Associations of environmental characteristics and lifestyle behaviours with donor blood parameters

de Groot, R.

2020

document version
Publisher's PDF, also known as Version of record

- CHAPTER SEVEN -

Association of iron status with physical activity and physical capacity: triangulation of new evidence from a randomised controlled trial, Mendelian randomization and an umbrella review of meta-analyses

Rosa de Groot, Steven Bell, Jeroen Lakerveld, Johannes Brug, Wim de Kort, Katja van den Hurk & Emanuele di Angelantonio, on behalf of the INTERVAL trial group

In preparation

Abstract

Background
The role of iron status on levels of physical activity and physical capacity in adults is not fully established.

Methods
We used three complementary methods to test the hypothesis that iron status is associated with physical activity levels and capacity. First, analyses of blood donation on self-reported and device-measured moderate to vigorous intensity physical activity (MVPA), total physical activity, physical activity energy expenditure, and sedentary behaviour using data from the randomised-controlled INTERVAL trial (N=24496, 48% female; ISCTN trial registration: ISRCTN24760606) of donors randomly assigned to different inter-donation intervals (women: 16, 14 and 12 weeks; men: 12, 10 and 8 weeks). Second, two-sample Mendelian randomisation analyses, taking genetic associations for haemoglobin and ferritin concentrations from one dataset, and genetic associations with device-measured physical activity from another. Third, an umbrella review of meta-analyses of randomised controlled trials and/or prospective observational studies in humans of iron status or supplementation with measures of physical capacity (PROSPERO registration: CRD42019123885).

Findings
No significant differences in physical activity or sedentary behaviour outcomes were found across randomised groups in INTERVAL – using self-reported or device-measured data – including within *a priori* defined subgroups. We observed similar null findings between genetically-predicted haemoglobin or ferritin and device-measured MVPA. Only short-term effects of blood donation on physical capacity-related outcomes were found in our umbrella review, with additional evidence that iron supplementation may only be beneficial for a limited number of measures of physical capacity in anaemic populations.

Interpretation
Given the consistency of our findings across complementary research methodologies, we conclude that medium-to-long term lower iron status does not translate into differences in habitual physical activity levels or capacity.

Introduction

Iron is essential for the synthesis of haemoglobin, which is required for oxygen transport through the body[1]. Physical activity increases oxygen demand of muscle tissue. Thus reduced oxygen capacity – whether that be through lack of ability to deliver oxygen into the body, or poor oxygen carrying capacity in the blood - can negatively affect physical activity[2,3]. In turn, reduced levels of physical activity has implications for health beyond the consequences of lower iron status itself, e.g. increased risk of cardiovascular diseases, cancer and mortality due to a less physically active lifestyle[4,5].

The effect of iron status on measures of physical capacity such as VO_2 max has been extensively researched in various study populations[2,6]. Several studies have found a positive association of iron supplementation with improvements in exercise performance and a reduction in symptoms related to iron deficiency such as fatigue[6,7]. While there are summaries of research in this area, these have typically focussed on single sub-groups of the population (e.g. those with iron deficiency or specific illness), and the effects of iron status with physical activity levels in healthy populations is not clear. While these studies are informative, to understand the breadth by which iron status may influence physical capacity, a more comprehensive synthesis of evidence on this subject is required to understand if iron status influence physical capacity.

Further, physical capacity outcomes may not reflect habitual physical activity levels. Blood donors provide an ideal study population to examine this association as they are required to be in good health to donate and a standard whole blood donation causes immediate changes in haematological parameters, including declining haemoglobin levels and the loss of approximately 250 mg of iron[8]. Studies of donor populations to date have only investigated the effects of and recovery following one to three donations using physical capacity measures, while less is known about the long-term consequences of repeated blood donation on (habitual) physical activity[9].

The INTERVAL study, a large randomised controlled trial (RCT) that allocated donors to different donation intervals[10], found that although more intensive donation regimes were associated with lower mean haemoglobin and iron stores (ferritin) at 2 years, overall, no relation with total energy expenditure was found[11]. However, physical activity is only a component of total energy expenditure. In addition, these analyses were based on self-reported questionnaires predominantly capturing structured planned physical activity. These questionnaires have limited validity and reliability, because of recall and social desirability biases[12]. Accelerometers can provide objectively measured estimations of free-living activity that are less prone to these biases[13].

Furthermore, previous reports from INTERVAL did not consider whether this null association was universal across sub-populations. For example, it is possible that donors prone to iron deficiency might show significant decreases in their level of physical activity if assigned to a more demanding donation schedule.

It could be argued that blood donors are a selected group of the population in which the association between iron status and levels of physical activity may not be generalizable to the wider population. Replication of findings in other populations or using other study designs is therefore essential. One means of achieving this is to leverage natural genetic variation in individuals to emulate a clinical trial ("Mendelian randomisation") based on the principle that genotypes are randomly assigned at meiosis and therefore will generally not be affected by confounding or reverse causation[14].

The purpose of this paper is therefore to holistically assess whether iron status affects physical activity and physical capacity levels in adults. Accordingly, we examined the association of regular blood donation - and thus reduced iron status - on physical activity and sedentary behaviour outcomes. We did this using repeated self-reported questionnaire data and objective wrist-worn accelerometer data in the INTERVAL study[15]. Next, we utilised published genetic data to further assess the causal association of iron markers with levels of physical activity via Mendelian randomisation. Finally, to complement these analyses, we conducted an umbrella systematic review of published meta-analyses to provide a comprehensive overview of evidence on the topic of iron status and tests of physical capacity across different study populations.

Methods

We used three complementary approaches to study the association between iron status and physical activity and capacity: 1) randomised controlled trial, 2) Mendelian randomisation, and 3) umbrella systematic review. These approaches are outlined separately below.

1. Randomised controlled trial - INTERVAL

Study design and participants

For this study we used data from the INTERVAL study (ISRCTN registration number: ISRCTN24760606), a large, parallel group, pragmatic, randomised trial conducted at the NHS Blood and Transplant (NHSBT) that was initially scheduled for 2- and later extended to 4-years in England, UK. Details of the study rationale, design and recruitment have been described in-depth elsewhere[10,16]. Briefly, donors who were aged 18 years or older, were willing to be randomly assigned to any of the trial's sex-specific intervention groups and

met the eligibility criteria for blood donation were recruited between 2012 and 2014. The National Research Ethics Service approved (11/EE/0538) this study, and all donors provided informed consent prior to donation.

Randomisation

Men were randomly assigned to inter-donation intervals of 12-weeks (standard), 10-weeks or 8-weeks, and women to 16-weeks (standard), 14-weeks or 12-weeks. Randomisation was in a sex-specific ratio of 1:1:1 allocated by use of a computer program built into the trial database at the coordinating centre, with a minimisation algorithm to ensure that key characteristics (e.g. age, weight and donor status) were balanced across trial groups at baseline. Further, randomisation was stratified by donation centre and sex. Due to the nature of this trial, blinding to randomisation was not feasible, the assigned inter-donation interval was recorded in the NHSBT donor database and donors were informed via email.

Procedures

Blood samples

Prior to the donation at enrolment and the 2- and at the 4-year resurvey, a non-fasting research blood sample was taken and transported to a central laboratory for a full blood count analysis (Sysmex XN-2000 haematology analyser, UK BioCentre, Stockport, UK).

Aliquots of EDTA (edetic acid) plasma, serum, and buffy coat were stored at –80°C. Ferritin was measured in serum samples with an immunoturbidimetric assay (Roche/Hitachi chemistry analyser, Stichting Huisartsen Laboratorium, Etten-Leur, Netherlands). Laboratory analysts were blinded to trial allocation. The INTERVAL trial adhered to the standard NHSBT deferral policy. As such, only a research sample was taken if a donor was deferred or unable to donate.

Questionnaires

The baseline questionnaire ascertained information regarding lifestyle behaviours (i.e. diet, alcohol intake, smoking), medication and supplement use, and history of iron deficiency. At the 2-year and 4-year follow-up assessments, information on physical activity was recorded using the Recent Physical Activity Questionnaire (RPAQ)[17,18]. Summary variables including PAEE (kJ/kg/day), time spent in sedentary behaviour, moderate-to-vigorous and total physical activity (minutes per day) were derived using the RPAQ data processing guidelines[19]. We updated metabolic equivalent task (MET) values using the latest version of the Compendium of Physical Activity[20].

Accelerometers

Physical activity was also measured using a tri-axial AX3 accelerometer (Axivity, York, UK), from 2014 until 2016 in a subset of the INTERVAL cohort, as part of the 2-year follow-up data collection. A summary of the accelerometer study can be found in Appendix 1. Briefly, donors who expressed an interest in participating in the objective physical activity monitoring study were chosen at random from each trial group to ensure equal representation in the sub-study (Supplementary Figure 1). Participants were asked to wear the accelerometer continuously on their dominant wrist over a 7-day period. Data processing procedures used in the UK Biobank were adopted which have been detailed elsewhere[21]. Moderate-to-vigorous physical activity (MVPA) was defined as Euclidean Norm Minus One values greater than 125 milli-*gravity* (approximately 3 METS) and expressed in minutes per day[13]. A comparison of the accelerometer subsample with other INTERVAL participants and the NHSBT general donor population is shown in Supplementary Table 1. Bland and Altman analyses comparing self-reported and objectively measured MVPA showed proportional bias, that is, with a higher mean of the two measurements the difference between the two measurements increased (the bias) (Supplementary Figure 2). Modest correlations between MVPA as measured in this study and known indicators of health are shown in Supplementary Figure 3.

Outcome measures

Primary outcomes were between-group differences in self-reported time spent in MVPA, total energy expenditure, and sedentary behaviour at 2-years and changes between 2- and 4-years. Further primary outcomes were time spent in MVPA and total physical activity at 2-years as measured by accelerometer. Supporting analyses of the primary outcomes consisted of correlations between biological parameters and physical activity, and change in iron status and physical activity.

Statistical Analysis

The statistical analyses followed a pre-specified plan (Appendix 2). Briefly, data for men and women were analysed separately by the intention-to-treat principle according to their randomised groups with the longest sex-specific interval group as reference group (i.e., current practice). All primary outcomes are presented as (geometric) means with 95% confidence intervals. We compared randomised groups by calculating p-values for differences or linear trends using multilevel linear regression models. To account for multiple testing Benjamini and Hochberg p-values were calculated with a false-discovery rate of 10%[22]. Information on the supporting analyses of the main outcomes are presented in Appendix 2-3, Supplementary Figure 4.

Statistical power in INTERVAL was based on the primary and key secondary endpoints of the trial, the outcomes assessed in this study were predefined secondary outcomes of the initial protocol of the INTERVAL trial[10,11].

2. Mendelian randomisation

Data Sources and Variant Selection

Two-sample Mendelian randomisation was used to assess the causal association of haemoglobin and iron biomarkers with MVPA and total physical activity levels, using publicly available data. We obtained genetic associations of haemoglobin[23] (n=128,120) and ferritin[24] (n=48,972) using data on individuals of European ancestry. This included an initial list of 24 genetic variants associated with either haemoglobin or ferritin at $p < 5 \times 10^{-8}$. We used Phenoscanner[25,26] to cross-reference these variants (and their proxies, defined as having an $r^2 \geq 0.8$ in the 1000 Genomes project phase 3 release European sample) to ensure that they were not associated with known confounders of the association between iron and physical activity (e.g. smoking status). Thereafter, we obtained genetic associations of the with device-measured physical activity (Axivity AX3 wrist-worn triaxial accelerometer) in 91,105 UK Biobank participants of European descent[27].

Instrumental variable analyses

Univariable (two-sample) Mendelian randomisation using the inverse-variance weighted method was employed as our primary analysis. Sensitivity analyses comprised of the weighted median method to assess whether associations were robust to invalid genetic instruments[28] and MR-Egger analyses to test for unmeasured directional pleiotrophy[29]. To account for potential genetic pleiotropy between markers we also performed multivariable Mendelian randomisation analyses[30]. All analyses were conducted with R version 3.6.0 using Mendelian randomisation[31], version 0.4.1.

3. Umbrella systematic review

The protocol of this umbrella review was registered in PROSPERO under registration number CRD42019123885. See Appendix 4 for a detailed description of the search, article selection, data extraction, quality assessment, data synthesis, and analysis. In short, in March 2019 we searched 8 bibliographic databases in collaboration with a clinical librarian (LS) for peer-reviewed systematic reviews with meta-analysis that included randomised controlled trials or prospective observational studies in humans, with iron status and/or supplementation (also including erythropoiesis-stimulating agents [ESA]) as the exposure, and any measure of physical capacity as the outcome. We restricted the search to articles that were published in English after 1990. All identified publications were screened by two reviewers independently

(RdG and SB) based initially on title and abstract, and later by full-text. Subsequently, all included studies were independently assessed by the same authors for general methodological quality of the systematic review using the AMSTAR2 checklist. We extracted systematic review (year of publication, study design and population, number of included studies, effect sizes) and original study (study population and sample size per group and effect sizes) related information. Included studies were divided in two groups, those that involved blood withdrawal studies (e.g. immediate change in iron status) and those aimed at improving iron status (e.g. through iron supplementation or ESA). The former will be referred to as blood withdrawal studies and the latter as iron supplementation studies herein. A narrative synthesis was structured around the type of outcome for both blood withdrawal and iron supplementation studies (with sub-group analyses based on study population, e.g. general, diseased populations and iron deficient non-anaemic). Each meta-analysis was re-analysed with a random effects model, including estimations of heterogeneity (I^2). Publication bias was assessed in outcomes with more than 10 studies using contour enhanced funnel plots.

Role of the funding source
The academic investigators and representatives of NHSBT, a funder of the INTERVAL trial, participated in the study design and oversight. The investigators at the study's academic coordinating centre had sole access to the study database, and had final responsibility for data collection, data integrity, data analysis, and data interpretation, as well as manuscript drafting and the decision to submit the manuscript for publication. All authors gave approval to submit for publication.

Results

1. Randomised controlled trial - INTERVAL

Availability on primary outcomes ranged between 58.8 and 99.8% for RPAQ at 4-years and RPAQ at 2-years respectively. Accelerometers were worn by approximately a quarter of participants. Missing data were similar across randomised groups (Supplementary Figure 5). Trial arm was not associated with drop-out at 4-year in either men or women. Older donors, those with more donations 2 years prior to entering INTERVAL and those with higher baseline ferritin levels were more likely to be retained at 4-year, while new female donors (no donations prior to INTERVAL) were more likely to drop-out (Supplementary Table 2-4).

Table 1 shows characteristics of participants at the 2- and 4-year follow-ups with physical activity information (RPAQ and/or accelerometer data) and

Table 1: Participant characteristics INTERVAL by sex and intervention group.

	Female			Male		
	12 weeks	14 weeks	16 weeks	8 weeks	10 weeks	12 weeks
2-year follow-up						
No. of participants	3965	3969	3894	4206	4222	4240
Age[1], years	44.9 ± 13.6	45.0 ± 13.5	45.2 ± 13.4	47.8 ± 13.2	48.3 ± 13.3	48.2 ± 13.2
Body mass index[1]	25.9 ± 4.9	26.0 ± 5.0	26.1 ± 5.1	26.7 ± 4.1	26.6 ± 4.1	26.5 ± 3.9
New donors[2]	268 (6.8%)	268 (6.8%)	291 (7.5%)	194 (4.6%)	204 (4.8%)	210 (5.0%)
No. donation 2-year prior to INTERVAL	3.24 ± 1.64	3.26 ± 1.62	3.21 ± 1.64	3.94 ± 1.80	3.96 ± 1.78	3.90 ± 1.77
No. of donation attempts during INTERVAL[3,4]	6.77 ± 1.68	6.07 ± 1.37	5.34 ± 1.16	9.92 ± 2.33	8.42 ± 1.79	7.18 ± 1.40
Deferral rate for low haemoglobin during INTERVAL	7.19%	6.27%	5.06%	5.72%	3.74%	2.55%
4-year follow-up						
No. of participants[5]	2367	2335	2298	2558	2696	2602
New donors[1]	102 (5.0%)	101 (5.0%)	117 (5.7%)	79 (3.6%)	85 (3.6%)	91 (3.9)
No. of donation attempts during INTERVAL[4,6]	4.0 ± 1.8	3.5 ± 1.6	3.1 ± 1.4	5.9 ± 2.6	5.1 ± 2.1	4.3 ± 1.7

Data are presented as mean ± SD or number of participants (%) or % unless otherwise indicated. [1]At baseline INTERVAL. [2]Participants who had not previously provided a whole blood donation. [3]During the first 2-years of INTERVAL. [4]All donation attempts, including successful. [5]Includes donors who started the Recent Physical Activity Questionnaire at 4-year. [6]Between 2- and 4-year of INTERVAL.

iron status (haemoglobin and/or ferritin levels) available at 2 years by sex and trial arm.

2- and 4-year outcomes

Tests for trend showed that self-reported PAEE, time spent in MVPA and sedentary behaviour at 2- and 4-year did not differ across randomised groups (Figure 1 and 2; Supplementary Table 5).

Similarly, there was no evidence that device measured MVPA or total physical activity at 2-year differed across randomised groups. Sensitivity analyses with different MVPA thresholds (e.g. 100 and 150m*g*) showed similar results (Supplementary Figure 6). Our findings did not vary across any of the predefined potential effect modifiers (i.e. haemoglobin and ferritin levels at baseline, the use of iron supplementation, or age) at either 2- or 4-year. In total, 9 tests were nominally significant, i.e. p <0.05, however, after accounting for multiple testing with the Benjamini and Hochberg method none were considered statistically significant (Supplementary Table 5). Post-hoc analysis with average acceleration as a physical activity volume measure did not show differences across groups (Supplementary Figure 7).

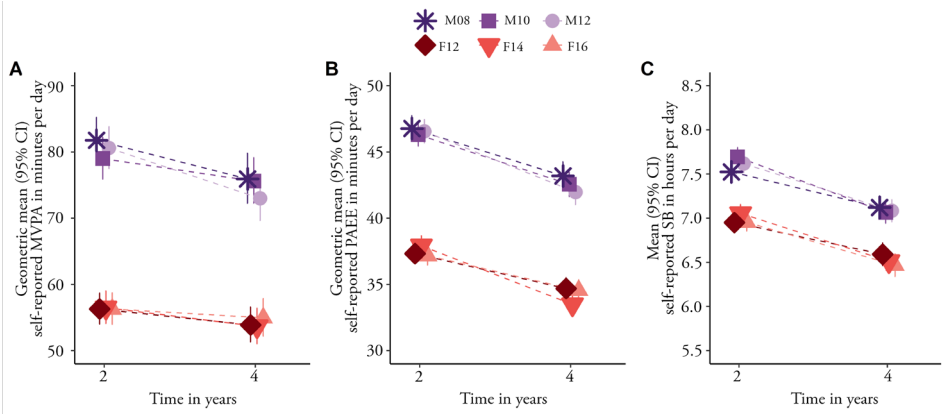


Figure 1: Levels of geometric mean (95% CI) self-reported MVPA (A), geometric mean (95% CI) self-reported PAEE (B) and mean (95% CI) sedentary behaviour (C) outcomes by sex (females red, males purple) and intervention group at 2- and 4-year.

MVPA: moderate-to-vigorous physical activity. PAEE: physical activity energy expenditure. SB: sedentary behaviour.

Change in physical activity between 2- and 4-year follow-up

In line with results obtained examining outcomes separately at 2- and 4-year, analyses of change between these two time points showed no differences in self-reported PAEE, MVPA, nor time spent in sedentary behaviour across donation-intervals or subgroups. Only a greater decrease in PAEE in women

assigned to the 14 week interval compared to women assigned to a 12 week interval was found (p = 0.006), however this was likely a false-positive outcome (Benjamini-Hochberg p-value = 0.42) (Figure 2; Supplementary Table 5).

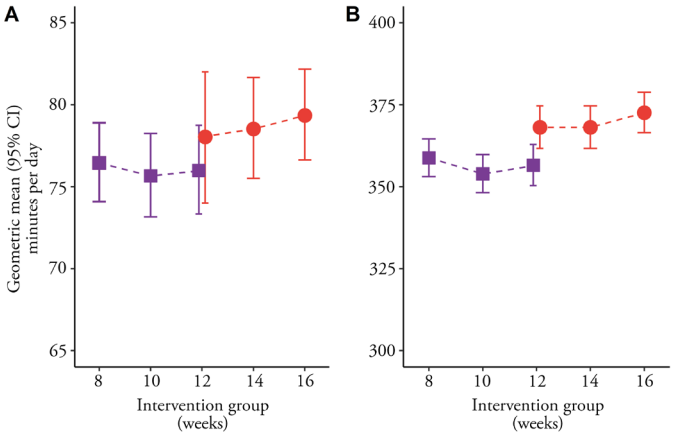


Figure 2: Levels of geometric (95% CI) accelerometer derived moderate-to-vigorous physical activity (A) and total physical activity (B) by sex (females red circles, males purple squares) and intervention group at 2-year.

Likewise, the bivariate dual change score models assessing individual level change in MVPA and iron (haemoglobin and ferritin separately) revealed no evidence to reject a model in which neither iron nor MVPA are associated with change in the other (Supplementary Table 6-9; Supplementary Figure 8).

2. Mendelian randomisation

Concordant with the above, we also found no association between genetically-predicted haemoglobin and amount of time spent in device-measured MVPA (standard deviation difference in MVPA per standard deviation higher level of haemoglobin, β 0.03, 95%CI -0.02 to 0.08, p=0.21) or ferritin (β 0.01, 95%CI -0.05 to 0.06, p=0.85; Figure 3). Our findings were consistent across sensitivity analyses, including when using multivariable methods (Supplementary Table 10).

3. Umbrella review

As shown in Supplementary Figure 9 our systematic search yielded a total of 2,076 potential studies. After assessing titles and abstracts, and subsequently full-texts, 15 systematic reviews with meta-analysis met our eligibility criteria.

We were unable to include two of these as the analyses were based on original

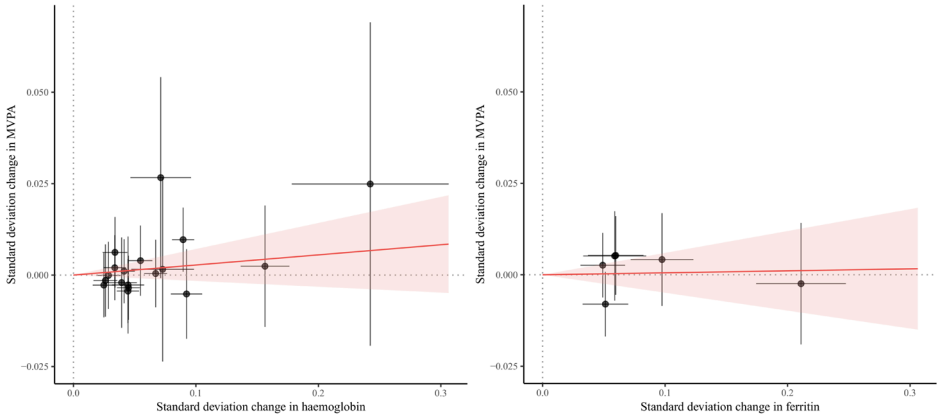


Figure 3: Cross-hair plots of univariable inverse-variance weighted Mendelian randomisation analyses of genetically-predicted haemoglobin and ferritin with device measured moderate-to-vigorous physical activity.

studies that were not reported, and our efforts to contact the review authors were unsuccessful. Supplementary Table 11 shows a complete list accompanied with reasons for not including papers of which the full-text was assessed.

7

General characteristics of the included studies are presented in Supplementary Table 12 and the bias assessment, which generally indicated a high risk, is outlined in Appendix 5. Six studies included heart disease patients (heart failure, coronary heart disease)[32-36], three iron deficient-non anaemic (IDNA) study populations[7,37-39], one women of reproductive age[40] and two included study populations following blood donation[9,41]. Results for the latter are presented separately as these involved pre- and post-blood withdrawal results thus investigating acute effects of blood donation rather than more general effects of iron status on physical capacity. In total 11 physical capacity outcomes were recorded by 61 original studies, of which 23 studies were reported by more than one included review (Supplementary Table 13).

