Klosterhoff R et al. "I put it in my head that the supplement would help me": Open-placebo improves exercise performance in female cyclists. *PloS one*. 2019;14(9):e0222982.
31. Saito T, Barreto G, Saunders B, Gualano B. Is Open-Label Placebo a New Ergogenic Aid? A Commentary on Existing Studies and Guidelines for Future Research. *Sports Med*. 2020;50:1225–9.
32. Grgic J, Grgic I, Pickering C, Schoenfeld BJ, Bishop DJ, Pedisic Z. Wake up and smell the coffee: caffeine supplementation and exercise performance-an umbrella review of 21 published meta-analyses. *Br J Sports Med*. 2019;54:681-8.
33. Beedie C, Benedetti F, Barbiani D et al. Consensus statement on placebo effects in sports and exercise: The need for conceptual clarity, methodological rigour, and the elucidation of neurobiological mechanisms. *Eur J Sport Sci*. 2018;18(10):1383-9.
34. Bridge CA, Jones MA. The effect of caffeine ingestion on 8 km run performance in a field setting. *J Sports Sci*. 2006;24(4):433-9.
35. Clarke N, Baxter H, Fajemilua E. Coffee and Caffeine Ingestion Have Little Effect on Repeated Sprint Cycling in Relatively Untrained Males. *Sports*. 2016;4(3).
36. Dolan P, Witherbee KE, Peterson KM, Kerksick CM. Effect of Carbohydrate, Caffeine, and Carbohydrate + Caffeine Mouth Rinsing on Intermittent Running Performance in Collegiate Male Lacrosse Athletes. *J Strength Cond Res*. 2017;31(9):2473-9.
37. Flinn S, Gregory J, McNaughton LR, Tristram S, Davies P. Caffeine ingestion prior to incremental cycling to exhaustion in recreational cyclists. *Int J Sports Med*. 1990;11(3):188-93.
38. Karayiğit R, Yaşli BÇ, Karabiyik H, Koz M, Ersöz G. Effect of serial caffeine mouth rinse on wingate anaerobic performance. *SPORMETRE*. 2017;15(4):191-6.
39. Mc Naughton LR, Lovell RJ, Siegler JC, Midgley AW, Sandstrom M, Bentley DJ. The effects of caffeine ingestion on time trial cycling performance. *J Sports Med Phys Fitness*. 2008;48(3):320-5.
40. Wiles JD, Coleman D, Tegerdine M, Swaine IL. The effects of caffeine ingestion on performance time, speed and power during a laboratory-based 1 km cycling time-trial. *J Sports Sci*. 2006;24(11):1165-71.

41. Rezaei S, Akbari K, Gahreman DE et al. Caffeine and sodium bicarbonate supplementation alone or together improve karate performance. *J Int Soc Sports Nutr*. 2019;16(1):44.
42. Grgic J, Venier S, Mikulic P. Both caffeine and placebo improve vertical jump performance as compared to a non-supplement, control condition. *Int J Sport Physiol Perf*. in press. DOI: 10.1123/ijspp.2019-1028
43. Bird SR, Wiles J, Robbins J. The effect of sodium bicarbonate ingestion on 1500-m racing time. *J Sports Sci*. 1995;13(5):399-403.
44. Carr AJ, Slater GJ, Gore CJ, Dawson B, Burke LM. Reliability and effect of sodium bicarbonate: buffering and 2000-m rowing performance. *Int J Sports Physiol Perform*. 2012;7(2):152-60.
45. Coombes J, McNaughton LR. Effects of Bicarbonate Ingestion on Leg Strength and Power During Isokinetic Knee Flexion and Extension. *J Strength Cond Res*. 1993;7(4):241-9.
46. Coppoolse R, Barstow TJ, Stringer WW, Carithers E, Casaburi R. Effect of acute bicarbonate administration on exercise responses of COPD patients. *Med Sci Sports Exerc*. 1997;29(6):725-32.
47. Goldfinch J, Mc Naughton L, Davies P. Induced metabolic alkalosis and its effects on 400-m racing time. *Eur J Appl Physiol Occup Physiol*. 1988;57(1):45-8.
48. Griffen C, Rogerson D, Ranchordas M, Ruddock A. Effects of Creatine and Sodium Bicarbonate Coingestion on Multiple Indices of Mechanical Power Output During Repeated Wingate Tests in Trained Men. *Int J Sport Nutr Exerc Metab*. 2015;25(3):298-306.
49. Lindh AM, Peyrebrune MC, Ingham SA, Bailey DM, Folland JP. Sodium bicarbonate improves swimming performance. *Int J Sports Med*. 2008;29(6):519-23.
50. Materko W, Santos E, Novaes JS. Effect of Bicarbonate Supplementation on the Muscular Strength. *Med Sci Sport Exer*. 2008;11(5):25-33.
51. McLellan T, Jacobs I, Lewis W. Acute altitude exposure and altered acid-base states. II. Effects on exercise performance and muscle and blood lactate. *Eur J Appl Physiol Occup Physiol*. 1988;57(4):445-51.
52. McNaughton LR. Sodium citrate and anaerobic performance: implications of dosage. *Eur J Appl Physiol Occup Physiol*. 1990;61(5-6):392-7.

53. McNaughton L, Curtin R, Goodman G, Perry D, Turner B, Showell C. Anaerobic work and power output during cycle ergometer exercise: Effects of bicarbonate loading. *J Sport Sci*. 1991;9(2):151-60.
54. McNaughton LR. Bicarbonate ingestion: Effects of dosage on 60 s cycle ergometry. *J Sport Sci*. 1992;10(5):415-23.
55. McNaughton LR. Sodium bicarbonate ingestion and its effects on anaerobic exercise of various durations. *J Sports Sci*. 1992;10(5):425-35.
56. McNaughton L, Cedaro R. Sodium citrate ingestion and its effects on maximal anaerobic exercise of different durations. *Eur J Appl Physiol Occup Physiol*. 1992;64(1):36-41.
57. McNaughton LR, Ford S, Newbold C. Effect of Sodium Bicarbonate Ingestion on High Intensity Exercise in Moderately Trained Women. *J Strength Cond Res*. 1997;11(2):98-102.
58. McNaughton L, Dalton B, Palmer G. Sodium bicarbonate can be used as an ergogenic aid in high-intensity, competitive cycle ergometry of 1 h duration. *Eur J Appl Physiol Occup Physiol*. 1999;80(1):64-9.
59. Miller P, Robinson AL, Sparks SA, Bridge CA, Bentley DJ, McNaughton LR. The Effects of Novel Ingestion of Sodium Bicarbonate on Repeated Sprint Ability. *J Strength Cond Res*. 2016;30(2):561-8.
60. Morris DM, Shafer RS, Fairbrother KR, Woodall MW. Effects of lactate consumption on blood bicarbonate levels and performance during high-intensity exercise. *Int J Sport Nutr Exerc Metab*. 2011;21(4):311-7.
61. Oliveira LF, de Salles Painelli V, Nemezio K et al. Chronic lactate supplementation does not improve blood buffering capacity and repeated high-intensity exercise. *Scand J Med Sci Sports*. 2017;27(11):1231-9.
62. Painelli VS, Roschel H, Jesus F et al. The ergogenic effect of beta-alanine combined with sodium bicarbonate on high-intensity swimming performance. *Appl Physiol Nutr Metab*. 2013;38(5):525-32.
63. Pierce EF, Eastman NW, Hammer WH, Lynn TD. Effect of induced alkalosis on swimming time trials. *J Sports Sci*. 1992;10(3):255-9.
64. Tiryaki GR, Atterbom HA. The effects of sodium bicarbonate and sodium citrate on 600 m running time of trained females. *J Sports Med Phys Fitness*. 1995;35(3):194-8.
65. Wilkes D, Gledhill N, Smyth R. Effect of acute induced metabolic alkalosis on 800-m racing time. *Med Sci Sports Exerc*. 1983;15(4):277-80.

