Non-local muscle fatigue effects on muscle strength, power, and endurance in healthy individuals: A systematic review and meta-analysis

David G Behm1, Shahab Alizadeh1, Saman Hadjizedah Anvar1,2, Courtney Hanlon1, Emma Ramsay1, Mohamed Mamdouh Ibrahim Mahmoud1, Joseph Whitten1, James P. Fisher3, Olaf Prieske4, Helmi Chaabene5, Urs Granacher5, James Steele3,6

1. School of Human Kinetics and Recreation, Memorial University of Newfoundland, St. John’s, Newfoundland and Labrador, Canada
2. University of Tehran
3. School of Sport, Health and Social Science, Solent University, Southampton UK
4. Division of Exercise and Movement, University of Applied Sciences for Sport and Management Potsdam, Potsdam, Germany
5. Division of Training and Movement Science, University of Potsdam, Potsdam, Germany
6. ukactive Research Institute, London, UK

Please cite as: Behm, D.G., Alizadeh, S., Hadjizedah, S., Hanlon, C., Ramsay, E., Mahmoud, M.M.I., ... Steele, J. (2020, September 28). Non-local muscle fatigue effects on muscle strength, power, and endurance in healthy individuals: A systematic review and meta-analysis. https://doi.org/10.31236/osf.io/w7x3v

ABSTRACT

Objective: To examine whether non-local muscle fatigue occurs following performance of a fatiguing bout of exercise of a different muscle(s).

Design: Systematic review and meta-analysis.

Search and Inclusion: A systematic literature search using a Boolean search strategy was conducted with PubMed, SPORTDiscus, Web of Science, and Google Scholar in April 2020 and was supplemented with additional ‘snowballing’ searches up to September 2020. To be included

All authors have read and approved this version of the manuscript. This article was last modified on Sept 28, 2020.
in our analysis, studies had to include at least one intentional performance measure (i.e., strength, endurance, or power), which if reduced could be considered evidence of muscle fatigue, and also had to include the implementation of a fatiguing protocol to a location (i.e., limb or limbs) that differed to those for which performance was measured. We excluded studies that measured only mechanistic variables such as electromyographic, or spinal/supraspinal excitability. After search and screening, 52 studies were eligible for inclusion including 57 groups of participants (median sample = 11) and a total of 303 participants.

**Results:** The main multilevel meta-analysis model including all effects sizes (278 across 50 clusters [median = 4, range = 1 to 18 effects per cluster) revealed a trivial point estimate with high precision for the interval estimate (-0.02 [95%CIs = -0.14 to 0.09]), yet with substantial heterogeneity ($Q_{(277)} = 642.3, p < 0.01, I^2 = 67.4\%$). Subgroup and meta-regression analyses showed that NLMF effects were not moderated by study design (between vs. within-participant), homologous vs. heterologous effects, upper or lower body effects, participant training status, sex, age, the time of post-fatigue protocol measurement, or the severity of the fatigue protocol. However, there did appear to be an effect of type of outcome measure where both strength (0.11 [95%CIs = 0.01 to 0.21]) and power outcomes had trivial effects (-0.01 [95%CIs = -0.24 to 0.22]), whereas endurance outcomes showed moderate albeit imprecise effects (-0.54 [95%CIs = -0.95 to -0.14]).

**Conclusions:** Overall, the findings do not support the existence of a general NLMF effect; however, when examining specific types of performance outcomes there may be an effect specifically upon endurance-based outcomes (i.e., time to task failure). However, there are relatively fewer studies that have examined endurance effects or mechanisms explaining this possible effect, in addition to fewer studies including women or younger and older participants, and considering causal effects of prior training history through the use of longitudinal intervention study designs. Thus, it seems pertinent that future research on NLMF effects should be redirected towards these still relatively unexplored areas.

**1.0 Introduction**

From the inception of the Harvard Fatigue laboratory in 1927 to the present period, muscle fatigue has had a long history of experimental research [1]. Talbott [2] initially characterised two types of fatigue: ‘physical’ fatigue due to physical tasks and ‘functional’ fatigue due to mental tasks. For approximately 90 years, fatigue-related research has investigated the transient decrease in the exercised muscles’ capacity to produce force or power [3]. The mechanisms underlying reductions in muscular capacity to meet task demands have been broadly divided into peripheral
Typically, an exercised muscle when subsequently examined shows evidence of fatigue (i.e. a reduction in capacity, often operationalised through measurements of maximal voluntary force). However, it has been shown that force and neuromuscular activation deficits may be incurred in a non-exercised muscle following fatigue of a contralateral muscle. This has been interpreted as strong evidence of a central neural impact upon fatigue, although other global factors such as metabolite dispersion and psychological factors are also thought to contribute [7]. At the start of the new millennium, researchers began to investigate ‘crossover’ fatigue, defined as changes in the performance output of a contralateral muscle following unilateral muscle fatigue [8-10]. While crossover fatigue primarily examined contralateral homologous muscles, it has been argued that global mechanisms effects upon fatigue are unlikely to be limited to just contralateral homologous muscles [11-15]. Hence, investigators began to examine heterologous muscles, incorporating the broader term of ‘non-local’ muscle fatigue (NLMF) to indicate a temporary deficit in performance of non-exercised homologous and/or heterologous muscle groups that could be located contralateral, or ipsilateral, as well as inferior or superior to the fatigued muscle groups [7, 14-17].

