

Inter-individual differences in weight change following exercise interventions: A systematic review and meta-analysis of randomised controlled trials

Philip J Williamson, Greg Atkinson, Alan M Batterham

Health and Social Care Institute

Teesside University

Middlesbrough

UK

TS1 3BA

Keywords: Inter-individual variation, Weight loss, Exercise, RCT, Systematic Review, Meta-Analysis

Running Title: Individual Variance in Weight Loss Response: A Systematic Review

Corresponding Author: Williamson, PJ (Health and Social Care Institute,

Teesside University, Middlesbrough, UK, TS1 3BA, email:

P.Williamson@tees.ac.uk, phone: 01642384143)

Greg.Atkinson@tees.ac.uk

A.Batterham@tees.ac.uk

Funding

No sources of funding were used in the preparation of this article.

Conflict of Interest

Philip Williamson, Greg Atkinson and Alan Batterham declare that they have no conflicts of interest that are relevant to the content of this review.

Abstract

Previous reports of substantial inter-individual differences in weight change following an exercise intervention are often based solely on the observed responses in the intervention group. Therefore, we aimed to quantify the magnitude of inter-individual differences in exercise-mediated weight change. We synthesized randomised controlled trials (RCT) of structured, supervised exercise interventions. Fourteen electronic databases were searched for relevant studies published up to March 2017. Search terms focused on structured training, RCTs and body weight. We then sifted these results for those RCTs ($n=12$, 1500 participants) that included relevant comparator group data. Standard deviations (SD) of weight change were extracted, thereby allowing the SD for true inter-individual differences in weight-loss to be calculated for each study. Using a random effects meta-analysis, the pooled SD (95% CI) for true individual responses was 0.8 (-0.9 to 1.4) kg. The 95% prediction interval (based on 2SDs) for true inter-individual responses was -2.8 to 3.6 kg. The probability (% chance) that the true individual response variability would be clinically meaningful (>2.5 kg) in a future study in similar settings was 23% ('unlikely'). Therefore, we conclude that evidence is limited for the notion that there are clinically important individual differences in exercise-mediated weight change.

1.1 Introduction

Interest in the individualised response to a treatment intervention, and its applicability to medical and exercise interventions, has been growing over the last three decades ^(1,2,3,4,5,6,7,8). There has been specific interest in the inter-individual differences in weight change in response to exercise training for around 20 years ^(9,10,11,12,13,14). Such interest has developed into a dedicated field of research; precision medicine – encompassing ‘tailor-made’ therapies based on the individual response of a patient ⁽⁵⁾. It is predicted that this individual approach to medicine will ultimately reduce costs and improve quality of healthcare ⁽¹⁵⁾. It has also been suggested that personalized medicine may revolutionize healthcare through utilization of individual genetic information, thereby improving drug safety and efficacy ⁽¹⁶⁾. Nevertheless, associations that have been reported between genotype and treatment responses are often small ⁽¹⁷⁾.

A limitation of published research on the efficacy of exercise training has been reported to be the focus on group mean data, with inter-individual variation in response often being overlooked ⁽¹¹⁾. Such a focus on mean effects could obfuscate important individual differences in response ^(11,18,19). If such individual differences are present, and predictors of individual response are identified, then targeted intervention strategies could be formulated to maximize weight loss for individuals.

1.2 Research Design and Data Analysis Issues

There have been reports of inter-individual variation in adiposity and weight response to exercise ^(9,10,11,12,13), including observations that exercise can cause a less-than-expected weight loss for some individuals ⁽²⁰⁾. It has been suggested that the response to exercise may be influenced by a multitude of individual characteristics, including sex ^(20,21), genetics ⁽²²⁾, age, and baseline status of the measured outcome ⁽²³⁾. Clinically-relevant inter-individual response variation should be quantified and judged properly ^(24,25) before the clinical relevance of these effect modifiers of response are appraised, relative to a robust minimal clinically important difference (MCID). Crucially, this quantification requires an appropriate control/ comparator group, preferably within a randomised trial design. Regrettably, substantial treatment response heterogeneity has been claimed from observations solely on the intervention group ^(11,13,26). When the comparator sample is absent or ignored, the interpretation of response heterogeneity is prone to all the philosophical issues highlighted by Stephen Senn, particularly the problem of the “counterfactual” ⁽²⁵⁾.

An appropriate method to quantify “true” individual response variability in a parallel group study involves the application of the following equation; $SD_{IR} = \sqrt{SD_I^2 - SD_C^2}$ ^(24,27), where SD_{IR} is the true inter-individual response variability, expressed as a standard deviation, and SD_I^2 and SD_C^2 are the standard deviations of the changes in the intervention and control samples, respectively. The SD_{IR} should be interpreted as the amount by which the net mean effect of the intervention (intervention minus control) differs typically

between individuals ⁽²⁷⁾.

1.3 Aims of the Review

In view of the above design and analysis issues, there is uncertainty about previously-drawn conclusions in weight-loss studies. To date, there has been no published quantitative synthesis of the evidence for individual response variation in studies on exercise-mediated weight loss. Therefore, we aimed to conduct a systematic review and meta-analysis of the available research to allow for quantification of 'true' inter-individual variation in weight change in response to an exercise intervention.

2. Methods

This study was undertaken in accordance with the ethics procedures and guidance of Teesside University. The review is reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement ⁽²⁸⁾. The review protocol was registered with PROSPERO, the International Prospective Register of Systematic Reviews (CRD42016049982). An initial scoping literature review was undertaken to gauge the likely number of eligible studies for inclusion in the meta-analysis.

2.1 Study Question

Our systematic review was designed to address the following question:

Across all the relevant studies that include a suitable comparator sample, are there substantial inter-individual differences in body mass loss in response to an exercise intervention?

2.2 Literature Search and Study Selection

This review involved a systematic electronic search of peer-reviewed original literature using the following commonly used databases: Centre for Reviews and Dissemination (York), CINAHL, Cochrane Central Register of Controlled Trials (CENTRAL), Cochrane Database of Systematic Reviews, Cochrane Methodology Register, Database of Abstract Reviews or Effects (DARE), Database of Promoting Health Effectiveness Reviews (DoPHER), EMBASE, Medline (Ovid), NHS Economic Evaluation Database (NHS EED), PROSPERO, PubMed, SCOPUS and Sport Discus. These databases were first searched in December 2016, before a secondary search in March 2017. The search strategy was designed to include all articles published in the English language. Search terms comprised of “exerc*” AND (“train*” OR “condition*”) AND (“structure” OR “supervised”) AND (“weight” OR “body compos**” OR “BMI*”) AND (“randomi*” OR “RCT”). Subsequently, additional searches of reference lists, Google Scholar and relevant bibliographic hand searches with no limit of language or publication date were also completed. Only studies conducted in humans were considered.

Studies were screened for those that would meet the inclusion criteria. Titles and abstracts were initially scrutinised to exclude those studies clearly beyond

the scope of this review. For potential studies that appeared to meet the inclusion criteria, or those for which a decision was unable to be made based upon the title and abstract alone, full, published articles were obtained for detailed assessment against the inclusion criteria. Where multiple papers from a single study have been published, these were treated as a single study. Included studies were randomized intervention studies, reporting the standard deviation of the change in body mass in both arms. All studies targeting specific populations (e.g. pregnant women, children, and individuals suffering from specific diseases) were excluded. The remaining full-text articles were included in the systematic review and meta-analysis. A complete overview of the process is presented at Fig. 1 and a comprehensive summary of the studies reviewed is presented in Table 1.

Two reviewers (PW and GA) independently assessed publications for eligibility. The decision to include studies was hierarchical and made initially upon the basis of the study title, abstract and presence of keywords. When a study could not be excluded with certainty, the full text was obtained for evaluation. Disagreements between reviewers were resolved through discussion with a third reviewer (AB) and a consensus approach was used.