66. McGuinness LA, Higgins JPT. Risk-of-bias VISualization (robvis): An R package and Shiny web app for visualizing risk-of-bias assessments. *Research Synthesis Methods*. 2020;n/a(n/a). DOI: 10.1002/jrsm.1411

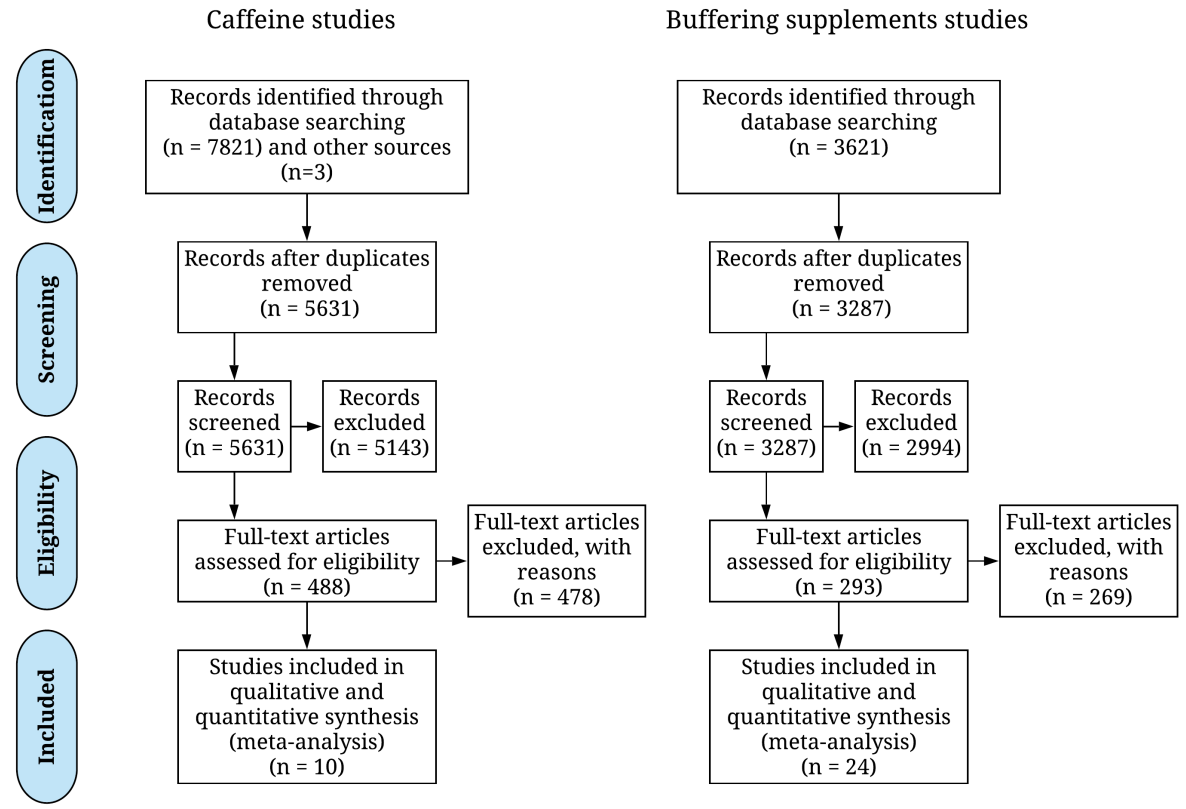
Figures Legend

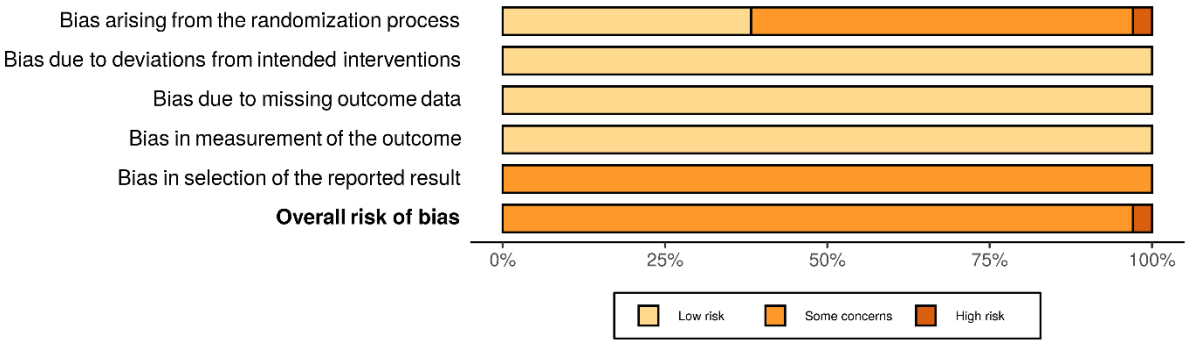
Figure 1. Flow chart of the search strategy and study selection.

Figure 2. Risk of bias presented as percentages across all included studies for the five main domains of evaluation. (Figure was created using *robvis* (66) and is in a colourblind-friendly colour scheme).

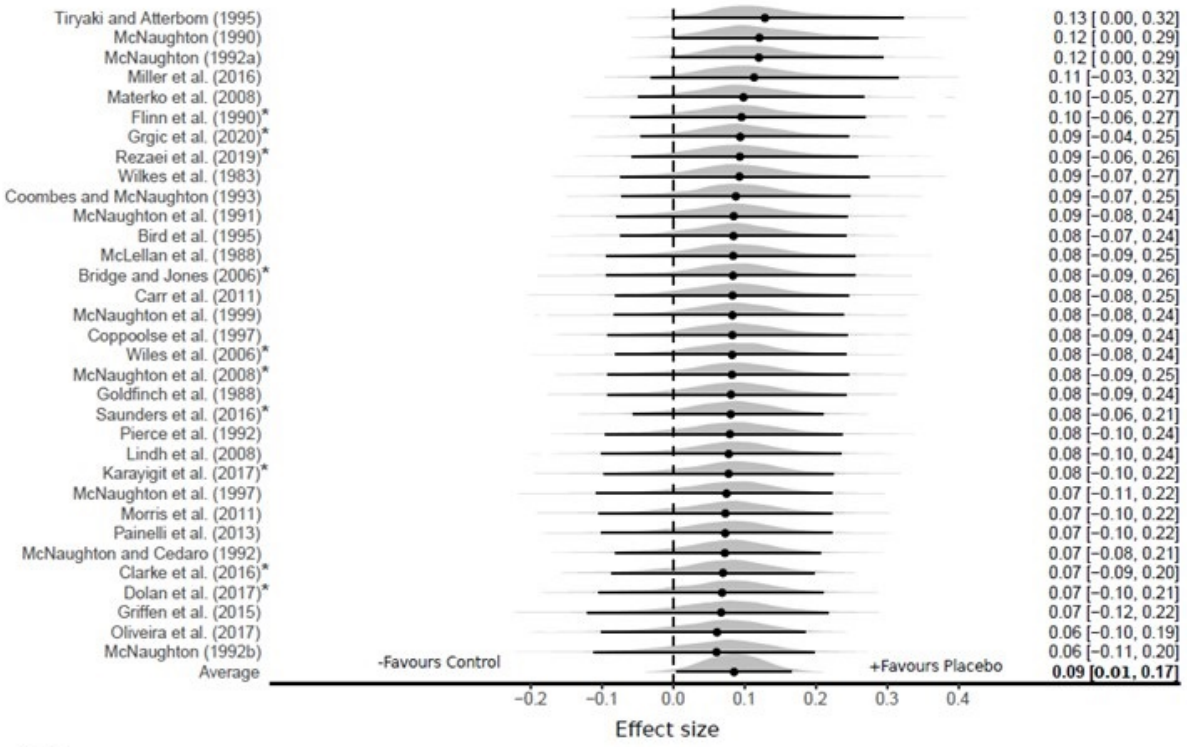
Figure 3. Forest plot of the effect sizes for the included buffering supplements and caffeine studies. Panel A shows effects sizes of placebo versus control and Panel B shows effect sizes of the active intervention versus control. Results from individual studies represent shrunken estimates based on the random effects model fitting and borrowing of information across studies to reduce uncertainty. Circles represent the pooled estimate from individual studies and across studies (average), generated with Bayesian inference along with the 95% credible intervals (95%CrI). Positive values favour the non-control condition. Included studies investigating caffeine are indicated with an asterisk (*).

Figure 4. Funnel plot of intervention (Panel A) and placebo (Panel B) effect sizes for the included buffering supplements and caffeine studies.