Outcomes

Overall, pooled effect estimates showed that iron supplementation increased/improved absolute and relative VO_2 max, VO_2 peak, exercise duration, VO_2 at anaerobic threshold and six minutes walking distance compared to placebo or control – although many of these findings were based on primarily anaemic populations. No differences were found for time to exhaustion, respiratory exchange ratio, lactate, or energy expenditure in pooled meta-analyses. In blood withdrawal studies, relative VO_2 max, time to exhaustion and peak work rate significantly decreased after bloodletting. Information about the assessment of the outcomes is provided in Appendix 6. A more in-depth

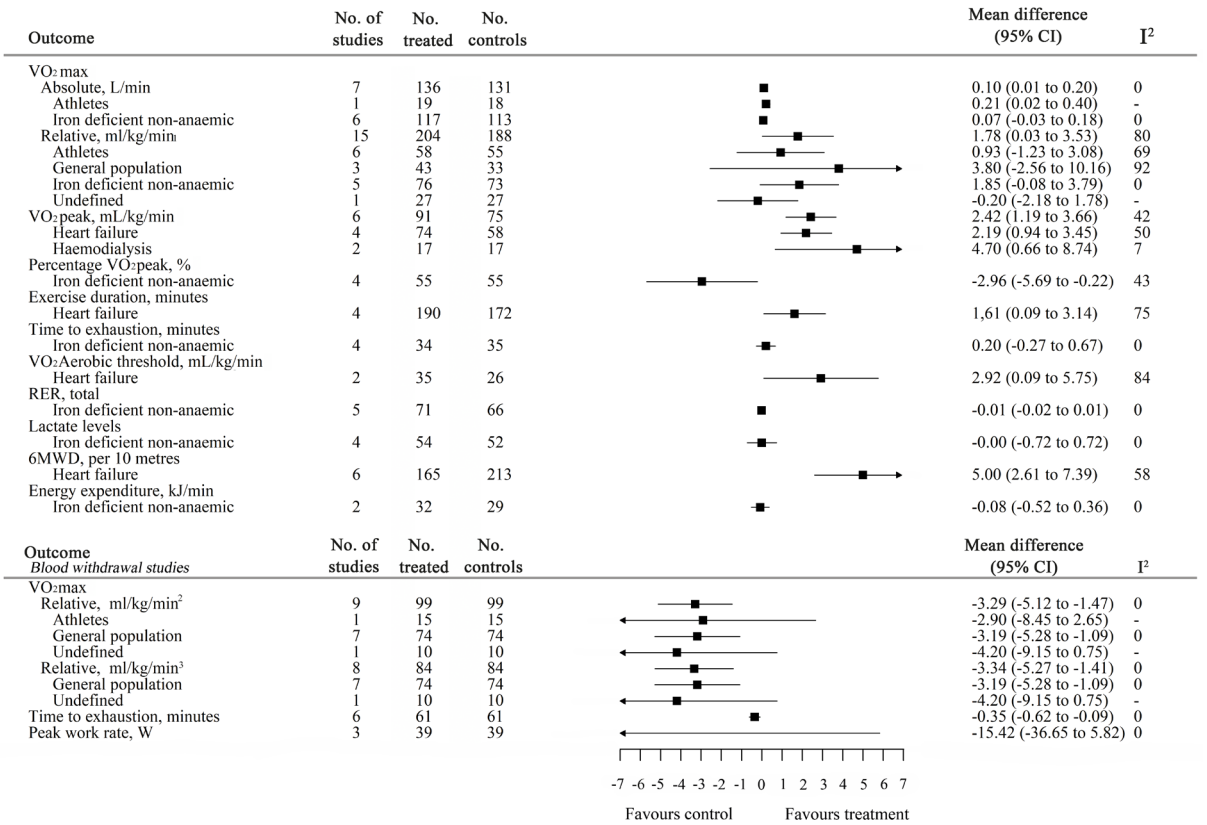


Figure 4: Overall and subgroup pooled effect estimates from meta-analyses of iron supplementation or erythropoietin stimulating agents (top panel) or blood withdrawal (lower panel) with measures of physical capacity.

1Sensitivity analysis omitting Gordon et al, (2014). IDNA: Iron deficient non-anaemic. RER: Respiratory exchange ratio. 6MWD: 6 Minutes walking distance.

overview is provided in Appendix 7. Pooled and sub-population effect estimates, together with heterogeneity index information, are shown in Figure 4. Forest plots for each outcome are provided in Supplementary Figures 12-25. No publication bias was detected (Supplementary Figure 26).

Discussion

We combined evidence from a large randomised controlled trial and Mendelian randomisation and found that lower iron status does not appear to affect self-reported or device-measured physical activity levels in generally healthy people. Our umbrella review suggests that only in anaemic populations (low haemoglobin and ferritin levels) iron supplementation seems beneficial in increasing physical capacity outcomes. Supplementation in iron-deficient non-anaemic populations characterized by low ferritin levels but still adequate haemoglobin levels did not result in improved physical capacity outcomes. Our umbrella review revealed only short-term effects of blood donation on physical capacity outcomes and that the beneficial effects of iron supplementation may be primarily limited to anaemic populations.

As donors in most Western countries do not donate *maximally* or as frequent as INTERVAL donors, our findings are reassuring for blood services as well as for donors about how blood donation may impact their level of physical activity. However, in our umbrella review we found evidence that relative VO_2 max can be reduced acutely (24-48 hours) following whole blood withdrawal. This may not be entirely attributable to an acute change in iron status, but also to physiological changes in blood volume and blood pressure[3]. The findings may partially explain results from observational studies where tiredness after blood donation is often reported by donors[42]. These short-term effects could have implications for certain populations, for example, competing athletes.

Only in anaemic patient populations it appears that iron treatment (usually over a longer time period) influences measures of physical capacity. Except for the percentage VO_2 peak aerobic capacity, no other physical capacity outcome that was studied exclusively in IDNA populations showed any difference between iron treatment and control. Operationalisation of non-anaemic iron deficiency differed between systematic reviews and even more between the original studies. It seems that providing the oxygen-carrying capacity is acceptable, reflected by a normal range of haemoglobin levels, differences in non-haem-iron stores do not seem to affect physical capacity in the general population. This may be one explanation for the fact that we did not see changes in physical activity levels during our trial. Presumably, primarily haemoglobin-loss (associated with blood donation) and thus reduced oxygen carrying capacity is of most importance for habitual physical activity. We and

others[11,15] found that mean haemoglobin values reduced during the course of the INTERVAL study and this was more pronounced in the more intense inter-donation interval groups. Although diminished, these levels were still within the normal range and thus likely adequate so as not to affect habitual physical activity.

It is probable that small changes in haemoglobin only become apparent at the anaerobic threshold[43]. This corresponds to intense vigorous physical activity, which typically reflects only a minor proportion of an individual's overall physical activity. This hypothesis is further supported by the findings from our umbrella review on physical capacity outcomes. There we showed that beneficial effects of iron treatment were predominantly present in patient populations with both low (anaemic) haemoglobin levels and iron stores, and thus had greater potential for improvement.

Our investigation has several major strengths including the triangulation of findings from robust research designs. The consistency of our findings across different methodologies gives us greater confidence in our overall conclusions. With over 24,000 participants at the 2-year follow-up, the INTERVAL study is the largest study so far documenting the effects of different blood donation intervals over a prolonged period. This enabled the investigation of long-term effects of iron status on physical activity in a healthy population, including within *a priori* defined subgroups. That the results obtained from self-reported physical activity were concordant with those obtained when using data from over 5,600 device-measured individuals is another merit.

However, it is also important that the results of our study are interpreted in context of the limitations. Although trial arm allocation was not associated with attrition, older donors and those with more donations two years prior to entering the study were less likely to drop-out of the trial, limiting extrapolation to the wider donor population. However, our null trial arm findings were consistent with results from Mendelian randomisation analyses of iron biomarkers and device-measured physical activity. Here we used large non-donor populations, which gives us confidence in our conclusions despite such selection biases within the trial. The robustness of the results in the umbrella review largely depended on the quality and availability of included reviews and, in turn the quality of the original studies. Risk of bias was found to be rather high, partly due to fact that included studies were published prior to the dissemination of the AMSTAR2 checklist. Given our wide-angled approach and consistency in results with other methods it could be argued that the effects of this bias may be limited, however, remain a possibility.

In the INTERVAL trial, the device-measured MVPA levels may seem rather high (geometric mean of approximately 1.25 hours per day), however these levels are comparable to other large population-based studies using accelerometers[44-46]. Different thresholds clearly influence reported mean activity levels, although our sensitivity analysis showed that more lenient or stringent cut-off points did not alter our conclusions.

As previously noted, our results on the impact of donation on physical activity levels may be reassuring to blood donors and blood services, however, lower iron status resulting from donation may affect other health-related outcomes both in a positive and negative way. This needs to be thoroughly investigated before shorter inter-donation intervals can be considered safe in the long-term.

Given our comprehensive efforts to use the largest and most robust available data sources, and consistency in our findings across methodologies, it can be concluded that medium-to-long term lower iron status (lower ferritin but still adequate haemoglobin levels) does not translate into differences in habitual physical activity levels.

Acknowledgments and funding
The academic coordinating centre would like to thank blood donor centre staff and blood donors for participating in the INTERVAL trial.

Participants in the INTERVAL randomised controlled trial were recruited with the active collaboration of NHS Blood and Transplant England (www.nhsbt.nhs.uk), which has supported field work and other elements of the trial. DNA extraction and genotyping was co-funded by the National Institute for Health Research (NIHR), the NIHR BioResource (http://bioresource.nihr.ac.uk) and the NIHR [Cambridge Biomedical Research Centre at the Cambridge University Hospitals NHS Foundation Trust] [*]. The academic coordinating centre for INTERVAL was supported by core funding from: NIHR Blood and Transplant Research Unit in Donor Health and Genomics (NIHR BTRU-2014-10024), UK Medical Research Council (MR/L003120/1), British Heart Foundation (SP/09/002; RG/13/13/30194; RG/18/13/33946) and the NIHR [Cambridge Biomedical Research Centre at the Cambridge University Hospitals NHS Foundation Trust] [*]. A complete list of the investigators and contributors to the INTERVAL trial is provided in reference [**]. The academic coordinating centre would like to thank blood donor centre staff and blood donors for participating in the INTERVAL trial.

This work was supported by Health Data Research UK, which is funded by the UK Medical Research Council, Engineering and Physical Sciences

Research Council, Economic and Social Research Council, Department of Health and Social Care (England), Chief Scientist Office of the Scottish Government Health and Social Care Directorates, Health and Social Care Research and Development Division (Welsh Government), Public Health Agency (Northern Ireland), British Heart Foundation and Wellcome.

*The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

**Di Angelantonio E, Thompson SG, Kaptoge SK, Moore C, Walker M, Armitage J, Ouwehand WH, Roberts DJ, Danesh J, INTERVAL Trial Group. Efficiency and safety of varying the frequency of whole blood donation (INTERVAL): a randomised trial of 45 000 donors. Lancet. 2017 Nov 25;390(10110):2360-2371.

Contributions

This work is currently in preparation. RdG, SBe, KvdH, JL, JB, WdK and EDA devised the research question and analyses plan. SB, EA, RS conducted the Mendelian Randomisation analyses. LS, RdG and SBe developed the systematic review search strategy. RdG performed all systematic review and INTERVAL analyses. Steven Bell supervised the analyses. RdG.and SBe wrote the manuscript with input of the manuscript from all authors. This is a project of the INTERVAL trial research group. All contributing authors are listed in alphabetical order: Elias Allara, Steven Bell, Soren Brage, Johannes Brug, Emanuele Di Angelantonio, Rosa de Groot, Mark Hamer, Katja van den Hurk, Stephen Kaptoge, Wim de Kort, Jeroen Lakerveld, Linda Schoonmade, Rebecca Smith and Maike Sweegers.

References

1 Abbaspour, N., Hurrell, R. & Kelishadi, R. Review on iron and its importance for human health. J Res Med Sci 19, 164-174 (2014).

2 Beard, J. & Tobin, B. Iron status and exercise. Am J Clin Nutr 72, 594S-597S, doi:10.1093/ajcn/72.2.594S (2000).

3 Calbet, J. A., Lundby, C., Koskolou, M. & Boushel, R. Importance of hemoglobin concentration to exercise: acute manipulations. Respir Physiol Neurobiol 151, 132-140, doi:10.1016/j.resp.2006.01.014 (2006).

4 Warburton, D. E. R. & Bredin, S. S. D. Health benefits of physical activity: a systematic review of current systematic reviews. Curr Opin Cardiol 32, 541-556, doi:10.1097/HCO.0000000000000437 (2017).

5 Arem, H. et al. Leisure time physical activity and mortality: a detailed pooled analysis of the dose-response relationship. JAMA Intern Med 175, 959-967, doi:10.1001/jamainternmed.2015.0533 (2015).

6 Pasricha, S. R., Low, M., Thompson, J., Farrell, A. & De-Regil, L. M. Iron supplementation benefits physical performance in women of reproductive age: a systematic review and meta-analysis. J Nutr 144, 906-914, doi:10.3945/jn.113.189589 (2014).

7 Houston, B. L. et al. Efficacy of iron supplementation on fatigue and physical capacity in non-anaemic iron-deficient adults: a systematic review of randomised controlled trials. BMJ Open 8, e019240, doi:10.1136/bmjopen-2017-019240 (2018).

8 Brittenham, G. M. Iron deficiency in whole blood donors. Transfusion 51, 458-461, doi:10.1111/j.1537-2995.2011.03062.x (2011).

9 Van Remoortel, H., De Buck, E., Compernolle, V., Deldicque, L. & Vandekerckhove, P. The effect of a standard whole blood donation on oxygen uptake and exercise capacity: a systematic review and meta-analysis. Transfusion 57, 451-462, doi:10.1111/trf.13893 (2017).

10 Moore, C. et al. The INTERVAL trial to determine whether intervals between blood donations can be safely and acceptably decreased to optimise blood supply: study protocol for a randomised controlled trial. Trials 15, 363, doi:10.1186/1745-6215-15-363 (2014).

11 Di Angelantonio, E. et al. Efficiency and safety of varying the frequency of whole blood donation (INTERVAL): a randomised trial of 45 000 donors. Lancet 390, 2360-2371, doi:10.1016/S0140-6736(17)31928-1 (2017).

12 Helmerhorst, H. J., Brage, S., Warren, J., Besson, H. & Ekelund, U. A systematic review of reliability and objective criterion-related validity of physical activity questionnaires. Int J Behav Nutr Phys Act 9, 103, doi:10.1186/1479-5868-9-103 (2012).

13 White, T., Westgate, K., Wareham, N. J. & Brage, S. Estimation of Physical Activity Energy Expenditure during Free-Living from Wrist Accelerometry in UK Adults. PLoS One 11, e0167472, doi:10.1371/journal.pone.0167472 (2016).

14 Burgess, S. et al. Using published data in Mendelian randomisation: a blueprint for efficient identification of causal risk factors. Eur J Epidemiol 30, 543-552, doi:10.1007/s10654-015-0011-z (2015).

15 Kaptoge, S. et al. Longer-term efficiency and safety of increasing the frequency of whole blood donation (INTERVAL): extension study of a randomised trial of 20 757 blood donors. Lancet Haematol, doi:10.1016/S2352-3026(19)30106-1 (2019).

16 Moore, C. et al. Recruitment and representativeness of blood donors in the INTERVAL randomised trial assessing varying inter-donation intervals. Trials 17, 458, doi:10.1186/s13063-016-1579-7 (2016).

17 Golubic, R. et al. Validity of electronically administered Recent Physical Activity Questionnaire (RPAQ) in ten European countries. PLoS One 9, e92829, doi:10.1371/journal.pone.0092829 (2014).

18 Besson, H., Brage, S., Jakes, R. W., Ekelund, U. & Wareham, N. J. Estimating physical activity energy expenditure, sedentary time, and physical activity intensity by self-report in adults. Am J Clin Nutr 91, 106-114, doi:10.3945/ajcn.2009.28432 (2010).

19 MRC Epidemiology Unit. Recent Physical Activity Questionnaire (RPAQ): RPAQ_ Guidelines and Syntax v1.5, <http://www.mrc-epid.cam.ac.uk/physical-activity-downloads/> (2018).

20 Ainsworth, B. E. et al. 2011 Compen-

dium of Physical Activities: a second update of codes and MET values. Med Sci Sports Exerc 43, 1575-1581, doi:10.1249/MSS.0b013e31821ece12 (2011).

21 Doherty, A. et al. Large Scale Population Assessment of Physical Activity Using Wrist Worn Accelerometers: The UK Biobank Study. PLoS One 12, e0169649, doi:10.1371/journal.pone.0169649 (2017).

22 Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B (Methodological) 57, 289– 300 (1995).

23 Chami, N. et al. Exome Genotyping Identifies Pleiotropic Variants Associated with Red Blood Cell Traits. Am J Hum Genet 99, 8-21, doi:10.1016/j.ajhg.2016.05.007 (2016).

24 Benyamin, B. et al. Novel loci affecting iron homeostasis and their effects in individuals at risk for hemochromatosis. Nat Commun 5, 4926, doi:10.1038/ncomms5926 (2014).

25 Staley, J. R. et al. PhenoScanner: a database of human genotype-phenotype associations. Bioinformatics 32, 3207-3209, doi:10.1093/bioinformatics/btw373 (2016).

26 Kamat, M. A. et al. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. Bioinformatics 35, 4851-4853, doi:10.1093/bioinformatics/btz469 (2019).

27 Doherty, A. et al. GWAS identifies 14 loci for device-measured physical activity and sleep duration. Nat Commun 9, 5257, doi:10.1038/s41467-018-07743-4 (2018).

28 Bowden, J., Davey Smith, G., Haycock, P. C. & Burgess, S. Consistent Estimation in Mendelian Randomisation with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol 40, 304-314, doi:10.1002/gepi.21965 (2016).

29 Haycock, P. C. et al. Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomisation studies. Am J Clin Nutr 103, 965-978, doi:10.3945/ajcn.115.118216 (2016).

30 Burgess, S. & Thompson, S. G. Multivariable Mendelian randomisation: the use of pleiotropic genetic variants to estimate causal effects. Am J Epidemiol 181, 251-260, doi:10.1093/aje/kwu283 (2015).

31 Yavorska, O. O. & Burgess, S. MendelianRandomisation: an R package for performing Mendelian randomisation analyses using summarized data. Int J Epidemiol 46, 1734-1739, doi:10.1093/ije/dyx034 (2017).

32 Jankowska, E. A. et al. Effects of intravenous iron therapy in iron-deficient patients with systolic heart failure: a meta-analysis of randomised controlled trials. European journal of heart failure 18, 786-795, doi:10.1002/ejhf.473 (2016).

33 Ngo, K. et al. Erythropoiesis-stimulating agents for anaemia in chronic heart failure patients. Cochrane Database Syst Rev, Cd007613, doi:10.1002/14651858.CD007613.pub2 (2010).

34 Tehrani, F. et al. Erythropoiesis stimulating agents in heart failure patients with anemia: a meta-analysis. Cardiovascular drugs and therapy 23, 511-518, doi:10.1007/s10557-009-6203-6 (2009).

35 Zhang, S., Zhang, F., Du, M., Huang, K. & Wang, C. Efficacy and safety of iron supplementation in patients with heart failure and iron deficiency: a meta-analysis. Br J Nutr 121, 841-848, doi:10.1017/S000711451900014X (2019).

36 Zhou, X., Xu, W., Xu, Y. & Qian, Z. Iron supplementation improves cardiovascular outcomes in patients with heart failure. Am J Med, doi:10.1016/j.amjmed.2019.02.018 (2019).

37 Burden, R. J., Morton, K., Richards, T., Whyte, G. P. & Pedlar, C. R. Is iron treatment beneficial in, iron-deficient but non-anaemic (IDNA) endurance athletes? A systematic review and meta-analysis. Br J Sports Med 49, 1389-1397, doi:10.1136/bjsports-2014-093624 (2015).

38 Pratt, J. J. & Khan, K. S. Non-anaemic iron deficiency - a disease looking for recognition of diagnosis: a systematic review. Eur J Haematol 96, 618-628, doi:10.1111/ejh.12645 (2016).

39 Kansagara, D. et al. Treatment of anemia in patients with heart disease: a systematic review. Ann Intern Med 159, 746-757, doi:10.7326/0003-4819-159-11-201312030-

00007 (2013).

40 Low, M. S., Speedy, J., Styles, C. E., De-Regil, L. M. & Pasricha, S. R. Daily iron supplementation for improving anaemia, iron status and health in menstruating women. Cochrane Database Syst Rev 4, CD009747, doi:10.1002/14651858.CD009747.pub2 (2016).

41 Johnson, D. M., Roberts, J. & Gordon, D. The acute effects of whole blood donation on cardiorespiratory and haematological factors in exercise: A systematic review. PLoS One 14, e0215346, doi:10.1371/journal.pone.0215346 (2019).

42 Newman, B. H. Blood donor complications after whole-blood donation. Curr Opin Hematol 11, 339-345 (2004).

43 Bassett, D. R., Jr. & Howley, E. T. Limiting factors for maximum oxygen uptake and determinants of endurance performance. Med Sci Sports Exerc 32, 70-84, doi:10.1097/00005768-200001000-00012 (2000).

44 Barker, J. et al. Physical activity of UK adults with chronic disease: cross-sectional analysis of accelerometer-measured physical activity in 96 706 UK Biobank participants. Int J Epidemiol, doi:10.1093/ije/dyy294 (2019).

45 Cassidy, S. et al. Accelerometer-derived physical activity in those with cardio-metabolic disease compared to healthy adults: a UK Biobank study of 52,556 participants. Acta Diabetol 55, 975-979, doi:10.1007/s00592-018-1161-8 (2018).

46 Matthews, C. E. et al. Accelerometer-measured dose-response for physical activity, sedentary time, and mortality in US adults. Am J Clin Nutr 104, 1424-1432, doi:10.3945/ajcn.116.135129 (2016).

Supplementary figures, tables and appendices

Appendix 1: Accelerometers

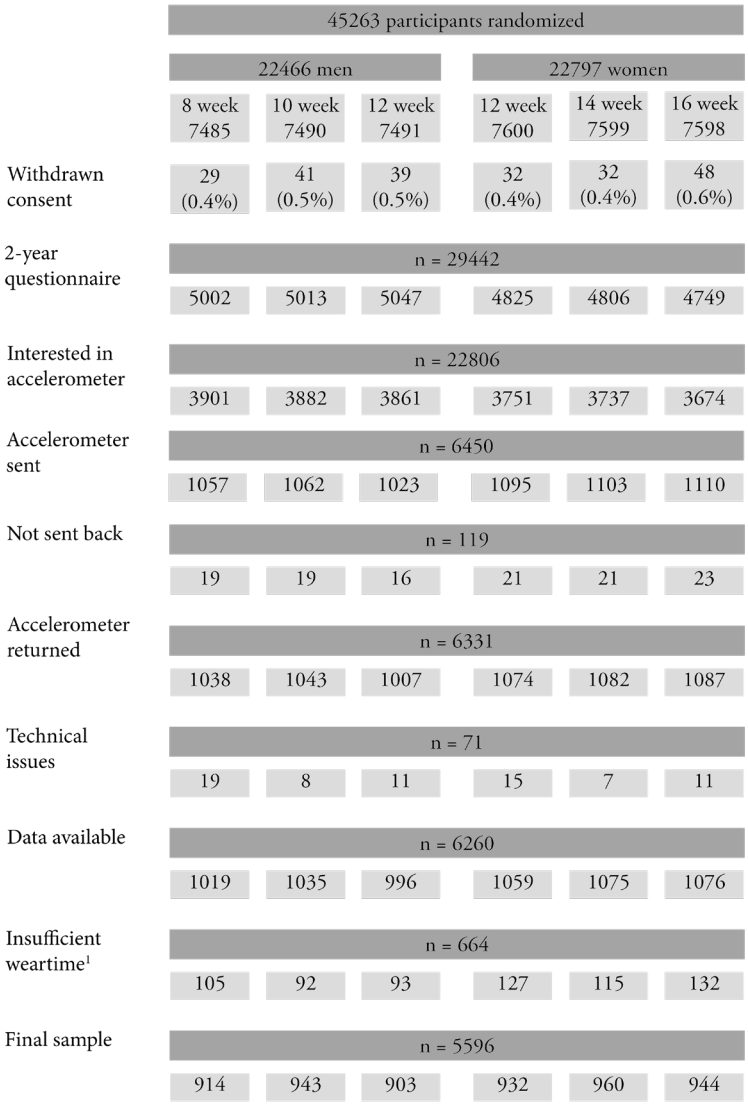
Recruitment and data collection

Participants were recruited when answering the 2-year INTERVAL questionnaire. Those interested and whom provided consent to retrieve their name and address from the NHSBT records in order to send the device were eligible. Accelerometers were sent by mail and set up to start at 00.01AM two days after postal dispatch and stopped recording after 8 days. Participants were requested to start wearing the accelerometer immediately after receiving it in the post and carry on with their normal activities. Supplementary Figure 2 provides an overview of information on donors who participated in this sub-study.