2.3 Study Eligibility

2.3.1 Inclusion Criteria

To be included for quantitative synthesis, studies were required to meet the following criteria: (1) participants were required to be aged 18 or over; (2) taking part in studies where the experimental arm was an exercise-based intervention; (3) which was designed to elicit weight loss; (4) reporting change in adiposity indices (body mass index, body fat or body weight); (5) with no history of diabetes, metabolic, cardiovascular, musculoskeletal or inflammatory disease; (6) the exercise intervention was required to be supervised; (7) the investigation had to be an RCT design; and (8) greater than six weeks in duration. Since the interventions were exercised-based, investigators and participants were not blinded. Studies were included if they were published in peer-reviewed journals or full manuscripts were available (i.e. theses and dissertations). Where several intervention arms were present, all data other than that from the control-only and exercise-only arms were excluded. Where more than one exercise intervention was present, results were combined to avoid double counting of the control sample ⁽²⁹⁾. The same procedure for combining groups was applied to studies with a single exercise intervention but with results reported separately for sub-groups.

2.3.2 Exclusion Criteria

Studies were excluded if they (1) included unsupervised exercise interventions, behaviour therapy, dietary modification, health education, surgical, drug or hormone treatment that did not include exercise; (2) if change in body mass/ composition was not a primary or secondary aim of the

study; (3) if no relevant comparator sample were present; or (4) the full-text manuscript was written in a language other than English.

2.4 Data Extraction and Synthesis

Digitizelt (Brunschweig, Germany) graph digitizer software was used in cases where data were only presented in Figures rather than text or tables.

Study characteristics such as study design, participant characteristics (age, sex, ethnicity), measurement methods, change scores, SD_{change} and information to assess the risk of bias were extracted by the lead author.

A standardized data extraction sheet was used to collect data on participants' characteristics, study methods, sample size, prescribed intervention (frequency, intensity, duration and type), outcomes assessed, loss to follow up and study type. The data for Table 1 and Fig 1 was collected by PW before GA verified its accuracy and the eligibility of studies for inclusion. Where data were incompletely or unclearly reported, the lead author contacted study authors for clarification. Effect sizes were calculated for the relevant measures.

2.5 Assessment of Study Quality

Methodological risk of bias was assessed and reported in accordance with the Cochrane Handbook ⁽²⁹⁾ and the guidelines of the Cochrane Consumers and Communication Review Group ⁽³⁰⁾, which recommend the explicit reporting of

the following elements for RCTs: random sequence generation; allocation sequence concealment; blinding (participants, personnel); blinding (outcome assessment); completeness of outcome data; selective reporting; and other sources of bias. Each item was judged as being at high, low or unclear risk of bias as set out in the criteria provided ⁽²⁹⁾. A summary of risk of bias is presented in Figs 2 and 3, produced using RevMan software (Review Manager. Version 5.3. Copenhagen: The Nordic Cochrane Centre. The Cochrane Collaboration, 2014).

Studies were deemed to be at highest risk of bias if they scored as high or unclear risk of bias for either the sequence generation or allocation concealment domains, based on growing empirical evidence that these factors are particularly important potential sources of bias ⁽²⁹⁾.

In all cases, risk of bias was independently assessed (by PW and GA), with any disagreements resolved by discussion to reach consensus. Risk of bias results were incorporated into the review using standard tables, systematic narrative discussion and commentary about each element, leading to an overall assessment of the risk of bias of those studies selected for inclusion and a judgement about the internal validity of the review's results.

2.6 Meta-Analysis

First, to put the results for individual response variance in context we conducted a random-effects meta-analysis for the mean difference in weight

loss across the included studies, using a restricted maximum likelihood (REML) model combined with the Knapp-Hartung method (t-distribution for between-study variance). Second, for each study we extracted the standard deviation of the changes in body mass for both control (C) and exercise intervention (I) groups. The true individual response variance (intervention minus control) was then derived as $SD_I^2 - SD_C^2$. The standard error (SE) for this variance was calculated using the following equation:

$SE = \sqrt{[2(SD_{Exp}^4/DF_{Exp} + SD_{Con}^4/DF_{Con})]}$, where DF_{Exp} and DF_{Con} are the degrees of freedom of the standard deviations in the exercise and control groups ⁽²⁸⁾. Note that a negative value for the individual response variance, for either the point estimate or lower bound of the confidence interval or prediction interval, implies greater variability in the changes in body mass in the control versus intervention groups.

The individual response variances with their SEs were meta-analysed using a REML model combined with the Knapp-Hartung method. It is important to note that the variances are unbiased, whereas the SD is not, and deriving a SE for the SD for individual responses is also problematic. Therefore, synthesising the individual response variances rather than the SDs for individual responses is imperative. We derived the point estimate for the pooled individual response variance together with its uncertainty expressed as a 95% Confidence Interval (CI). The point estimate and confidence limits were then converted to an SD metric by taking the square root. In the case that the lower limit of the interval was negative, we first ignored the sign, took the square root, and then re-applied the sign. This approach is consistent with the 'nobound' option in SAS/STAT® software, which permits negative variances

(SAS Institute Inc. 2017. *SAS/STAT 14.3 User's Guide*. Cary, NC: SAS Institute Inc.).

For both meta-analyses, between-study heterogeneity was quantified through the tau statistic (τ) – a SD describing the typical variability in the mean effect between studies ⁽³¹⁾. Using the SE for the pooled mean effect and the tau, a 95% prediction interval was derived to quantify the expected range of true effects in future studies in similar settings ⁽³²⁾. For the individual response variability, this prediction interval was derived for $2 \times \text{SD}_{\text{IR}}$, as the SD_{IR} should be doubled before evaluating its magnitude to reflect a comparison between a typically high (mean + SD_{IR}) and typically low (mean – SD_{IR}) responder ⁽²⁷⁾. The magnitude of both the mean weight loss and the individual response variability ($2 \times \text{SD}_{\text{IR}}$) was evaluated against a minimum clinically important difference for weight loss of 2.5 kg ⁽³³⁾ by calculating the probability that the effect in a future study in similar settings would exceed this threshold ⁽³²⁾. This probability was interpreted using the qualitative probabilistic anchors advanced by Hopkins *et al.* ⁽³⁴⁾. Inasmuch as we must work with the response variances, rather than the SDs, we first halved the minimal clinically important difference (equivalent to doubling the SD for individual responses), squared it (to express it in variance metric) and then derived the probability that the response variance in a new study would be clinically relevant, as described above. The threshold of 2.5 kg for the minimum clinically important weight loss was chosen, conservatively, as the lowest value from the range of clinically relevant effects presented by Jensen *et al.* ⁽³³⁾. By definition, effects smaller than this threshold are defined as trivial (not clinically relevant). Effects >2.5 kg but <7.5 kg are defined as 'small' (yet clinically important). We

define 'moderate' effects as >7.5 kg but <15 kg, and 'large' effects as >15 kg (34).

All statistical analyses were conducted using Comprehensive Meta-Analysis software, version 3 (Biostat Inc., Englewood, NJ, USA).

3. Results

3.1 Study Selection

The initial search generated 3187 results (Fig. 1). 3061 of these were excluded based on titles and abstracts alone, and 66 duplicates were rejected. The complete text was obtained for 60 articles. A further 10 were identified from relevant reference lists and hand searches. Following examination of these articles, 12 were identified that met the eligibility criteria and are summarized in Table 1. A further 20 met all selection criteria, apart from the reporting of SD_{change} . The authors of these papers were contacted, but only four responses were received, and full data were not provided in these instances (35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54). Contact was made by email. If, after four weeks, no response was received, a further email was sent. Following a further four-week period, papers from these authors were excluded. One paper met all inclusion criteria (55), except for the fact that median and interquartile range values were presented for changes in body mass, rather than means and SDs. No non-published studies (i.e., dissertations) were found to be eligible for inclusion.

The included studies encompassed a 17-year publication period between 1999 and 2016. Included studies involved a total of 1500 participants (EX: $n=922$, CON: $n=578$). Three trials involved outcomes of aerobic training interventions ^(56,57,58), three involved the outcomes of resistance training interventions ^(59,60,61), one study involved the outcomes on separate aerobic and resistance training interventions ⁽⁶²⁾ and five studies involved the outcomes of combined/concurrent training ^(63,64,65,66,67). The duration of studies ranged from 8 to 52 weeks, study sample sizes ranged from 24 to 411 and reported pre-intervention mean body mass ranged from 65.5 to 128.0 kg.