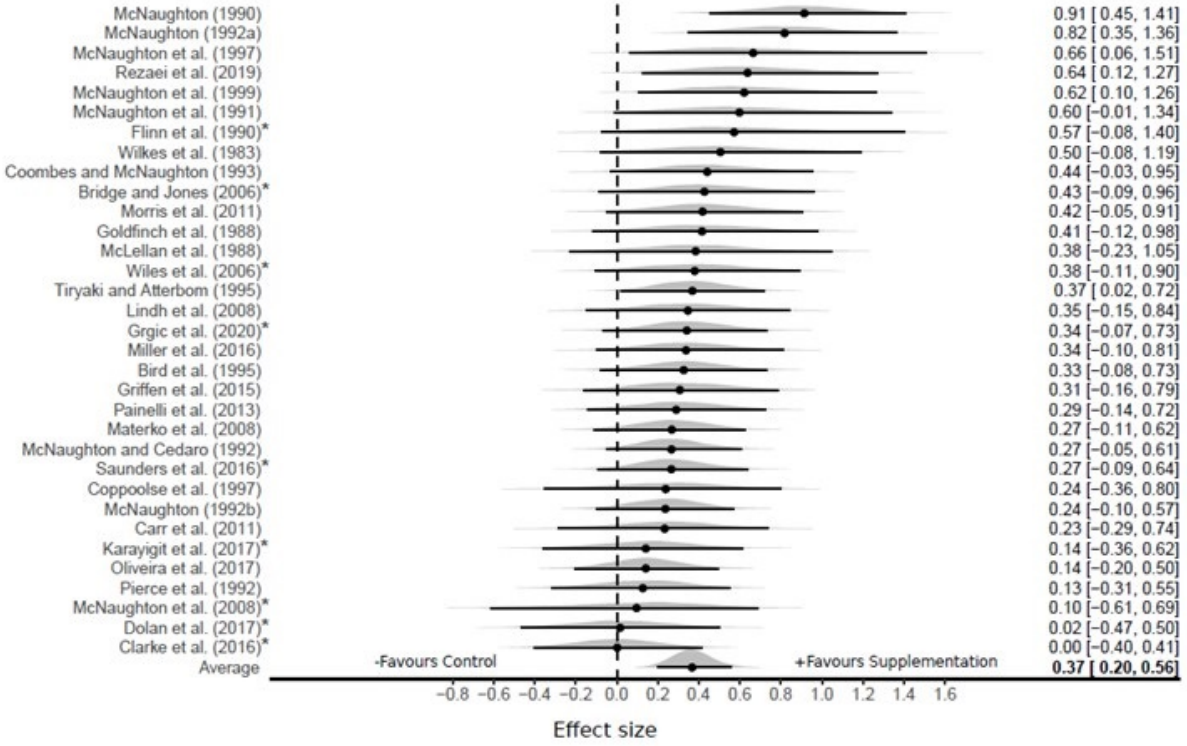

Caffeine studies
Buffering supplements studies
Identification
Screening
Eligibility
Included
Records identified through database searching (n = 7821) and other sources (n=3)
Records after duplicates removed (n = 5631)
Records screened (n = 5631)
Records excluded (n = 5143)
Full-text articles assessed for eligibility (n = 488)
Full-text articles excluded, with reasons (n = 478)
Studies included in qualitative and quantitative synthesis (meta-analysis) (n = 10)
Records identified through database searching (n = 3621)
Records after duplicates removed (n = 3287)
Records screened (n = 3287)
Records excluded (n = 2994)
Full-text articles assessed for eligibility (n = 293)
Full-text articles excluded, with reasons (n = 269)
Studies included in qualitative and quantitative synthesis (meta-analysis) (n = 24)

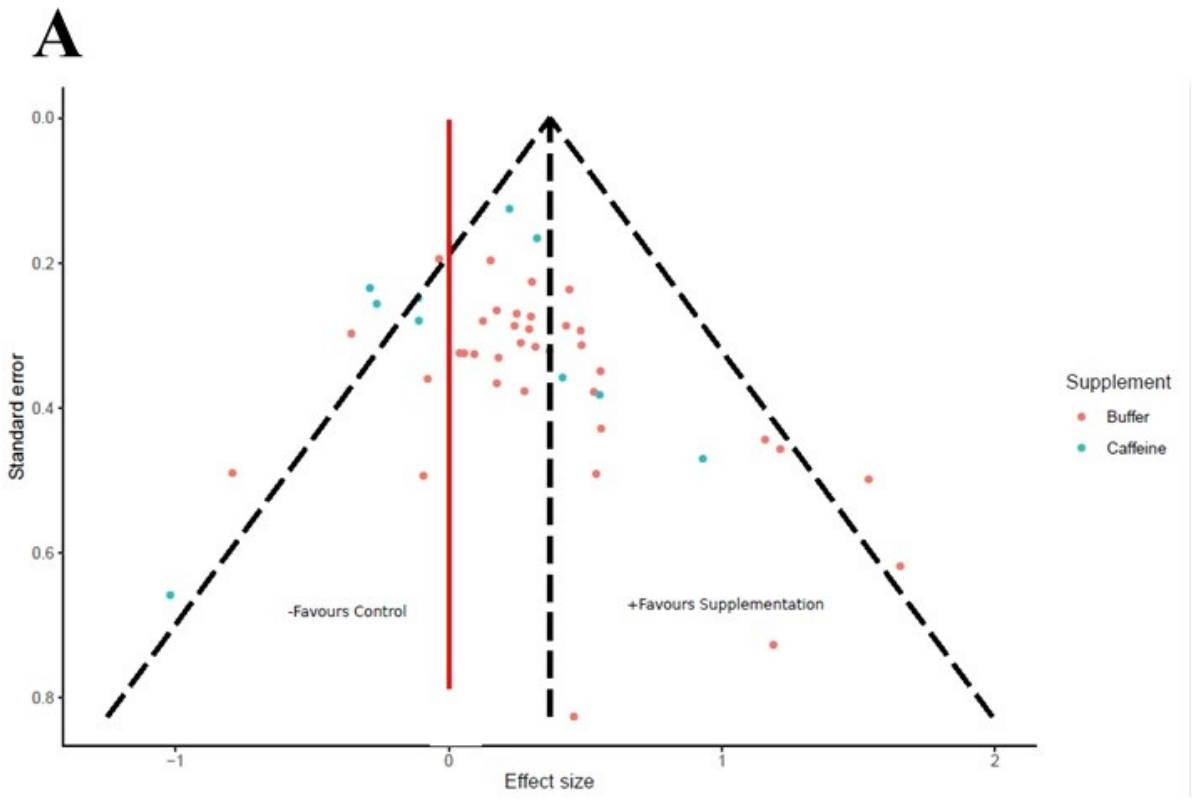

Bias arising from the randomization process
Bias due to deviations from intended interventions
Bias due to missing outcome data
Bias in measurement of the outcome
Bias in selection of the reported result
Overall risk of bias
0%
25%
50%
75%
100%
Low risk
Some concerns
High risk

A


Tiryaki and Atterbom (1995) 0.13 [0.00, 0.32]
McNaughton (1990) 0.12 [0.00, 0.29]
McNaughton (1992a) 0.12 [0.00, 0.29]
Miller et al. (2016) 0.11 [−0.03, 0.32]
Materko et al. (2008) 0.10 [−0.05, 0.27]
Flinn et al. (1990)* 0.10 [−0.06, 0.27]
Grgic et al. (2020)* 0.09 [−0.04, 0.25]
Rezaei et al. (2019)* 0.09 [−0.06, 0.26]
Wilkes et al. (1983) 0.09 [−0.07, 0.27]
Coombes and McNaughton (1993) 0.09 [−0.07, 0.25]
McNaughton et al. (1991) 0.09 [−0.08, 0.24]
Bird et al. (1995) 0.08 [−0.07, 0.24]
McLellan et al. (1988) 0.08 [−0.09, 0.25]
Bridge and Jones (2006)* 0.08 [−0.09, 0.26]
Carr et al. (2011) 0.08 [−0.08, 0.25]
McNaughton et al. (1999) 0.08 [−0.08, 0.24]
Coppoolse et al. (1997) 0.08 [−0.09, 0.24]
Wiles et al. (2006)* 0.08 [−0.08, 0.24]
McNaughton et al. (2008)* 0.08 [−0.09, 0.25]
Goldfinch et al. (1988) 0.08 [−0.09, 0.24]
Saunders et al. (2016)* 0.08 [−0.06, 0.21]
Pierce et al. (1992) 0.08 [−0.10, 0.24]
Lindh et al. (2008) 0.08 [−0.10, 0.24]
Karayigit et al. (2017)* 0.08 [−0.10, 0.22]
McNaughton et al. (1997) 0.07 [−0.11, 0.22]
Morris et al. (2011) 0.07 [−0.10, 0.22]
Painelli et al. (2013) 0.07 [−0.10, 0.22]
McNaughton and Cedaro (1992) 0.07 [−0.08, 0.21]
Clarke et al. (2016)* 0.07 [−0.09, 0.20]
Dolan et al. (2017)* 0.07 [−0.10, 0.21]
Griffen et al. (2015) 0.07 [−0.12, 0.22]
Oliveira et al. (2017) 0.06 [−0.10, 0.19]
McNaughton (1992b) 0.06 [−0.11, 0.20]
-Favours Control
+Favours Placebo
Average 0.09 [0.01, 0.17]
−0.2 −0.1 0.0 0.1 0.2 0.3 0.4
Effect size