Data processing

Data processing procedures adopted by the UK Biobank were used for data collected from INTERVAL participants. These have been described in detail elsewhere[1]. Briefly, acceleration was recorded at 100 Hz with a dynamic range of ±8*g*. Epochs of 10 seconds for which the standard deviation was <13 milli-gravity (m*g)* in all three axis were used for auto-calibration[1,2]. Calibration coefficients of the previous accelerometer record from the same device but from a different participant were used if errors occurred. To distinguish the physical activity component, the gravitational component and machine noise needs to be addressed. The former was removed by the Euclidian Norm Minus One (ENMO), with negative values rounded to zero[3]. A fourth order Butterworth low pass filter with a cut-off filter of 20Hz was used to remove the machine noise component[4]. Further, non-wear time was defined as stationary episodes for at least 60 minutes where all three axes had a standard deviation of less than 13 mg^{1}. Non-wear time was imputed by using personal mean values for each participant, e.g. data from similar time-of-day but on different days of measurement[1]. Moderate-to-vigorous physical activity and total physical activity were defined respectively, as ENMO values greater than 124 and greater than 29 m*g*, and expressed in minutes per day[5].

Supplementary Figure 1: CONSORT flowchart



[1]A minimum of 72 hours

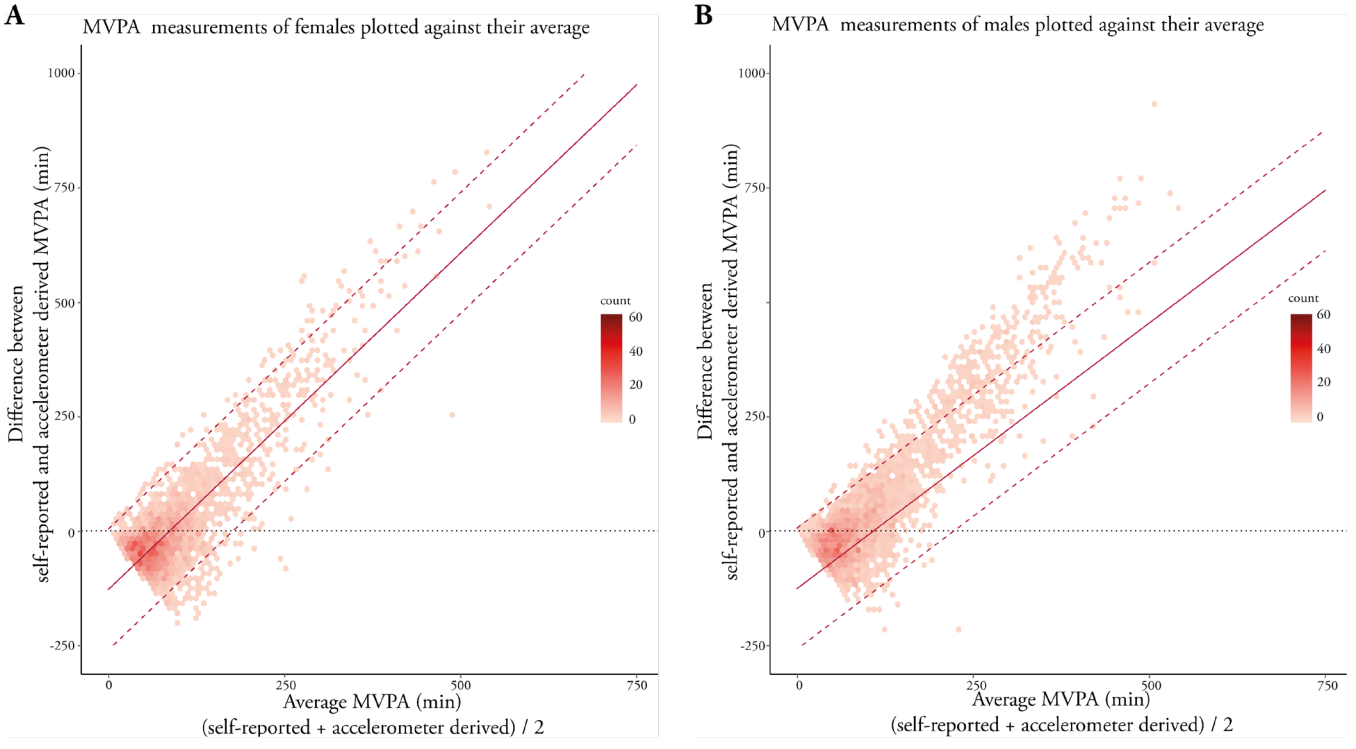
Supplementary Figure 1: CONSORT flowchart: overview of participants including those with valid accelerometer data.

Supplementary table 1: Comparison subsamples.

	Accelerometer	RPAQ	INTERVAL	NHSBT	Accelerometer	RPAQ	INTERVAL	NHSBT
No. of participants and donors	2980	11809	11828	744134	2885	12646	12668	586372
Age, years	46.8 ± 13.4	47 ± 13.5	47.9 ± 13.5	41.0 ± 14.9	49.7 ± 13.1	50.1 ± 13.2	47.9 ± 13.5	44.1 ± 15.0
BMI[1]	25.8 ± 4.8	26 ± 5	26 ± 5	n/a	26.7 ± 4	26.6 ± 4	26.6 ± 4	n/a
New donors	246 (8%)	1008 (7%)	827 (7%)	177,215 (24%)	157 (5%)	717 (5%)	608 (2%)	111,438 (19%)
No. of donations 2 year prior to INTERVAL	3.1 ± 1.6	3.2 ± 1.6	3.2 ± 1.6	1.9 ± 1.8	3.8 ± 1.8	3.9 ± 1.8	3.9 ± 1.8	2.3 ± 2.0
No. of donations attempts during INTERVAL[1]	6.1 ± 1.5	6.1 ± 1.5	6.1 ± 1.5	n/a	8.6 ± 2.2	8.5 ± 2.2	8.5 ± 2.2	n/a
Haemoglobin at 2-year, g/L	13.1 ± 1.1	13.1 ± 1.1	13.1 ± 1.1	n/a	14.4 ± 1.2	14.4 ± 1.2	14.4 ± 1.2	n/a
Haemoglobin at 4-year, g/L	13.0 ± 1.0	13.1 ± 1.0	13.1 ± 1.0	n/a	14.3 ± 1.1	14.3 ± 1.1	14.3 ± 1.1	n/a
Ferritin at 2-year, µg/L	26.0 (14.0 - 44.0)	25.0 (13.0 - 43.0)	25.0 (13.0 - 43.0)	n/a	31.0 (17.0 - 55.0)	31.0 (17.0 - 54.0)	31.0 (17.0 - 54.0)	n/a
Ferritin at 4-year, µg/L	25.0 (14.0 - 44.0)	26.0 (15.0 - 45.0)	26.0 (15.0 - 45.0)	n/a	31.0 (17.0 - 51.0)	31.0 (17.0 - 53.0)	31.0 (17.0 - 53.0)	n/a
Physical activity energy expenditure, kJ/kg per day at 2-year[2]	36.4 (25.2 - 56.6)	36.3 (24.7 - 56.8)	36.3 (24.7 - 56.8)	n/a	45.6 (30.8 - 72.9)	45.0 (29.9 - 73.3)	45.0 (29.9 - 73.3)	n/a
Physical activity energy expenditure, kJ/kg per day at 4year[2]	33.4 (24 - 49.1)	33.3 (23.6 - 50.6)	33.3 23.6 - 50.6)	n/a	42.1 (29.2 - 66.9)	41.0 (28.2 - 64.8)	41.0 (28.2 - 64.8)	n/a

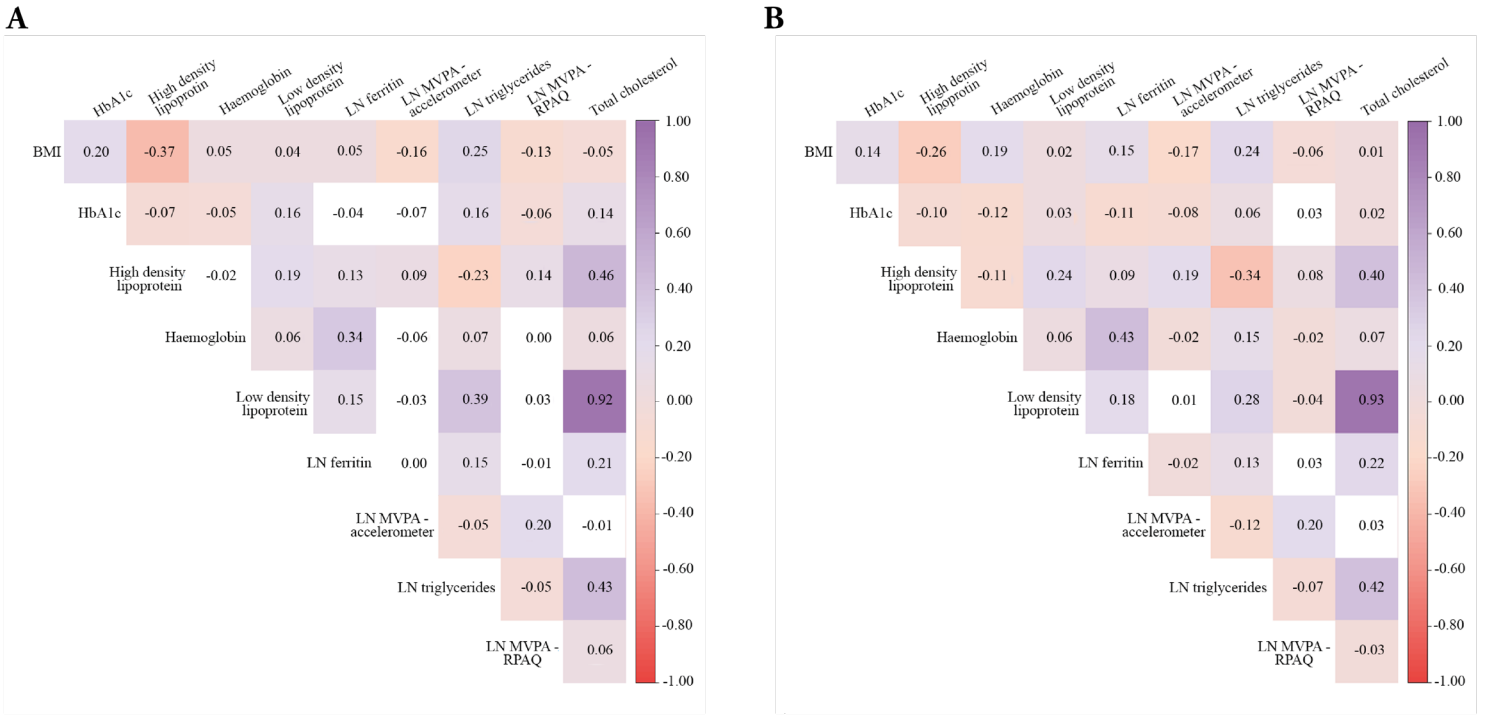
Data are mean (SD), number of participants (%) or median (interquartile range), unless otherwise specified. [1]First two years of INTERVAL. [2]As measured with the RPAQ. RPAQ: Recent physical activity questionnaire. NHSBT: National Health Service Blood and Transplant. BMI: Body mass index.

Supplementary Figure: 2 Bland-Altman plots of moderate-to-vigorous physical activity as self-reported and device-measured by sex



Supplementary Figure 2: Bland and Altman plots for women (A) and men (B) using 2-year data.
Solid red lines represent mean bias and dashed red lines regression based 95% limits of agreement.
For women (A) mean bias: -126.49 + (1.47 * average), for men (B) mean bias: -124.52 + (1.16*average).

Supplementary Figure 3: Pearson correlation heatmap of physical activity variables with biological parameters



Supplementary Figure 2: Pearson correlation plot for women (A) and men (B) using 2-year data.

White squares indicate non-statistically significant results. BMI: Body mass index, LN: natural logarithm, MVPA, moderate-to-vigorous physical activity, RPAQ: recent physical activity questionnaire.

Appendix 2: Pre-specified statistical analysis plan for an INTERVAL trial study on iron stores and physical activity

1. Introduction
This statistical analysis plan describes the *a priori* planned analysis and reporting for the iron stores and physical activity and capacity paper. This paper concerns one of the sub-studies as described in the statistical analysis plan of the principal INTERVAL trial paper.

2. Overall analysis strategy
Analysis will be conducted on an intention-to-treat basis (i.e. donors who did not adhere to their allocated donation interval are analysed per assigned donation interval). Men and women will always be analysed separately. Only participants of whom iron status (haemoglobin and/or ferritin level at 2-year) and physical activity (Recent physical activity questionnaire (RPAQ) and/or valid accelerometer data, i.e. at least 72 hours wear time at 2-year) is available will be included. To minimize the number of statistical tests performed, analyses will assess linear trend across sex-specific randomised groups (i.e. 8 weeks, 10 weeks and 12 weeks for men; 12 weeks, 14 weeks and 16 weeks for women). All effect estimates will be adjusted for donation centre and in case of change over time analysis also for baseline covariates (see section 5). This study will be reported according to CONSORT guidelines.

3. Descriptive analyses of outcomes
Continuous variables will be described as mean and standard deviation or median and 25th percentile – 75th percentile, as appropriate. Categorical variables will be given as number and percentage. Regression coefficients and 95% confidence intervals (95%CI) will be presented. Missing data per variable included in the analyses will be presented.

Baseline characteristics by randomised group and sex including:
Age, past donation history before INTERVAL (no. of donation previous two years), no. of donation (attempts) during INTERVAL, new vs. existing donor status (participants who had and who had not provided a full blood donation before enrolment in INTERVAL).

A CONSORT flow diagram will be provided with information on donors participating at baseline, 2-year and 4-year, withdrawals and missing data. A separate CONSORT diagram will document the flow of participants through the accelerometer sub-study.

4. Primary outcomes and supporting analyses of the primary outcomes
Primary outcomes:

- Self-reported time spent in moderate-to-vigorous physical activity (MVPA), physical activity energy expenditure, sedentary behaviour at 2 years, 4 years and change between 2- and 4-year.
- Objectively measured time spent in MVPA and total physical activity at 2-year.

Supporting analyses:

- Correlation between biological parameters (total cholesterol, HDL cholesterol, LDL cholesterol, triglycerides, haemoglobin, ferritin and HbA1c) and physical activity.
- Dynamic change among self-reported MVPA with haemoglobin and ferritin separately (between 2- and 4-year)
-Agreement between MVPA assessed by accelerometers and the RPAQ measurements.
-Attrition at 4-year.

Post-hoc analysis:

- Objectively measured average acceleration at 2-year.

5. Analysis of primary and supporting analyses of the primary outcomes
For the primary analyses multilevel linear regression with random intercepts for centre will be conducted, if the models do not converge, linear regression with centre as dummy variable will be conducted instead.

Subgroup analyses will be carried out if assessment of the relevant statistical interaction with baseline characteristics by linear regression is statistically significant (see controlling for false positive rates). Continuous baseline characteristics will be analysed as linear terms in the regression models, but results presented in groups as specified below. The subgroups to be compared will be:

- Haemoglobin: linear in regression models, but presented as below / above the sex-specific median at baseline.
- Ferritin levels (above/below 15µg/L of baseline measurement)
- Use of iron supplements (Between baseline and end of main trial)
- Age, linear in regression models, but presented as <50 vs. 50+ years.
- 'Active/not active' based on 2-year data, linear in regression models but presented as above/below sex- specific median physical activity energy expenditure (only for change analyses)

Analyses of the primary outcomes at 4-year will be adjusted for baseline covariates. These will include the baseline (2-year) version of the outcome.

Supporting analyses of the primary outcomes:

- Pearson correlation coefficients will be presented in a table.
- Bivariate dual change score models (BDCSM) will be used to explore dynamics among self-reported MVPA with haemoglobin and ferritin separately (between 2- and 4-year). Four models were estimated: 1) a baseline model (neither haemoglobin nor MVPA influences the other); 2) changes in haemoglobin associated with changes in MVPA model; 3) changes in MVPA associated with changes in haemoglobin model; and 4) dynamic change model in which both changes in MVPA and haemoglobin associate with changes in the other. See Appendix 3 and Supplementary Figure 4 for a brief description of the BDCSM methodology, the main parameters of interest and fit statistics. The BDCSM models are presented as vector fields[6].
- Bland and Altman plots will show the level of agreement between the two measurements of MVPA.
- Multilevel logistic regression with random intercepts for centre will be used to assess attrition at 4-year.

Sensitivity analysis

To investigate robustness of our findings, sensitivity analyses will be conducted with different intensity cut-point for MVPA (i.e. 100 m*g* and 150 m*g*[5,7]).

6. Controlling for false positives

Physical activity related outcomes were secondary outcomes of the main INTERVAL trial paper. In total this paper has 11 main outcomes (5 at 24 months, 3 concerning change between 24 and 48 months and three at 48 months, see section 4). To take into account multiple testing Benjamini and Hochberg p-values were calculated with a false-discovery rate of 10%. We acknowledge that this might lead to some false positives, however, consider this preferential to overly conservative approaches that may result in important false negatives. In particular the analyses with objectively measured physical activity may be underpowered as this was not one of the main outcomes of the INTERVAL trial. Significant subgroup interactions would indicate that further investigation could be considered because some donors might be more prone to change their physical activity. The standard p-value (if <0.05) will also be reported so the reader may make their own judgements of the evidence.

7. Missing data

As continued participation in the INTERVAL trial was optional after the initial 2 years, most missing data will be at the online questionnaire at 4 years. Followed by the non-completion of the online questionnaires at 2 years. Attrition bias will be investigated with logistic mixed model analyses with random intercepts for centre at 48 months. Characteristics that will be assessed include trial arm, age, number of donations 2 years prior INTERVAL donor status, and haemoglobin and ferritin levels at start of the INTERVAL trial.

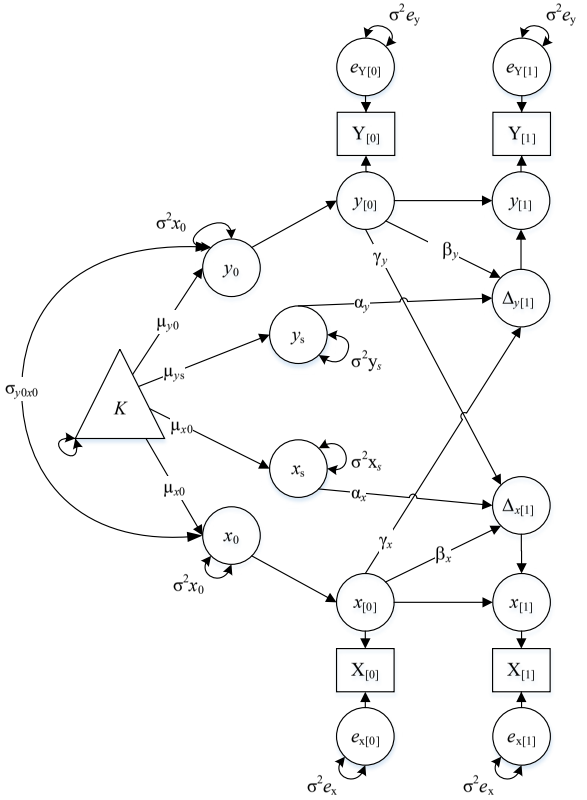
Appendix 3: Description Bivariate Dual Change Score Models

Bivariate dual change score models (BDCSM) are dynamic structural equation models (SEM)[8] and in our case allow us to model, and compare, four hypothetical scenario of the moderate-to-vigorous physical activity (MVPA) and iron biomarker (both haemoglobin and ferritin, for ease of reading this appendix only refers to haemoglobin) relationship over time. A brief overview of the models and the main components are outlined below. A more elaborate discussion inclusive of mathematical formulae can be found elsewhere[8,9]. The basic model estimates the situation in which neither MVPA nor haemoglobin influence change in the other and was used as reference to compare the more complex models (ie, baseline model). The second model estimates the scenario in which haemoglobin affects MVPA but not vice-versa. This is followed by a model in which MVPA can influence haemoglobin but not the alternative. In the final model both MVPA and haemoglobin levels are able to affect change in the other between time points, assuming that they are bi-directionally related. Results for crude as well as models adjusted for age, sex and intervention arm are presented in Supplementary table 8-9.

In BDCSM models, change is assumed to consist of three elements: 1) an additive component of change of y_s and x_s which is the sum of the latent changes over time, α; 2) an auto-proportional component reflecting a quantity proportional to the previous value of the same variable, β; 3) an amount proportional to the previous value of the alternative variable (if included in the model), also known as the coupling parameter (γ). Constraining the coupling parameter of MVPA to haemoglobin to zero while estimating the coupling parameter of haemoglobin to MVPA models the hypothesis that MVPA does not affect haemoglobin at the consecutive time point (at 48 months) but haemoglobin influences MVPA (ie, model 2 outlined above). A graphical representation of the BDCSM models is shown in Supplementary Figure 4.

Fit statistics including the Root Mean Square Error of Approximation (RMSEA), Akaike Information Criterion (AIC), Sample-size adjusted Bayesian (SSA BIC), Comparative Fit Index (CFI) and the Tucker-Lewis Index (TLI) are presented for all models, the model parameters of the adjusted model are only shown for the model with the best fit (based on log likelihood and x^2 test). All models were estimated using R version 3.6.1 lavaan, MLR. Vector fields were plotted using MATLAB version R2019b Update 1 (9.7.0.1216025) 64-bit (glnxa64).

Supplementary Figure 4: Graphical representation of Bivariate Dual Change Score Models



Parameter	Definition
Y_t, X_t	Observed values for Y(RPAQ) and X(Haemoglobin)
y_t, x_t	Latent scores for variables Y and X
$\Delta y_1, \Delta x_1$	Latent changes in y and x
y_0, x_0	Latent intercept for variables y and x
y_s, x_s	Latent slope for variables y and x
$\alpha y_t, \alpha x$	Additive component of change for variables y and x
β_x, β_y	Autoproportional change parameters for variables y and x
γ_y, γ_x	Coupling parameters for variables y and x
$\mu y_0, \mu x_0$	Mean of initial conditions for variables y and x
$\mu y_s, \mu x_s$	Mean of additive scores for variables y and x
$\sigma^2 y_0, \sigma^2 x_0$	Variance of inital conditions of y and x
$\sigma^2 y_s, \sigma^2 x_s$	Variance of additive scores of y and x
$\sigma^2 e_Y, \sigma^2 e_X$	Residual variance of initial conditions of y and x
σ_{x0y0}	Covariance of intercepts for variables y and x
K	Constant to estimate means and intercepts (set to 1)

Supplementary Figure 3: Bivariate Dual Change Score Model representation.

Appendix 4: PROSPERO protocol

The protocol of this umbrella review was registered in PROSPERO under registration number CRD42019123885.

Search strategy

We reviewed systematic reviews with meta-analysis that examined associations between iron stores and physical capacity. In March 2019 we searched 8 electronic databases (MEDLINE via PUBMED interface, EMBASE, Cochrane Library, Scopus, Web of Science, CINAHL plus, Epistemonikos and Transfusion Evidence Library) to identify eligible studies. Search terms related to 'iron stores' (e.g. iron, haemoglobin, ferritin) were used in combination with search terms including 'physical capacity' (e.g. VO_2 max, physical endurance) and 'systematic review' (see Search strategy per database below). Forward and backward screening was conducted on all papers included at the full-text screening stage to identify additional potentially eligible articles.

Article selection

Peer-reviewed systematic reviews with meta-analysis that included randomised controlled trials or prospective observational studies in humans with iron stores or supplementation as the exposure and any measure of physical capacity as the outcome that were published after 1990 in the English language were considered eligible for inclusion.

No restrictions with regards to nationality, race or sex were applied. Systematic reviews including adolescents and/or children were excluded, unless adult (≥18 years of age) subgroups results were reported separately. Older versions of systematic reviews that contained exactly the same studies as more recent ones were also excluded, as were conference abstracts.

After duplicates were removed, all identified publications were screened by two reviewers independently (RdG and SB) based initially on title and abstract and later on full report. Any disagreements were discussed by RdG and SB, and in cases of non-agreement, a third author, KvdH, decided.

Data extraction and quality assessment

From each systematic review (1) year of publication; (2) study design; (3) search strategy and results (e.g. number of data sources, date range on included studies); (4) number of included studies; (5) intervention/exposure; (6) outcomes reported which are relevant to this umbrella review; (7) total number of participants; (8) eligibility criteria; (9) summary effect estimate (RR, OR, SMD), 95% confidence intervals, standard error, sample size; (10) fixed or random analysis; (11) heterogeneity index (I^2 or Q); (12)

data necessary for risk-of-bias assessment, and (13) conflict of interest and funding data were extracted by one author (RdG) and agreement was checked in a subsample (SB). In the event of missing information, the authors of the systematic review in question were contacted up to three times to seek clarification. All included studies were independently assessed by two authors (RdG and SB) for general methodological quality of the systematic review using the Assessment of Multiple Systematic Reviews 2 (AMSTAR). Initially the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) methodology planned to be conducted, however upon reflection of the aim of the current systematic review we decided not to construct a GRADE score. This review aims to provide an overview of the associations between iron status and physical capacity outcomes and does not seek to provide recommendations on treatments in patient populations.