Insert Fig 1 here

Insert Table 1 here

3.2 Study Outcomes

The pooled mean group difference in pre/post changes in weight (intervention minus control) was -1.4 kg (95% CI -0.3 to -2.5 kg). Substantial between study heterogeneity was observed ($\tau=1.5$ kg: -0.4 to 2.2 kg). The prediction interval revealed that, were investigators to undertake a future trial, the 95% plausible range for mean weight change vs. control would be -5.0 to 2.1 kg. The probability (% chances) that the mean weight loss (intervention minus control) in a future study in similar settings would exceed the minimum clinically important difference of a reduction of 2.5 kg was 26% ('possibly' clinically important).

The pooled point estimate for the inter-individual variability in weight change in response to an exercise intervention (SD_{IR}) was 0.8 (-0.9 to 1.4) kg. The between-study heterogeneity (τ) was 1.0 (-1.7 to 2.2) kg. The 95% prediction interval for $2 \times SD$ for true inter-individual responses was -2.8 to 3.6 kg. The probability (% chances) that the individual response variability ($2 \times SD$) in a future study in similar settings would be clinically meaningful (>2.5 kg) is 23% - 'unlikely' to be clinically important. Therefore, the odds are greater than 3:1 against the notion that there is clinically relevant individual response variance.

3.3 Study Quality and Risk of Bias

Table 2 and Figs 2 and 3 present a summary of risk of bias within included studies. Overall, risk of bias was mostly low or of unclear risk in the outcome of interest.

Insert Table 2 here

Insert Figs 2 & 3 here

4. Discussion

The aim of our review was to synthesise the available evidence for inter-individual variation in weight change following an exercise-focussed intervention. This is the first systematic review and meta-analysis designed to address this specific aim. We found that the evidence is limited for clinically relevant 'true' inter-individual variation in weight change in response to an

exercise intervention, once the random variability in weight over time in the control group is accounted for. Also, the observed pooled inter-individual response variability, when compared to the pooled mean change in weight was small. The prediction interval ranged from small negative (more response variability in control group) to small positive (more variability in the exercise arm), and revealed that the magnitude of the true individual response variability in a future study in similar settings is unlikely to be clinically important. Similarly, the prediction interval for the mean weight loss ranged from moderate reduction to trivial weight gain, and indicated that the magnitude of mean weight loss in a future study in similar settings was only possibly clinically relevant.

4.1 Aerobic training interventions

Aerobic training has been reported to provide positive changes in body mass and body composition ^(68,69). In the current review, three studies were designed to investigate the effect of aerobic training interventions on weight loss, amongst other outcomes ^(56,57,58). Although all three studies showed greater variability of changes in weight in the intervention arm, only one showed substantial true individual response variability. As part of the large-scale Mid-West Exercise Trial 2 (MET-2), a control sample ($n=18$) were compared with groups engaging in 5 days per week of aerobic exercise eliciting 400 Kcal ($n=37$) and 600 Kcal ($n=37$) of energy expenditure per session ⁽⁵⁸⁾. While group means evidenced substantial changes in body weight (400 Kcal: -3.9 kg, 600 Kcal: -5.2 kg, control: 0.5 kg), greater variability

of changes (SD) was observed in the two intervention groups (400kcal: 4.9 kg, 600kcal: 5.6 kg, pooled SD: 5.27 kg) than in the control sample (3.5 kg). The SD for individual response for this study was therefore 3.9 kg (95% CI, 1.8 to 5.3 kg). The individual response variability in this study is clearly clinically relevant: $2 \times \text{SD for individual response} > \text{minimal clinically important difference for the lower confidence limit}$. Indeed, the true individual response variance in this study was at least 7-fold greater than any other included study. Nevertheless, removal of this study from the meta-analysis had no material effect on the pooled SD for inter-individual variation in response (0.7 kg, vs. 0.8 kg with all studies included), and a negligible effect on the heterogeneity. This finding is due in part to the low weight afforded to this study in the analysis – just 1.03% - primarily due to relatively small sample size.

4.2 Resistance training interventions

Three of the included papers were designed to investigate the effects of resistance training on body weight ^(59,60,61). Of these, one study showed a larger SD of body mass changes over 15 weeks of resistance training in intervention versus control ⁽⁶⁰⁾. This study reported trivial increases in mean body mass in both groups (Exercise: 0.54 kg, Control: 0.49 kg). The SD of the changes was 1.87 kg in intervention vs. 1.82 in control, resulting in a trivial SD for individual response of 0.4 kg.

4.3 Separate aerobic and resistance training interventions

A single paper reported upon the impact of separate training modalities ⁽⁶²⁾. The SD of the change in body mass was 1.3 kg in control, 1.5 kg in resistance training, and 1.9 kg in aerobic training (pooled intervention SD of changes = 1.89 kg). The SD for individual response in this study was therefore 1.4 kg, representing small individual response variability.

4.4 Combined/concurrent training

The effects of concurrent training on body composition are equivocal. Weight loss ⁽⁷⁰⁾ and weight gain ⁽⁶⁸⁾ have been reported, but other health outcomes are often also positively influenced ⁽⁷¹⁾. Five studies included in the present review were designed to examine the effects of combined or concurrent aerobic and resistance exercise protocols ^(63,64,65,66,67). Clinically relevant individual response variability was present in just one trial of an intervention involving 12 months of 1 hour per week combined aerobic and circuit-style training ($n=193$), alongside recommendations to undertake 30 minutes of exercise, 6 days per week, compared to a non-exercise control group ($n=194$) ⁽⁶⁷⁾. Mean weight change was -0.49 kg in the intervention group vs. 0.08 kg in control, with SD of the changes of 3.32 and 2.97 kg, respectively. The SD for individual response was therefore 1.5 kg.

4.5 Limitations

We synthesised 12 studies involving a total of 1500 participants. Small sample size is common in supervised exercise-based intervention trials ⁽⁷²⁾, but our review included 4 larger ($N \geq 100$) studies ^(56,61,62,67). Six studies

recruited fewer than 20 participants for one or more of the groups (57,58,59,63,64,66), and might be prone to small study bias at the individual study level.

We restricted our search to RCTs incorporating exercise-only interventions; included studies differed by exercise mode, intensity, frequency and duration, and length of intervention. This intervention heterogeneity might influence mean effects and/ or individual response variance. There are too few studies to compare the effects in, for example, aerobic versus resistance versus combined interventions.

Given the substantial heterogeneity of the true individual response variance, we derived and presented a prediction interval capturing the plausible range for the true individual response variability, consistent with the data and model, in a future study in similar settings. The prediction interval has been described as providing “potentially the most relevant and complete statistical inferences to be drawn from random effects meta-analyses” (73). However, we exercise due caution in inferences drawn from the prediction interval given the coverage issues identified in the simulations conducted by Partlett and Riley (74). These authors reported that the coverage of the interval was particularly poor in cases of low effect heterogeneity and/or markedly variable sample size. With the specific combination of number of studies, between-study heterogeneity of individual response variance and mixture of study sizes in the current review (with REML and Knapp-Hartung estimation) these simulations indicate a maximum under-coverage of our derived prediction

interval of 1%. Such under-coverage would have no material affect on the derived probability of individual response variance in a future trial being clinically relevant. However, we still consider it prudent to view our prediction interval as approximate, as recommended by Partlett and Riley⁽⁷⁴⁾.

Where multiple exercise arms were present in a study, these were combined to avoid double counting of the control arm. This may obscure the effect of different exercise doses; however, analysis of each individual exercise condition vs control, revealed no material difference in individual response variability.

In advance of the study, we proposed various potential effect modifiers (moderators) to account for heterogeneity in individual response variance, including baseline body weight, age, and sex. However, we elected not to conduct any secondary meta regression analyses, as we only had access to study-level covariates (e.g., mean baseline weight, mean age, and proportion of males/females). Fisher et al.⁽⁷⁵⁾ describe this type of analysis as 'daft', as it has a high risk of ecological bias⁽⁷⁶⁾; the 'deft' approach advocated by Fisher et al.⁽⁷⁵⁾ requires either study level analysis of the effects of putative effect modifiers (e.g., treatment interaction effects with sex, age, weight etc.), or an individual-participant data meta-analysis, with relevant interaction terms included in the model. However, obtaining individual participant data from study authors would likely prove to be a major undertaking in this, or indeed any, review. This contention is underscored by the difficulties we experienced

in communicating with authors merely to obtain a simple standard deviation of change scores from the data.

