Resistance training prescription for muscle strength and hypertrophy in healthy adults: a systematic review and Bayesian network meta-analysis


ABSTRACT

Objective To determine how distinct combinations of resistance training prescription (RTx) variables (load, sets and frequency) affect muscle strength and hypertrophy. Data sources MEDLINE, Embase, Emcare, SPORTDiscus, CINAHL, and Web of Science were searched until February 2022. Eligibility criteria Randomised trials that included healthy adults, compared at least 2 predefined conditions (non-exercise control (CTRL) and 12 RTx, differentiated by load, sets and/or weekly frequency), and reported muscle strength and/or hypertrophy were included. Analyses Systematic review and Bayesian network meta-analysis methodology was used to compare RTxs and CTRL. Surface under the cumulative ranking curve values were used to rank conditions. Confidence was assessed with threshold analysis. Results The strength network included 178 studies (n=5097; women=45%). The hypertrophy network included 119 studies (n=3364; women=47%). All RTxs were superior to CTRL for muscle strength and hypertrophy. Higher-load (>80% of single repetition maximum) prescriptions maximised strength gains, and all prescriptions comparatively promoted muscle hypertrophy. While the calculated effects of many prescriptions were similar, higher-load, multiset, thrice-weekly training (standardised mean difference (95% credible interval); 1.60 (1.38 to 1.82) vs CTRL) was the highest-ranked RTx for strength, and higher-load, multiset, twice-weekly training (0.66 (0.47 to 0.85) vs CTRL) was the highest-ranked RTx for hypertrophy. Threshold analysis demonstrated these results were extremely robust. Conclusion All RTx promoted strength and hypertrophy compared with no exercise. The highest-ranked prescriptions for strength involved higher loads, whereas the highest-ranked prescriptions for hypertrophy included multiple sets. PROSPERO registration number CRD42021259663 and CRD42021258902.

INTRODUCTION

Skeletal muscle is critical for numerous functional and metabolic processes essential to good health. Resistance training (RT), muscle contraction against external weight, potently increases muscle strength and mass (hypertrophy), improves physical performance, provides a myriad of metabolic-health benefits and combats chronic disease risk.1-4 Although endogenous biological and physiological factors are pertinent to maximising RT-induced skeletal muscle adaptations,5-6 RT programming variables can affect RT adaptations.7-13 Therefore, a RT prescription (RTx) should be determined appropriately. Each RTx is comprised of a distinct combination of RT variables, and the most-studied RTx variables include the load lifted per repetition, sets per exercise (generally

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Resistance training with varying numbers of variables (load, sets, weekly frequency) potently increases muscle strength and mass.
⇒ Resistance training prescription involves multiple variables, but the optimal resistance training prescription remains contentious.
⇒ Network meta-analysis allows simultaneous comparisons between multiple resistance training prescriptions.

WHAT THIS STUDY ADDS

⇒ This network meta-analysis is the largest synthesis of resistance training prescription data from randomised trials.
⇒ All resistance training prescriptions are better than no exercise for strength and hypertrophy in healthy adults.
⇒ The top-ranked prescriptions for strength were characterised by higher loads and the top-ranked prescriptions for hypertrophy were characterised by multiple sets.
⇒ All resistance training prescriptions increased strength and hypertrophy, suggesting that healthy adults can adopt a resistance training prescription of their choice and preference.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Since all protocols increased strength and hypertrophy, rather than determining an ‘optimal’ protocol, future work could determine minimal ‘doses’ of resistance exercise and practices to promote engagement and adherence in this health-promoting form of exercise.
Systematic review

involving a single RT manoeuvre or muscle group) and weekly frequency (the number of RT sessions completed per week).

Guideline developers rely on systematic reviews and meta-analyses for determining recommendations, as these study designs are, in most cases, the most robust forms of evidence. Indeed, various meta-analyses have provided seminal evidence on the univariate impact of load, sets or frequency to improve muscle strength, mass and physical function. However, these univariate analyses limit RT guideline development because individual RT variables are neither mutually exclusive nor prescribed independently; rather, several variables are collectively inherent to any RTx. Comparisons between multivariate RT prescriptions are needed to advance optimal RTx guidelines.

Pairwise meta-analyses are methodologically constrained to only comparing two RTxs. Several RTxs are conceivable, and multiple pairwise meta-analyses are unlikely to yield congruent insights. Network meta-analysis (NMA) expands on pairwise meta-analysis by permitting the simultaneous comparison of multiple treatments. NMA leverages direct and indirect evidence to produce enhanced effect estimates between all treatments, even when some comparisons have never been tested in randomised trials. Additionally, NMA permits the rank-ordering of all included treatments and the incorporation of data from multi-arm trials. Within exercise science, NMA has been used to compare different types of exercise, within RT, NMA has only been used to compare different load doses. Importantly, NMA can compare several multivariate RTxs.

The purpose of this systematic review and NMA was to determine how different RTxs affect muscle strength, hypertrophy and physical function in healthy adults. Specifically, we sought to compare distinct combinations of RTx variables—load, sets and frequency—and non-exercising control groups. For each outcome, we used NMA to integrate data from randomised trials.