Data synthesis and analysis

A narrative synthesis was structured around the type of outcome and the main study population (e.g. blood donation related study population or studies that involved participants from IDNA, general or diseased populations). Each meta-analysis was re-analysed with a random effects model, including estimations of heterogeneity (I^2). Publication bias was assessed in outcomes with more than 10 studies, using contour enhanced funnel plots.

Search strategy per database

Database	Results
PubMed	1207
Embase	665
SportDiscus	124
Cinahl	124
Scopus	542
Web of Science	180
Cochrane Library	15
Epistemonikos	35
Transfusion Evidence Library	17
Total	2909
After duplicate removal	2078

Search	PubMed Query 5-3-2019	Items found
#5	#1 AND #2 AND #3 Filters: Publication date from 1990/01/01	1207
#4	#1 AND #2 AND #3	1529

Search	PubMed Query 5-3-2019	Items found
#3	(("Review Literature as Topic"[Mesh] OR "Review"[Publication Type] OR "Meta-Analysis as Topic"[Mesh] OR review[tiab] OR meta-analys*[tiab] OR "Meta-Analysis "[Publication Type] OR systematic[sb]) NOT ("Letter"[Publication Type] OR "Editorial"[Publication Type] OR "Comment"[Publication Type]))	3053107
#2	"Oxygen Consumption"[Mesh] OR "Physical Fitness"[Mesh] OR "Athletic Performance"[Mesh] OR "Physical Endurance"[Mesh] OR "Physical Exertion"[Mesh] OR "Ergometry"[Mesh] OR "Exercise test"[MeSH Terms] OR "Muscle Fatigue"[Mesh] OR oxygen consumption*[tiab] OR oxygen demand*[tiab] OR oxygen requirement*[tiab] OR oxygen uptake[tiab] OR anaerobic threshold*[tiab] OR metabolic equivalent*[tiab] OR physical endurance*[tiab] OR energy expenditure*[tiab] OR physical fit*[tiab] OR physically fit*[tiab] OR physical condition*[tiab] OR cardiorespiratory fit*[tiab] OR physical activit*[tiab] OR physical performanc*[tiab] OR athletic performanc*[tiab] OR aerobic capacit*[tiab] OR exercise capacit*[tiab] OR exercise performanc*[tiab] OR exercise endurance*[tiab] OR muscle fatigue[tiab] OR vo2 max[tiab] OR vo2max[tiab] OR physical exertion*[tiab] OR ergometr*[tiab] OR exercise test*[tiab] OR fitness test*[tiab] OR step test*[tiab] OR walk test*[tiab] OR walking test*[tiab] OR walking distance*[tiab] OR walk distance*[tiab] OR sport performance*[tiab] OR muscular fatigue[tiab]	404440
#1	"Iron"[Mesh] OR "Iron Compounds"[Mesh] OR "ferritins"[MeSH Terms] OR "Hemoglobins"[Mesh] OR "Transferrin"[Mesh] OR "hepcidins"[MeSH Terms] OR "Anemia"[Mesh] OR iron[tiab] OR ferrous compound*[tiab] OR ferritin*[tiab] OR ferric compound*[tiab] OR hemoglobin*[tiab] OR haemoglobin*[tiab] OR hb[tiab] OR transferrin*[tiab] OR hepcidin*[tiab] OR anemia*[tiab] OR anaemia*[tiab] OR non-iron[tiab] OR noniron[tiab]	629026

Search	Embase.com Query 5-3-2019	Items found
#5	#1 AND #2 AND #3 AND [1990-2019]/py	665
#4	#1 AND #2 AND #3	665
#3	(('systematic review'/exp OR 'meta analysis'/exp OR (systematic* NEAR/3 review):ab,ti,kw OR meta-analys*:ab,ti,kw) NOT ('conference abstract'/it OR 'conference review'/it OR 'editorial'/it OR 'erratum'/it OR 'letter'/it OR 'note'/it OR 'short survey'/it))	277383
#2	'oxygen consumption'/exp OR 'fitness'/exp OR 'athletic performance'/exp OR 'endurance'/exp OR 'exercise'/exp OR 'ergometry'/exp OR 'exercise test'/exp OR 'muscle fatigue'/exp OR 'oxygen consumption*':ab,ti,kw OR 'oxygen demand*':ab,ti,kw OR 'oxygen requirement*':ab,ti,kw OR 'oxygen uptake':ab,ti,kw OR 'anaerobic threshold*':ab,ti,kw OR 'metabolic equivalent*':ab,ti,kw OR 'physical endurance*':ab,ti,kw OR 'energy expenditure*':ab,ti,kw OR 'physical* fit*':ab,ti,kw OR 'physical condition*':ab,ti,kw OR 'cardiorespiratory fit*':ab,ti,kw OR 'physical activit*':ab,ti,kw OR (physical NEAR/3 performanc*):ab,ti,kw OR (athletic NEAR/3 performanc*):ab,ti,kw OR (sport NEAR/3 performance*):ab,ti,kw OR (exercise NEAR/3 performanc*):ab,ti,kw OR 'aerobic capacit*':ab,ti,kw OR 'exercise capacit*':ab,ti,kw OR 'exercise endurance*':ab,ti,kw OR 'muscle fatigue':ab,ti,kw OR 'muscular fatigue':ab,ti,kw OR 'vo2-max':ab,ti,kw OR vo2max:ab,ti,kw OR 'physical exertion*':ab,ti,kw OR ergometr*:ab,ti,kw OR 'exercise test*':ab,ti,kw OR 'fitness test*':ab,ti,kw OR 'step test*':ab,ti,kw OR 'walk* test*':ab,ti,kw OR 'walk* distance*':ab,ti,kw	965464
#1	'iron'/exp OR 'iron derivative'/exp OR 'ferritin'/exp OR 'hemoglobin'/exp OR 'transferrin'/exp OR 'hepcidin'/exp OR 'anemia'/exp OR iron:ab,ti,kw OR 'ferrous compound*':ab,ti,kw OR ferritin*:ab,ti,kw OR 'ferric compound*':ab,ti,kw OR hemoglobin*:ab,ti,kw OR haemoglobin*:ab,ti,kw OR hb:ab,ti,kw OR transferrin*:ab,ti,kw OR hepcidin*:ab,ti,kw OR anemia*:ab,ti,kw OR anaemia*:ab,ti,kw OR non-iron:ab,ti,kw OR noniron:ab,ti,kw	967604

Search	SportDiscus (Ebsco) Query 5-3-2019	Items found
S5	S1 AND S2 AND S3 AND Limiters - Published Date: 19900101-20191231	124
S4	S1 AND S2 AND S3	155
S3	TI (review OR meta-analys*) OR AB (review OR meta-analys*) OR KW (review OR meta-analys*)	108175
S2	DE ("OXYGEN consumption" OR "EXCESS post-exercise oxygen consumption" OR "OXYGEN consumption" OR "EXCESS post-exercise oxygen consumption" OR "ANAEROBIC threshold" OR "CALORIC expenditure" OR "PHYSICAL fitness" OR "EXERCISE" OR "METABOLIC equivalent" OR "ERGOMETRY" OR "EXERCISE tests" OR "TREADMILL exercise tests" OR DE "AEROBIC capacity" OR "ANAEROBIC capacity") OR TI ("oxygen consumption*" OR "oxygen demand*" OR "oxygen requirement*" OR "oxygen uptake" OR "anaerobic threshold*" OR "metabolic equivalent*" OR "physical endurance*" OR "energy expenditure*" OR "physical* fit*" OR "physical condition*" OR "cardiorespiratory fit*" OR "physical activit*" OR (physical N3 performanc*) OR (athletic N3 performanc*) OR (sport N3 performance*) OR (exercise N3 performanc*) OR "aerobic capacit*" OR "exercise capacit*" OR "exercise endurance*" OR "muscle fatigue" OR "muscular fatigue" OR "vo2-max" OR vo2max OR "physical exertion*" OR ergometr* OR "exercise test*" OR "fitness test*" OR "step test*" OR "walk* test*" OR "walk* distance*") OR AB ("oxygen consumption*" OR "oxygen demand*" OR "oxygen requirement*" OR "oxygen uptake" OR "anaerobic threshold*" OR "metabolic equivalent*" OR "physical endurance*" OR "energy expenditure*" OR "physical* fit*" OR "physical condition*" OR "cardiorespiratory fit*" OR "physical activit*" OR (physical N3 performanc*) OR (athletic N3 performanc*) OR (sport N3 performance*) OR (exercise N3 performanc*) OR "aerobic capacit*" OR "exercise capacit*" OR "exercise endurance*" OR "muscle fatigue" OR "muscular fatigue" OR "vo2-max" OR vo2max OR "physical exertion*" OR ergometr* OR "exercise test*" OR "fitness test*" OR "step test*" OR "walk* test*" OR "walk* distance*") OR KW ("oxygen consumption*" OR "oxygen demand*" OR "oxygen requirement*" OR "oxygen uptake" OR "anaerobic threshold*" OR "metabolic equivalent*" OR "physical endurance*" OR "energy expenditure*" OR "physical* fit*" OR "physical condition*" OR "cardiorespiratory fit*" OR "physical activit*" OR (physical N3 performanc*) OR (athletic N3 performanc*) OR (sport N3 performance*) OR (exercise N3 performanc*) OR "aerobic capacit*" OR "exercise capacit*" OR "exercise endurance*" OR "muscle fatigue" OR "muscular fatigue" OR "vo2-max" OR vo2max OR "physical exertion*" OR ergometr* OR "exercise test*" OR "fitness test*" OR "step test*" OR "walk* test*" OR "walk* distance*")	214217
S1	DE ("IRON in the body" OR "FERRITIN" OR "TRANSFERRIN" OR "IRON deficiency diseases" OR "IRON deficiency anemia" OR "PHYSIOLOGICAL effects of iron" OR "HEMOGLOBINS" OR "HEMOGLOBIN" OR "CARBOXYHEMOGLOBIN" OR "OXYHEMOGLOBIN" OR "ANEMIA") OR TI (iron OR "ferrous compound*" OR ferritin* OR "ferric compound*" OR hemoglobin* OR haemoglobin* OR hb OR transferrin* OR hepcidin* OR anemia* OR anaemia* OR non-iron OR noniron) OR AB (iron OR "ferrous compound*" OR ferritin* OR "ferric compound*" OR hemoglobin* OR haemoglobin* OR hb OR transferrin* OR hepcidin* OR anemia* OR anaemia* OR non-iron OR noniron) OR KW (iron OR "ferrous compound*" OR ferritin* OR "ferric compound*" OR hemoglobin* OR haemoglobin* OR hb OR transferrin* OR hepcidin* OR anemia* OR anaemia* OR non-iron OR noniron)	12707

Search	Cinahl (Ebsco) Query 5-3-2019	Items found
S4	S1 AND S2 AND S3	124
S3	MH ("Systematic Review" OR "Meta Analysis") OR TI ((systematic* N3 review) OR meta-analys*)) OR AB ((systematic* N3 review) OR meta-analys*))	134619

Search	Cinahl (Ebsco) Query 5-3-2019	Items found
S2	MH ("Oxygen Consumption+" OR "Physical Fitness+" OR "Athletic Performance" OR "Physical Endurance+" OR "Exercise+" OR "Ergometry" OR "Exercise Test+" OR "Muscle Fatigue") OR TI ("oxygen consumption*" OR "oxygen demand*" OR "oxygen requirement*" OR "oxygen uptake" OR "anaerobic threshold*" OR "metabolic equivalent*" OR "physical endurance*" OR "energy expenditure*" OR "physical* fit*" OR "physical condition*" OR "cardiorespiratory fit*" OR "physical activit*" OR (physical N3 performanc*) OR (athletic N3 performanc*) OR (sport N3 performance*) OR (exercise N3 performanc*) OR "aerobic capacit*" OR "exercise capacit*" OR "exercise endurance*" OR "muscle fatigue" OR "muscular fatigue" OR "vo2-max" OR vo2max OR "physical exertion*" OR ergometr* OR "exercise test*" OR "fitness test*" OR "step test*" OR "walk* test*" OR "walk* distance*") OR AB ("oxygen consumption*" OR "oxygen demand*" OR "oxygen requirement*" OR "oxygen uptake" OR "anaerobic threshold*" OR "metabolic equivalent*" OR "physical endurance*" OR "energy expenditure*" OR "physical* fit*" OR "physical condition*" OR "cardiorespiratory fit*" OR "physical activit*" OR (physical N3 performanc*) OR (athletic N3 performanc*) OR (sport N3 performance*) OR (exercise N3 performanc*) OR "aerobic capacit*" OR "exercise capacit*" OR "exercise endurance*" OR "muscle fatigue" OR "muscular fatigue" OR "vo2-max" OR vo2max OR "physical exertion*" OR ergometr* OR "exercise test*" OR "fitness test*" OR "step test*" OR "walk* test*" OR "walk* distance*") OR SU ("oxygen consumption*" OR "oxygen demand*" OR "oxygen requirement*" OR "oxygen uptake" OR "anaerobic threshold*" OR "metabolic equivalent*" OR "physical endurance*" OR "energy expenditure*" OR "physical* fit*" OR "physical condition*" OR "cardiorespiratory fit*" OR "physical activit*" OR (physical N3 performanc*) OR (athletic N3 performanc*) OR (sport N3 performance*) OR (exercise N3 performanc*) OR "aerobic capacit*" OR "exercise capacit*" OR "exercise endurance*" OR "muscle fatigue" OR "muscular fatigue" OR "vo2-max" OR vo2max OR "physical exertion*" OR ergometr* OR "exercise test*" OR "fitness test*" OR "step test*" OR "walk* test*" OR "walk* distance*")	200596
S1	MH ("Iron" OR "Iron Compounds+" OR "Ferritin" OR "Transferrin" OR "Hemoglobins" OR "Anemia+") OR TI (iron OR "ferrous compound*" OR ferritin* OR "ferric compound*" OR hemoglobin* OR haemoglobin* OR hb OR transferrin* OR hepcidin* OR anemia* OR anaemia* OR non-iron OR noniron) OR AB (iron OR "ferrous compound*" OR ferritin* OR "ferric compound*" OR hemoglobin* OR haemoglobin* OR hb OR transferrin* OR hepcidin* OR anemia* OR anaemia* OR non-iron OR noniron) OR SU (iron OR "ferrous compound*" OR ferritin* OR "ferric compound*" OR hemoglobin* OR haemoglobin* OR hb OR transferrin* OR hepcidin* OR anemia* OR anaemia* OR non-iron OR noniron)	74969

Search	Scopus Query 5-3-2019	Items found
#5	#4 AND (PUBYEAR > 1990)	542
#4	#1 AND #2 AND #3	544
#3	TITLE-ABS-KEY ((systematic* W/3 review) OR meta-analys*)	374751
#2	TITLE-ABS-KEY ("oxygen consumption*" OR "oxygen demand*" OR "oxygen requirement*" OR "oxygen uptake" OR "anaerobic threshold*" OR "metabolic equivalent*" OR "physical endurance*" OR "energy expenditure*" OR "physical* fit*" OR "physical condition*" OR "cardiorespiratory fit*" OR "physical activit*" OR (physical W/3 performanc*) OR (athletic W/3 performanc*) OR (sport W/3 performance*) OR (exercise W/3 performanc*) OR "aerobic capacit*" OR "exercise capacit*" OR "exercise endurance*" OR "muscle fatigue" OR "muscular fatigue" OR "vo2-max" OR vo2max OR "physical exertion*" OR ergometr* OR "exercise test*" OR "fitness test*" OR "step test*" OR "walk* test*" OR "walk* distance*")	621909
#1	TITLE-ABS-KEY (iron OR "ferrous compound*" OR ferritin* OR "ferric compound*" OR hemoglobin* OR haemoglobin* OR hb OR transferrin* OR hepcidin* OR anemia* OR anaemia* OR non-iron OR noniron)	1449476

Search	Web of Science Core Collection Query 5-3-2019	Items found
#4	#1 AND #2 AND #3	180
#3	TS=((systematic* NEAR/3 review) OR meta-analys*)	264630
#2	TS= ("oxygen consumption*" OR "oxygen demand*" OR "oxygen requirement*" OR "oxygen uptake" OR "anaerobic threshold*" OR "metabolic equivalent*" OR "physical endurance*" OR "energy expenditure*" OR "physical* fit*" OR "physical condition*" OR "cardiorespiratory fit*" OR "physical activit*" OR (physical NEAR/3 performanc*) OR (athletic NEAR/3 performanc*) OR (sport NEAR/3 performance*) OR (exercise NEAR/3 performanc*) OR "aerobic capacit*" OR "exercise capacit*" OR "exercise endurance*" OR "muscle fatigue" OR "muscular fatigue" OR "vo2-max" OR vo2max OR "physical exertion*" OR ergometr* OR "exercise test*" OR "fitness test*" OR "step test*" OR "walk* test*" OR "walk* distance*")	363289
#1	TS= (iron OR "ferrous compound*" OR ferritin* OR "ferric compound*" OR hemoglobin* OR haemoglobin* OR "hb" OR transferrin* OR hepcidin* OR anemia* OR anaemia* OR "non-iron" OR "noniron")	844570

Search	Cochrane Library (Wiley) Query 5-3-2019	Items found
#4	Cochrane Reviews	15
#3	#1 AND #2	2354
#2	(oxygen NEXT consumption* OR oxygen NEXT demand* OR oxygen NEXT requirement* OR oxygen NEXT uptake OR anaerobic NEXT threshold* OR metabolic NEXT equivalent* OR physical NEXT endurance* OR energy NEXT expenditure* OR physical* NEXT fit* OR physical NEXT condition* OR cardiorespiratory NEXT fit* OR physical NEXT activit* OR (physical NEAR/3 performanc*) OR (athletic NEAR/3 performanc*) OR (sport NEAR/3 performance*) OR (exercise NEAR/3 performanc*) OR aerobic NEXT capacit* OR exercise NEXT capacit* OR exercise NEXT endurance* OR muscle NEXT fatigue OR muscular NEXT fatigue OR vo2 NEXT max OR vo2max OR physical NEXT exertion* OR ergometr* OR exercise NEXT test* OR fitness NEXT test* OR step NEXT test* OR walk* NEXT test* OR walk* NEXT distance*):ti,ab,kw (Word variations have been searched)	57731
#1	(iron OR ferrous NEXT compound* OR ferritin* OR ferric NEXT compound* OR hemoglobin* OR haemoglobin* OR hb OR transferrin* OR hepcidin* OR anemia* OR anaemia* OR non-iron OR noniron):ti,ab,kw (Word variations have been searched)	44931

Search	Transfusion Evidence Library Query 5-3-2019	Items found
#4	#3 Filter study design: Systematic review	17
#3	#1 AND #2	

Search	Transfusion Evidence Library Query 5-3-2019	Items found
#2	title: ("oxygen consumption" OR fitness OR "athletic performance" OR endurance OR exercise OR ergometry OR "exercise test" OR "muscle fatigue" OR "oxygen consumption*" OR "oxygen demand*" OR "oxygen requirement*" OR "oxygen uptake" OR "anaerobic threshold*" OR "metabolic equivalent*" OR "physical endurance*" OR "energy expenditure*" OR "physical* fit*" OR "physical condition*" OR "cardiorespiratory fit*" OR "physical activit*" OR "physical performanc*" OR "athletic performanc*" OR "sport performance*" OR "exercise performanc*" OR "aerobic capacit*" OR "exercise capacit*" OR "exercise endurance*" OR "muscle fatigue" OR "muscular fatigue" OR "vo2-max" OR vo2max OR "physical exertion*" OR ergometr* OR "exercise test*" OR "fitness test*" OR "step test*" OR "walk* test*" OR "walk* distance*") OR article_abstract: ("oxygen consumption" OR fitness OR "athletic performance" OR endurance OR exercise OR ergometry OR "exercise test" OR "muscle fatigue" OR "oxygen consumption*" OR "oxygen demand*" OR "oxygen requirement*" OR "oxygen uptake" OR "anaerobic threshold*" OR "metabolic equivalent*" OR "physical endurance*" OR "energy expenditure*" OR "physical* fit*" OR "physical condition*" OR "cardiorespiratory fit*" OR "physical activit*" OR "physical performanc*" OR "athletic performanc*" OR "sport performance*" OR "exercise performanc*" OR "aerobic capacit*" OR "exercise capacit*" OR "exercise endurance*" OR "muscle fatigue" OR "muscular fatigue" OR "vo2-max" OR vo2max OR "physical exertion*" OR ergometr* OR "exercise test*" OR "fitness test*" OR "step test*" OR "walk* test*" OR "walk* distance*") OR keywords: ("oxygen consumption" OR fitness OR "athletic performance" OR endurance OR exercise OR ergometry OR "exercise test" OR "muscle fatigue" OR "oxygen consumption*" OR "oxygen demand*" OR "oxygen requirement*" OR "oxygen uptake" OR "anaerobic threshold*" OR "metabolic equivalent*" OR "physical endurance*" OR "energy expenditure*" OR "physical* fit*" OR "physical condition*" OR "cardiorespiratory fit*" OR "physical activit*" OR "physical performanc*" OR "athletic performanc*" OR "sport performance*" OR "exercise performanc*" OR "aerobic capacit*" OR "exercise capacit*" OR "exercise endurance*" OR "muscle fatigue" OR "muscular fatigue" OR "vo2-max" OR vo2max OR "physical exertion*" OR ergometr* OR "exercise test*" OR "fitness test*" OR "step test*" OR "walk* test*" OR "walk* distance*")	
#1	title: (iron OR "iron derivative" OR ferritin OR hemoglobin OR transferrin OR hepcidin OR anemia OR "ferrous compound*" OR ferritin* OR "ferric compound*" OR hemoglobin* OR haemoglobin* OR hb OR transferrin* OR hepcidin* OR anemia* OR anaemia* OR "non-iron" OR noniron) OR article_abstract: (iron OR "iron derivative" OR ferritin OR hemoglobin OR transferrin OR hepcidin OR anemia OR "ferrous compound*" OR ferritin* OR "ferric compound*" OR hemoglobin* OR haemoglobin* OR hb OR transferrin* OR hepcidin* OR anemia* OR anaemia* OR "non-iron" OR noniron) OR keywords: (iron OR "iron derivative" OR ferritin OR hemoglobin OR transferrin OR hepcidin OR anemia OR "ferrous compound*" OR ferritin* OR "ferric compound*" OR hemoglobin* OR haemoglobin* OR hb OR transferrin* OR hepcidin* OR anemia* OR anaemia* OR "non-iron" OR noniron)	

Supplementary Figure 5: CONSORT flowchart for INTERVAL iron and physical activity study

	45263 participants randomized					
	22466 men			22797 women		
	8 week 7485	10 week 7490	12 week 7491	12 week 7600	14 week 7599	16 week 7598
Withdrawn consent	29 (0.4%)	41 (0.5%)	39 (0.5%)	32 (0.4%)	32 (0.4%)	48 (0.6%)
Met inclusion criteria	4206 (56.4%)	4222 (56.7%)	4240 (56.9%)	3965 (52.4%)	3969 (52.5%)	3894 (51.6%)
RPAQ 2-year	4193 (99.7%)	4205 (99.6%)	4228 (99.7%)	3955 (99.6%)	3951 (99.6%)	3885 (99.8%)
RPAQ 4-year	2558 (60.8%)	2696 (63.9%)	2602 (61.4%)	2367 (59.7%)	2335 (58.8%)	2298 (59.0%)
Accelerometer study[1]	914 (21.7%)	943 (22.3%)	944 (24.2%)	932 (23.5%)	960 (24.2%)	944 (24.2%)
MVPA	914 (21.7%)	943 (22.3%)	944 (24.2%)	932 (23.5%)	960 (24.2%)	944 (24.2%)
Total physical activity	914 (21.7%)	943 (22.3%)	944 (24.2%)	932 (23.5%)	960 (24.2%)	944 (24.2%)

[1]Valid accelerometer data, see Supplementary Figure 2 for the flow chart of the accelerometer study. RPAQ: recent physical activity questionnaire. MVPA: moderate-to-vigorous physical activity.

Supplementary Figure 5: CONSORT flowchart: recruitment, participation and completeness of main outcomes.

Supplementary Table 2: Logistic mixed regression of attrition at 4-year.