B


McNaughton (1990) 0.91 [0.45, 1.41]
McNaughton (1992a) 0.82 [0.35, 1.36]
McNaughton et al. (1997) 0.66 [0.06, 1.51]
Rezaei et al. (2019) 0.64 [0.12, 1.27]
McNaughton et al. (1999) 0.62 [0.10, 1.26]
McNaughton et al. (1991) 0.60 [−0.01, 1.34]
Flinn et al. (1990)* 0.57 [−0.08, 1.40]
Wilkes et al. (1983) 0.50 [−0.08, 1.19]
Coombes and McNaughton (1993) 0.44 [−0.03, 0.95]
Bridge and Jones (2006)* 0.43 [−0.09, 0.96]
Morris et al. (2011) 0.42 [−0.05, 0.91]
Goldfinch et al. (1988) 0.41 [−0.12, 0.98]
McLellan et al. (1988) 0.38 [−0.23, 1.05]
Wiles et al. (2006)* 0.38 [−0.11, 0.90]
Tiryaki and Atterbom (1995) 0.37 [0.02, 0.72]
Lindh et al. (2008) 0.35 [−0.15, 0.84]
Grgic et al. (2020)* 0.34 [−0.07, 0.73]
Miller et al. (2016) 0.34 [−0.10, 0.81]
Bird et al. (1995) 0.33 [−0.08, 0.73]
Griffen et al. (2015) 0.31 [−0.16, 0.79]
Painelli et al. (2013) 0.29 [−0.14, 0.72]
Materko et al. (2008) 0.27 [−0.11, 0.62]
McNaughton and Cedaro (1992) 0.27 [−0.05, 0.61]
Saunders et al. (2016)* 0.27 [−0.09, 0.64]
Coppoolse et al. (1997) 0.24 [−0.36, 0.80]
McNaughton (1992b) 0.24 [−0.10, 0.57]
Carr et al. (2011) 0.23 [−0.29, 0.74]
Karayigit et al. (2017)* 0.14 [−0.36, 0.62]
Oliveira et al. (2017) 0.14 [−0.20, 0.50]
Pierce et al. (1992) 0.13 [−0.31, 0.55]
McNaughton et al. (2008)* 0.10 [−0.61, 0.69]
Dolan et al. (2017)* 0.02 [−0.47, 0.50]
Clarke et al. (2016)* 0.00 [−0.40, 0.41]
-Favours Control
+Favours Supplementation
Average 0.37 [0.20, 0.56]
−0.8 −0.6 −0.4 −0.2 0.0 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6
Effect size


A
Standard error
0.0
0.2
0.4
0.6
0.8
−1
0
1
2
Effect size
-Favours Control
+Favours Supplementation
Supplement
Buffer
Caffeine

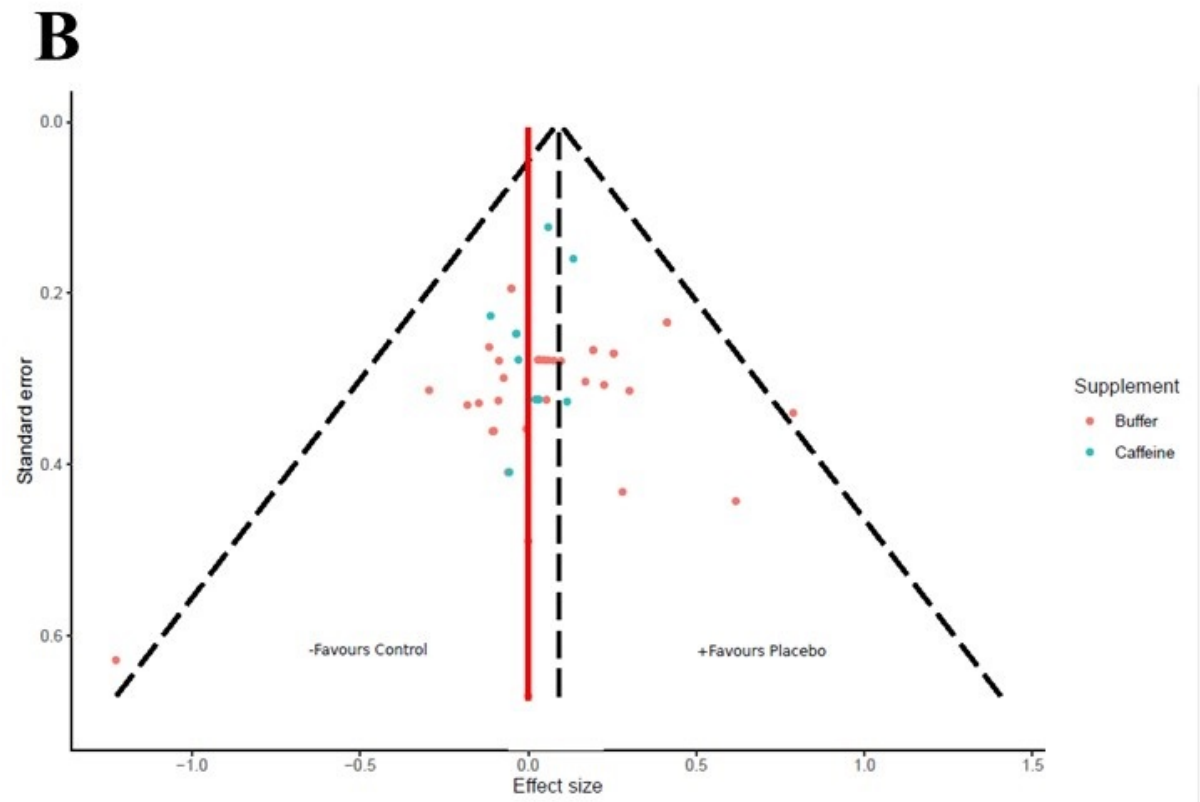

B
Standard error
0.0
0.2
0.4
0.6
−1.0
−0.5
0.0
0.5
1.0
1.5
Effect size
-Favours Control
+Favours Placebo
Supplement
Buffer
Caffeine

Table 1. Studies included in the meta-analysis.

Authors and location	Participants	Supplementation protocol	Exercise protocol(s)
CAFFEINE STUDIES			
Bridge and Jones (2006) UK	Trained male distance runners (n=8)	Capsule containing 3 mg/kg of caffeine or placebo (glucose) 60 min before exercise	8 km running TT
Clarke et al. (2016) UK	Recreationally active males (n=12)	Beverage containing 3 mg/kg of caffeine or placebo (flavoured water) 45 min before exercise	Eighteen x 4 s cycling sprints separated by 116 s recovery
Dolan et al. (2017) USA	Male lacrosse players (n=14)	Caffeine solution or placebo (flavoured water) mouth rinse 10 seconds before exercise	YoYo (Level 1)
Flinn et al. (1990) Australia	Recreational male cyclists (n=9)	Beverage containing 10 mg/kg of caffeine or placebo (p-flour) 180 min before exercise	Incremental cycle test
Karayiğit et al. (2017) Turkey	Physically active males (n=10)	8 x 25 ml mouth rinse with caffeine solution (2%) or placebo (water) at 30-s intervals during 5 min warm-up	30-s cycling Wingate anaerobic test
McNaughton et al. (2008) UK	Trained male cyclists (n=6)	Beverage containing 6 mg/kg of caffeine or placebo (flavoured water) 60 min before exercise	1-h cycling TT
Wiles et al. (2006) UK	Trained male cyclists (n=8)	Beverage containing 5 mg/kg of caffeine or placebo (flavoured water) 60 min before exercise	1-km cycling TT
Saunders et al. (2017) Brazil	Trained male cyclists (n=42)	Capsule containing 6 mg/kg of caffeine or placebo (dextrose) 60 min before exercise	30 min cycling TT