METHODS
Protocol and registration
This review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses extension statement for network meta-analyses (PRISMA-NMA) and Cochrane Handbook for Systematic Reviews of Interventions. The PRISMA-NMA checklist is provided in online supplemental appendix 1. This review combines NMAs registered in the International Prospective Register of Systematic Reviews (https://www.crd.york.ac.uk/PROSPERO/).

Eligibility criteria
The eligibility criteria are detailed in table 1. Only trials that included healthy adults ≥18 years old, were randomised, compared at least 2 of 13 unique conditions (box 1), and measured muscle strength, size and/or physical function were included. Physical function was subdivided into three domains: mobility, the ability to physically move; balance, the ability to maintain a body position during a task; and gait speed, the time taken to locomote over a given distance. Trials that included athletes, persons with comorbidities or military persons; spanned <6 weeks; involved unsupervised RT (eg, home-based exercise); were reported in a non-English language; or were non-randomised were excluded.

Condition coding framework
Arms of included studies were classified as 1 of 12 RTxs or non-exercise control (CTRL). Each RTx was classified based on the load, set and frequency prescription (box 1). RTxs were denoted with a three-character acronym—XY#—where X is load (H, ≥80% one-repetition maximum (1RM); L, <80% 1RM); Y is sets (M, multiset; S, single-set); and # is the weekly frequency (3, ≥3 days/week; 2, 2 days/week; 1, 1 day/week), respectively. For example, HM2 denotes higher-load, multiset, twice-weekly RT within this framework. CTRL was comprised of subjects who received no intervention.

Search strategy
MEDLINE, Embase, Emcare, SPORTDiscus, CINAHL and Web of Science were systematically searched until 7 February 2022. Multiple experts developed the search strategy, which included subject headings and keywords specific to the research question and each database. No language nor study design limits were used in the search strategy. The complete search strategy is provided in online supplemental appendix 2. Relevant systematic reviews (online supplemental appendix 3) were manually selected, and the references were scrutinised for eligibility.

Study selection and data extraction
All records underwent title/abstract screening by two independent reviewers, with discrepancies resolved by a third reviewer. The full text of potentially eligible reports was then assessed for inclusion by two independent reviewers, with discrepancies resolved by a third reviewer. Reports deemed eligible for inclusion then underwent data extraction.

Data from included studies were extracted independently by pairs of reviewers, with any discrepancies resolved by consensus with a third reviewer (BSC or JCM). Extracted data included study and participant characteristics, RTx details and measurements of muscle strength and/or size (online supplemental appendix 4). Measures of mobility, balance and/or gait speed were extracted when the mean participant age was ≥55 years old. Authors of studies with missing data were contacted twice with a request for the missing data. The systematic review software Covidence (Veritas Health Innovation, Melbourne, Australia. Available at www.covidence.org) was used for record screening and data extraction.

Mean change from baseline and SD change (SDchange) from baseline were the outcomes of interest and extracted when reported. When unreported, SD was calculated with SEs, CIs, p values or t-statistics, and SDchange was imputed from pre-SD and post-SD values with a correlation coefficient of 0.5. RT loads reported as repetition maximum (RM) were converted to a percentage of one-repetition maximum (%1RM) with the equation: %1RM = 100−(RM(2.5)). The highest-ranked measurement was extracted, per predetermined hierarchy (online supplemental appendix 5), when multiple measurements were reported for a single outcome (eg, MRI and ultrasonography for muscle size). The longest period that all conditions were unchanged from baseline was analysed when the outcome(s) of interest were measured at multiple time points. Cohorts randomised separately but reported together (eg, young and old) were analysed independently. Within-group outcomes reported by participant sex were grouped by condition.

Risk of bias
Reviewers independently evaluated the within-study risk of bias using the Cochrane Risk of Bias V2.0. tool. Signalling questions and criteria were followed to inform the risk of bias appraisals for the intention-to-treat effect. Articles were assessed in duplicate at the strength and hypertrophy outcome level for bias: (1)
arising from the randomisation process, (2) due to deviations from intended interventions, (3) due to missing outcome data, (4) in the measurement of the outcome and (5) in the selection of reported result. Every domain was determined to be of high, moderate (some concerns) or low risk of bias, and studies were subsequently given an overall classification of high, moderate or low risk of bias. Any disagreement was resolved by consensus (BSC and JCM).

**Statistical analysis**

Standardised mean differences (SMD), adjusted for small-sample size bias, were calculated as the summary statistic because each outcome was measured with various tools. The direction of effect was standardised to analyse mobility, gait speed and balance to ensure consistency of desirable outcomes. When multiple studies compared two conditions, random-effects pairwise meta-analyses were conducted to identify comparison-level heterogeneity, publication bias, outliers and influential cases. To account for within-trial correlations in multi-arm trials (≥3 conditions), the SE in the base/reference arm was calculated as the square root of the covariance between calculated effects, assuming a correlation of 0.5 between effect sizes.

NMA integrated all direct evidence, with one network meta-analysis review; meta-analyses review; meta-regressions; umbrella reviews; network meta-analyses). Narrative reviews. Observational studies (eg, retrospective, prospective or longitudinal).