	Female			Male		
	OR	95% CI	p-value	OR	95% CI	p-value
Trial arm						
16 (f) / 12 (m) weeks		Reference			Reference	
14 (f) / 10 (m) weeks	1.01	0.99 to 1.11	0.82	0.94	0.86 to 1.03	0.19
12 (f) / 18 (m) weeks	1.01	0.91 to 1.11	0.90	1.09	1.00 to 1.19	0.06
Age[1]	0.98	0.98 to 0.98	0.00	0.98	0.98 to 0.99	0.00
No. donations 2 years before INTERVAL	0.89	0.87 to 0.92	0.00	0.88	0.86 to 0.90	0.00
Donor status						
Repeat		Reference			Reference	
New donor[1]	1.20	1.01 to 1.42	0.04	1.04	0.86 to 1.26	0.69
Haemoglobin g/dL, baseline	1.00	0.96 to 1.05	0.97	0.97	0.93 to 1.01	0.10
Ferritin ug/L, baseline	1.003	1.00 to 1.00	0.00	1.004	1.00 to 1.00	0.01

[1]At start INTERVAL. [2]Participant who had not before INTERVAL provided a whole blood donation. [3]Unrounded beta: 0.9976 95%CI: 0.9963 to 0.9999. [4] Unrounded beta: 0.9991 95%CI: 0.9983 to 0.9999. OR: Odds ratio. CI: Confidence interval. (f) females. (m) males.

Supplementary Table 3: Baseline characteristics of all female participants, those lost to follow up, and those remaining until 4-year .

	All female participants			All female participants lost to follow up (4 -year)			Remaining female participants (4-year)		
	12	14	16	12	14	16	12	14	16
No. of participants	3965	3969	3894	1914	1937	1861	2051	2032	2033
Age, years[1]	44.9 ± 13.6	45.0 ± 13.5	45.2 ± 13.4	42.8 ± 13.8	42.4 ± 13.7	42.9 ± 13.8	46.8 ± 13.0	47.3 ± 13.0	47.3 ± 12.8
No. donation 2 year prior to INTERVAL	3.24 ± 1.64	3.26 ± 1.62	3.21 ± 1.64	3.01 ± 1.63	3.05 ± 1.62	3.01 ± 1.61	3.45 ± 1.62	3.45 ± 1.59	3.39 ± 1.65
New donors[2]	268 (6.8%)	268 (6.8%)	291 (7.5%)	166 (8.7%)	167 (8.6%)	174 (9.4%)	102 (5.0%)	101 (5.0%)	117 (5.7%)
Haemoglobin, g/dL	13.4 ± 0.9	13.4 ± 0.9	13.4 ± 0.9	13.4 ± 0.9	13.4 ± 0.89	13.4 ± 0.9	13.4 ± 0.9	13.4 ± 0.8	13.4 ± 0.9
Ferritin[3], µg/L	28.0 (15.4 -46.0)	26.0 (15.2 - 44.0)	27.0 (15.0 - 43.0)	27.0 (15.0 - 44.2)	25.2 (15.0 - 42.0)	26.0 (14.0 - 42.0)	29.0 (16.0 - 47.0)	27.1 (16.0 - 45.0)	27.1 (16.0 -45.0)

Data are mean (SD), number of participants % or median (interquartile range). [1]At baseline INTERVAL trial. [2]Participant who had not previously provided a full blood donation. [3]Ferritin at baseline.

Supplementary Table 4: Baseline characteristics of all male participants, those lost to follow up, and those remaining until 4-year

	All male participants			All male participants lost to follow up (4 -year)			Remaining male participants (4-year)		
	8	10	12	8	10	12	8	10	12
No. of participants	4206	4222	4240	2014	1849	1933	2192	2373	2307
Age, years[1]	47·8 ± 13·2	48·3 ± 13·3	48·2 ± 13·2	46·2 ± 13·7	46·7 ± 13·5	46·2 ± 13·7	49·3 ± 12·6	49·6 ± 12·9	49·8 ± 12·5
No. donation 2 year prior to INTERVAL	3·94 ± 1·80	3·96 ± 1·78	3·90 ± 1·77	3·72 ± 1·81	3·73 ± 1·81	3·65 ± 1·80	4·15 ± 1·76	4·14 ± 1·74	4·11 ± 1·72
New donors[2]	194 (4·6%)	204 (4·8%)	210 (5·0%)	115 (5·7%)	119 (6·4%)	119 (6·2%)	79 (3·6%)	85 (3·6%)	91 (3·9%)
Haemoglobin, g/dL	15·0 ± 1·0	14·9 ± 1·0	14·9 ± 1·0	15·0 ± 1·0	15·0 ± 1·0	14·9 ± 1·0	15·0 ± 1·0	15·0 ± 1·0	15·0 ± 1·0
Ferritin[3], µg/L	46·0 (27·0 - 74·0)	44·0 (27·0 - 72·0)	45·0 (27·1 - 72·0)	46·0 (26·8 - 74·9)	44·0 (26·0 - 75·0)	46·0 (28·0- 76·0)	46·0 (27·0 - 74·0)	44·0 (27·3 - 69·0)	44·0 (27·0 - 70·0)

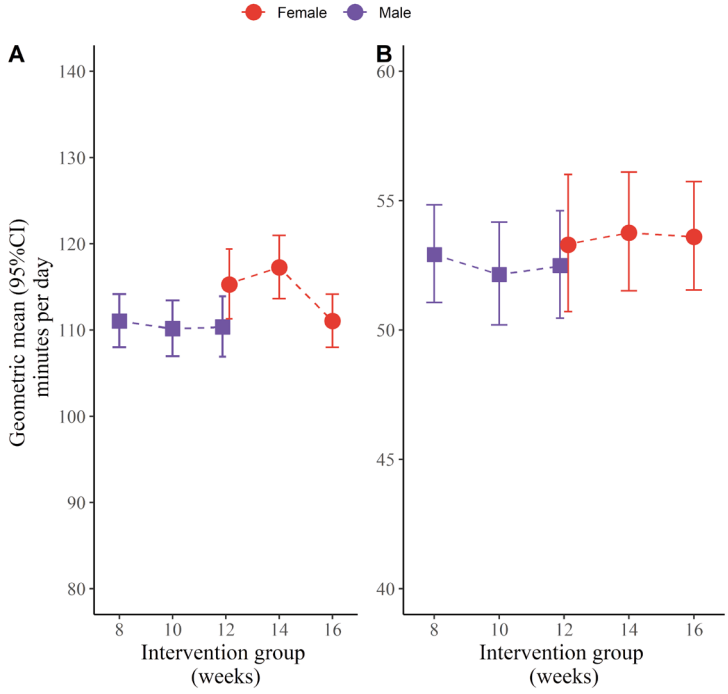
Data are mean (SD), number of participants % or median (interquartile range). [1]At baseline INTERVAL trial. [2]Participant who had not previously provided a full blood donation. [3]Ferritin at baseline.

Supplementary Table 5: Physical activity and sedentary behaviour primary outcomes by sex and intervention group.

	Female				Male			
	12 weeks	14 weeks	16 weeks	p[1]	8 weeks	10 weeks	12 weeks	p[1]
2-year								
Time spent in MVPA, minutes per day[2]	86·6 (62·8 to 110·4)	84·8 (63·8 to 110·6)	86·1 (61·5 to 109·8)	0·60	79·9 (59·1 to 104·6)	80·4 (58·9 to 106·5)	80·4 (56·6 to 107·6)	0·71
Time spent in TPA, minutes per day[2]	389·9 (331·1 to 440·6)	383·5 (332·5 to 434·8)	384·7 (327·4 to 439·1)	0·69	363·1 (307·8 to 421·6)	360·0 (310·2 to 415·4)	363·9 (309·9 to 427)	0·83
Time spent in MVPA, minutes per day[3]	64·0 (29·5 to 131·6)	64·9 (28·7 to 134·2)	63·7 (30·3 to 126·5)	1·00	89·0 (41·6 to 197·0)	86·1 (39·5 to 194)	86·7 (40·3 to 191·3)	0·65
PAEE, kJ/kg per day[3]	36·2 (24·8 to 56·2)	37·0 (24·9 to 57·3)	35·7 (24·7 to 56·7)	0·77	45·1 (30 to 73·2)	44·7 (29·8 to 72·6)	45·2 (29·7 to 74)	0·74
Time spent in SB, hours per day[3]	6·9 ± 3·2	7·1 ± 3·2	7·0 ± 3·2	0·95	7·5 ± 3·4	7·7 ± 3·4	7·6 ± 3·4	0·24
Change between 2- and 4 year[4]								
PAEE, kJ/kg per day[3]	-3·6 ± 28	-5·5 ± 27·6	-3·4 ± 26·6	0·86	-4·5 ± 34·5	-4·7 ± 34·3	-5 ± 32·6	0·16
MVPA, minutes per day[3]	-1·6 (-42·3 to 27·9)	-1·9 (-45·1 to 28·8)	-0·9 (-41·7 to 28·7)	0·63	-6·2 (-57·3 to 31·6)	-5·2 (-53·7 to 32·5)	-5·1 (-50·4 to 31·3)	0·41
SB, hours per day[3]	-0·4 ± 2·6	-0·5 ± 2·4	-0·5 ± 2·4	0·10	-0·5 ± 2·6	-0·5 ± 2·7	-0·5 ± 2·7	0·52
4-year								
PAEE, kJ/kg per day[3]	33·6 (23·7 to 50·8)	33·0 (23·2 to 49·6)	33·1 (23·8 to 51·3)	0·82	42·1 (28·2 to 67·3)	40·8 (28·4 to 64·5)	40·3 (28 to 63·1)	0·09
Time spent in MVPA, minutes per day[3]	61·3 (31·5 to 113·7)	62·7 (30·5 to 119·3)	63·1 (32·2 to 116·5)	0·56	82·3 (38·5 to 173·6)	81·3 (41·2 to 155·6)	78·5 (38·5 to 158·3)	0·28
Time spent in SB, hours per day[3]	6·6 ± 3·1	6·5 ± 3·1	6·5 ± 3·1	0·17	7·1 ± 3·4	7·1 ± 3·3	7·1 ± 3·3	0·62

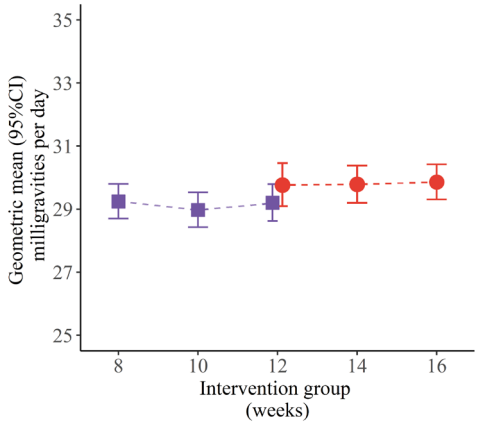
Data are means (SD) or geometric means (interquartile range). [1]P-values are from linear trend across groups. [2]As measured with accelerometers at 2 years. [3]As measured with the Recent Physical Activity Questionnaire at 4 years. [4]Data presented are crude changes, p-values are adjusted for 2-year measurements values of the outcome. MVPA: moderate-to-vigorous physical activity: TPA: total physical activity. PAEE: physical activity energy expenditure. SB: sedentary behaviour.

Supplementary Figure 6: Moderate-to-vigorous physical activity with different cut-off points



Supplementary Figure 6: Moderate-to-vigorous physical activity using 100 milli-gravity (A) and 150 milli-gravity per trial arm and sex using objectively measured 2-year data.

Supplementary Figure 7: Average acceleration



Supplementary Figure 7: Average acceleration per trial arm and sex. These were post-hoc analyses.

Supplementary Table 6: Model fit indices and comparison of adjusted BDCSM models with haemoglobin.

	Model 1	Model 2	Model 3	Model 4
Fit statistics				
Log likelihood	-113034·118	-113032·650	-113031·146	-113029·676
RMSEA	0·013	0·014	0·013	0·015
AIC	226322·236	226321·301	226318·293	226317·353
SSA BIC	226948·129	226952·122	226949·114	226953·103
CFI	0·999	0·999	1·000	1·000
TLI	0·975	0·969	0·976	0·965
Model comparison				
Versus baseline[1]		1·468	2·972	4·442

[1]Difference in log likelihood. BDCSM, bivariate dual change score models. MVPA, moderate-to-vigorous physical activity. Model 1: neither MVPA nor haemoglobin effects the other. Model 2: haemoglobin affects MVPA. Model 3: MVPA affects haemoglobin. Model 4: Both haemoglobin and MVPA affect the other. RMSEA: Root Mean Square Error of Approximation. AIC: Akaike Information Criterion. SSA BIC: Sample-size adjusted Bayesian Information Criterion. CFI: Comparative Fit Index. TLI: Tucker-Lewis Index. Models adjusted for age, sex and trial arm.

Supplementary Table 7: Model fit indices and comparison of adjusted BDCSM models with ferritin.

	Model 1	Model 2	Model 3	Model 4
Fit statistics				
Log likelihood	-98044·826	-98044·548	-98044·322	-98044·045
RMSEA	0·000	0·000	0·000	0·000
AIC	196151·653	196153·096	196152·644	196154·090
SSA BIC	196304·430	196310·802	196310·349	196316·724
CFI	1·000	1·000	1·000	1·000
TLI	1·001	1·001	1·001	1·002
Model comparison				
Versus baseline[1]		0·278	0·504	0·781

[1]Difference in log likelihood. BDCSM, bivariate dual change score models. MVPA, moderate-to-vigorous physical activity. Model 1: neither MVPA nor ferritin effects the other. Model 2: ferritin affects MVPA. Model 3: MVPA affects ferritin. Model 4: Both ferritin and MVPA affect the other. RMSEA: Root Mean Square Error of Approximation. AIC: Akaike Information Criterion. SSA BIC: Sample-size adjusted Bayesian Information Criterion. CFI: Comparative Fit Index. TLI: Tucker-Lewis Index. Models adjusted for age, sex and trial arm.

Supplementary Table 8: Parameter estimates (95% confidence intervals) for the crude and age, sex and intervention arm adjusted BDCSM models in which neither haemoglobin nor MVPA affect the other.

	Crude model		Adjusted model	
	Haemoglobin	MVPA[1]	Haemoglobin	MVPA[1]
Fixed effects				
Intercept	13.81 (13.79 to 13.82)	1.92 (1.90 to 1.93)	14.38 (14.31 to 14.45)	2.07 (1.99 to 2.16)
Slope (α)	5.68 (5.46 to 5.90)	0.95 (0.90 to 0.99)	7.70 (7.43 to 7.96)	1.03 (0.93 to 1.12)
Autoproportional (β)	-0.42 (-0.44 to -0.40)	-0.52 (-0.54 to -0.50)	-0.54 (-0.56 to -0.52)	-0.54 (-0.56 to -0.52)
Coupling (γ)	-	-	-	-
Random effects				
Intercept variance	1.74 (1.70 to 1.78)	1.75 (1.70 to 1.80)	1.29 (1.25 to 1.32)	1.07 (1.02 to 1.13)
Slope variance	0.89 (0.87 to 0.92)	1.10 (1.05 to 1.15)	0.82 (0.80 to 0.85)	1.08 (1.03 to 1.13)
Intercept covariance	0.08 (0.06 to 0.10)		-0.04 (-0.06 to -0.02)	

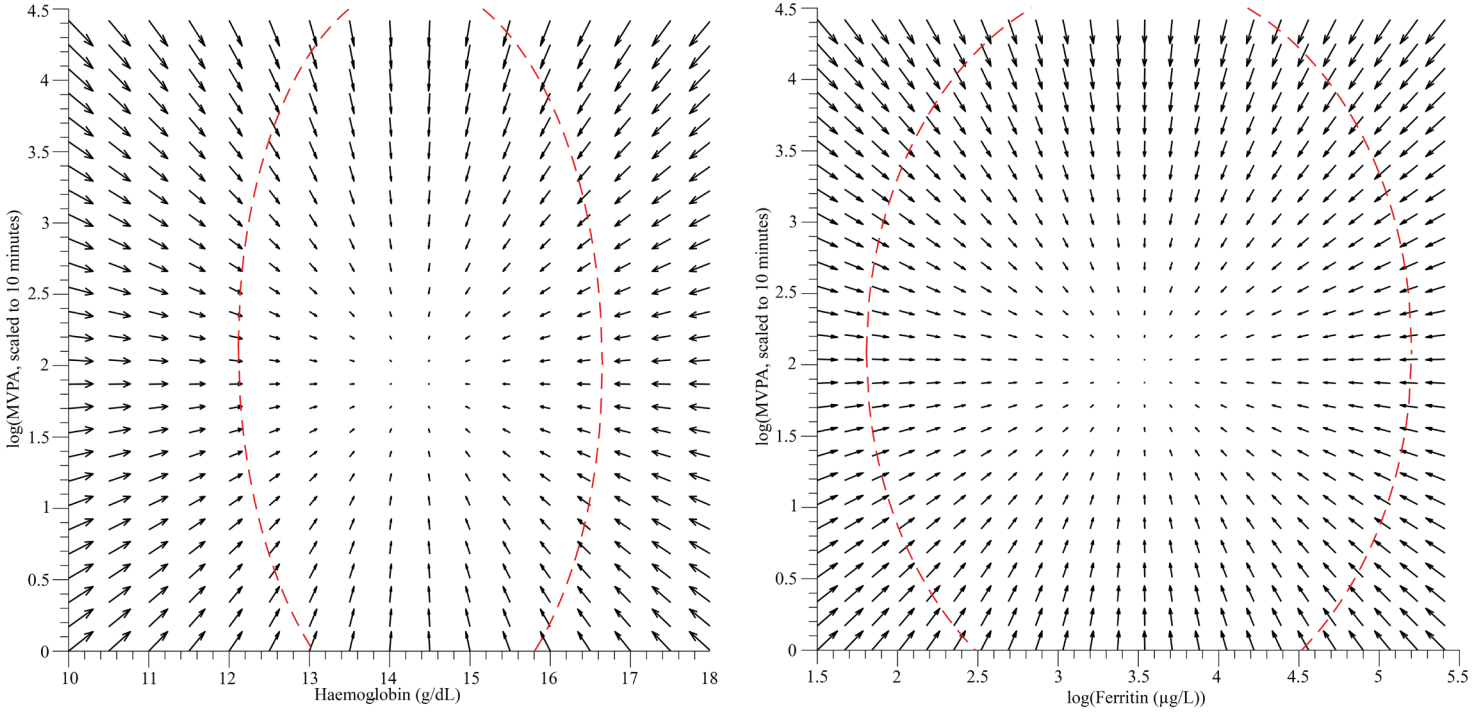
BDCSM, bivariate dual change score models. MVPA, moderate-to-vigorous physical activity reported per 10 minutes. [1]Log transformed values are presented.

Supplementary Table 9: Parameter estimates (95% confidence intervals) for the crude and age, sex and intervention arm adjusted BDCSM models in which neither ferritin nor MVPA affect the other

	Crude model		Adjusted model	
	Haemoglobin	MVPA[1]	Haemoglobin	MVPA[1]
Fixed effects				
Intercept	3·28 (3·26 to 3·29)	1·92 (1·90 to 1·93)	3·50 (3·48 to 3·53)	2·10 (2·06 to 2·13)
Slope (α)	1·58 (1·53 to 1·64)	0·95 (0·90 to 0·99)	1·71 (1·64 to 1·77)	1·05 (0·99 to 1·11)
Autoproportional (β)	-0·47 (-0·48 to -0·45)	-0·52 (-0·54 to -0·50)	-0·49 (-0·51 to -0·47)	-0·54 (-0·56 to -0·52)
Coupling (γ)	-	-	-	-
Random effects				
Intercept variance	0·76 (0·74 to 0·78)	1·75 (1·70 to 1·80)	0·72 (0·71 to 0·73)	1·71 (1·66 to 1·76)
Slope variance	0·47 (0·45 to 0·48)	1·1 0 (1·05 to 1·15)	0·46 (0·45 to 0·47)	1·09 (1·04 to 1·14)
Intercept covariance	0·03 (0·02 to 0·05)		0·00 (-0·01 to 0·02)	

BDCSM, bivariate dual change score models. MVPA, moderate-to-vigorous physical activity reported per 10 minutes. [1]Log transformed values are presented.

Supplementary Figure 8: Vector fields showing individual change in iron status and objectively measured MVPA



Supplementary Figure 8: Vector field showing joint movements between haemoglobin and log moderate-to-vigorous physical activity (MVPA) of INTERVAL participants between 2- and 4-year b follow-up.

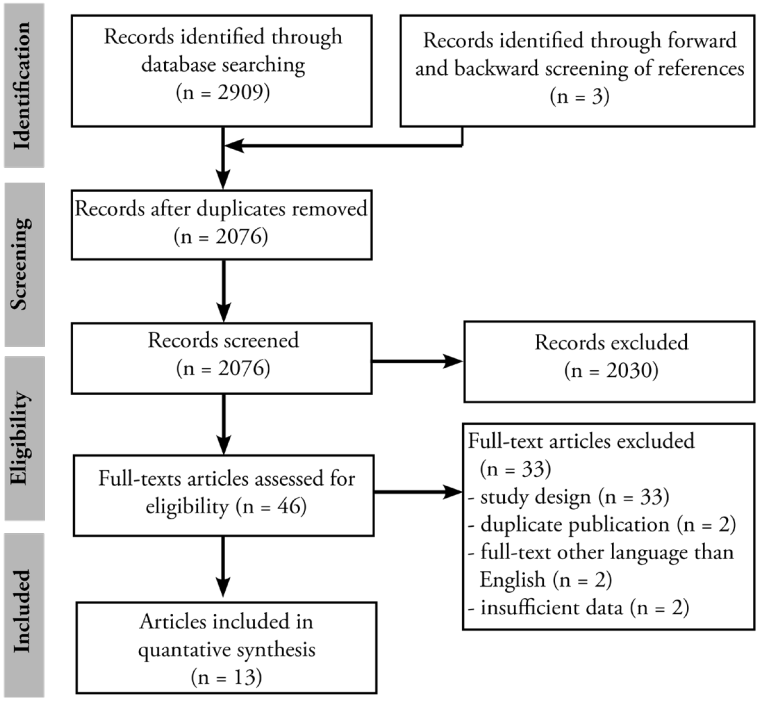
Ellipsoid reflects 95% of the data.

Supplementary Table 10: Sensitivity analyses with multivariable models.

Exposure	No. exposure variants	Out-come	MR unit	MR method	Beta (95%CI)	p-value	p-value heteroge-neity
Ferritin	6	MVPA	SD Δper SDΔ in exposure	IVW	0.01 (-0.05 to 0.06)	0.845	0.351
Ferritin	6	MVPA	SDΔ per SDΔ in exposure	IVW (random model)	0.01 (-0.05 to 0.06)	0.853	0.351
Ferritin	6	MVPA	SDΔ per SDΔ in exposure	MR Egger beta	-0.00 (-0.13 to 0.12)	0.974	
Ferritin	6	MVPA	SDΔ per SDΔ in exposure	MR Egger intercept	0.00 (-0.01 to 0.01)	0.892	
Ferritin	6	MVPA	SDΔ per SDΔ in exposure	Weighted median	0.02 (-0.05 to 0.08)	0.640	
Haemoglobin	18	MVPA	SDΔ per SDΔ in exposure	IVW	0.03 (-0.02 to 0.07)	0.213	0.700
Haemoglobin	18	MVPA	SDΔ per SDΔ in exposure	IVW (random model)	0.03 (-0.02 to 0.07)	0.213	0.700
Haemoglobin	18	MVPA	SDΔ per SDΔ in exposure	MR Egger beta	0.07 (-0.02 to 0.16)	0.141	
Haemoglobin	18	MVPA	SDΔ per SDΔ in exposure	MR Egger intercept	-0.00 (-0.01 to 0.00)	0.323	
Haemoglobin	18	MVPA	SDΔ per SDΔ in exposure	Weighted median	0.02 (-0.05 to 0.08)	0.600	
Haemoglobin	19	MVPA	SDΔ per SDΔ in exposure	MV IVW (fixed model)	0.03 (-0.02 to 0.08)	0.219	
Ferritin	19	MVPA	SDΔ per SDΔ in exposure	MV IVW (fixed model)	0.01 (-0.06 to 0.08)	0.830	
Haemoglobin	19	MVPA	SDΔ per SDΔ in exposure	MV IVW (random model)	0.03 (-0.02 to 0.08)	0.219	0.506
Ferritin	19	MVPA	SDΔ per SDΔ in exposure	MV IVW (random model)	0.01 (-0.06 to 0.08)	0.830	0.506

MVPA: Moderate-to-vigorous physical activity. MR: Mendelian Randomisation. 95%CI: 95% confidence interval). Δ: change. IVW: inverse variance weighted. MV: multivariable.

Supplementary Figure 9: Flowchart



Supplementary Figure 9: Flowchart for the umbrella review on iron supplementation/ status and physical capacity.