Rezaei et al. (2019) Iran	Trained males and females karatekas (n=8)	Capsule containing 6 mg/kg of caffeine or placebo (cellulose) 50 min before exercise	Karate Specific Aerobic Test
Grgic et al. (2020) Australia	Recreationally trained males (n=26)	Capsule containing 6 mg/kg of caffeine or placebo (dextrose) 60 min before exercise	Countermovement jump
BUFFERING SUPPLEMENT STUDIES			
Bird et al. (1995) UK	Male distance runners (n=10)	0.3 mg/kg of sodium bicarbonate solution or placebo (sodium chloride and calcium carbonate)	1500 m running
Carr et al. (2012) Australia	Well-trained rowers (n=8)	0.3 mg/kg of sodium bicarbonate in gelatine capsules or placebo (corn flour) 90 min before exercise	2000 m rowing ergometer TT
Coombes and McNaughton (1993) Australia	Healthy physical education university students (n=9)	0.3 mg/kg of sodium bicarbonate solution or placebo (calcium carbonate) 90 min before exercise	Isokinetic leg extension/flexion exercise
Coppoolse et al. (1997) USA	Healthy (n=5)	0.3 mg/kg of sodium bicarbonate solution or placebo 60 min before exercise	Cycling test with a work rate increment of 25 or 30 W/min
Goldfinch et al. (1988) Australia	Athletes (n=6)	0.4 mg/kg of sodium bicarbonate solution or placebo (calcium carbonate) 60 min before exercise	400 m run
Griffen et al. (2015) UK	Well-trained (n=9)	Chronic supplementation of 0.3 mg/kg of sodium bicarbonate solution or placebo (maltodextrin)	6 x 10 s cycling sprints 7.5% BM

Lindh et al. (2008) UK	Elite-standard swimmers (n=9)	0.3 mg/kg of sodium bicarbonate in gelatine capsules or placebo (calcium carbonate) 90 min before exercise	200m freestyle swim
Materko et al. (2008) Brazil	Strength trained (n=11)	0.3 mg/kg of sodium bicarbonate solution or placebo (sodium chloride) 120 min before exercise	Bench press test Pull press test
McLellan et al. (1988) Canada	Healthy (n=4)	Chronic supplementation of 0.2 mg/kg of sodium bicarbonate in gelatine capsules or placebo (calcium carbonate)	Cycling: 10 min at 50 and 70% and 90% of VO_{2max} until exhaustion
McNaughton (1990) Australia	Healthy (n=11)	0.1, 0.2, 0.3, 0.4 and 0.5 mg/kg of sodium citrate solution or placebo (calcium carbonate) 90 min before exercise	Maximal 1-min cycle effort
McNaughton et al. (1991) Australia	Cyclists (n=8)	0.4 mg/kg of sodium bicarbonate solution or placebo (calcium carbonate) 60 min before exercise	Maximal 1-min cycle effort
McNaughton (1992a) Australia	Healthy (n=9)	0.1, 0.2, 0.3, 0.4 and 0.5 mg/kg of sodium bicarbonate solution or placebo (calcium carbonate) 90 min before exercise	Maximal 1-min cycle effort
McNaughton (1992b) Australia	Healthy (n=8)	0.3 mg/kg of sodium bicarbonate solution or placebo (calcium carbonate) 90 min before exercise	Maximal 10-s cycle effort Maximal 30-s cycle effort Maximal 120-s cycle effort Maximal 240-s cycle effort

McNaughton and Cedaro (1992) Australia	Healthy (n=10)	0.5 mg/kg of sodium citrate solution or placebo (calcium carbonate) 90 min before exercise	Maximal 10-s cycle effort Maximal 30-s cycle effort Maximal 120-s cycle effort Maximal 240-s cycle effort
McNaughton et al. (1997) Australia	Physical active women (n=10)	0.3 mg/kg of sodium bicarbonate solution or placebo 90 min before exercise	Maximal 1-min cycle effort
McNaughton et al. (1999) UK	Cyclists (n=10)	0.3 mg/kg of sodium bicarbonate solution or placebo (sodium chloride) 90 min before exercise	60 min cycling
Miller et al. (2016) UK	Active team and individual sports (n=11)	0.3 mg/kg of sodium bicarbonate solution or placebo (sodium chloride)	10 x 6 s cycle sprints with 60 recovery
Morris et al. (2011) USA	Competitive cyclists (n=11)	0.1 mg/kg of lactate in gelatine capsules or placebo (aspartame) 90 min before exercise	Cycling test until exhaustion starting at 3 W/BM and increases of 0.3 W/BM
Oliveira et al. (2017) Brazil	Athletes of rugby, judo and jiu-jitsu at university level (n=18)	Chronic supplementation of 0.5 mg/kg/day of sodium bicarbonate or calcium lactate in gelatine capsules or placebo (calcium carbonate)	4 bouts of the 30-s Wingate upper body anaerobic test with 3 min recovery
Painelli et al. (2013) Brazil	Junior-standard swimmers (n=7)	0.3 mg/kg of sodium bicarbonate in gelatine capsules or placebo (dextrose) 90 min before exercise	100 m swimming 200 m swimming
Pierce et al. (1992) USA	Varsity swimmers (n=7)	0.2 mg/kg of sodium bicarbonate solution or placebo (sodium chloride) 60 min before exercise	100-yard (91,4 m) swim freestyle

			Individual 200-yard swims
Rezaei et al. (2019) Iran	Karatekas (n=8)	Chronic supplementation of 0.3 mg/kg of sodium bicarbonate in gelatine capsules or placebo (cellulose)	Karate Specific Aerobic Test
Tiryaki and Atterbom (1995) Turkey	Track athletes and non-athletes (n=15)	0.3 mg/kg of sodium bicarbonate, sodium citrate or placebo solution 120 min before exercise	600 m running test
Wilkes et al. (1983) Canada	Varsity track athletes (n=6)	0.3 mg/kg of sodium bicarbonate solution or placebo (calcium carbonate) 120 min before exercise	800 m run race

TT = time-trial

Table 2. Moderator analyses conducted for placebo and supplement effect sizes.

Moderator		Placebo vs. Non-placebo Parameter Estimate [95% CrI]	Supplement vs. Non-placebo Parameter Estimate [95% CrI]
[*Supplement type*]	Caffeine (n=11)	0.06 [-0.11 to 0.23]	0.13 [-0.14 to 0.44]
	Buffers (n=45)	0.12 [0.01 to 0.20]	0.46 [0.26 to 0.68]
[*Supplement delivery*]	Capsules (n=13)	0.02 [-0.12 to 0.16]	0.33 [0.04 to 0.66]
	Solution (n=41)	0.13 [0.03 to 0.23]	0.43 [0.23 to 0.70]
[*Training status*]	Untrained (n=25)	0.09 [-0.04 to 0.22]	0.46 [0.16 to 0.81]
	Trained (n=31)	0.09 [0.03 to 0.23]	0.33 [0.13 to 0.57]
[*Exercise duration*]	<30s (n=8)	0.08 [-0.12 to 0.28]	0.15 [-0.20 to 0.52]
	30s–10min (n=34)	0.09 [-0.02 to 0.20]	0.46 [0.25 to 0.20]
	+10min (n=14)	0.07 [-0.08 to 0.21]	0.31 [-0.01 to 0.68]
[*Exercise type*]	Capacity (n=10)	0.09 [-0.09 to 0.27]	0.33 [0.01 to 0.71]
	Performance (n=46)	0.09 [0.00 to 0.17]	0.38 [0.19 to 0.61]
[*Study year*]	Before 2000 (n=34)	0.14 [0.02 to 0.25]	0.58 [0.35 to 0.89]
	After 2000 (n=22)	0.03 [-0.08 to 0.15]	0.18 [-0.02 to 0.40]

n: Number of outcomes for factor level; CrI: Bayesian credible interval.

Supplemental Digital Content

Supplemental Digital Content 1 - .doc. – Table. PRISMA guidelines checklist

Supplemental Digital Content 2 - .doc. – Table. Extracted Study Data

Supplemental Digital Content 3 - .doc – Tables. Adapted Downs & Black questionnaire and scores

Supplemental Digital Content 4 - .jpg – Figure. Panel A: Bayesian Forest Plot illustrating effect sizes comparing placebo versus control. Panel B: Bayesian Forest Plot illustrating effect sizes comparing supplementation versus control. Distributions represent "shrunken estimates" based on all effect sizes obtained from the study, the random effects model fitted and borrowing of 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.