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Study inclusion and exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inclusion criteria</td>
<td>Exclusion criteria</td>
</tr>
<tr>
<td>Population</td>
<td>Population</td>
</tr>
<tr>
<td>► Humans ≥18 years old.</td>
<td>► Non-human species.</td>
</tr>
<tr>
<td>► Generally healthy (no disease condition indicated other than sarcopenia).</td>
<td>► ≤18 years old.</td>
</tr>
<tr>
<td>► Community-dwelling adults.</td>
<td>► Persons with or at risk for comorbidities (eg, cardiovascular disease, type II diabetes, type I diabetes, cancer, peripheral artery disease, osteoarthritis).</td>
</tr>
<tr>
<td>Intervention</td>
<td>Intervention</td>
</tr>
<tr>
<td>► Upper-body, lower-body and/or whole-body resistance training.</td>
<td>► Persons that are injured (eg, musculoskeletal-related fracture and/or repair).</td>
</tr>
<tr>
<td>► RTx aligns with one predefined node; specifically, exercises performed:</td>
<td>► Athletes or military personnel.</td>
</tr>
<tr>
<td>with high(H; ≥80% 1 RM or ≥8 RM) or low(L; &lt;80% 1 RM or &gt;8 RM) load, AND</td>
<td>► Explicitly mentions obese and/or overweight participants.</td>
</tr>
<tr>
<td>for a single (S) or multiple (M) sets, AND</td>
<td>► Individuals that are hospitalised (inpatient/outpatient rehabilitation).</td>
</tr>
<tr>
<td>once-weekly (1), twice-weekly (2) or at least thrice-weekly (3).</td>
<td>► Individuals living in long-term care homes.</td>
</tr>
<tr>
<td>Intervention duration ≥6 weeks.</td>
<td>Intervention</td>
</tr>
<tr>
<td>Comparison.</td>
<td>► Resistance training involved added intervention (eg, blood flow restriction).</td>
</tr>
<tr>
<td>► RTx variable (load, sets or frequency) differentially prescribed between training groups</td>
<td>► RTx does not align with one node (eg, load 60–90% 1RM).</td>
</tr>
<tr>
<td>► Eligible RTx compared with CTRL.</td>
<td>► Explicitly mentions unsupervised resistance training.</td>
</tr>
<tr>
<td>Outcome</td>
<td>Outcome</td>
</tr>
<tr>
<td>► Eligible outcome(s) assessed pre-intervention and post-intervention.</td>
<td>► Resistance training familiarisation/lead-in &gt;4 weeks.</td>
</tr>
<tr>
<td>Muscle strength:</td>
<td>► CTRL received treatment beyond habitual lifestyle (eg, nutritional advice, lifestyle consultation).</td>
</tr>
<tr>
<td>► 1RM test.</td>
<td>Comparison</td>
</tr>
<tr>
<td>► Isometric maximum voluntary contraction.</td>
<td>► Eligible RTx not compared with another eligible RTx nor CTRL.</td>
</tr>
<tr>
<td>► Isokinetic maximum voluntary contraction.</td>
<td>Outcome</td>
</tr>
<tr>
<td>Muscle size: Fat-free mass, fat-free and bone-free mass, lean mass, whole-muscle cross-sectional area or volume or thickness of muscle fibres cross-sectional area. Eligible measurement instruments:</td>
<td>► No measure of muscle strength, size, mobility, gait speed or balance.</td>
</tr>
<tr>
<td>► Ultrasonography.</td>
<td>Study design</td>
</tr>
<tr>
<td>► MRI.</td>
<td>► Non-randomised trials.</td>
</tr>
<tr>
<td>► CT.</td>
<td>Systematic reviews (ie, systematic reviews; meta-analyses review; meta-regressions; umbrella reviews; network meta-analyses).</td>
</tr>
<tr>
<td>► Bioelectrical impedance.</td>
<td>► Narrative reviews.</td>
</tr>
<tr>
<td>► Dual-energy X-ray absorptiometry.</td>
<td>Observational studies (eg, retrospective, prospective or longitudinal).</td>
</tr>
<tr>
<td>► Hydrostatic weighing.</td>
<td></td>
</tr>
<tr>
<td>► Air displacement plethysmography.</td>
<td></td>
</tr>
<tr>
<td>► Microscopy.</td>
<td></td>
</tr>
<tr>
<td>Physical function: Assessed physical function in older adults (mean age ≥55 years old) in the domain(s):</td>
<td></td>
</tr>
<tr>
<td>► Mobility: (defined as a person’s ability to move physically, eg, Timed Up and Go Test, Chair Rise Sit to Stand).</td>
<td></td>
</tr>
<tr>
<td>► Balance: (defined as the ability to maintain a controlled body position during a given task, eg, Berg Balance Test, Sit and Reach Test).</td>
<td></td>
</tr>
<tr>
<td>► Gait speed: (defined as the time it takes to cover a given distance, eg, 6 Minute Walk Test, or 25 Foot Walk Test).</td>
<td></td>
</tr>
<tr>
<td>Study design</td>
<td>Study design</td>
</tr>
<tr>
<td>► Randomised trial.</td>
<td>► Non-randomised trials.</td>
</tr>
<tr>
<td>► Reported in English.</td>
<td>Systematic reviews (ie, systematic reviews; meta-analyses review; meta-regressions; umbrella reviews; network meta-analyses).</td>
</tr>
</tbody>
</table>