Supplementary Table 11: Reasons for exclusion of full-text articles.

Authors	Year	Journal	Decision	Reason
Alaunyte, et al.	2015	J Int Soc Sports Nutr	Exclude	Wrong study design
Avni et al.,	2012	Eur J Heart Fail	Exclude	Insufficient data provided
Bailie	2010	Arzneimittelforschung	Exclude	Wrong study design
Caetano Júnior, et al.	2014	Revista Andaluza de Medicina del Deporte	Exclude	Full-text not English
Calbet, et al.	2006	Respir Physiol Neurobiol	Exclude	Wrong study design
Cody, et al.	2016	Cochrane Database Syst Rev	Exclude	Wrong outcome
Deldicque, L., et al.	2015	Front Nutr	Exclude	Wrong study design
DellaValle	2013	Curr Sports Med Rep	Exclude	Wrong study design
Economos, et al.	1993	Sports Med	Exclude	Wrong study design
Ekblom	1997	World Rev Nutr Diet	Exclude	Wrong study design
Jin, et al.	2010	Eur J Heart Fail	Exclude	Insufficient data provided
Kotecha, et al	2011	Am Heart J	Exclude	Based on review that is included this umbrellla review
Koenig, et al.	1998	Exercise Immunology Review	Exclude	Wrong study design
Lawler, et al.	2010	J Card Fail	Exclude	Wrong study design
Mancini, et al.	2003	Kidney Int Suppl	Exclude	Wrong study design
McClung	2012	J Trace Elem Med Biol	Exclude	Wrong study design
McClung	2019	Biol Trace Elem Res	Exclude	Wrong study design
McClung	2013	Annu Rev Nutr	Exclude	Wrong study design
Michalczyk, et al.	2016	Nutrients	Exclude	Wrong study design
Nielsen, et al.	1998	Sports Med	Exclude	Wrong study design
Otto, et al.	2013	Extrem Physiol Med	Exclude	Wrong study design
Pasricha, et al.	2014	Cochrane Database Syst Rev	Exclude	More recent review included
Pedlar, et al.	2018	Eur J Sport Sci	Exclude	Wrong study design
Pompano, et al.	2019	J Nutr	Exclude	Wrong study design
Gandra, et al.	2010	American Journal of Kidney Diseases	Exclude	Wrong study design
Rubeor, et al.	2018	Sports Health	Exclude	Wrong study design
Stugiewicz, et al.	2016	Eur J Heart Fail	Exclude	Wrong study design
van Dronkelaar, et al.	2018	J Am Med Dir Assoc	Exclude	Wrong study design
VanHeest, et al.	2007	Curr Sports Med Rep	Exclude	Wrong study design
Vitale, et al.	2018	Card Fail Rev	Exclude	Wrong study design
Xue, et al.	2006	Chinese Journal of Clinical Rehabilitation	Exclude	Full-text not English
Zimmermann	2003	Schweizerische Zeitschrift fuer Sportmedizin & Sporttraumatologie	Exclude	Wrong study design
Zourdos, et al.	2015	J Strength Cond Res	Exclude	Wrong study design

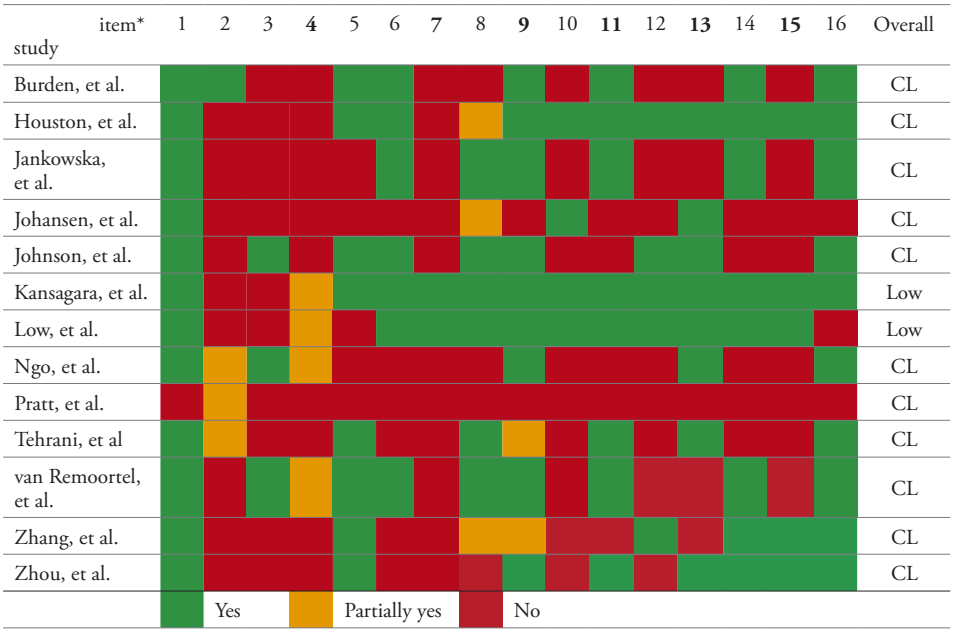
Supplementary Table 12: Characteristics of systematic reviews with meta-analysis included in the umbrella review.

Author	Year	Design	Population	VO_2 max	VO_2 peak	$\%VO_2$ peak	Exercise duration	Time to exhaustion	VO_2 anaerobic threshold	Energy expenditure	RER	Lactate levels	Peak work rate	6MWD
				n studies (n study participants)										
Burden, et al.[10]	2015	RCTs & nRCTs	IDNA	4 (86)	-		-	-						-
Houston, et al.[11]	2018	RCTs	IDNA	8 (204)	-	-	-	4 (69)	-	-	-	-	-	-
Jankowska, et al.[12]	2016	RCTs	HF	-	-	-	-	-	-	-	-	-	-	1 (431)
Johansen, et al.[13]	2010	RCT & obs.	KD	-	2 (34)	-	-	-	-	-	-	-	-	-
Johnson, et al.[14]	2019	RCTs & obs.	BD	4 (45)	-	-	-	5 (71)	-	-	-	-	-	-
Kansagara, et al.[15]	2013	RCT & obs.	HD	-	-	-	-	-	-	-	-	-	-	3 (96)
Low, et al.[16]	2016	RCT & quasi-RCT	WRA	7[1] (267)/ 17[2] (491)	-	4 (110)	-	2 (38)	-	2 (61)	5 (137)	4 (106)		
Ngo, et al.[17]	2010	RCTs	HF	-	3 (102)	-	4 (362)	-	2 (61)	-	-	-	-	3 (96)
Pratt & Khan[18]	2016	RCT & obs.	IDNA	3 (99)	-	-	-	-	-	-	3 (99)	-	-	-
Tehrani, et al.[19]	2009	RCTs	HF	-	3 (102)	-	4 (362)	-	-	-	-	-	-	3 (95)
Van Remoortel, et al.[20]	2017	RCT & obs.	BD	8 (84)	-	-	-	4 (37)	-	-	-	-	3 (39)	-
Zhang, et al.[21]	2019	RCTs	HF	-	3 (416)	-	-	-	-	-	-	-	-	2 (471)
Zhou[22]	2019	RCTs	HF	-	3 (416)		-	-						2 (471)

6MWD: 6 minute walking distance. RER: Respiratory exchange ratio. RCT: randomised controlled trials. nRCT: non randomised controlled trials. IDNA: iron deficient non aneamic. WRA: women of reproductive age. HF: heart failure. KD: Kidney disease. HD: Heart disease. BD: Blood donation. [1]Relative VO_2 max included studies. [2]Absolute VO_2 max included studies

Appendix 5: Risk of bias assessment

The AMSTAR2 was used to assess the risk of bias of the included studies. We constructed an overall scoring as proposed by the developers of the AMSTAR2[78]. As well as an adjusted AMSTAR2 score (see below) as part of the overall scoring outcomes are partly attributable to the fact that the AMSTAR2 was not yet disseminated when the majority of the included systematic reviews were published. For example, all studies reported a flow chart of the search process as recommended by the PRISMA statement. However, without justification of all excluded studies during the full-text screening phase this item should be scored as no according to the AMSTAR2. The lack of a justification may not reflect a higher risk of bias, as such we slightly adjusted the scoring.



Supplementary Figure 10: AMSTAR2 risk of bias scores.

Items in bold represent critical domains. High: No or one non-critical weakness. Moderate: More than one non-critical weakness. Low: One critical flaw with or without non-critical weaknesses.Critically low (CL): more than one critical flaw with or without non-critical weaknesses.

*AMSTAR-2 items:

Item 1: Did the research questions and inclusion criteria for the review include the components of PICO?

Item 2: Did the report of the review contain an explicit statement that the review methods were established prior to conduct of the review and did the report justify any significant deviations from the protocol?

Item 3: Did the review authors explain their selection of the study designs for inclusion in the review?

Item 4: Did the review authors use a comprehensive literature search strategy?

Item 5: Did the review authors perform study selection in duplicate?

Item 6: Did the review authors perform data extraction in duplicate?

Item 7: Did the review authors provide a list of excluded studies and justify the exclusions?

Item 8: Did the review authors describe the included studies in adequate detail?

Item 9: Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?

Item 10: Did the review authors report on the sources of funding for the studies included in the review?

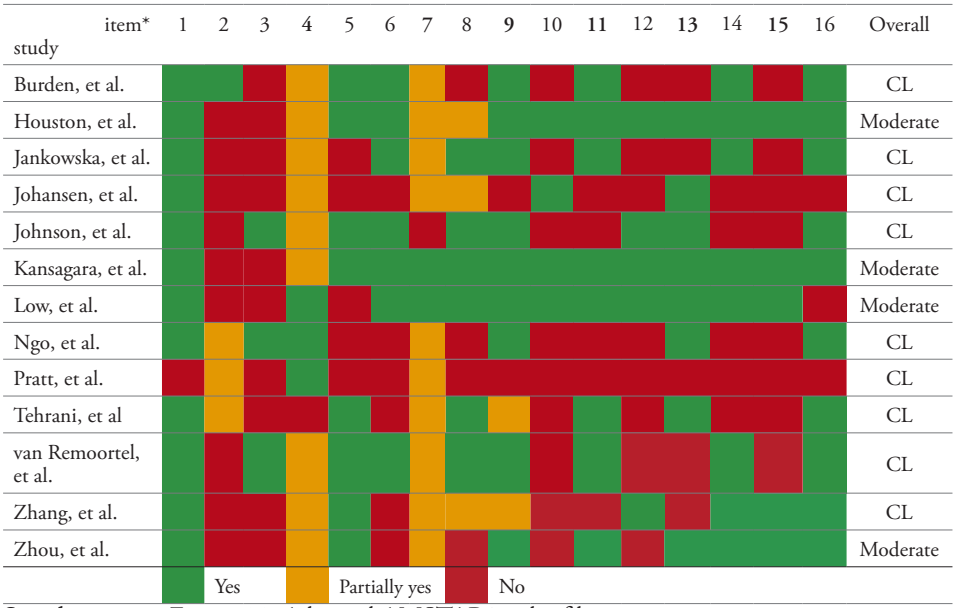
Item 11: If meta-analysis was justified did the review authors use appropriate methods for statistical combination of results?

Item 12: If meta-analysis was performed did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?

Item 13: Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?

Item 14: Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?

Item 15: If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?



Supplementary Figure 11: Adjusted AMSTAR2 risk of bias scores.

Items in bold represent critical domains. High: No or one non-critical weakness. Moderate: More than one non-critical weakness. Low: One critical flaw with or without non-critical weaknesses. Critically low (CL): more than one critical flaw with or without non-critical weaknesses.

The following adjustments were made:
Item 4: A justification about publication restriction was no longer required in order to score partially yes. Neither was the consultation of an expert necessary as to be graded with a yes.
Item7: Reporting of a PRISMA flow chart was considered sufficient to score a partially yes.

Supplementary Table 13: Original studies that were included in the eligible systematic reviews with meta-analyses.

	Burden, 2016	Houston, 2018	Jankowska, 2016	Johansen, 2010	Johnson, 2019	Kansagara, 2013	Low, 2016	Ngo, 2010	Pratt,2016	Tehrani, 2009	van Remoortel 2017	Zhang, 2019	Zhou, 2019
Studies													
Akiba, et al.[23]				x									
Anker, et al.[24]			x									x	x
Barany, et al.[25]				x									
Barany, et al.[26]				x									
Birnbaum, et al.[27]											x		
Blee, et al.[28]	x												
Brownlie, et al.[29]		x					x						
Brownlie, et al.[30]		x					x		x				
Burden, et al.[31]		x											
Burnley, et al.[32]					x						x		
Christian & Mancini[33]								x					
Delano, et al.[34]				x									
DellaValle & Haas[35]	x						x						
Dellweg, et al.[36]											x		
Duda, et al.[37]											x		
Ekblom, et al.[38]											x		
Foster, et al.[39]											x		
Friedman, et al.[40]	x												
Fritsch, et al.[41]											x		
Ghali, et al. [42]								x		x			
Gordon, et al.[43]					x								
Grunze, et al.[44]				x									
Hill, et al.[45]					x						x		
Hinton & Sinclair[46]	x	x					x		x				
Hinton, et al.[47]	x						x		x				
Hollman, et al.[48]											x		
Jensen, et al.[49]							x						
Judd, et al.[50]					x						x		
Klingshirn, et al.[51]	x	x					x						
Kourea, et al.[52]						x		x					

Studies	Burden, 2016	Houston, 2018	Jankowska, 2016	Johansen, 2010	Johnson, 2019	Kansagara, 2013	Low, 2016	Ngo, 2010	Pratt, 2016	Tehrani, 2009	van Remoortel 2017	Zhang, 2019	Zhou, 2019
LaManca & Haymes[53]	x	x					x						
LaManca.[54]							x						
Li, et al.[55]							x						
Lundin, et al. [56]				x									
Lyle, et al.[57]							x						
Macdougall, et al.[58]				x									
Magazanik, et al.[80]	x												
Mancini, et al.[81]						x		x		x			
Marrades, et al.[59]				x									
Mayer, et al.[60]				x									
Metra, et al.[61]				x									
Newhouse, et al.[62]							x						
Okonko, et al.[63]												x	x
Palazzuoli, et al.[64]								x		x			
Parissis, et al.[65]						x		x		x			
Peeling, et al.[66]	x	x											
Ponikowski, et al.[67]								x		x			
Ponikowski, et al.[68]			x									x	x
Radjen, et al.[82]	x						x						
Rajaram, et al.[69]							x						
Robertson, et al.[70]				x									
Rosenlof, et al.[71]				x									
Schoene, et al.[72]	x												
Taniguchi, et al.[73]							x						
Toblli, et al.[74]												x	x
Tsutsui, et al.[75]				x									
Walsh & McNaughton.[83]	x												
Yoshida, et al.[76]							x						
Zhu & Haas.[77]	x	x					x		x				

Appendix 6: Assessment methods of outcomes meta-analyses

VO_2 peak and, absolute and relative VO_2 max

All were assessed using maximal cardiopulmonary exercise test with incremental workload until subjects were unable to maintain a pre-specified cadence. Seven tests were conducted on a cycle ergometer[1-7], two on a rowing ergo meter [8,9], four on treadmill[10-13] and in one study subjects had the opportunity to either choose a cycle ergo meter or a treadmill[14]. Two studies did not provide information on the device used (cycle or rowing meter or treadmill)[15,16]. Nine studies reported that subjects were requested to restrain from any strenuous physical activity 2 days to 2 hours prior to testing[1,3-5,8,10,11]. We were unable to retrieve five original studies[17-21] and three studies were reported in a non-English language[17,22]. We refer to the original articles for more specific information. Outcomes were based on reporting by the included systematic reviews.

Percentage VO_2 peak

Relative intensity of the workload was determined at baseline by all four studies[3,4,11,23] reporting percentage VO2 peak. The baseline resistance level was also used for the post-treatment tests.

Exercise duration

Modified Naughton protocols were used to determine exercise duration[24-26]. No specific information could be retrieved from one study[27].

Time to exhaustion – normal

Three studies used input from either a previous VO_2 max[3,14] or recent 10K race[11] to determine the intensity of the time to exhaustion test. Time to exhaustion was determined during a discontinuous incremental treadmill test, the end point was volitional exhaustion[10].

Time to exhaustion – blood withdrawal

Two studies used an incremental exercise test[28,29,] the endpoint was defined as the moment when participants were unable to maintain the workload. Hill et al., used a predefined work rate which was held constant, the end point of the test was the moment participants were unable to maintain a minimum cadence[30]. Verbal encouragement was reported by two studies[28,30], others did not provide information on this point. No information on this outcome measure could be found for two articles[31,32], one article was in a non-English language[22].

VO_2 anaerobic threshold
The VO_2 anaerobic threshold was defined by both studies as the point at which the rate of carbon dioxide production exceeded oxygen consumption[24,27].

Respiratory exchange ratio
Oxygen and CO_2 measurements[4] were averaged over 15-30 second intervals by several authors[2,3]. No details were provided concerning the exact timing of the measurements by the majority[2,4], other than pre- and post-test[3,23]. One study reported the workload at the moment the respiratory exchange ratio was measured to be near the individual's lactate threshold[11].

Lactate levels
Lactate concentration was measured at the end of the cardiopulmonary exercise test by one study using a capillary sample[4]. Another reported lactate to be measured three minutes after completion of the test with blood from a venipuncture[11]. The third study only reported the use of capillary samples but did not provide a time indication[3].

Peak work rate
Three studies presented peak work rates but only one original study was available. This study described peak work rate as the maximal work rate during an exercise test which started at power output of 20W, with gradual increases of 20W every 3 minutes until exhaustion[33]. The other full-texts were not retrievable[34] and/or in another language than English[22].

6 Minute walking distance
Original studies reported the use of the "6 minute walking distance test"[24,35-37]. Only one study specified participants not to be encouraged during the test[27] and one reported the assessor to be blinded to the treatment[38].

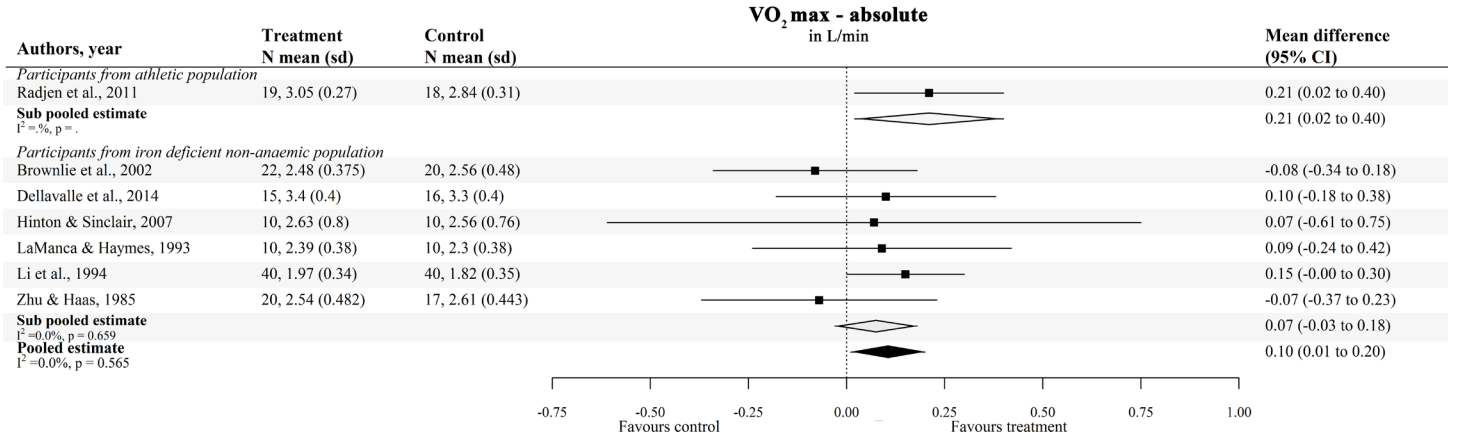
Energy expenditure
Two studies reported energy expenditure. Energy expenditure in kJ/min was determined from a 60 min exercise bout on a cycle ergometer set at 60% of each subjects individual VO2peak, using metabolic equivalent values from the Compendium of Physical Activities by Ainsworth with 1 MET = 4.186 kJ/kg body weight/h by the first study[2]. In the other study subjects performed a 15 km time trial on a cycle ergometer at a predetermined resistance to allow subjects to achieve 70% of VO2max at 60 rpm. Average VO2 and RER values were used to calculate energy expenditure using the equation by McArdle et al (i.e. VO2 x Resparatiory exchange ratio (RER) + 16040)[23].

Appendix 7: Elaboration on findings of the meta-analyses

The pooled results of seven studies examining the effect of iron supplementation on absolute VO_2 max involving 276 participants showed a beneficial effect of iron supplementation relative to control (mean difference [MD] 0.10 L/min; 95% confidence interval [95%CI] 0.01 to 0.20; I^2=0%), see Figure 3; Supplementary Figures 10-23. Fifteen studies enrolling 392 participants reported relative VO_2 max. Although none of the pooled subgroup estimates showed a significant beneficial effect of iron supplementation compared to control, the overall pooled result did (MD 1.78 ml/kg/min; 95%CI 0.03 to 3.53; I^2=80%). Visual inspection of the relative VO2 max contour-enhanced funnel plot did not indicate publication bias. The effect of blood withdrawal on relative VO_2 max was examined in 9 studies including 99 participants. Meta-analysis identified a statistically significant reduction (MD -3.34 mL/min/kg; 95%CI -5.27 to -1.41; I^2=0%). The pooled effect estimates were not affected by excluding one study that examined the outcome after 48-72 hours and not after 24-48 hours (MD -3.34 mL/min/kg; 95%CI -5.24 to -1.41)[43]. Six studies, of which four included heart failure and two haemodialysis patients, involving 166 participants assessed peak VO_2. Pooled effect estimates showed beneficial effects of iron supplementation compared to control (MD 2.42 mL/kg/min; 95%CI 1.19 to 3.66; I^2=42%). Pooled effect estimates of heart failure studies[63,79] that did not provide iron to the control group was stronger and remained significant while studies[64,67] providing iron to the control group showed a non-significant improvement. Four studies involving 110 IDNA participants examined the effect of iron supplementation versus placebo on percentage VO_2 peak. Levels were lower in the treatment group, indicating that the point at which physical activity efforts increase marginally while blood lactates rises fast is at a higher percentage of VO_2 max (MD -2.96%; 95%CI -5.69 to -0.22; I^2=43%). Exercise duration was measured in four studies including 362 heart failure patients and showed beneficial effects of erythropoiesis-stimulating agents (ESAs) and iron supplements compared with control (MD 96.82 seconds; 95%CI 5.22 to 188.42; I^2=75%). With exception of one study[79], all control treatment consisted of placebo and iron (standard practice). Heterogeneity reduced considerably after sub-group analyses based on type of study population (symptomatic heart failure patients I^2=58% and heart failure patients I^2=15%). Beneficial effects were only seen in study populations that included functional class III-IV New York Heart Association heart failure patients (MD 139.86 seconds; 95%CI 55.51 to 224.22). The inclusion criteria for the other two studies were more lenient (symptomatic heart failure). There was no statistical difference in time to exhaustion in individuals (69 participants, four studies) who received iron supplement compared with controls (MD 0.20 minutes; 95%CI -0.27 to 0.67; I^2 =0%). In six studies (122 participants) time to exhaustion was