Additionally, the energy expenditure induced by the exercise interventions undertaken in the included studies – and whether this would be sufficient, in theory, to induce weight loss above the minimal clinically important difference – is unknown. Whilst beyond the scope of this systematic review and meta-analysis, it is therefore unknown what effects exercise protocols with larger energy expenditures would elicit.

To make inferences in the current study we adopted a threshold for the minimum clinically important weight loss of 2.5 kg – the smallest threshold of absolute weight loss for clinical benefit reported by Jensen et al. ⁽³³⁾. Readers who disagree with this choice may consider our reported prediction intervals in relation to their own belief in the minimum clinically important difference to make inferences.

Finally, we acknowledge that 20 possibly eligible studies were excluded due to their authors not providing the data requested by e-mail communication.

We assume that these studies are missing at random, as we have no reason to believe that authors would withhold data pertaining to response variance.

4.6 Findings in Relation to Current Recommendations and Future Research

Directions

This is the first systematic review to focus on the true inter-individual variation in weight loss in response to exercise interventions. We conducted a comprehensive literature search over 14 databases. Evidence in relation to the inter-individual response to various treatments/ interventions is growing rapidly. However, based on the findings of this systematic review, we find limited evidence for the presence of clinically important 'true' inter-individual variation in body mass in response to exercise training. Therefore, further investigation of underpinning mechanisms is likely not warranted, as the prediction interval reveals that individual response variance in a future study in similar settings is unlikely to be clinically important. A caveat here, as acknowledged above, is that we only synthesised 12 effects from heterogeneous exercise interventions. If individual differences in response to interventions targeting body weight are considered important from a precision medicine standpoint, then future randomised trials should be sufficiently sized to afford adequate precision of estimation for both mean intervention effects and the SD for individual responses. The latter would require at least 4× the sample size required to define the mean intervention effect with adequate power and precision, and even larger samples if individual response variance is trivial-small.

4.7 Conclusions

To date, much of the research claiming to evidence substantial inter-individual differences in response to an exercise intervention has been conducted in the absence of a suitable comparator sample ^(11,13,14). To quantify the true inter-individual response to an exercise intervention, studies should include a

comparator arm, preferably in a randomised controlled trial. Future work should employ this research design, and incorporate sound statistical quantification of the response variance in each arm, combined with a threshold for the minimal clinically important difference, to determine the presence of clinically important individual variation in response. In summary, our findings constitute limited evidence for the notion of substantial inter-individual differences in weight loss responses to exercise interventions; individual response variability in a future trial in similar settings is unlikely to be clinically important. Our findings, if replicated, confirmed, and extended, might prevent researchers wasting valuable resources searching for explanations of treatment heterogeneity that does not exist or is clinically trivial.

References

1. Prud'Homme D, Bouchard C, LeBlanc C, Landry F, Fontaine E. Sensitivity of maximal aerobic power to training is genotype-dependent. *Med Sci Sports Exerc* 1984; **16**: 89-93.
2. Lortie G, Simoneau JA, Hamel P, Boulay MR, Landry F, Bouchard CI. Responses of maximal aerobic power and capacity to aerobic training. *Int J Sports Med* 1984; **5**: 232-236.
3. Hamel P, Simoneau J, Lortie G, Boulay MR, Bouchard C. Heredity and muscle adaptation to endurance training. *Med Sci Sports Exerc* 1986; **18**: 690-696.

4. Rose EA, Partfitt G. A quantitative and qualitative explanation of the individual differences in affective responses to prescribed and self-selected exercise intensities. *J Sport Exerc Psychol* 2007; **29**: 281-309.
5. Senn S, Rolfe K, Julious SA. Investigating variability in patient response to treatment – a case study from a replicate cross-over study. *Stat Methods Med Res* 2011; **20**: 657-666.
6. Bouchard C. Genomic predictors of trainability. *Exp Physiol* 2012; **97**: 347-352.
7. Bouchard C, Antunbes-Correa LM, Ashley EA, et al. Personalized preventive medicine: genetics and the response to regular exercise in preventive interventions. *Prog Cardiovasc Dis* 2014; **57**: 337-346.
8. Mann TN, Lamberts RP, Lambert MI. High responders and low responders: factors associated with individual variation in response to standardized training. *Sports Med* 2014; **44**: 1113-1124.
9. Snyder KA, Donnelly JE, Jabobsen DJ, Hertner G, Jakicic JM. The effects of long-term, moderate intensity, intermittent exercise on aerobic capacity, body composition, blood lipids, insulin and glucose in overweight females. *Int J Obes* 1997; **21**: 1180-1189.
10. Barbeau P, Gutin B, Litaker M, Owens S, Riggs S, Okuyama T. Correlates of individual differences in body-composition changes resulting from physical training in obese children. *Am J Clin Nutr* 1999; **69**: 705-711.
11. King NA, Hopkins M, Caudwell P, Stubbs RJ, Blundell JE. Individual variability following 12 weeks of supervised exercise: identification and characterization of compensation for exercise-induced weight loss. *Int J Obes* 2008; **32**: 177-184.

12. Barwell ND, Malkova D, Leggate M, Gill JM. Individual responsiveness to exercise-induced fat loss is associated with change in resting substrate utilization. *Metabolism* 2009; **58**: 1320-1328.
13. Caudwell P, Hopkins M, King NA, Stubbs RJ, Blundell JE. Exercise alone is not enough: weight loss also needs a healthy (Mediterranean) diet? *Public Health Nutr* 2009; **12**: 1663-1666.
14. Caudwell P, Gibbons C, Hopkins M, King N, Finlayson G, Blundell JI. No sex difference in body fat in response to supervised and measured exercise. *Med Sci Sports Exerc* 2013; **45**: 351-358.
15. Spear BB, Heath-Chiozzi M, Huff J. Clinical applications of pharmacogenetics. *Trends Mol Med* 2001; **7**: 201-204.
16. Katsanis SH, Javitt G, Hudson K. A case study of personalized medicine. *Science* 2008; **320**: 53-55.
17. Khoury MJ, Galea S. Will precision medicine improve population health? *JAMA* 2016; **316**: 1357-1358.
18. Bouchard C. Human adaptability may have a genetic basis. In: Landry F, (ed). *Health and risk estimation, risk reduction and health promotion. Proceedings of the 18th annual meeting of the society of prospective medicine*. Canadian Public Health Association: Ottawa, 1983. pp.463-476.
19. Bouchard C, Rankinen T. Individual differences in response to regular physical activity. *Med Sci Sports Exerc* 2001; **33**, S6: S452-453.
20. Donnelly JE, Smith BK. Is exercise effective for weight loss with ad libitum diet? Energy balance, compensation, and gender differences. *Exerc Sport Sci Rev* 2005; **33**: 169-174.

21. Ballor DL, Keeseey RE. A meta-analysis of the factors affecting exercise-induced changes in body mass, fat mass and fat-free mass in males and females. *Int J Obes* 1991; **15**: 717-726.
22. Simoneau JA, Lortie G, Boulay MR, Marcotte M, Thibault MC, Bouchard C. Inheritance of human skeletal muscle and anaerobic capacity adaptation to high-intensity intermittent training. *Int J Sports Med* 1986; **7**: 167–171.
23. Sisson SB, Katzmaryk PT, Earnest CP, Bouchard C, Blair SN, Church TS. Volume of exercise and fitness non-response in sedentary post-menopausal women. *Med Sci Sports Exerc* 2009; **41**: 539–545.
24. Atkinson G, Batterham A. True and false interindividual differences in the physiological response to an intervention. *Exp Physiol* 2015; **100**: 577-588.
25. Williamson PJ, Atkinson G, Batterham AM. Inter-individual responses of maximal oxygen uptake to exercise training: a critical review. *Sports Med* 2017; **47**: 1501-1513.
26. King NA, Horner K, Hills AP, *et al.* Exercise, appetite and weight management: understanding the compensatory responses in eating behaviour and how they contribute to variability in exercise-induced weight loss. *Br J Sports Med* 2012; **46**: 315-322.
27. Hopkins W. Individual responses made easy. *J Appl Physiol (1985)* 2015; **118**: 1444-1446.
28. Liberati A, Altman DG, Tetzlaff J, *et al.* The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009; **339**: b2700.