Systematic review

Box 1 Description of predefined conditions

| Condition acronym – condition description | CTRL – non-exercise control. |
| LS1 – lower load, single set/exercise, 1 day/week/day/week resistance training. |
| LS2 – lower load, single set/exercise, 2 days/week/day/week resistance training. |
| LS3 – lower load, single set/exercise, ≥3 days/week/resistance training. |
| LM1 – lower load, multiple sets/exercise, 1 day/week/day/week resistance training. |
| LM2 – lower load, multiple sets/exercise, 2 days/week/day/week resistance training. |
| LM3 – lower load, multiple sets/exercise, ≥3 days/week/resistance training. |
| HS1 – higher load, single set/exercise, 1 day/week/day/week resistance training. |
| HS2 – higher load, single set/exercise, 2 days/week/day/week resistance training. |
| HS3 – higher load, single set/exercise, ≥3 days/week/resistance training. |
| HM1 – higher load, multiple sets/exercise, 1 day/week/day/week resistance training. |
| HM2 – higher load, multiple 2 sets/exercise, 2 days/week/day/week resistance training. |
| HM3 – higher load, multiple sets/exercise, ≥3 days/week/resistance training. |

Confidence in recommendations

The robustness of recommendations was assessed with threshold analysis. Several factors, including bias and sampling error, can influence NMA results. Threshold analysis determines how much the included evidence could change—for any reason—before treatment recommendations differ and identifies the subsequent treatment recommendation. Identifying the robustness of results with threshold analysis permits guideline developers to have appropriate confidence levels in the reported recommendations.

Sensitivity analysis and network meta-regression

Sensitivity analyses were conducted to explore the impact of outliers, influential cases and sources of network inconsistency on model fit, relative effects and treatment rankings. The first sensitivity analysis excluded studies identified during pairwise meta-analyses and node-splitting, and the second sensitivity analysis excluded node(s) comprised of only one study. Network meta-regression (NMR), assuming independent treatment interactions, was performed to determine if additional factors improved model fit and altered treatment effects. NMR covariates included age, training status, the proportion of females, duration, volitional fatigue, relative weekly volume load, outcome measurement tool, outcome measurement region and publication year. Missing data on covariates were managed through multivariate imputation by chained equations (n imputations=20). NMR is detailed in online supplemental appendix 6.

All analyses were performed in R V 4.0.4 using the packages: ‘esc’ to calculate SMD; ‘dmetar’ to conduct pairwise meta-analyses and assess comparison-level heterogeneity; ‘multinma’ to conduct NMA, NMR and consistency testing; ‘mmeta’ to perform thresholding; and ‘mice’ to perform multiple imputation. Figures were created with ‘multinma’, ‘metafor’ ggplot2, and GraphPad Prism (V9.1.0 for Windows, GraphPad Software, San Diego, California, USA, www.graphpad.com). All code was made publicly available (see Data Sharing Statement).

Equity, diversity and inclusion statement

Our author group comprises various disciplines, career stages and genders. Data collection, analysis and reporting methods were not altered based on regional, educational or socioeconomic differences of the community in which the included studies were conducted. The only consistently reported equity, diversity and inclusion-relevant variable on which we have analyzed the data is biological sex.

RESULTS

Included studies

The systematic search yielded 16,880 records after duplicates were removed. Following title/abstract screening, 1,051 full texts were assessed for inclusion. A total of 192 articles were included in this review (figure 1). Characteristics of included studies are detailed in the online supplemental appendix 6.

Network geometry

Network geometry for strength is displayed in figure 2A. The strength NMA (178 studies, n=5,097) included 13 conditions and 32 direct comparisons. The three largest nodes were CTRL (n=1,321), LM3 (n=1,133) and LM2 (n=710), and the three smallest nodes were HM1 (n=54), LS1 (n=34), and HS1 (n=13). The most common comparisons were LM3 versus CTRL (51 studies), HM3 versus LM3 (32 studies), HM3 versus CTRL (30 studies) and LM2 versus CTRL (30 studies).

Network geometry for hypertrophy is displayed in figure 2B. The hypertrophy NMA (119 studies, n=3,364) included 11 conditions—no studies included HS1 or LS1—and 24 direct comparisons. The three largest nodes were CTRL (n=847), LM3 (n=810) and LM2 (n=548), and the three smallest nodes were HS3 (n=60), HS2 (n=21) and HM1 (n=11). The most...
common comparisons were LM3 versus CTRL (35 studies), HM3 versus LM3 (22 studies), LM2 versus CTRL (18 studies) and HM3 versus CTRL (17 studies).