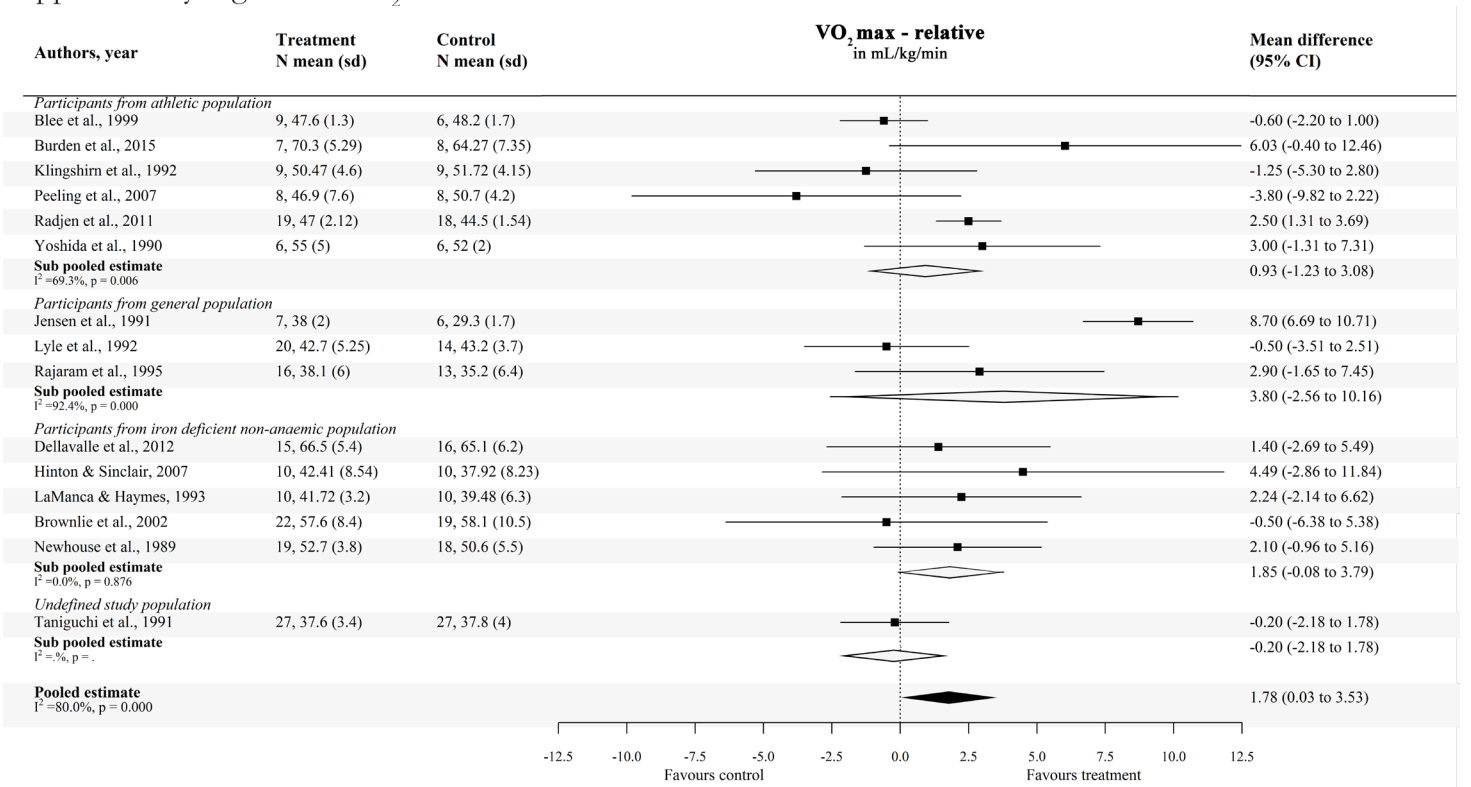
reduced 24-72 hours after withdrawal of 450 – 500 ml blood (MD -0.35 minutes; 95%CI -0.62 to -0.09; I^2=0%). Omitting the study[43] in which time to exhaustion was measured after 48-72 hours did not change the results (MD -0.35 minutes; 95%CI -0.62 to -0.09). Two studies including 61 participants reported peak VO_2 at anaerobic threshold, which improved by 2.92 mL/kg/min in the ESA treatment group compared to control (95%CI 0.09 to 5.75; I^2=84%). The effect of iron supplementation on energy expenditure was studied in 61 (recreationally) trained participants in two studies and showed no effect. (MD -0.08 kJ/min; 95%CI -0.52 to 0.36; I^2 = 0%). Respiratory exchange ratio (RER) was reported by five studies including 137 IDNA participants. There was no evidence of a significant change in RER in the group receiving iron supplementation (MD -0.01 total; 95%CI -0.02 to 0.01; I^2 =0%). Three studies enrolling 75 participants showed no change in lactate at levels (MD 0.00 mmol/L; 95%CI -0.72 to 0.72;I^2 =0%). In three blood donation studies peak work rate was assessed in 78 participants. Meta-analysis identified a decrease in peak work rate 24-48 hours after blood withdrawal (MD -15.42W; 95%CI -36.7 to 5.82; I^2 =0%). 6MWD was measured in six studies (378 participants) and improved the distance with 49.96 meters (95%CI 26.08 to 73.85; I^2 =58%) in the group assigned to receive iron treatment. Energy expenditure was measured in two studies including 61 participants and showed no beneficial effect of iron supplementation (MD -0.08 kJ/min; 95%CI -0.52 to 0.36; I^2 = 0%).

Supplementary Figure 12: VO_2 - max - absolute



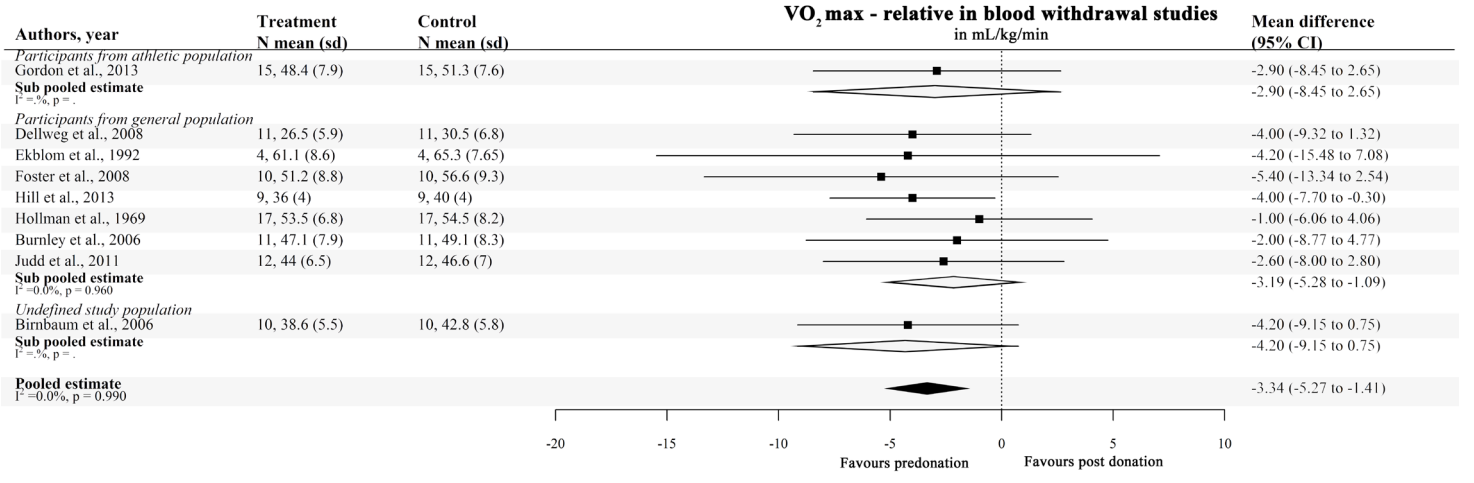
Supplementary Figure 12: VO_2 max - absolute.

Supplementary Figure 13: VO_2 max – relative



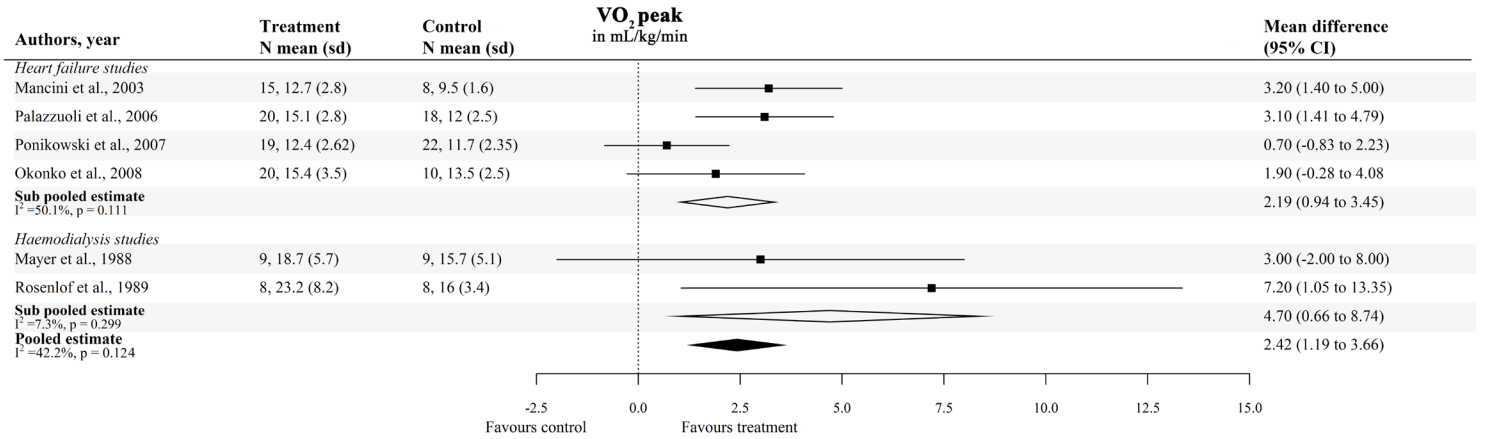
Supplementary Figure 14: VO_2 max - relative in blood withdrawal studies.

Supplementary Figure 14: VO_2 max – relative in blood withdrawal studies



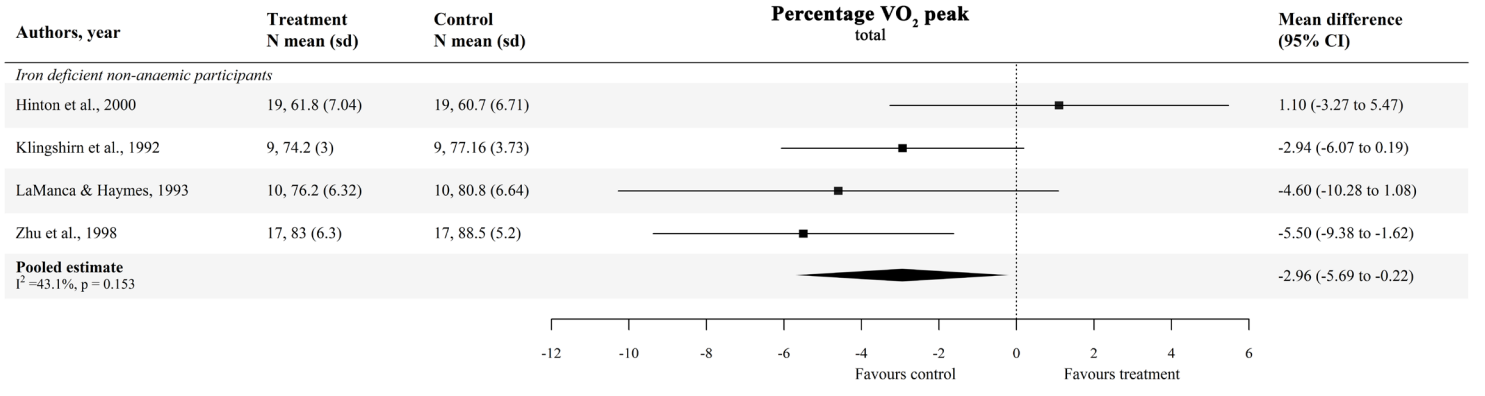
Supplementary Figure 14: VO_2 max - relative in blood withdrawal studies.

Supplementary Figure 15: VO_2 peak



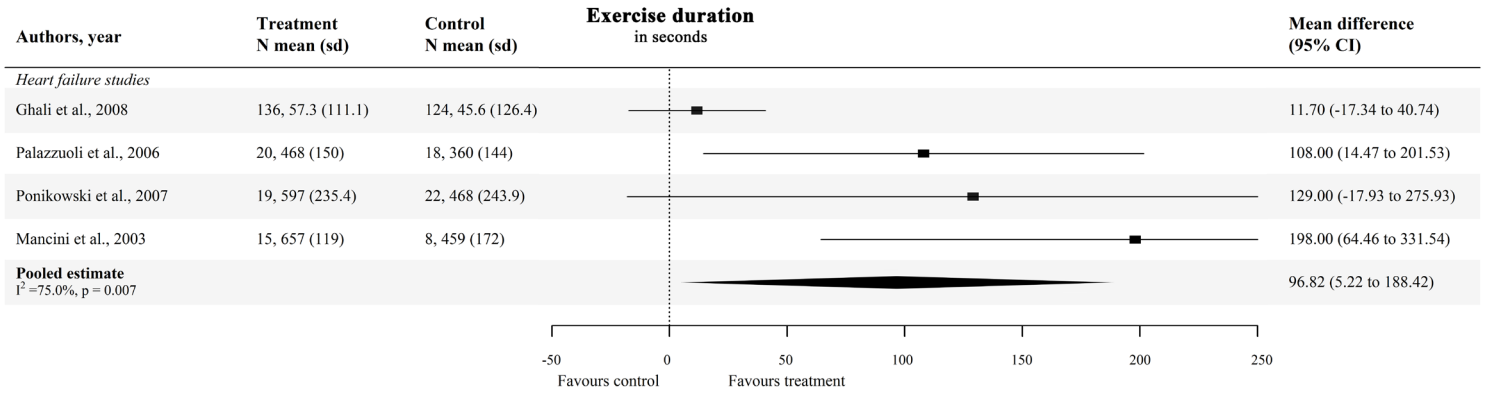
Supplementary Figure 15: VO_2 peak.

Supplementary Figure 16: percentage VO_2 peak



Supplementary Figure 16: percentage VO_2 peak.

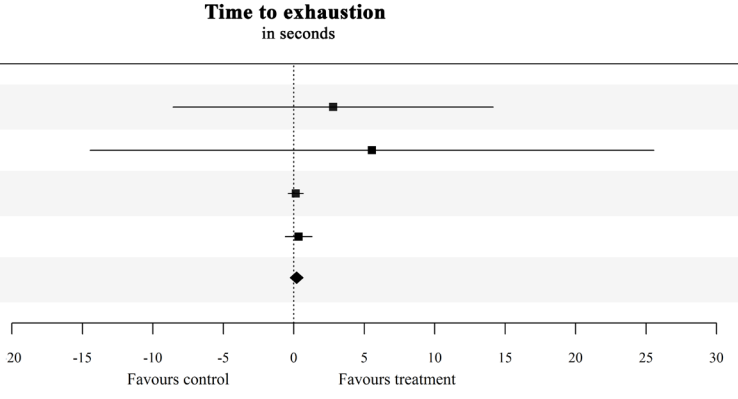
Supplementary Figure 17: Exercise duration



Supplementary Figure 17: Exercise duration.

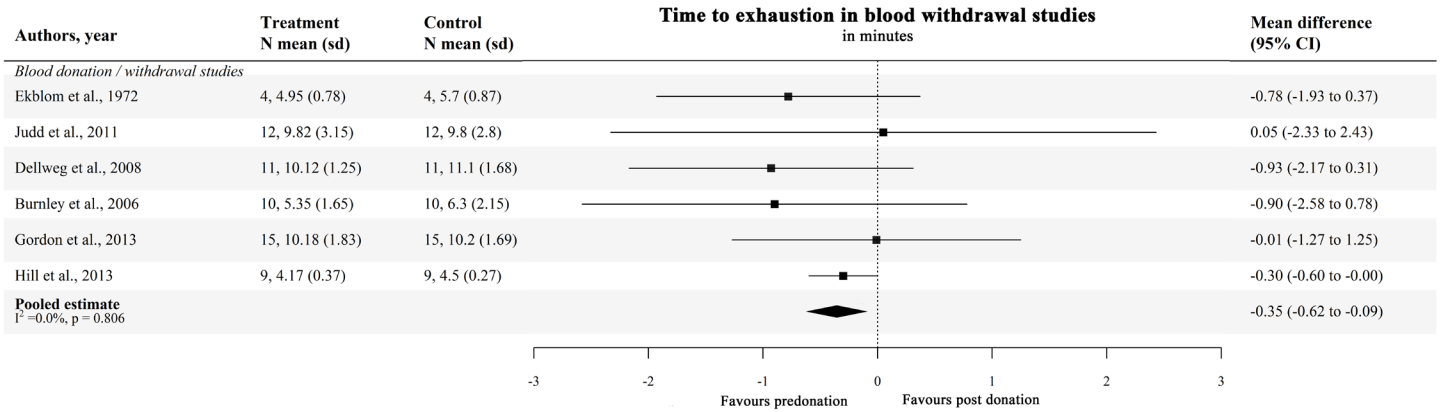
Supplementary Figure 18: Time to exhaustion

Authors, year	Treatment N mean (sd)	Control N mean (sd)	Mean difference (95% CI)
Iron deficient non-anaemic participants			
Klingshirn et al., 1991	9, 83.23 (13.62)	9, 80.4 (10.82)	2.79 (-8.57 to 14.15)
LaManca et al, 1993	10, 51.4 (23.6)	10, 45.9 (22.04)	5.55 (-14.46 to 25.56)
Peeling et al., 2007	8, 3.36 (0.54)	8, 3.2 (0.57)	0.14 (-0.40 to 0.68)
Burden et al., 2015	7, 6.22 (0.8)	8, 5.9 (1.07)	0.34 (-0.61 to 1.29)
Pooled estimate I^2 =0.0%, p = 0.895			0.20 (-0.27 to 0.67)



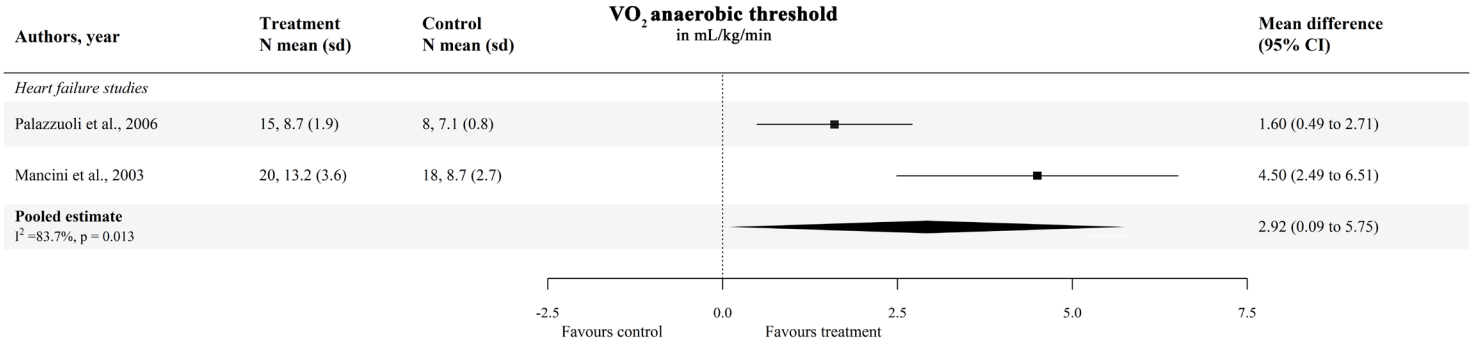
Supplementary Figure 18: Time to exhaustion.

Supplementary Figure 19: Time to exhaustion in blood withdrawal studies



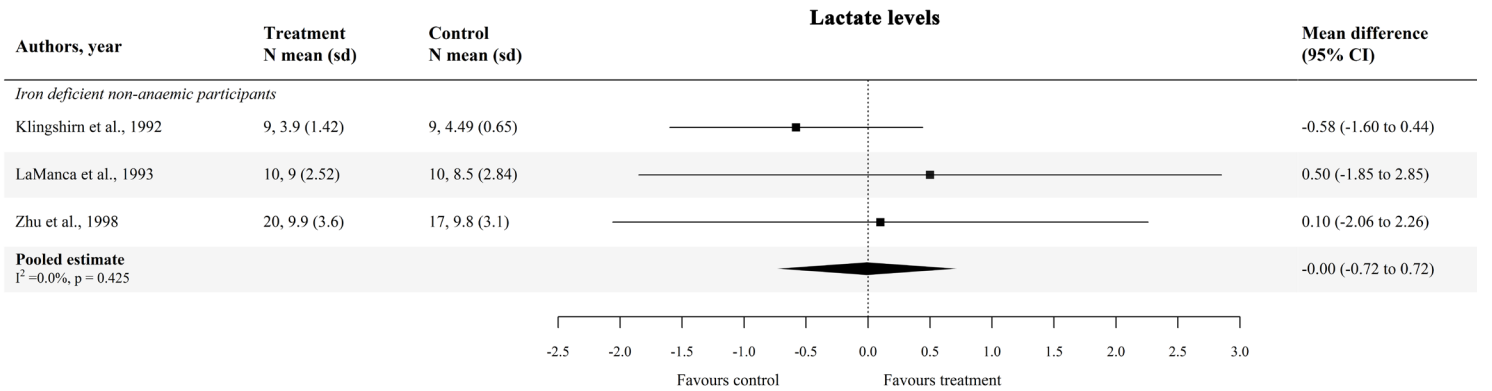
Supplementary Figure 19: Time to exhaustion in blood withdrawal studies.

Supplementary Figure 20: VO_2 anaerobic threshold



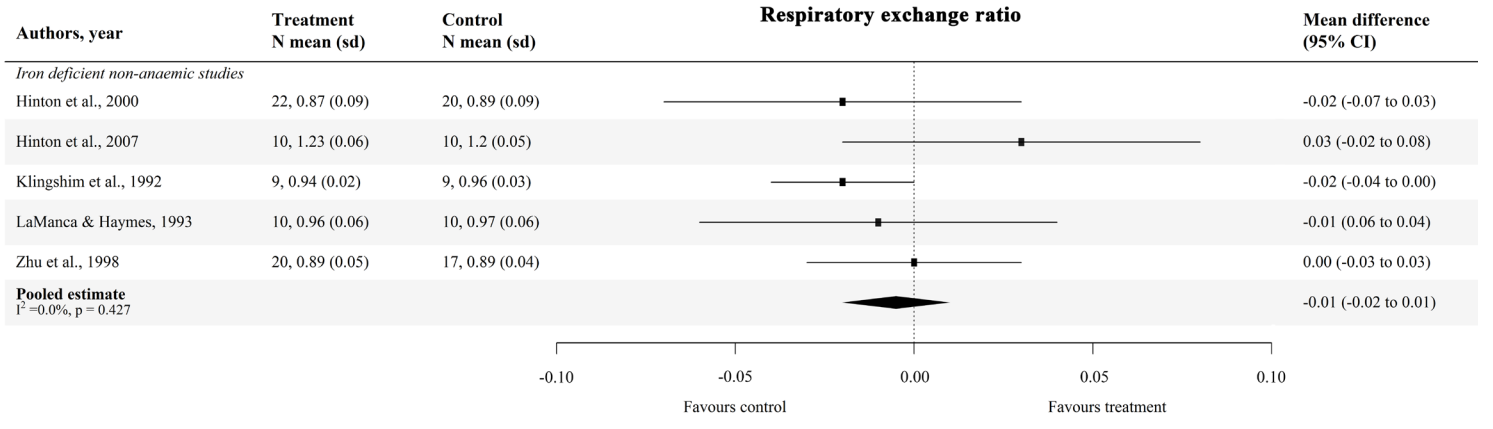
Supplementary Figure 20: VO_2 anaerobic threshold.

Supplementary Figure 21: Lactate levels



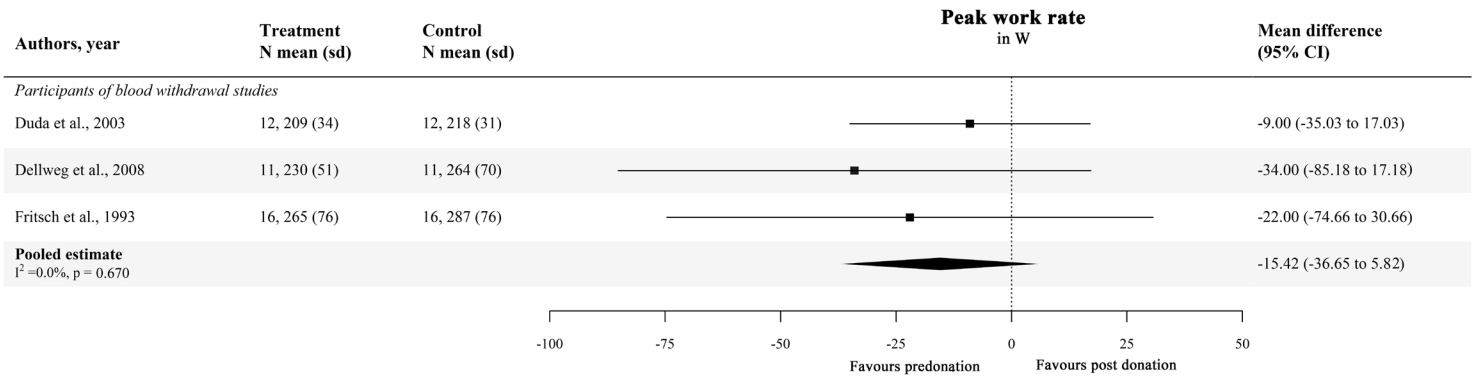
Supplementary Figure 21: Lactate levels.