29. Higgins JPT, Green S (eds). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from <http://handbook.cochrane.org>.
30. Ryan R. Cochrane Consumers and Communication Review Group. Cochrane Consumers and Communication Review Group reviews: how to consider equity issues. <http://cccr.org>, February 2013 (accessed 15th April 2017).
31. Higgins JPT. Commentary: heterogeneity in meta-analysis should be expected and appropriately quantified. *Int J Epidemiol* 2008; **37**: 1158-1160.
32. IntHout J, Ioannidis JP, Rovers MM, Goeman JJ. Plea for routinely presenting prediction intervals in meta-analyses. *BMJ Open* 2016; **6**: e010247.
33. Jensen MD, Ryan DH, Apovian CM, et al. 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults. A report of the American Cardiology/American Heart Association task force on practice guidelines and the Obesity Society. *Circulation*. 2013; **01**-cir.
34. Hopkins WG, Marshall SW, Batterham AM, Hanin J. Progressive statistics for studies in sports medicine and exercise science. *Med Sci Sports Exerc* 2009; **41**: 3-12.
35. Schuit AJ, Schouten EG, Miles TP, Evans WJ, Saris WH, Kok FJ. The effect of six months training on weight, body fatness and serum lipids in apparently healthy elderly Dutch men and women. *Int J Obes* 1998; **22**: 847-853.

36. Maiorana A, O'Driscoll G, Dembo L, Goodman C, Taylor R, Green D. Exercise training, vascular function, and functional capacity in middle-aged subjects. *Med Sci Sports Exerc* 2001; **33**: 2022-2028.
37. Donnelly JE, Hill JO, Jacobsen DJ, *et al.* Effects of a 16-month randomized controlled exercise trial on body weight and composition in young, overweight men and women: the Midwest Exercise Trial. *Arch Intern Med* 2003; **163**: 1343-1350.
38. Potteiger JA, Jacobsen DJ, Donnelly JE, Hill JO. Glucose and insulin responses following 16 months of exercise training in overweight adults: the Midwest Exercise Trial. *Metabolism* 2003; **52**: 1175-1181.
39. Schmitz KH, Jensen MD, Kugler KC, Jeffrey RW, Leon AS. Strength training for obesity prevention in midlife women. *Int J Obes* 2003; **27**: 326-333.
40. Takeshima N, Rogers ME, Islam MM, Yarmauchi T, Watanabe E, Okada A. Effect of concurrent aerobic and resistance circuit exercise training on fitness in older adults. *Eur J Appl Physiol* 2004; **93**: 173-182.
41. Toraman NF, Erman A, Agyar E. Effects of multicomponent training on functional fitness in older adults. *J Aging Phys Act* 2004; **12**: 538-553.
42. Shojaee-Moradie F, Baynes KCR, Pentecost C, *et al.* Exercise training reduces fatty acid availability and improves the insulin sensitivity of glucose metabolism. *Diabetologia* 2007; **50**: 404-413.
43. McTiernan A, Sorensen B, Irwin ML, *et al.* Exercise effect on weight and body fat in men and women. *Obesity* 2007; **15**: 1496-1512.

44. Nalbant Ö, Toraman NF, Ögü C, *et al.* Vitamin E and aerobic exercise: effects on physical performance in older adults. *Aging Clin Exp Res* 2009; **21**: 111-121.
45. Coker RH, Williams RH, Yeo SE, *et al.* The impact of exercise training compared to caloric restriction on hepatic and peripheral insulin resistance in obesity. *J Clin Endocrinol Metab* 2009; **94**: 4258-4266.
46. Kerksick CM, Wismann-Bunn J, Fogt D, *et al.* Changes in weight loss, body composition and cardiovascular disease risk after altering macronutrient distributions during a regular exercise program in obese women. *Nutr J* 2010; **9**: 59.
47. Atashak S, Peeri M, Jafari A, Ali Azarbayjani M. Effects of ginger supplementation and resistance training on lipid profiles and body composition in obese men. *Journal of Medicinal Plants Research* 2011; **5**: 3827-3832.
48. Sheikholeslami Vatani D, Ahmadi S, Ahmadi Dehrashid K, Gharibi F. Changes in cardiovascular risk factors and inflammatory markers of young, healthy, men after six weeks of moderate or high intensity resistance training. *J Sports Med Phys Fitness* 2011; **51**: 696-700.
49. Donges CE, Duffield R. Effects of resistance or aerobic exercise training on total and regional body composition in sedentary overweight middle-aged adults. *Appl Physiol, Nutr Metab* 2012; **37**: 499-509.
50. Tracy BL, Hart CE. Bikram yoga training and physical fitness in healthy young adults. *J Strength Cond Res* 2013; **27**: 822-830.
51. Herring LY, Wagstaff C, Scott A. The efficacy of 12 weeks supervised exercise in obesity management. *Clin Obes* 2014; **4**: 220-227.

52. Kim YS, Nam JS, Yeo DW, Kim KR, Suh SH, Ahn CW. The effects of aerobic exercise training on serum osteocalcin, adipocytokines and insulin resistance on obese young males. *Clin Endocrinol* 2015; **82**: 686-694.
53. Bittari ST, Maeda SS, Marone MM, Santili C. Physical exercises with free weights and elastic bands can improve body composition parameters in postmenopausal women: WEB protocol with a randomized controlled trial. *Menopause* 2016; **23**: 383-389.
54. Tan S, Wang J, Cao L, Guo Z, Wang Y. Positive effect of exercise training at maximal fat oxidation intensity on body composition and lipid metabolism in overweight middle-aged women. *Clin Physiol Functional Imaging* 2016; **36**: 225-230.
55. Atlantis E, Chow CM, Kirby A, Fiatarone Singh MA. Worksite intervention effects on physical health: a randomised controlled trial. *Health Promot Int* 2006; **21**: 191-200.
56. Church TS, Martin CK, Thompson AM, Earnest CP, Mikus CR, Blair SN. Changes in weight, waist circumference and compensatory responses with different doses of exercise amongst sedentary, overweight postmenopausal women. *PLoS One* 2009; **4**: e4515. Doi: 10.1371/journal.pone.0004515.
57. Tan S, Wang X, Wang J. Effects of supervised exercise training at the intensity of maximal fat oxidation in overweight young women. *J Exerc Sci Fitness* 2012; **10**: 64-69.
58. Donnelly JE, Honas JJ, Smith BK, *et al*. Aerobic exercise alone results in clinically significant weight loss for men and women: midwest exercise trial 2. *Obesity* 2013; **21**: E219-E228.

59. Prabhakaran B, Dowling EA, Branch JD, Swain DP, Leuthotz BC. Effect of 14 weeks of resistance training on lipid profile and body fat percentage in premenopausal women. *Br J Sports Med* 1999; **33**: 190-195.
60. Schmitz KH, Ahmed RL, Yee D. Effects of a 9-month strength training intervention on insulin, insulin-like growth factor (IGF)-I, IGF-binding protein (IGFBP)-I, and IGFBP-3 in 30-50-year old women. *Cancer Epidemiol Biomarkers Prev* 2002; **11**: 1597-1604.
61. Teixeira PJ, Going SB, Houtkooper LB, *et al*. Resistance training in postmenopausal women with and without hormone therapy. *Med Sci Sports Exerc* 2003; **35**: 555-562.
62. Dognes CE, Duffield R, Drinkwater EJ. Effects of resistance training or aerobic training on interleukin-6, c-reactive protein, and body composition. *Med Sci Sports Exerc* 2010; **42**: 3014-3013.
63. Lockwood CM, Moon JR, Tobkin SE, *et al*. Minimal nutrition intervention with high-protein/low-carbohydrate and low fat, nutrient-dense food supplement improves body composition and exercise benefits in overweight adults: a randomized controlled trial. *Nutri Metabol* 2008;**5**: doi: 10.1186/1743-7075-5-11.
64. Burtcher M, Gatterer H, Kunczicky H, Brandstatter E, Ulmer H. Supervised exercise in patients with impaired fasting glucose: impact on exercise capacity. *Clin J Sports Med* 2009; **19**: 394-398.
65. Vilela BL, Silva AABS, de Lira CAB, Andrade Mdos S. Workplace exercise and educational program for improving fitness outcomes related to health in workers: a randomized controlled trial. *J Occup Environ Med* 2015; **57**: 235-240.