**Risk of bias**

Within-study risk of bias was moderate–high for both strength and hypertrophy outcomes. In the strength network, 22%, 67% and 1% of studies had a high, moderate or low risk of bias, respectively. In the hypertrophy network, 18%, 82% and 0% of studies had a high, moderate or low risk of bias, respectively. Study-level risk of bias assessments for both strength and hypertrophy is detailed in online supplemental appendix 7.

**RTxs versus CTRL**

The relative effect of each RTx compared with CTRL on muscle strength is displayed in figure 3A. The posterior SMD for all prescriptions ranged from 0.75 to 1.60, with the largest relative effect from HM3 (1.60 (1.38 to 1.82)). Compared with CTRL, the relative effect of LS1 (0.75 (−0.16 to 1.68)) and HS1 (0.79 (−0.88 to 2.45)) were the only comparisons that the 95% CrI crossed zero.

The relative effect of each RTx compared with CTRL on muscle hypertrophy is displayed in figure 3B. The posterior SMD for all RTx ranged from 0.10 to 0.66, with the largest relative effect from HM2 (0.66 (0.47 to 0.85)). Compared with CTRL, the relative effect of HS2 (0.10 (−0.57 to 0.80)), HS3

---

**Figure 2** Network geometry for all available studies evaluating strength (A) and hypertrophy (B). Each node represents a unique condition, and the size of each node is proportional to the sample size per condition. Each edge represents direct evidence, and the width of each edge is proportional to the number of studies comparing connected nodes. Resistance training prescriptions are denoted with a three-character acronym—XY#—where X is load (H, ≥80% 1-repetition maximum (1RM); L, <80% 1 RM); Y is sets (M, multiset; S, single-set); and # is the weekly frequency (3, ≥3 days/week; 2, 2 days/week; 1, 1 day/week), respectively. For example, ‘HM2’ denotes higher-load, multiset, twice-weekly training. CTRL, non-exercising control group.

**Figure 3** Forest plots displaying network estimates for relative effects of resistance training prescriptions versus non-exercising control for strength (A) and hypertrophy (B). Each resistance training prescription (RTx) is denoted with a three-character acronym—XY#—where X is load (H, ≥80% 1-repetition maximum (1RM); L, <80% 1 RM); Y is sets (M, multiset; S, single-set); and # is the weekly frequency (3, ≥3 days/week; 2, 2 days/week; 1, 1 day/week), respectively. For example, ‘HM2’ denotes higher-load, multiset, twice-weekly training. CTRL, non-exercising control; SMD, standardised mean difference; 95% CrI, 95% credible interval.
(0.34 (−0.02 to 0.71)) and HM1 (0.40 (−0.35 to 1.17)) were the only comparisons that the 95% CrI crossed zero.

### Comparing RTxs

The relative effects from all 133 network comparisons for muscle strength and hypertrophy are displayed in table 2. For comparisons between RTxs (ie, not CTRL), the 95% CrI excluded zero for 13.6% (9/66) and 2.2% (1/45) of comparisons in the strength and hypertrophy NMA, respectively. For muscle strength, there was a 95% probability that HM2 yields a larger relative effect than LS1, LS2, LS3, LM2 and LM3 and that HM3 yields a larger relative effect than LS2, LS3, LM2 and LM3. There was a 95% probability for muscle hypertrophy that HM2 yields a larger relative effect than LS3.

### Ranking conditions

Figure 4 displays the probability that each condition would rank in the top three best interventions for muscle strength and hypertrophy, such that scores closer to 100% indicate a greater chance of ranking in the top three. HM3 (85.5%), HM2 (83.5%) and HM1 (60.5%) were most likely to rank in the top three for muscle strength. HM2 (86.9%), LM1 (48.7%) and LM2 (48.3%) were most likely to rank in the top three for muscle hypertrophy. CTRL was the only condition with a 0% chance for strength and hypertrophy. Posterior rankings and distribution curves for all conditions are reported in the online supplemental appendix 8.

### Network inconsistency

Model fit outputs and node-splitting plots are reported in the online supplemental appendix 9. In the strength network, the UME model (DIC=402.3) was not meaningfully different than the random-effects NMA model (DIC=400.8). Node-splitting was performed on 29 comparisons; the only significant difference was LM1 versus HM1 (p<0.01). In the hypertrophy network, the UME model (DIC=143.1) was meaningfully different than the random-effects NMA model (DIC=137.8). Node-splitting was performed on 22 comparisons; the only significant difference was LS2 versus CTRL (p<0.01).

### Threshold analysis

Threshold analysis results for strength and hypertrophy are shown in online supplemental appendix 10. HM3 was the top-ranked condition for strength; however, 65 comparisons indicated some sensitivity to the level of uncertainty and potential biases in the evidence. The revised top-ranked strength condition was HM2 in 92% (60/65) or HM1 in 8% (5/65) of comparisons. HM2 was the top-ranked condition for hypertrophy, and this finding was robust. Two comparisons indicated some sensitivity to the level of uncertainty and potential biases in the evidence, and HM1 was the revised top-ranked condition in both cases.