Supplemental Digital Content 5 - .jpg – Figure. Plots illustrating the relative intervention (top density) and placebo (bottom density) effects for sodium bicarbonate studies. Densities illustrate the shrunken posterior estimates for each study, with the x-axis representing the standardised mean difference.

Supplemental Digital Content 6 - .jpg – Figure. Plots illustrating the relative intervention (top density) and placebo (bottom density) effects for caffeine studies. Densities illustrate the shrunken posterior estimates for each study, with the x-axis representing the standardised mean difference.

Text S1 - Checklist of items to include when reporting a systematic review or meta-analysis

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4-5
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	6
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	6
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	6
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	6
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	6-7

Section/topic	#	Checklist item	Reported on page #
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	6-7
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	7-8
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	7-8
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	8
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	8-9
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	8-9
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	8-9
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	8-9
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	10
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	10
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome-level assessment (see Item 12).	10

Section/topic	#	Checklist item	Reported on page #
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group and (b) effect estimates and confidence intervals, ideally with a forest plot.	10-11
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	10-11
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	10-11
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression) (see Item 16).	10-11
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., health care providers, users, and policy makers).	12-15
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review level (e.g., incomplete retrieval of identified research, reporting bias).	12-15
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	15
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	16

Supplementary Table. Extracted Study Data

Author	Year	Supp	Direction	IMean	ISd	PMean	PSd	CMean	CSd	N
Bird (1995)	1995	SB	-1	252	10	257	13	258	13	10
Bird (1995)	1995	SB	-1	256	15	257	11	258	13	10
Carr (2011)	2011	SB	1	348	67	346	61	344	59	8
Coombes (1993)	1993	SB	1	11	1	10	1	10	2	9
Coppoolse (1997)	1997	SB	1	296	50	302	53	302	53	5
Goldfinch (1988)	1988	SB	-1	57	2	59	2	58	2	6
Griffen (2015)	2015	SB	1	8140	488	7818	488	7988	552	9
Lindh (2008)	2008	SB	-1	112	5	114	4	114	4	9
Materko (2008)	2008	SB	1	103	18	104	18	100	18	11
Materko (2008)	2008	SB	1	70	11	70	11	68	1	11
McLellan (1988)	1988	SB	1	12	4	10	2	10	2	4
McNaughton (1990)	1990	SB	1	37	3	36	4	35	3	11
McNaughton (1990)	1990	SB	1	39	2	36	4	35	3	11
McNaughton (1990)	1990	SB	1	41	2	36	4	35	3	11
McNaughton (1990)	1990	SB	1	41	1	36	4	35	3	11
McNaughton (1990)	1990	SB	1	45	2	36	4	35	3	11
McNaughton (1991)	1991	SB	1	9940	344	9288	278	9263	246	8
McNaughton (1992)	1992	SB	1	36	4	36	3	35	3	9
McNaughton (1992)	1992	SB	1	39	2	36	4	35	3	9
McNaughton (1992)	1992	SB	1	41	2	36	4	35	3	9
McNaughton (1992)	1992	SB	1	41	1	36	4	35	3	9
McNaughton (1992)	1992	SB	1	45	2	36	4	35	3	9
McNaughton (1992)	1992	SB	1	7	0	7	0	7	0	8
McNaughton (1992)	1992	SB	1	20	2	19	3	20	2	8
McNaughton (1992)	1992	SB	1	74	8	68	6	70	8	8
McNaughton (1992)	1992	SB	1	122	64	113	59	109	66	8
McNaughton (1992)	1992	SB	1	8	0	8	0	8	0	10
McNaughton (1992)	1992	SB	1	19	2	19	2	19	2	10
McNaughton (1992)	1992	SB	1	77	8	68	6	68	7	10
McNaughton (1992)	1992	SB	1	129	61	113	59	109	60	10
McNaughton (1997)	1997	SB	1	27	1	25	1	25	1	10
McNaughton (1999)	1999	SB	1	951	81	839	89	836	100	10
Miller (2016)	2016	SB	1	63	8	70	12	60	12	11
Morris (2011)	2011	SB	1	62	18	52	19	54	12	11
Oliveira (2017)	2017	SB	1	34244	5160	33192	4338	33436	4928	18
Oliveira (2017)	2017	SB	1	33244	5379	33192	4338	33436	4928	18
Painelli (2013)	2013	SB	-1	62	5	64	6	63	6	7
Painelli (2013)	2013	SB	-1	135	10	139	11	137	10	7
Pierce (1992)	1992	SB	-1	54	1	54	1	53	1	7
Pierce (1992)	1992	SB	-1	120	3	120	4	127	14	7
Pierce (1992)	1992	SB	-1	140	14	139	16	139	14	7
Tiryaki (1995)	1995	SB	-1	121	11	120	10	125	9	15
Tiryaki (1995)	1995	SB	-1	120	11	120	10	125	9	15
Wilkes (1983)	1983	SB	-1	123	2	125	2	126	2	6
Rezaei (2019)	2019	SB	1	689	13	632	17	628	15	8

Bridge (2006)	2006	Caf	-1	1901	43	1924	43	1926	35	8
Clarke (2016)	2016	Caf	1	873	172	887	119	892	143	12
Clarke (2016)	2016	Caf	1	862	44	887	119	892	143	12
Dolan (2017)	2017	Caf	1	1342	320	1397	360	1436	292	14
Flinn (1990)	1990	Caf	1	206	8	167	8	164	9	9
Karayigit (2017)	2017	Caf	1	646	74	652	74	655	68	10
McNaughton (2008)	2008	Caf	-1	28	1	26	2	26	2	6
Wiles (2006)	2006	Caf	1	523	43	505	46	504	38	8
Saunders (2016)	2016	Caf	1	234	37	228	38	226	38	42
Rezaei (2019)	2019	Caf	1	674	44	636	39	631	38	8
Grgic (2020)	2020	Caf	1	37	7	36	6	35	7	26

Data Legend: SB = Sodium Bicarbonate; Caf = Caffeine; I = Intervention; P = Placebo; C = Control; N = Sample Size.

Quality Appraisal – modified Downs & Black checklist.

Q1. Is the hypothesis/aim/objective of the study clearly described? (Yes = 1; No = 0)

Q2. Are the main outcomes to be measured clearly described in the introduction or methods section? If the main outcomes are first mentioned in the results section, answer no. (Yes = 1; No = 0).

Q3. Are the characteristics (age, height, weight, training status, healthy) of the participants included in the study clearly described? In observational studies, inclusion and/or exclusion criteria should be given. In case-control studies, inclusion and/or exclusion and the source of controls should be given. (Yes = 1; No = 0).

Q4. Are the interventions of interest clearly described? For exercise interventions, the type, intensity and duration should be described. If they provide a nutritional supplement the exact type, dose and duration should be provided. Treatments and placebo (where relevant) that are to be compared should be clearly described. (Yes = 1; No = 0).

Q5. Are the main findings of the study clearly described? Simple outcome data should be reported for all major findings so the reader can check the major analyses and conclusions. This does not cover statistical tests which are addressed in other questions. (Yes = 1; No = 0).

Q6. Does the study provide estimates of the random variability in the data for the main outcomes? In non-normal data, inter-quartile range should be reported. In normal data, standard deviation, standard error or confidence intervals should be reported. (Yes = 1; No = 0).

Q7. Have all important adverse events that may be a consequence of the intervention been reported. Answer yes if they confirm they have ethical approval (Yes = 1; No = 0).

Q8. Was at least one familiarization trial conducted prior to exercise testing? (Yes = 1; No = 0; Unable to determine = 0).

Q9. Were the exercise test conditions adequately standardised (factors including time of day; prior nutritional intake (including caffeine) and prior exercise)? (Yes (all relevant factors standardised) = 2; Yes (some relevant factors standardised) = 1; No = 0; Unable to determine = 0).

Q10. Was an attempt made to blind study subjects to the intervention they have received? For studies where the patients would have no way of knowing which intervention they received, this should be answered yes. (Yes = 1; No = 0; Unable to determine = 0).

Q11. Was the order of phase testing randomised or counterbalanced? (Yes = 1; No = 0; Unable to determine = 0).