66. Baillot A, Mampuya WM, Dionne IJ, Corneau E, Mexiat-Burdin A, Langlois MF. Impacts of supervised exercise training in addition to interdisciplinary lifestyle management in subjects awaiting bariatric surgery: a randomized controlled study. *Obes Surg* 2016; **26**: 2602-2610.
67. Dalager T, Justesen JB, Murray M, Boyle E, Siogaard GI. Implementing intelligent physical exercise training at the workplace: health effects among office workers – a randomized controlled trial. *Eur J Appl Physiol* 2016; **116**: 1433-1442.
68. Glowacki SP, Martin SE, Maurer A, Baek W, Green JS, Crouse SF. Effects of resistance, endurance, and concurrent exercise on training outcomes in men. *Med Sci Sports Exerc* 2004; **36**: 2119-2127.
69. Boutcher S. High-intensity intermittent exercise and fat loss. *J Obes* 2011.
70. Libardi CA, De Souza GV, Cavaglieri CR, Madruga VA, Chacon-Mikahil MP. Effect of resistance, endurance, and concurrent training on TNF- α , IL-6, and CRP. *Med Sci Sport Exerc* 2012; **44**: 50-56.
71. Dolzeal BA, Potteiger JA. Concurrent resistance training and endurance training influence basal metabolic rate in nondieting individuals. *J Appl Physiol* (1985) 1998; **85**: 695-700.
72. Keating SE, Johnson NA, Mielke GI, Coombes JS. A systematic review and meta-analysis of interval training versus moderate-intensity continuous training on body adiposity. *Obes Rev* 2017; Doi:10.1111/obr.12536.
73. Higgins JPT, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *J R Soc Ser A Stat Soc* 2009; **172**:137–159.

74. Partlett C, Riley RD. Random effects meta-analysis: Coverage performance of 95% confidence and prediction intervals following REML estimation. *Stat Med* 2017; **36**: 301-317.
75. Fisher DJ, Carpenter R, Morris TP, Freeman SC, Tierney JF. Meta-analytical methods to identify who benefits most from treatments: daft, deluded or deft approach? *BMJ* 2017; **356**: j573.
76. Petkova E, Tarpey T, Huang L, Deng L. Interpreting meta-regression: application to recent controversies in antidepressants' efficacy. *Stat Med* 2013; **32**: 2875-2892.

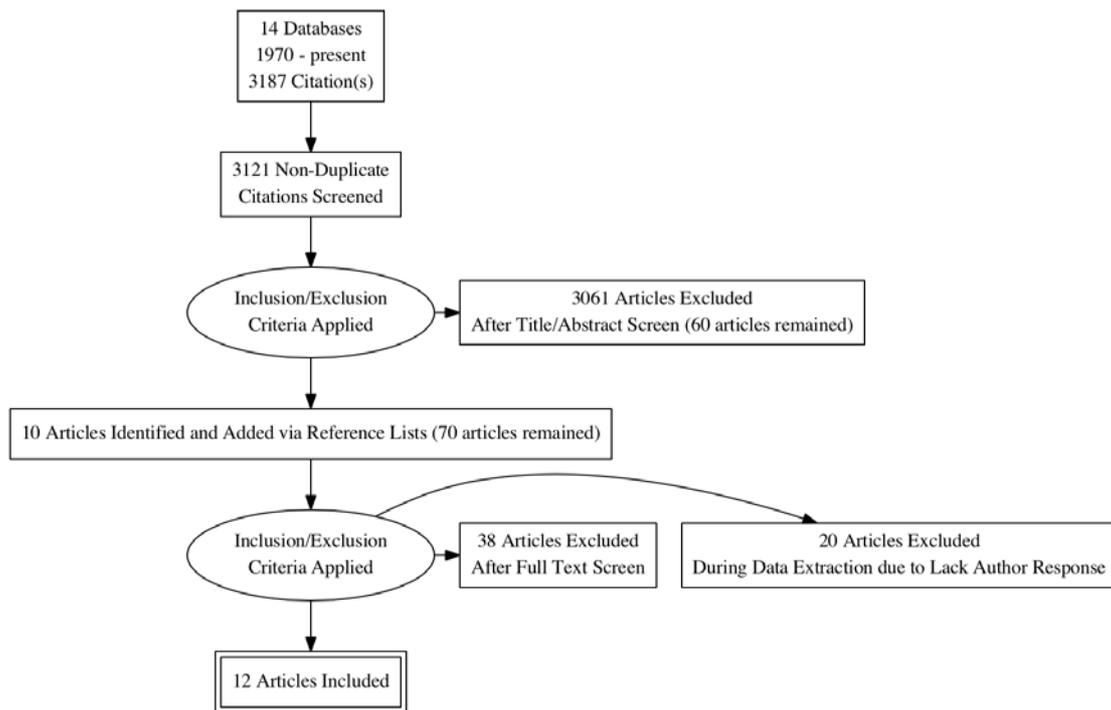


Figure 1 Flow diagram of the systematic review process.

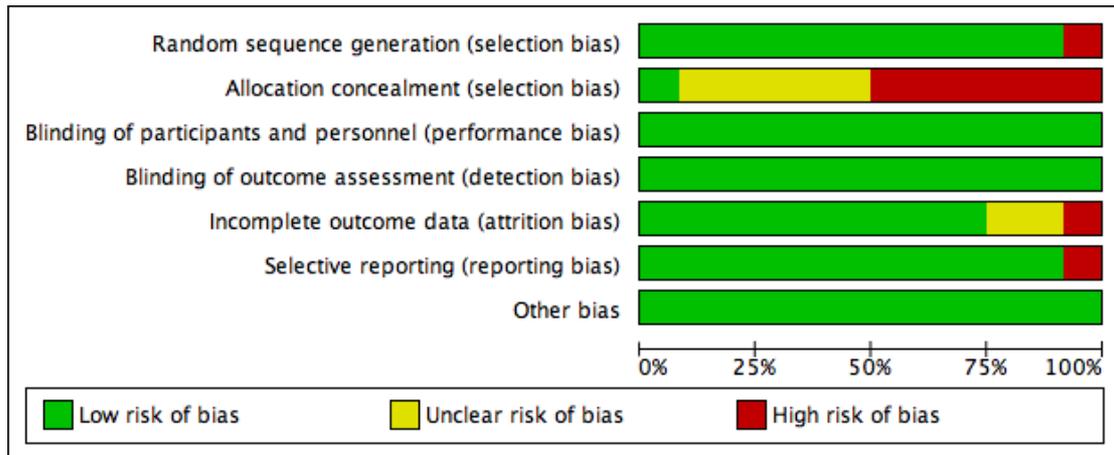


Fig 2. Graph (visual summary of Table 2) detailing breakdown of risk of study bias, stratified by risk category. (Risk of bias determined using Cochrane guidelines²⁹)

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
Baillot et al., 2016	+	?	+	+	+	-	+
Burtscher et al., 2009	+	-	+	+	+	+	+
Church et al., 2009	+	?	+	+	+	+	+
Dalager et al., 2016	+	?	+	+	-	+	+
Donges et al., 2010	-	-	+	+	+	+	+
Donnelly et al., 2013	+	+	+	+	?	+	+
Lockwood et al., 2008	+	-	+	+	?	+	+
Prabhakaran et al., 1999	+	-	+	+	+	+	+
Schmitz et al., 2002	+	?	+	+	+	+	+
Tan et al., 2012	+	-	+	+	+	+	+
Teixera et al., 2003	+	-	+	+	+	+	+
Vilela et al., 2015	+	?	+	+	+	+	+

Fig 3. Graph detailing breakdown of risk of study bias, stratified by study and specific risk factor