### Sensitivity analyses

Sensitivity analysis results are displayed in the online supplemental appendix 11. For both the strength and hypertrophy NMA, the second sensitivity analysis (discussed herein) most improved model fit. The strength network included 155 studies (n=4397) and 11 conditions (LS1 and HS1 excluded). The relative effects for all RTxs versus CTRL were tempered, such that posterior SMDs ranged from 0.77 to 1.49, with the largest relative effect from HM2 (1.49 (1.29 to 1.70)) and smallest from LS3 (0.77 (0.56 to 0.98)). The 95% CrI for each RTx versus CTRL excluded zero. There was a 95% probability that HM2 yields larger relative effects than LS2, LS3, LM1, LM2, LM3 and LS3; that HM3 was superior to LS2, LS3, LM1, LM2 and LM3; and that LM2 was superior to LS3. HM2 (99.9%) and HM3 (95.7%) remained most likely to rank in the top three for muscle strength.

The hypertrophy network included 115 studies (n=3240) and 9 conditions (HS2 and HM1 excluded). The relative effect for each RTx versus CTRL was roughly unchanged, with the largest relative effect from HM2 (0.59 (0.39 to 0.78)) and the smallest from HS3 (0.30 (−0.05 to 0.66)). Between prescriptions, there was a 95% probability that LM2 was superior to LS3. HM2 (82.8%) and LM2 (80.4%) remained most likely to rank in the top three for muscle hypertrophy.

### Network meta-regression

Network meta-regression results are displayed in the online supplemental appendix 12. Model fit was not meaningfully different than the unadjusted model for all covariates, except relative weekly volume load, which worsened model fit. Age, training status, proportion of females, duration, volitional fatigue, relative weekly volume load, outcome measurement tool, outcome measurement region and publication year did not yield any obvious modifying effect on the relative effect for each RTx versus CTRL, and data-sparse nodes reduced estimate precision.

### Physical function

Physical function results are reported in the online supplemental appendix 13. Few studies assessed mobility (25 studies, n=859, age (mean)=68 years), gait speed (15 studies, n=488, 68 years) and balance/flexibility (11 studies, n=323, 68 years). Compared with CTRL, there was a 95% probability that LM2, LM3 and HM3 improved mobility and gait speed, while HM3 was the only condition that improved balance/flexibility (figure 5). No differences were found between RT prescriptions for any physical function outcome.

### DISCUSSION

Twelve distinct RT prescriptions and non-exercising control groups were compared using network meta-analysis to determine their effect on gains in muscle strength, hypertrophy and improvements in physical function in healthy adults. Compared with no exercise, most load, sets and frequency combinations increased muscle strength and hypertrophy, indicating that several RTx resulted in beneficial skeletal muscle adaptations. RT with higher loads characterised the top-ranked strength prescriptions, and RT with multiple sets characterised the top-ranked hypertrophy prescriptions. A diverse range of RT prescriptions improved physical function, but evidence scarcity limited insights. Guideline developers and practitioners may consider these results when forming recommendations and prescribing RT for healthy adults.

Network meta-analysis has previously been used to compare different types of exercise31–34 and doses of RT load.35 In the NMA by Lopez et al.,35 23 (n=582) and 24 (n=604) studies were included in the strength and hypertrophy networks, respectively. The present strength (178 studies, n=3097) and hypertrophy (119 studies, n=3364) networks were much larger, and this is likely attributable to Lopez et al.’s excluding studies not including RT to momentary muscular failure and our more comprehensive search strategy (2629 vs 16 880 records identified). This NMA, to our knowledge, represents the largest synthesis of RT data from randomised trials.
Network estimates for all relative effects of resistance training prescriptions are displayed for strength (column header vs row header; values >0 favour the column condition) and hypertrophy (row header vs column header; values >0 favour the row condition). Data are displayed as posterior standardised mean difference (95% credible interval). Bolded numbers indicate a 95% probability one intervention yields a larger relative effect. Resistance training prescriptions are denoted with a three-character acronym—XYZ—where X is load (H, ≥80% 1-repetition maximum (1RM); L, <80% 1 RM); Y is sets (M, multiset; S, single-set); and # is weekly frequency (3, ≥3 days/week; 2, 2 days/week; 1, 1 day/week), respectively. For example, ‘HM2’ denotes higher-load, multiset, twice-weekly training.

CTRL, non-exercise control; N.D., no data.

<table>
<thead>
<tr>
<th>Strength</th>
<th>CTRL</th>
<th>HM1</th>
<th>HM2</th>
<th>HM3</th>
<th>HS1</th>
<th>HS2</th>
<th>HS3</th>
<th>LM1</th>
<th>LM2</th>
<th>LM3</th>
<th>LS1</th>
<th>LS2</th>
<th>LS3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypertrophy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTRL</td>
<td>1.54 (0.81 to 2.30)</td>
<td>1.59 (1.28 to 1.90)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HM1</td>
<td>0.40 (−0.35 to 1.17)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HM2</td>
<td>0.66 (0.47 to 0.85)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HM3</td>
<td>0.51 (0.35 to 0.67)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS1</td>
<td>N.D.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS2</td>
<td>0.10 (−0.57 to 0.80)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS3</td>
<td>0.34 (−0.02 to 0.71)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LM1</td>
<td>0.55 (0.19 to 0.90)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LM2</td>
<td>0.56 (0.42 to 0.71)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LM3</td>
<td>0.50 (0.39 to 0.61)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LS1</td>
<td>N.D.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LS2</td>
<td>0.48 (0.19 to 0.78)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LS3</td>
<td>0.37 (0.16 to 0.57)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 2** League table of all relative effects