Supplementary Figure 22: Respiratory exhange ratio



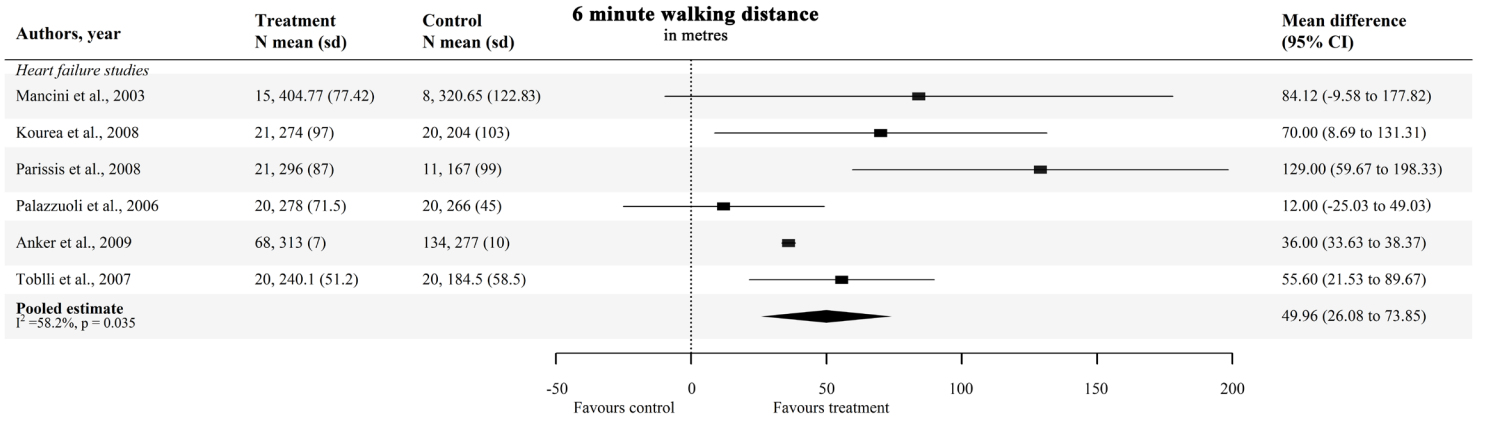
Supplementary Figure 22: Respiratory exchange ratio.

Supplementary Figure 23: Peak work rate



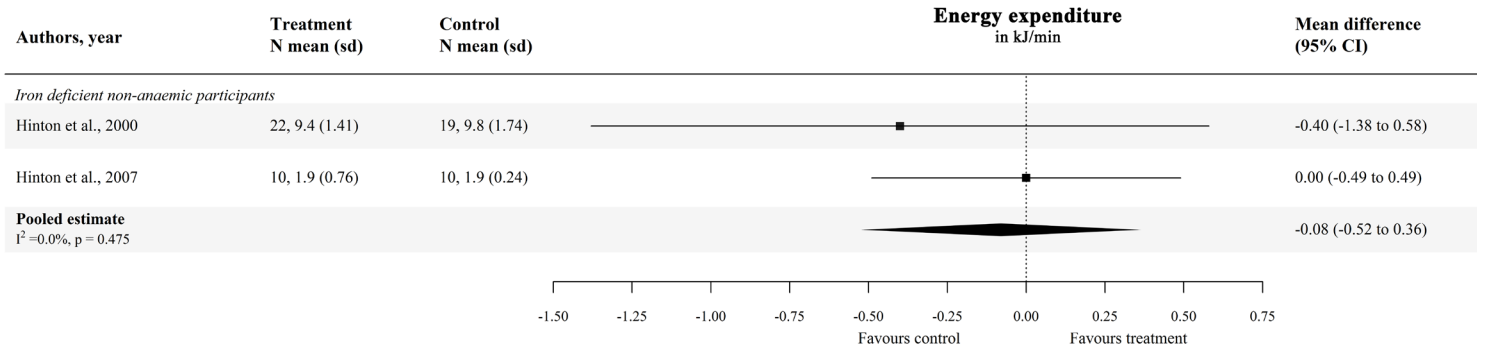
Supplementary Figure 23: Peak work rate.

Supplementary Figure 24: 6 minute walking distance



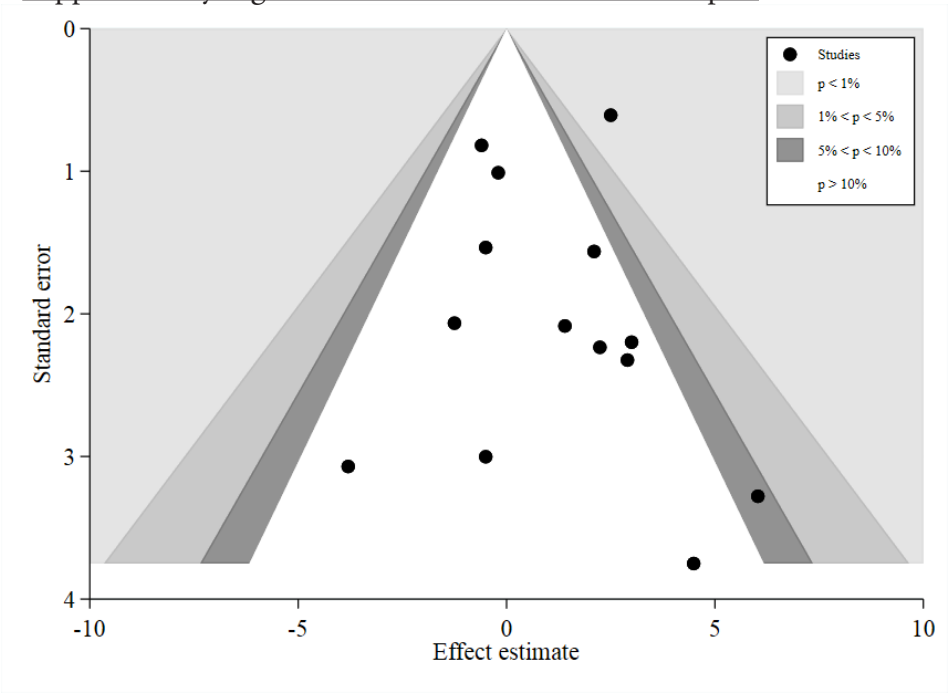
Supplementary Figure 24: 6 minute walking distance.

Supplementary Figure 25: Energy expenditure



Supplementary Figure 25: Energy expenditure.

Supplementary Figure 26: Contour enhanced funnel plot



Supplementary Figure 26: Funnel plot relative VO_2 max.

References Supplementary files Chapter 7

1.. Doherty, A. *et al.* Large Scale Population Assessment of Physical Activity Using Wrist Worn Accelerometers: The UK Biobank Study. PLoS One 12, e0169649, doi:10.1371/journal.pone.0169649 (2017).

2. van Hees, V. T. et al. Autocalibration of accelerometer data for free-living physical activity assessment using local gravity and temperature: an evaluation on four continents. J Appl Physiol (1985) 117, 738-744, doi:10.1152/japplphysiol.00421.2014 (2014).

3. van Hees, V. T. et al. Separating movement and gravity components in an acceleration signal and implications for the assessment of human daily physical activity. PLoS One 8, e61691, doi:10.1371/journal.pone.0061691 (2013).

4. Doherty, A. R. et al. Using wearable cameras to categorise type and context of accelerometer-identified episodes of physical activity. Int J Behav Nutr Phys Act 10, 22, doi:10.1186/1479-5868-10-22 (2013).

5. White, T., Westgate, K., Wareham, N. J. & Brage, S. Estimation of Physical Activity Energy Expenditure during Free-Living from Wrist Accelerometry in UK Adults. PLoS One 11, e0167472, doi:10.1371/journal.pone.0167472 (2016).

6. Boker, S. M. & McArdle, J. J. Statistical vector field analysis applied to mixed cross-sectional and longitudinal data. Exp Aging Res 21, 77-93, doi:10.1080/03610739508254269 (1995).

7. White, T. et al. Estimating energy expenditure from wrist and thigh accelerometry in free-living adults: a doubly labelled water study. Int J Obes (Lond), doi:10.1038/s41366-019-0352-x (2019).

8. McArdle, J. J. Latent variable modeling of differences and changes with longitudinal data. Annu Rev Psychol 60, 577-605, doi:10.1146/annurev.psych.60.110707.163612 (2009).

9. Ferrer, E. & JJ., M. Longitudinal modeling of developmental changes in psychological research. Curr Dir Psychol Sci 19, 149–154 (2010).

10. Burden, R. J., Morton, K., Richards, T., Whyte, G. P. & Pedlar, C. R. Is iron treatment beneficial in, iron-deficient but non-anaemic (IDNA) endurance athletes? A systematic review and meta-analysis. Br J Sports Med 49, 1389-1397, doi:10.1136/bjsports-2014-093624 (2015).

11. Houston, B. L. et al. Efficacy of iron supplementation on fatigue and physical capacity in non-anaemic iron-deficient adults: a systematic review of randomised controlled trials. BMJ Open 8, e019240, doi:10.1136/bmjopen-2017-019240 (2018).

12. Jankowska, E. A. et al. Effects of intravenous iron therapy in iron-deficient patients with systolic heart failure: a meta-analysis of randomised controlled trials. European journal of heart failure 18, 786-795, doi:10.1002/ejhf.473 (2016).

13. Johansen, K. L. et al. Systematic review and meta-analysis of exercise tolerance and physical functioning in dialysis patients treated with erythropoiesis-stimulating agents. American journal of kidney diseases : the official journal of the National Kidney Foundation 55, 535-548, doi:10.1053/j.ajkd.2009.12.018 (2010).

14. Johnson, D. M., Roberts, J. & Gordon, D. The acute effects of whole blood donation on cardiorespiratory and haematological factors in exercise: A systematic review. PLoS One 14, e0215346, doi:10.1371/journal.pone.0215346 (2019).

15. Kansagara, D. et al. Treatment of anemia in patients with heart disease: a systematic review. Ann Intern Med 159, 746-757, doi:10.7326/0003-4819-159-11-201312030-00007 (2013).

16. Low, M. S., Speedy, J., Styles, C. E., De-Regil, L. M. & Pasricha, S. R. Daily iron supplementation for improving anaemia, iron status and health in menstruating women. Cochrane Database Syst Rev 4, CD009747, doi:10.1002/14651858.CD009747.pub2 (2016).

17. Ngo, K. et al. Erythropoiesis-stimulating agents for anaemia in chronic heart failure patients. Cochrane Database Syst Rev, CD007613, doi:10.1002/14651858.CD007613.pub2 (2010).

18. Pratt, J. J. & Khan, K. S. Non-anaemic iron deficiency - a disease looking for recognition of diagnosis: a systematic review. Eur J Haematol 96, 618-628, doi:10.1111/ejh.12645 (2016).

19. Tehrani, F., Dhesi, P., Daneshvar, D., Phan, A., Rafique, A., Siegel, R.J. & Cerecek, B. Erythropoiesis Stimulating Agents in Heart Failure. Carduivasc Drugs Ther 23:511-518 (2009).

Patients with Anemia: A Meta-Analysis

20. Van Remoortel, H., De Buck, E., Compernolle, V., Deldicque, L. & Vandekerckhove, P. The effect of a standard whole blood donation on oxygen uptake and exercise capacity: a systematic review and meta-analysis. Transfusion 57, 451-462, doi:10.1111/trf.13893 (2017).

21. Zhang, S., Zhang, F., Du, M., Huang, K. & Wang, C. Efficacy and safety of iron supplementation in patients with heart failure and iron deficiency: a meta-analysis. Br J Nutr 121, 841-848, doi:10.1017/S000711451900014X (2019).

22. Zhou, X., Xu, W., Xu, Y. & Qian, Z. Iron supplementation improves cardiovascular outcomes in patients with heart failure. Am J Med, doi:10.1016/j.amjmed.2019.02.018 (2019).

23. Akiba, T. et al. Effects of recombinant human erythropoietin and exercise training on exercise capacity in hemodialysis patients. Artificial organs 19, 1262-1268 (1995).

24. Anker, S. D. et al. Ferric carboxymaltose in patients with heart failure and iron deficiency. N Engl J Med 361, 2436-2448, doi:10.1056/NEJMoa0908355 (2009).

25. Barany, P., Wibom, R., Hultman, E. & Bergstrom, J. ATP production in isolated muscle mitochondria from haemodialysis patients: effects of correction of anaemia with erythropoietin. Clin Sci (Lond) 81, 645-653 (1991).

26. Barany, P., Pettersson, E. & Konarski-Svensson, J. K. Long-term effects on quality of life in haemodialysis patients of correction of anaemia with erythropoietin. Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association 8, 426-432 (1993).

27. Birnbaum, L., Dahl, T. & Boone, T. Effect of blood donation on maximal oxygen consumption. J Sports Med Phys Fitness 46, 535-539 (2006).

28. Blee, T., Goodman, C., Dawson, B. & Stapff, A. The effect of intramuscular iron injections on serum ferritin levels and physical performance in elite netballers. J Sci Med Sport 2, 311-321 (1999).

29. Brownlie, T. t., Utermohlen, V., Hinton, P. S., Giordano, C. & Haas, J. D. Marginal iron deficiency without anemia impairs aerobic adaptation among previously untrained women. Am J Clin Nutr 75, 734-742, doi:10.1093/ajcn/75.4.734 (2002).

30. Brownlie, T. t., Utermohlen, V., Hinton, P. S. & Haas, J. D. Tissue iron deficiency without anemia impairs adaptation in endurance capacity after aerobic training in previously untrained women. Am J Clin Nutr 79, 437-443, doi:10.1093/ajcn/79.3.437 (2004).

31. Burden, R. J. et al. Effect of Intravenous Iron on Aerobic Capacity and Iron Metabolism in Elite Athletes. Med Sci Sports Exerc 47, 1399-1407, doi:10.1249/MSS.0000000000000568 (2015).

32. Burnley, M., Roberts, C. L., Thatcher, R., Doust, J. H. & Jones, A. M. Influence of blood donation on O2 uptake on-kinetics, peak O2 uptake and time to exhaustion during severe-intensity cycle exercise in humans. Exp Physiol 91, 499-509, doi:10.1113/expphysiol.2005.032805 (2006).

33. Christian, A. J. M. D. M. Does erythropoietin increase exercise capacity in anemic heart failure? Cardiology Review 20, 42-44 (2003).

34. Delano, B. G. Improvements in quality of life following treatment with r-HuEPO in anemic hemodialysis patients. American journal of kidney diseases : the official journal of the National Kidney Foundation 14, 14-18 (1989).

35. DellaValle, D. M. & Haas, J. D. Iron supplementation improves energetic efficiency in iron-depleted female rowers. Med Sci Sports Exerc 46, 1204-1215, doi:10.1249/MSS.0000000000000208 (2014).

36. Dellweg, D., Siemon, K., Mahler, F., Appelhans, P., Klauke, M., & Köhler, D. Spiroergometrische Leistungsdiagnostik vor und nach Blutspende. Pneumologie. 62, 372–377 (2008).

37. Zoladz, J. A. et al. Effect of moderate incremental exercise, performed in fed and fasted state on cardio-respiratory variables and leptin and ghrelin concentrations in young healthy men. J Physiol Pharmacol 56, 63-85 (2005).

38. Ekblom, B., Wilson, G. & Astrand, P. O. Central circulation during exercise after venesection and reinfusion of red blood cells. J Appl Physiol 40, 379-383, doi:10.1152/jappl.1976.40.3.379 (1976).

39. Foster, C., Porcari, J. P., Anderson, J., Paulson, M., Smaczny, D., Webber, H., … Udermann, B. . The talk test as a marker of exercise training intensity. Journal of Cardiopulmonary Rehabilitation and Prevention 28, 24-30 (2008).

40. Friedmann, B., Weller, E., Mairbaurl, H. & Bartsch, P. Effects of iron repletion on blood volume and performance capacity in young athletes. Med Sci Sports Exerc 33, 741-746 (2001).

41. Fritsch, J. et al. [Effect of a single blood donation on ergo-spirometrically determined cardiopulmonary performance capacity of young healthy probands]. Z Kardiol 82, 425-431 (1993).

42. Ghali, J. K. et al. Randomised double-blind trial of darbepoetin alfa in patients with symptomatic heart failure and anemia. Circulation 117, 526-535, doi:10.1161/CIRCULATIONAHA.107.698514 (2008).

43. Gordon, D. et al. Influence of blood donation on the incidence of plateau at VO2max. Eur J Appl Physiol 114, 21-27, doi:10.1007/s00421-013-2743-3 (2014).

44. Grunze, M. et al. Mechanisms of improved physical performance of chronic hemodialysis patients after erythropoietin treatment. Am J Nephrol 10 Suppl 2, 15-18; discussion 18-23, doi:10.1159/000168211 (1990).

45. Hill, D. W., Vingren, J. L. & Burdette, S. D. Effect of plasma donation and blood donation on aerobic and anaerobic responses in exhaustive, severe-intensity exercise. Appl Physiol Nutr Metab 38, 551-557, doi:10.1139/apnm-2012-0361 (2013).

46. Hinton, P. S. & Sinclair, L. M. Iron supplementation maintains ventilatory threshold and improves energetic efficiency in iron-deficient nonanemic athletes. Eur J Clin Nutr 61, 30-39, doi:10.1038/sj.ejcn.1602479 (2007).

47. Hinton, P. S., Giordano, C., Brownlie, T. & Haas, J. D. Iron supplementation improves endurance after training in iron-depleted, nonanemic women. J Appl Physiol (1985) 88, 1103-1111, doi:10.1152/jappl.2000.88.3.1103 (2000).

48. Hollmann, W., Chirdel, K., Forsberg, S. & Speer, K. [Studies on the effect of blood donation on cardiopulmonary performance]. Med Welt 20, 1158-1161 (1969).

49. Jensen, A. e. a. Iron supplementation and iron status in exercising young women. (1991).

50. Judd, T. B., Cornish, S. M., Barss, T. S., Oroz, I. & Chilibeck, P. D. Time course for recovery of peak aerobic power after blood donation. J Strength Cond Res 25, 3035-3038, doi:10.1519/JSC.0b013e3182132df7 (2011).

51. Klingshirn, L. A., Pate, R. R., Bourque, S. P., Davis, J. M. & Sargent, R. G. Effect of iron supplementation on endurance capacity in iron-depleted female runners. Med Sci Sports Exerc 24, 819-824 (1992).

52. Kourea, K. et al. Effects of darbepoetin-alpha on quality of life and emotional stress in anemic patients with chronic heart failure. Eur J Cardiovasc Prev Rehabil 15, 365-369, doi:10.1097/HJR.0b013e3282f849d0 (2008).

53. LaManca, J. J. & Haymes, E. M. Effects of iron repletion on VO2max, endurance, and blood lactate in women. Med Sci Sports Exerc 25, 1386-1392 (1993).

54. LaManca, J. Effects of Iron Supplementation on Aerobic Power, Endurance Performance, Blood Lactate, and Body Iron Stores in Women [Post doctoral thesis]. Florida State University. (1989).

55. Li, R. et al. Functional consequences of iron supplementation in iron-deficient female cotton mill workers in Beijing, China. Am J Clin Nutr 59, 908-913, doi:10.1093/

ajcn/59.4.908 (1994).

56. Lundin, A. P. Recombinant erythropoietin and chronic renal failure. Hospital practice (Office ed.) 26, 61-69 (1991).

57. Lyle, R. M. et al. Iron status in exercising women: the effect of oral iron therapy vs increased consumption of muscle foods. Am J Clin Nutr 56, 1049-1055, doi:10.1093/ajcn/56.6.1049 (1992).

58. Macdougall, I. C. et al. The treatment of renal anaemia in CAPD patients with recombinant human erythropoietin. Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association 5, 950-955, doi:10.1093/ndt/5.11.950 (1990).

59. Magazanik, A. et al. Effect of an iron supplement on body iron status and aerobic capacity of young training women. Eur J Appl Physiol Occup Physiol 62, 317-323 (1991).

60. Mayer, G., Thum, J., Cada, E. M., Stummvoll, H. K. & Graf, H. Working capacity is increased following recombinant human erythropoietin treatment. Kidney Int 34, 525-528 (1988).

61. Metra, M. et al. Improvement in exercise capacity after correction of anemia in patients with end-stage renal failure. Am J Cardiol 68, 1060-1066, doi:10.1016/0002-9149(91)90496-8 (1991).

62. Newhouse, I. J., Clement, D. B., Taunton, J. E. & McKenzie, D. C. The effects of prelatent/latent iron deficiency on physical work capacity. Med Sci Sports Exerc 21, 263-268 (1989).

63. Okonko, D. O. et al. Effect of intravenous iron sucrose on exercise tolerance in anemic and nonanemic patients with symptomatic chronic heart failure and iron deficiency FERRIC-HF: a randomised, controlled, observer-blinded trial. J Am Coll Cardiol 51, 103-112, doi:10.1016/j.jacc.2007.09.036 (2008).

64. Palazzuoli, A. et al. Erythropoietin improves anemia exercise tolerance and renal function and reduces B-type natriuretic peptide and hospitalization in patients with heart failure and anemia. American heart journal 152, 1096 e1099-1015, doi:10.1016/j.ahj.2006.08.005 (2006).

65. Parissis, J. T. et al. Effects of darbepoetin alpha on right and left ventricular systolic and diastolic function in anemic patients with chronic heart failure secondary to ischemic or idiopathic dilated cardiomyopathy. American heart journal 155, 751 e751-757, doi:10.1016/j.ahj.2008.01.016 (2008).

66. Peeling, P. et al. Effect of iron injections on aerobic-exercise performance of iron-depleted female athletes. Int J Sport Nutr Exerc Metab 17, 221-231 (2007).

67. Ponikowski, P. et al. Effect of darbepoetin alfa on exercise tolerance in anemic patients with symptomatic chronic heart failure: a randomised, double-blind, placebo-controlled trial. J Am Coll Cardiol 49, 753-762, doi:10.1016/j.jacc.2006.11.024 (2007).

68. Ponikowski, P. et al. Beneficial effects of long-term intravenous iron therapy with ferric carboxymaltose in patients with symptomatic heart failure and iron deficiencydagger. Eur Heart J 36, 657-668, doi:10.1093/eurheartj/ehu385 (2015).

69. Rajaram, S. et al. Effects of long-term moderate exercise on iron status in young women. Med Sci Sports Exerc 27, 1105-1110 (1995).

70. Robertson, H. T. et al. Recombinant erythropoietin improves exercise capacity in anemic hemodialysis patients. American journal of kidney diseases : the official journal of the National Kidney Foundation 15, 325-332 (1990).

71. Rosenlof, K. et al. Beneficial effects of erythropoietin on haematological parameters, aerobic capacity, and body fluid composition in patients on haemodialysis. J Intern Med 226, 311-317 (1989).

72. Schoene, R. B. et al. Iron repletion decreases maximal exercise lactate concentrations in female athletes with minimal iron-deficiency anemia. The Journal of laboratory and clinical medicine 102, 306-312 (1983).

73. Taniguchi al., Improvement in iron deficiency anemiathrough therapy with ferric ammonium citrate and vitaminC and the effects of aerobic exercise. (1991).

74. Toblli, J. E., Lombrana, A., Duarte, P. & Di Gennaro, F. Intravenous iron reduces NT-pro-brain natriuretic peptide in anemic

patients with chronic heart failure and renal insufficiency. J Am Coll Cardiol 50, 1657-1665, doi:10.1016/j.jacc.2007.07.029 (2007).

75. Tsutsui, M. Renewed cardiovascular dynamics induced by recombinant erythropoietin administration.

76. Yoshida, T. e. a. Dietary iron supplement during severe physical training in competitive female distance runners. (1990).

77. Zhu, Y. I. & Haas, J. D. Altered metabolic response of iron-depleted nonanemic women during a 15-km time trial. J Appl Physiol (1985) 84, 1768-1775, doi:10.1152/jappl.1998.84.5.1768 (1998).

78. Shea, B. J. et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 358, j4008, doi:10.1136/bmj.j4008 (2017).

79. Mancini, D. M. & Kunavarapu, C. Effect of erythropoietin on exercise capacity in anemic patients with advanced heart failure. Kidney international. Supplement, S48-52 (2003).

80. Magazanik, A. et al. Effect of an iron supplement on body iron status and aerobic capacity of young training women. Eur J Appl Physiol Occup Physiol 62, 317-323 (1991).

81. Mancini, D. M. et al. Effect of erythropoietin on exercise capacity in patients with moderate to severe chronic heart failure. Circulation 107, 294-299 (2003).

82. Radjen, S. et al. Effect of iron supplementation on maximal oxygen uptake in female athletes. Vojnosanit Pregl 68, 130-135 (2011).

83. Walsh, R. M. Effects of iron supplementation on iron status of young female swimmers during the pre-season phase of competition.