Literature Citation	Subjects/Groups	Exercise training program			Results	
		Mode	Length	Intensity/Frequency/Duration/Volume	Δ BW (kg) \pm SD	Other
Baillet <i>et al.</i> , 2016 ⁶⁶	<i>n</i> = 15 (EX), <i>n</i> = 14 (CON)	Endurance and circuit style with 9stations	12 wk	3/wk, 80 mins - 10WU, 50-60MB (30mins endurance, including treadmill, elliptical, arm ergo cycle, 20-30mins strength), 10CD. Endurance at 55-85% HRR	EX -0.92 (3.55), CON -0.3 (4.72)	Pre-Surgical Exercise Training (PreSET) intervention also improved social interaction and PA barriers
Burtscher <i>et al.</i> , 2009 ⁶⁴	<i>n</i> = 18 (EX), <i>n</i> = 18 (CON)	Aerobic training, circuit training	12 months	2/wk, 60mins, aerobic exercise (dancing, walking, running, skating, swimming) eliciting lactate response of 2-3mmol/L, interspersed with higher intensity efforts. Circuits included 6-8 exercises, 8-12 reps. All participants also advised to exercise for 30mins/day	EX -2.58 (4.12), CON 0.79 (4.93)	Counselling & supervised exercise maintained exercise capacity vs counselling alone. In EX, dietary goals (<BW by 5%) not achieved
Church <i>et al.</i> , 2009 ⁵⁶	<i>n</i> = 317 (EX), <i>n</i> = 94 (CON)	Aerobic training alternating treadmill and cycle ergometer	26 wk	3-4/wk, CON + 3 EX groups – 4, 8, 12 Kcal/kg BW, 50% VO ₂ alternating between semi-recumbent cycling and treadmills.	EX - 4 Kcal -1.4 (3.6), 8Kcal -2.1 (3.5), 12 Kcal -1.5 (3.4) Combined intervention -1.62 (3.5), CON -0.9 (3.37)	No difference between predicted and actual weight loss at 4 & 8 Kcal/kg, 12 Kcal/kg lost only half predicted amount
Dalager <i>et al.</i> , 2016 ⁶⁷	<i>n</i> = 89 (EX), <i>n</i> = 195 (CON)	Aerobic and resistance training	1 yr	1/wk, 20 mins aerobic exercise (running, rowing, ball games) 77-95% HR _{max} , 30 mins resistance training 60-80% 1RM for three sets of 8 reps, recommendations to undertake 30mins exercise/day at 64-76% HR _{max}	EX -0.49 (3.32), CON 0.08 (2.97)	5% (ITT) and 10% (PPA) > Δ VO ₂ max in EX than INT, 2.8% ∇ in SBP
Donges <i>et al.</i> , 2010 ⁶²	<i>n</i> = 76 (EX), <i>n</i> = 26 (CON)	Aerobic and resistance training	10 wk	RT 30-50 mins, 2-4 sets of 8-10 reps @ 70-75% of 10RM, AT 30-50 mins cycle ergometer 70-75% MHR	RT 0.8 (1.5), AT - 0.8 (1.9), Combined – -0.06 (1.89) CON 0.6 (1.3)	AT > Δ in body composition than RT & CON. CRP reduced in RT, IL6 unchanged in all groups
Donnelly <i>et al.</i> , 2013 ⁵⁸	<i>n</i> = 74 (EX), <i>n</i> = 18 (CON)	Aerobic training	10 months	5/wk, aerobic exercise – walking/jogging on treadmill (20% of sessions were undertaken on alternative activities such as stationary cycling, elliptical or walking/jogging	400 Kcal -3.9 (4.9), 600 Kcal - 5.2 (5.6), Combined EX -	No significant difference between exercise intervention, suggested some compensatory mechanisms, or when stratified by gender

				outside), expending 400 & 600 Kcal/session	4.55 (5.27), CON 0.5 (3.5)	
Lockwood <i>et al.</i> , 2008 ⁶³	<i>n</i> = 14 (EX), <i>n</i> = 10 (CON)	Aerobic and resistance training	10 weeks	AT 3/wk, self-selected exercise 15-35 mins @ 40-70% HRR, RT 2/wk, 1 set of 8-12 reps (or to failure)	EX -0.3 (1.87), CON -0.3 (1.58)	Individual variation in <i>ad libitum</i> EI reported to be linked with compensatory EI in EX
Prabhakaran <i>et al.</i> , 1999 ⁵⁹	<i>n</i> = 12 (EX), <i>n</i> = 12 (CON)	Resistance Training	14 wk	3/wk, 45-50 mins/session, 85% 1RM, loading major muscle groups, 2 sets of 8 reps plus 1 set to failure, 30-60 seconds rest	EX -0.7 (1.35), CON 0.49 (2.01)	Reduction in lipids and body fat % in EX
Schmitz <i>et al.</i> , 2002 ⁶⁰	<i>n</i> = 27 (EX), <i>n</i> = 27 (CON)	Resistance training	15 wk	2/wk, 50 mins, 3 sets of 8-10 reps, 9 exercises	EX 0.54 (1.87), CON 0.49 (1.82)	Strength training produced favourable Δ in fasting glucose, insulin and cancer risk factors
Tan <i>et al.</i> , 2012 ⁵⁷	<i>n</i> = 29 (EX), <i>n</i> = 19 (CON)	Track running	8 wk	5/wk, 40 mins of running at individualized Fat _{max} HR on outdoor track	EX -4.1 (1.6), CON 0.3 (1.2)	Fat _{max} also decreased fat mass, waist-hip ratio (both possibly related to change in fat oxidation rates), fasting plasma concentration (increased use of fat as fuel) and increased VO _{2max}
Teixeira <i>et al.</i> , 2003 ⁶¹	<i>n</i> = 117 (EX), <i>n</i> = 116 (CON)	RT, circuit and weight bearing aerobic exercise	12 months	3/wk, RT 6-70 mins, 2 sets of 6-8 reps at 70-80% 1RM, AT included walking, jogging, skipping, hopping, 10 mins as WU, then 20-25 mins @ 60% HR _{max}	EX (with HRT/without HRT) -0.2 (2.6)/0.34 (2.5) combined SD 2.55, CON (with HRT/without HRT) 0.8 (2.7)/-0.4 (3.3), combined SD 3.05. Total EX 0.07 (2.55), CON 0.23 (3.05)	Δ LST in all who exercised and non-exercisers not taking HRT, decreased FT on women on HRT. HRT appeared to protect against loss of LST
Vilela <i>et al.</i> , 2015 ⁶⁵	<i>n</i> = 30 (EX), <i>n</i> = 30 (CON)	RT, sporting activity	4 months	5/wk, RT including 2 days upper body exercises and 2 days lower body exercises. 4 x 10mins 3 sets of 30secs work, 30secs recovery, 5 mins flexibility, 1 x 15 mins sporting activity (soccer, volleyball, basketball)	EX 0.0 (2.6), CON 0.4 (2.6)	EX reduced body fat by 4.8 (1.8) %, in the absence of weight loss, suggesting increased lean tissue

Table 1. Studies presenting weight loss response to supervised exercise interventions.

BW body weight, kg kilograms, SD standard deviation, EX exercise condition, CON control condition, wk weeks, mins minutes, WU warm-up, MB main body of exercise session, CD cool-down, HRR heart rate reserve, PA physical activity, Reps repetitions, mmol/L millimole per litre, Kcal Kilocalorie, VO_2 oxygen uptake, Yr year, HR_{max} maximal heart rate, ITT intention to treat, PPA per protocol analysis, VO_{2max} maximal oxygen uptake, SBP systolic blood pressure, RT resistance training, RM repetition maximum, AT aerobic training, CRP C-reactive protein, IL6 – interleukin 6, EI energy intake, Fat_{max} intensity of maximal fat oxidation, VO_{2max} maximal oxygen uptake, HRT hormone replacement therapy, LST lean soft tissue, FT fat tissue, secs seconds.

Literature Citation	Random Sequence Generation		Allocation Concealment		Blinding of Participants		Blinding of Outcome Assessment		Incomplete Outcome Data Addressed		Selective Reporting		Other	
	Risk		Risk		Risk		Risk		Risk		Risk		Risk	Comment
Baillet <i>et al.</i> , 2016 ⁶⁶	Low	Quote "Patients were randomly allocated" Comment: Likely done	Unclear	Quote "Allocation was generated by a computer random sequence and kept in sealed envelopes" Comment: Likely done	Low	Comment: Exercise interventions preclude the blinding of participants to allocated group during the study. It is judged that this would not influence outcomes	Low	Comment: No mention of blinding of outcome assessment. It is judged that this would not influence outcomes	Low	Quote: "the only subject who abandoned the research project was in the usual care group and excluded from analyses". Comment: Likely done	High	Six domains for WRQL in methods, only one reported in written format; others presented in table format.	Low	The study appears free from other sources of bias.
Burtscher <i>et al.</i> , 2009 ⁶⁴	Low	Quote "Patients were randomly assigned" Comment: Likely done	High	Comment: No information provided on method of randomization. Comment: Possibly not done	Low	Comment: Exercise interventions preclude the blinding of participants to allocated group during the study. It is judged that this would not influence outcomes	Low	Comment: No mention of blinding of outcome assessment. It is judged that this would not influence outcomes	Low	Quote: "Due to financial problems, we had to terminate the exercise program at Month 12. To minimize possible bias, 18 patients were then compared to age- and gender-matched patients in a nested cohort approach".	Low	Comment: Study protocol available and all pre-specified outcomes reported in pre-specified way.	Low	The study appears free from other sources of bias.