Systematic review
Performing RT to momentary muscular failure may, however, be increasingly important for trained individuals. For both strength and hypertrophy, though, there was a large credible interval surrounding the non-significant effect estimate for many comparisons between RTx, so a wide range of different effects are possible for these comparisons. The available evidence does not permit definitive, statistically valid conclusions about the equivalency of each RTx, despite most comparisons between RTx not being statistically significantly different from each other.

Prescriptions for RT with higher loads were more likely to rank in the top three for strength than all lower-load prescriptions, and RT prescriptions with multiple sets per exercise were most likely to rank in the top three for hypertrophy (figure 4). Rankings are sensitive to uncertainties within the network, but posterior ranking credible intervals supported higher-load, multiset programmes being the highest-ranked for strength and multiple sets or multiple sessions being the highest-ranked for hypertrophy. Notably, sets and frequency are major components of RT volume, a key factor for hypertrophy. The probability of each condition ranking in the top three was calculated because the top-ranked RTx does not necessarily reflect the best intervention for all individuals. Personal preferences, including disliking higher loads or time constraints, including an inability to train more than once weekly, can be observed while still benefiting from RT. In our view, especially given the low participation rates in RT, practitioners should not avoid prescribing, nor should individuals be discouraged from completing non-top-ranked RTx. While all prescriptions increased muscle strength and mass, the top-ranked prescriptions involved higher loads for strength and higher volume for hypertrophy. We do not know how these RTx affect relevant health outcomes. Some data suggest that health benefits exist with low time commitment (30–60 min/week) to RT and greater time commitment with reduced health benefits.

Ours is the first review to assess confidence in RTx recommendations with threshold analysis. Several factors can influence NMA results, and the robustness of treatment recommendations should be considered when interpreting results. Previous methods to evaluate the confidence of meta-analytical findings do not consider how potentially influencing factors can change treatment recommendations or are not yet developed for Bayesian NMA. Threshold analysis determines how much the available evidence could change before recommendations differ and identifies a new top-ranked treatment. Sixty-five direct comparisons were identified that could potentially impact the

Systematic review

recommendation of HM3 as the top-ranked strength treatment; however, the revised treatment recommendation was HM2 in 60 of these cases and HM1 in the other five cases (online supplemental appendix 10), suggesting that performing RT with higher loads and multiple sets/exercise are robust recommendation for optimising RT-induced strength gains. The top-ranked RTx for hypertrophy—HM2—was sensitive to the uncertainty of only two comparisons, and HM1 was the revised recommendation because both comparisons were from the same multi-arm study. Furthermore, 127 of the 161 direct comparisons would need to change by more than four SDs to alter HM2 as the top recommendation for hypertrophy. The optimised recommendations of higher load, multiple-set programmes for strength and HM2 for hypertrophy were extremely robust.

Current guidelines collectively advise healthy adults to complete RT at least twice weekly. The results herein support these recommendations and should not deter practitioners from promoting existing guidelines to improve strength and hypertrophy, nor do these results contradict the effectiveness of guidelines incorporating additional RTx variables, such as rest intervals and contraction type and velocity. However, our results support RT at less than recommended often-cited levels for enhancing strength and hypertrophy. Most individuals do not meet current guidelines, and RTx complexities may impede the adoption of RT. Minimal-dose approaches have been proposed to reduce barriers to RT, and our results strongly support the WHO’s claim, ‘Doing some activity is better than none’. While others attempt to optimise RTx, we propose that, for most adults, regularly engaging in any RTx is more important than training to optimise strength and hypertrophy outcomes. Our analysis found multiple RTx comparable for healthy adults to increase muscle strength and mass. Thus, adults should engage in RT, even if they cannot meet existing recommendations.

Limitations

Risk of bias was frequently introduced by protocol deviations, randomisation procedures and selection of the reported result for both outcomes (online supplemental appendix 7). All three domains were regularly rated “Some concerns” because participants were aware of the intervention, appropriate analyses to estimate the effect of assignment were not performed and randomisation, concealment and prespecified analysis procedures were rarely reported. Double-blinding RT is unfeasible, but the remaining issues are prevalent and reoccurring in RT research. Researchers should preregister analysis plans and report randomisation procedures to reduce bias.