Q12. Were the main outcome measures used accurate (valid and reproducible)? Answer yes for tests that have been externally validated (Yes = 1; No = 0; Unable to determine = 0).

Q13. Were statistical tests used to assess the main outcomes appropriate? The statistical techniques used must be appropriate to the data and the research question. (Yes = 1; No = 0; Unable to determine = 0).

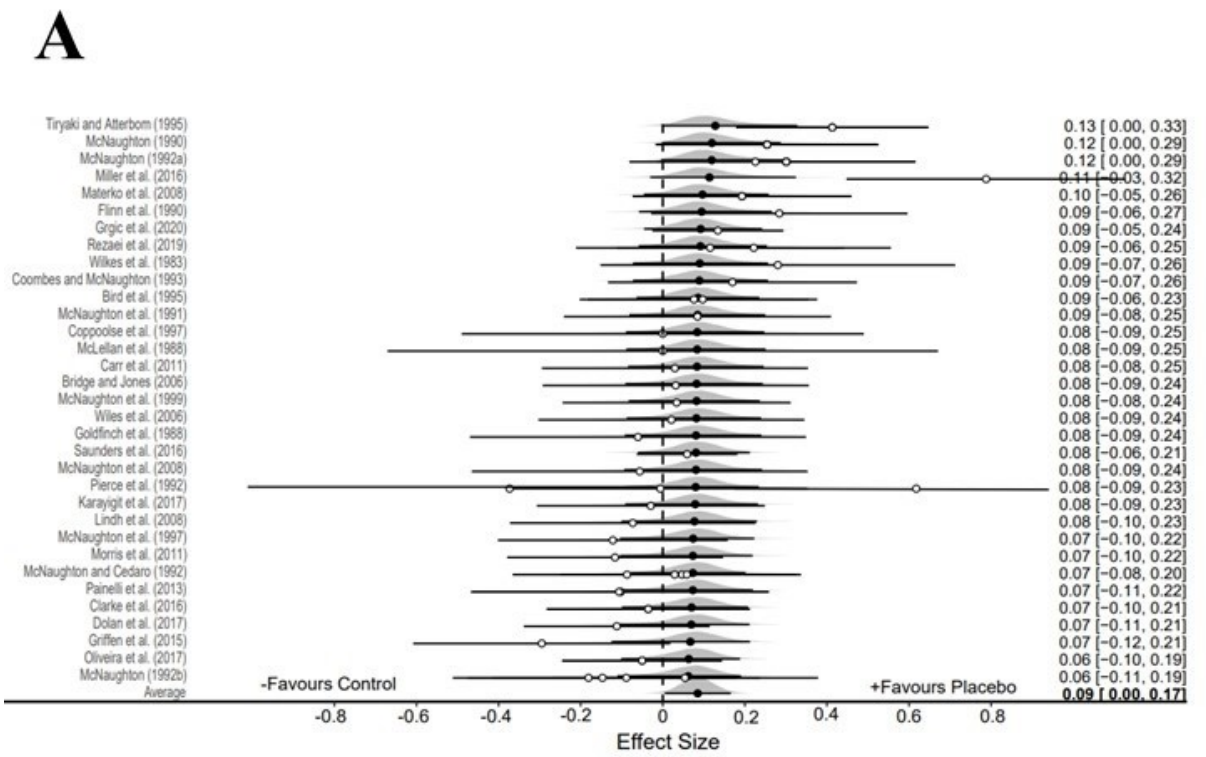
Q14. If any of the results of the study were based on 'data dredging' was this made clear? Any analyses that had not been planned at the outset should be clearly indicated. If no retrospective subgroup analyses were reported, then answer yes. (Yes = 1; No = 0; Unable to determine = 0).

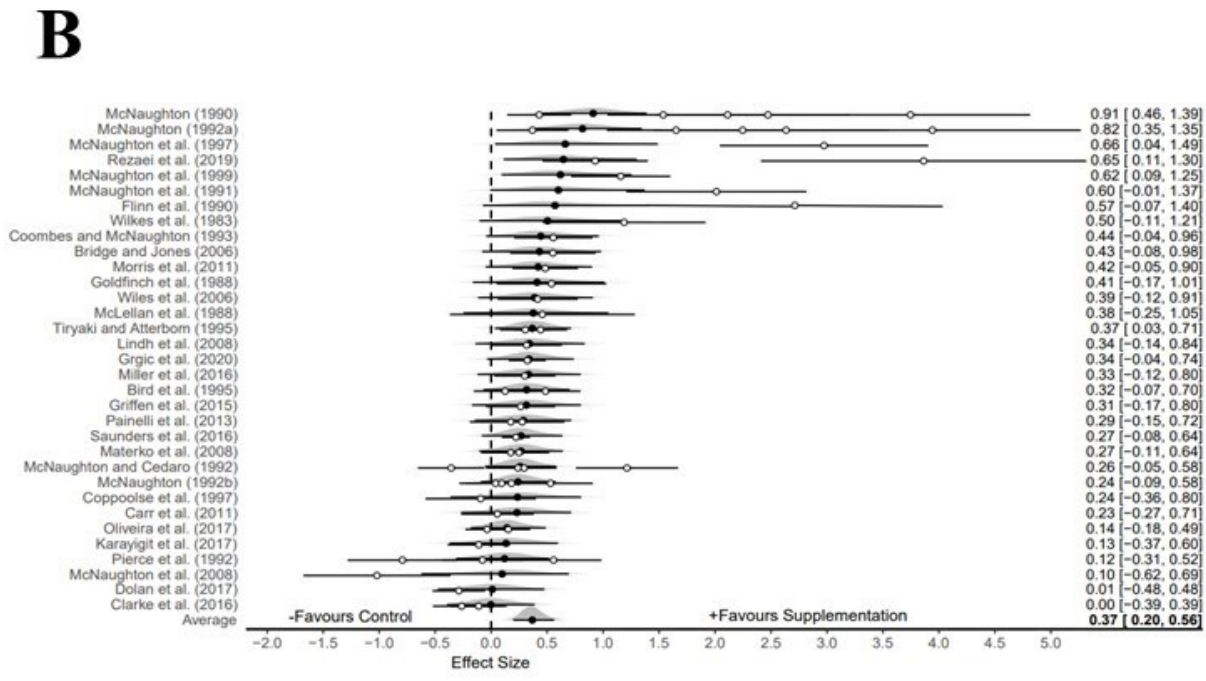
Note: The maximum attainable score for these studies was 14 and the categories were: High (12 – 14); Moderate (9 – 11); Low (6 – 8); Very Low (< 6).

TABLE 1. QUALITY APPRAISAL RESULTS FROM THE STUDIES INCLUDED IN THE META-ANALYSIS.