										Comment: Likely done				
Church <i>et al.</i> , 2009 ⁵⁶	Low	Quote "Patients were randomized to 1 of 3 exercise groups or a non-exercise control" Comment: Likely done	Unclear	Quote "The randomization sequence is computer generated by the study statistician" Comment: Statement found in published rationale paper. Possibly done	Low	Comment: Exercise interventions preclude the blinding of participants to allocated group during the study. It is judged that this would not influence outcomes	Low	Comment: No mention of blinding of outcome assessment. It is judged that this would not influence outcomes	Low	Comment: Missing data relatively balanced across intervention groups. Additionally, missing data were imputed by carrying forward from previous observation (1 week)	Low	Comment: Study protocol available and all pre-specified outcomes reported in pre-specified way.	Low	The study appears free from other sources of bias.
Dalager <i>et al.</i> , 2016 ⁶⁷	Low	Quote "Office workers were randomized 1:1 to a training group or a control group" Comment: Likely done	Unclear	Quote: "The participants were assigned with an arbitrary ID number and randomized individually, using a random number computer algorithm". Comment: Possibly done	Low	Comment: Exercise interventions preclude the blinding of participants to allocated group during the study. It is judged that this would not influence outcomes	Low	Quote: "The study was a 2-year, parallel group, examiner blinded RCT". Comment: Likely done	High	Quote: "Missing values in either baseline or follow-up measurement were substituted with data carried forwards or backwards". Comment: Missing data unbalanced across intervention groups. It is unknown as to what	Low	Comment: Study protocol available and all pre-specified outcomes reported in pre-specified way.	Low	The study appears free from other sources of bias.

										impact this might have on effect sizes.				
Donges <i>et al.</i> , 2010 ⁶²	High	Quote “Participants were semi randomly assigned....80% were randomly assigned, however 20% were allocated according to matching or preference”.	High	Comment: No information provided on method of randomization, other describing it as ‘semi-random’	Low	Comment: Exercise interventions preclude the blinding of participants to allocated group during the study. It is judged that this would not influence outcomes	Low	Comment: No mention of blinding of outcome assessment. It is judged that this would not influence outcomes	Low	Comment: No missing data apparent.	Low	Comment: Study protocol available and all pre-specified outcomes reported in pre-specified way.	Low	The study appears free from other sources of bias.
Donnelly <i>et al.</i> , 2013 ⁵⁸	Low	Quote: “Participants were randomized (2:2:1) to exercise or non-exercise”. Comment: Likely done.	Low	Quote: “Participants were stratified by gender and randomized by an independent statistician”. Comment: Possibly done.	Low	Comment: Exercise interventions preclude the blinding of participants to allocated group during the study. It is judged that this would not influence outcomes	Low	Quote: “Investigators and research assistants were blinded at the level of outcome assessments”. Comment: Likely done.	Unclear	Comment: No methodology for approaching missing data. Missing data relatively balanced across intervention groups.	Low	Comment: Study protocol available and all pre-specified outcomes reported in pre-specified way.	Low	The study appears free from other sources of bias.
Lockwood <i>et al.</i> , 2008 ⁶³	Low	Quote: “Subjects were randomly assigned” Comment: Likely done.	High	Comment: No information provided on method of concealment.	Low	Comment: Exercise interventions preclude the blinding of participants to allocated	Low	Comment: No mention of blinding of outcome assessment. It is judged that this would not	Unclear	Comment: No methodology for approaching missing data. Missing data	Low	Comment: Study protocol available and all pre-specified	Low	The study appears free from other sources of bias.

						group during the study. It is judged that this would not influence outcomes		influence outcomes		relatively balanced across intervention groups.		outcomes reported in pre-specified way.		
Prabhakaran <i>et al.</i> , 1999 ⁵⁹	Low	Quote: "Subjects were randomly assigned to either a non-exercising control group or a resistance exercise training group". Comment: Likely done.	High	Comment: No information provided on method of concealment.	Low	Comment: Exercise interventions preclude the blinding of participants to allocated group during the study. It is judged that this would not influence outcomes	Low	Comment: No mention of blinding of outcome assessment. It is judged that this would not influence outcomes	Low	Comment: Missing data relatively balanced across intervention groups.	Low	Comment: Study protocol available and all pre-specified outcomes reported in pre-specified way.	Low	The study appears free from other sources of bias.
Schmitz <i>et al.</i> , 2002 ⁶⁰	Low	Quote: "Randomized to no-contact control or treatment". Comment: Likely done.	Unclear	Comment: Randomization stratified by decade (30-39, 40-50) due to concerns regarding effects of hormonal changes.	Low	Comment: Exercise interventions preclude the blinding of participants to allocated group during the study. It is judged that this would not influence outcomes	Low	Quote: "Body weight and height measurements, blood draws and DEXA (body composition) were performed by clinical research nurses, blinded to treatment groups". Comment: Likely done.	Low	Comment: Missing data relatively balanced across intervention groups.	Low	Comment: Study protocol available and all pre-specified outcomes reported in pre-specified way.	Low	The study appears free from other sources of bias.

Tan <i>et al.</i> , 2012 ⁵⁷	Low	Quote: “Participants were randomly allocated into two groups”. Comment: Likely done.	High	Comment: No information provided on method of randomization.	Low	Comment: Exercise interventions preclude the blinding of participants to allocated group during the study. It is judged that this would not influence outcomes	Low	Comment: No mention of blinding of outcome assessment. It is judged that this would not influence outcomes	Low	Comment: Missing data relatively balanced across intervention groups. Additionally, reasons unlikely to affect outcome measures.	Low	Comment: Study protocol available and all pre-specified outcomes reported in pre-specified way.	Low	The study appears free from other sources of bias.
Teixeira <i>et al.</i> , 2003 ⁶¹	Low	Quote: “Subjects were randomly allocated to assigned to one year of weight-lifting and weight-bearing exercise or to a group with no exercise.” Comment: Likely done.	High	Comment: Subjects stratified by HRT status.	Low	Comment: Exercise interventions preclude the blinding of participants to allocated group during the study. It is judged that this would not influence outcomes	Low	Quote: “DEXA technicians were blind to participants group assignments”. Comment: Likely done.	Low	Comment: No missing data apparent.	Low	Comment: Study protocol available and all pre-specified outcomes reported in pre-specified way.	Low	The study appears free from other sources of bias.
Vilela <i>et al.</i> , 2015 ⁶⁵	Low	Quote: “Randomly distributed in control and experimental groups”. Comment: Likely done.	Unclear	Quote: “Randomly assigned drawing an opaque envelope”, with “names written on them”.	Low	Comment: Exercise interventions preclude the blinding of participants to allocated group during the study. It is judged that this	Low	Comment: No mention of blinding of outcome assessment. It is judged that this would not influence outcomes	Low	Comment: No missing data apparent.	Low	Comment: Study protocol available and all pre-specified outcomes reported in pre-	Low	The study appears free from other sources of bias.

				Comment: Likely done.		would not influence outcomes						specified way.		
--	--	--	--	--------------------------	--	------------------------------------	--	--	--	--	--	-------------------	--	--

Table 2. Summary descriptives of risk of bias for each of the included studies, in accordance with Cochrane guidelines²⁹. If study methodology did not explicitly state allocation was randomized, then it was deemed 'high risk' of bias for allocation concealment. Only those studies using central randomization, sequentially numbered drug containers or sequentially numbered, opaque, sealed envelopes were deemed 'low risk'.