Several limitations require acknowledgement and consideration when interpreting the findings of this review. Well-trained elite athletes/military persons and individuals with chronic disease were excluded, so the results should be translated to these populations with caution and additional insights. Mobility, gait speed and balance/flexibility findings should also be interpreted with caution due to the limited evidence available, which could be attributed to including only healthy older (>55 year) adults (eg, not frail). The coding framework for RTx prescriptions prevented the inclusion of periodized RT programmes

Figure 5 Forest plots displaying network estimates for relative effects of resistance training prescriptions versus non-exercising control for mobility (A), gait speed (B) and balance/flexibility. Each resistance training prescription (RTx) is denoted with a three-character acronym—XY#—where X is load (H, ≥80% 1-repetition maximum (1RM); L, <80% 1RM); Y is sets (M, multiset; S, single-set); and # is the weekly frequency (3, ≥3 days/week; 2, 2 days/week; 1, 1 day/week), respectively. For example, ‘HM2’ denotes higher-load, multiset, twice-weekly training. CTRL, non-exercising control; SMD, standardised mean difference; 95% CrI, 95% credible interval.
overlapping conditions (eg, loads ranging from 60–90% 1RM) from being captured in the network. Initially, our objective was to further divide the load and set prescriptions; however, this yielded sparse, disconnected networks, violating a critical assumption of NMA.⁴⁹ The continuous RTx variables investigated herein (load, sets, frequency) were classified categorically, so future work could use dose-response/model-based NMA methods to explore these RTx variables as continuous predictors.⁸⁰ ⁸¹ Several acute RT variables were not factored into the included RT prescriptions (eg, inter-set rest, time under tension, repetition velocity, volitional fatigue, tempo); where possible, NMR was used to explore if these factors improved model fit and altered effects. Results from NMR are correlative, however, and should be interpreted cautiously.¹² Nonetheless, many variables (inter-set rest, tempo, time under tension) were reported too infrequently for inclusion as covariates. Calculating the relative weekly volume load (ie, load × repetitions/set × number of sets × number of exercises × weekly frequency), which should impact results,¹³ also required approximations that hindered model fit. The principle of specificity¹⁴ (ie, the similarity between training and testing movement) and approximations of muscle mass (eg, lean mass) could infringe on transitivity assumptions⁸² when integrating results from multiple studies and NMR with the covariates measurement tool and region were imperfect solutions. Including one measurement per outcome for each study may limit the totality of evidence captured by this review, so future methodological work could explore the integration of multiple correlated effect sizes in NMA, as in recent pairwise meta-analyses.¹³ ¹⁴ Increasingly, within-subject models are used due to their increased statistical power.⁸⁵ To our knowledge, however, no methods are available to account for the additional correlation when including within-subject and between-subject comparisons in NMA. With consideration for these limitations, guideline developers and practitioners can obtain meaningful insights from this analysis.

Conclusion

This NMA represents the largest synthesis of RTx data from randomised trials. Most RTx increased muscle strength and mass compared with no exercise. Top-ranked prescriptions for muscle strength were characterised by lifting heavier loads, and multiple sets characterised top-ranked prescriptions for muscle hypertrophy. Guideline developers and practitioners should encourage the adoption of RT since all RTx can increase muscle strength and mass in healthy adults. The effects on health outcomes of various RTx remain largely unknown.

Twitter Brad S Currier @brad_currier and Stuart M Phillips @mackinprof

Acknowledgements The authors thank Alicia Wong for assistance with the risk of bias assessments. Components of this review were presented, in abstract form, at Europhysiology 2022, the 2022 Canadian Society for Exercise Physiology and the 2023 American College of Sports Medicine conferences.

Contributors BSC and JCM contributed equally to this paper. BSC, JCM and SMP conceived the review. BSC, JCM and LB designed and executed the systematic search. BSC, JCM, ACD, SMP, JAJK, LL, GC, LC-S, KIL and AV screened articles and extracted data. BSC, JCM, AY and AW completed within-study risk of bias assessments. BSC and JCM conducted the statistical analysis with assistance from JB and NJW. BSC, JCM and SMP drafted the manuscript. All authors critically revised the manuscript. All authors approved the submission of this manuscript. SMP accepts full responsibility for the work and/or the conduct of the study, had access to the data and controlled the decision to publish.

Funding No funding was received for this review. BSC is supported by an Alexander Graham Bell Canada Graduate Scholarship-Doctoral. JCM is supported by an Ontario Graduate Scholarship. NJW was supported by the NHIR Biomedical Research Centre at University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care. ACD and JAJK are supported by a Canadian Graduate Scholarship – Master’s Program offered by the Natural Sciences and Engineering Research Council. SMP is supported by the Canada Research Chairs programme.

Competing interests SMP reports grants or research contracts from the US National Dairy Council, Canadian Institutes for Health Research, Dairy Farmers of Canada, Roquette Freres, Ontario Centre of Innovation, Nestle Health Sciences, Myos, National Science and Engineering Research Council and the US NIH during the conduct of the study; personal financial support from Nestle Health Sciences, non-financial support from Enhanced Recovery, outside the submitted work. SMP has patents licensed to Exerkine but reports no financial gains from any patent or related work. The remaining authors report no competing interests.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

ORCID iDs Brad S Currier http://orcid.org/0000-0003-3955-3174 Stuart M Phillips http://orcid.org/0000-0002-1956-4098

REFERENCES


46 Borenstein M. Multiple comparisons within a study, in introduction to meta-analysis. 2009: 239–42.