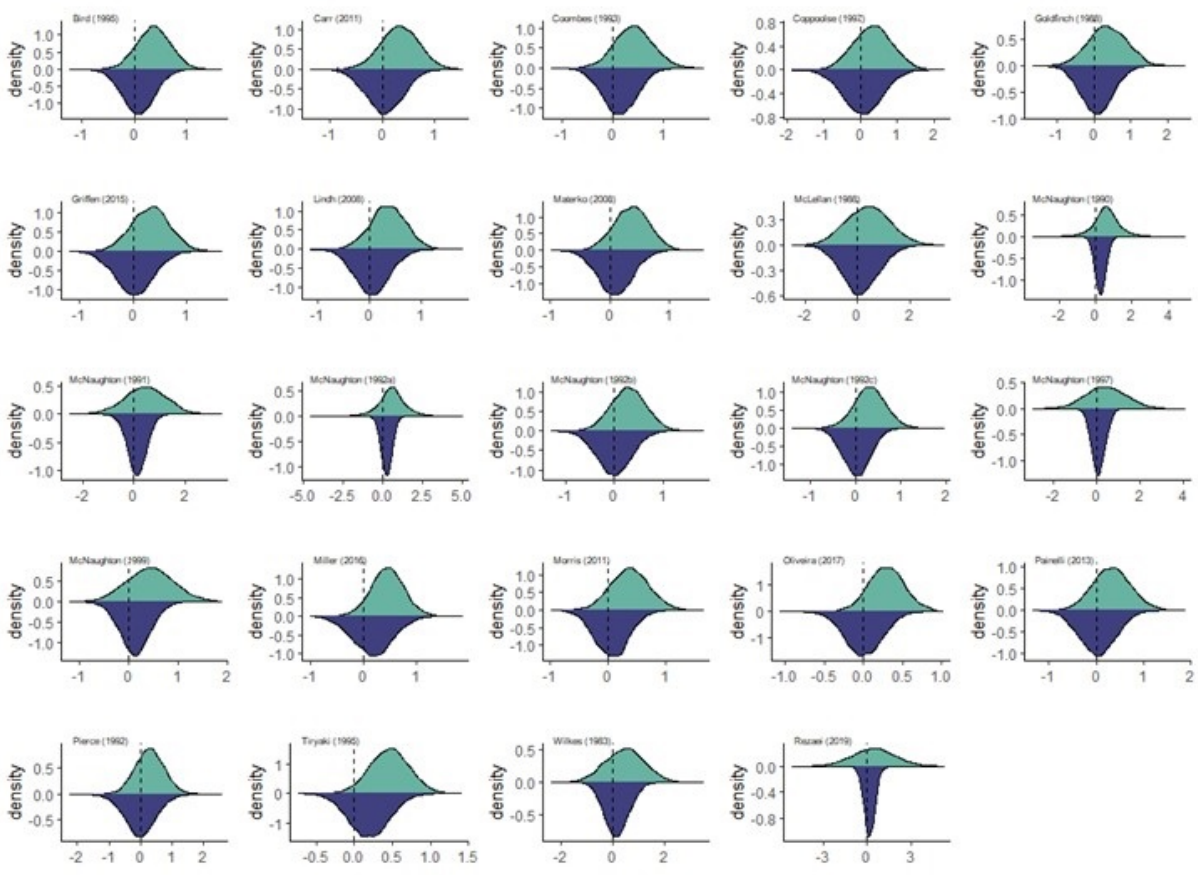
CAF	Author (year)	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Total
1	Bridge & Jones (2006)	1	1	1	1	1	0	1	0	1	1	1	1	1	1	**12**
2	Brietzke et al. (2017)	1	1	1	1	1	1	1	1	1	0	1	1	1	1	**13**
3	Clarke et al. (2016)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**14**
4	Dolan et al. (2017)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**14**
5	Flinn et al. (1990)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**14**
6	Karayigit et al. (2017)	1	1	1	1	1	1	0	1	1	1	1	1	1	1	**13**
7	McNaughton et al. (2008)	1	1	1	1	1	1	1	0	1	1	1	1	1	1	**13**
8	Wiles et al. (2006)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**14**
9	Saunders et al. (2017)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**14**
10	Rezaei et al. (2019)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**14**
11	Grgic et al. (2020)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**14**
Buffers																
1	Bird et al. (1995)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**14**
2	Carr et al. (2011)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**14**
3	Coombes et al. (1993)	1	1	1	1	1	1	1	0	1	1	1	1	1	1	**13**
4	Coppoolse et al. (1997)	1	1	1	1	1	1	1	1	1	0	1	1	1	1	**13**
5	Goldfinch et al. (1988)	1	1	1	1	1	1	1	0	1	1	1	1	1	1	**13**
6	Griffen et al. (2015)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**14**
7	Lindh et al. (2008)	1	1	1	1	1	1	1	0	1	1	1	1	1	1	**13**
8	Materko et al. (2008)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**14**

9	McLellan et al. (1988)	1	1	1	1	1	1	1	0	1	0	1	1	1	1	**12**
10	McNaughton (1990)	1	1	1	1	1	1	1	0	1	1	1	1	1	1	**13**
11	McNaughton et al. (1991)	1	1	1	1	1	1	1	0	1	1	1	1	1	1	**13**
12	McNaughton (1992)	1	1	1	1	1	1	1	0	1	1	1	1	1	1	**13**
13	McNaughton (1992)	1	1	1	1	1	1	1	0	1	1	1	1	1	1	**13**
14	McNaughton & Cedaro (1992)	1	1	1	1	1	1	1	0	1	1	1	1	1	1	**13**
15	McNaughton et al. (1997)	1	1	1	1	1	1	1	0	1	1	1	1	1	1	**13**
16	McNaughton et al. (1999)	1	1	1	1	1	1	1	0	1	1	1	1	1	1	**13**
17	Miller et al. (2016)	1	1	1	1	1	1	1	0	1	1	1	1	1	1	**13**
18	Morris et al. (2011)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**14**
19	Oliveira et al. (2017)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**14**
20	Painelli et al. (2013)	1	1	1	1	1	1	1	0	1	1	1	1	1	1	**13**
21	Pierce et al. (1992)	1	1	1	1	1	1	1	0	1	1	1	1	1	1	**13**
22	Tiryaki et al. (1995)	1	1	1	1	1	1	1	0	1	0	1	1	1	1	**12**
23	Wilkes et al. (1983)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**14**
24	Rezaei et al. (2019)	1	1	1	1	1	1	1	0	1	1	1	1	1	1	**13**

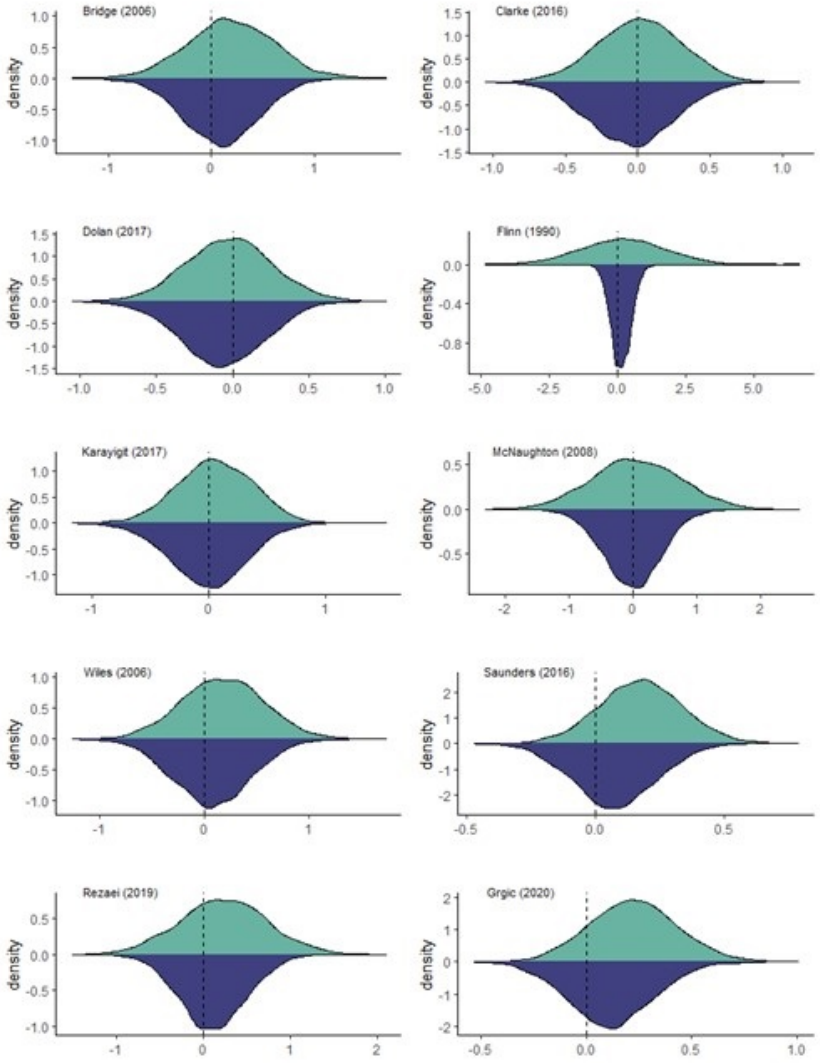




Supplemental Digital Content 4 - .jpg – Figure. Panel A: Bayesian Forest Plot illustrating effect sizes comparing placebo versus control. Panel B: Bayesian Forest Plot illustrating effect sizes comparing supplementation versus control. Distributions represent "shrunken estimates" based on all effect sizes obtained from the study, the random effects model fitted and borrowing of 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.



Supplemental Digital Content 5 - .jpg – Figure. Plots illustrating the relative intervention (top density) and placebo (bottom density) effects for sodium bicarbonate studies. Densities illustrate the shrunken posterior estimates for each study, with the x-axis representing the standardised mean difference.



Supplemental Digital Content 6 - .jpg – Figure. Plots illustrating the relative intervention (top density) and placebo (bottom density) effects for caffeine studies. Densities illustrate the shrunken posterior estimates for each study, with the x-axis representing the standardised mean difference.