Halperin et al. [7] most recently reviewed this area and identified several trends and inconsistencies in the NLMF literature. In particular, few NLMF studies investigated heterologous muscles preventing strong comparisons of the responses of homologous and heterologous muscles. A comparison of homologous and heterologous muscle responses could provide insights into the interconnectivity of the neuromuscular sensory and motor systems and further mechanisms underlying NLMF. Some of the conclusions Halperin et al. [7] drew in their review included: 1) NLMF effects were more predominant when the lower body (primarily the quadriceps) rather than the upper body was tested, 2) NLMF effects appeared greater when performance was examined in prolonged continuous and intermittent muscle action tasks (i.e., endurance) compared to single, short-duration (i.e., 3-5 s) maximal voluntary contractions (MVC; i.e., muscle strength or power), and 3) sex and training background of the participants could also influence the responses. Within the last five years since the Halperin et al. [7] review, there have been more than 20 articles published and thus an updated review would help better understand this area. Further, except for Miller et al. [18], who only focused on a small number of studies (k = 6) on heterologous effects, no prior reviews have conducted a quantitative synthesis using meta-
analysis to examine the existence and overall magnitude of overall NLMF effects, nor the moderating effects of the variables previously noted as impacting it. Thus, the objective of this systematic review with meta-analysis was to explore NLMF effects through studies examining non-local performance outcomes (including muscular strength, endurance, and power measures) in response to fatiguing protocols, as well as to investigate moderating variables; namely, homologous vs. heterologous and upper vs. lower body non-exercised tested muscles, participant training status, sex, and age, in addition to aspects of the study designs (between vs. within participant)\(^1\) and methods employed (time of post-fatigue protocol that performance measurements were taken, and the severity of the fatigue protocol used)\(^2\).

### 2.0 Methods

#### 2.1 Search strategy

A literature search following PRISMA-P review guidelines was performed by six of the co-authors separately and independently using PubMed, SPORTDiscus, Web of Science, and Google Scholar databases. The topic was systematically searched in April 2020 using a Boolean search strategy with the operator “AND” and a combination of the following title keywords: crossover, cross-over, non-local, nonlocal, contralateral, ipsilateral, homologous, heterologous AND fatigue. Based upon our knowledge of the area we also contributed additional studies which we had knowledge of but where not picked up in systematic searches; further, we conducted searches of our personal computer databases for related articles, and conducted additional ‘snowballing’ searches\(^{20}\) throughout the process of conducting the review and analysis, which located some newer studies not available when we conducted the initial systematic search. The search ended in September 2020. The search strategy aimed to locate all relevant articles that had examined NLMF whether as a primary, or secondary, aim of the study. For the sake of brevity, muscular strength, power, and endurance will generally be referred to without the descriptor “muscular” throughout the manuscript, and refers to where the output was measured.

#### 2.2 Inclusion and exclusion criteria (study selection)

We followed PICOS (Population, Intervention, Comparison, Outcome, Study Design) for selecting studies for inclusion. We included studies in any healthy (i.e. non-clinical) population without age limits and including both recreationally active or trained participants (including athletes). Studies also had to include the implementation of an intervention involving a fatiguing

\(^1\) Within participant designs are typically more statistically powerful [see 19].

\(^2\) It seemed reasonable to think that performance measurements taken soon after a non-local fatiguing protocol would exhibit greater deficits, and similarly that more severe fatigue protocols would elicit greater deficits.

DOI: 10.31236/osf.io/w7x3v

SportRxiv is free to access, but not to run. Please consider donating at www.storkinesiology.org/annual
protocol to a location (i.e., limb or limbs) that differed to those for which performance was measured. The comparison conditions included either pre- and post-, or just post-fatiguing protocol from non-fatiguing protocol control conditions, or just pre-fatiguing protocols for studies without control conditions. To be included in our analysis, studies had to have included at least one intentional performance measure as an outcome (i.e., a measure of strength, endurance, or ‘power’\(^3\)) and if the outcome was reduced, it could be considered as evidence of muscle fatigue. This included measures indicative of a voluntary reduction in the ability of a muscle to produce force or power (e.g., an MVC, or countermovement jump [CMJ]; [3]); however, studies that examined time to task failure (TTF) exercises at a constant level (e.g., cycling with a given resistance, or repetitions in an exercise performed with a given load) or a fatigue index were also included. Both between and within-participant study designs were included. We excluded studies that measured only mechanistic variables such as electromyographic, or spinal/supraspinal excitability. Further, we excluded studies where we were unable to obtain absolute scores for performance measures either as summary statistics or raw data (i.e., those reporting relative fatigue; see below). Related articles were included up to September 2020. To limit potential publication bias effects, we did not limit ourselves to only including peer review articles and, where they were identified and it was possible to either extract data or contact the authors to gain access to data, we also included studies presented as part of conference proceedings and unpublished studies (e.g., pre-prints and MSc/Ph.D. theses).

2.3 Data extraction

A data extraction table was prepared to map: (a) author and year of publication; (b) study design (within or between participant); (c) sample mean age; (d) percentage of males in sample; (e) whether participants were trained or just healthy/recreationally active; (f) whether homologous or heterologous effects were examined; (g) which muscle/muscle groups were subject to the fatiguing protocol; (h) which limbs were subjected to the fatiguing protocols (e.g., dominant/non-dominant, bilateral); (i) characteristics of the fatiguing protocol (i.e., the relative (i.e., 0-100% of max) and/or absolute loads, the time under loading, the number of sets/bouts performed); (j) whether the fatiguing task involved maximum effort (i.e., MVC or resulted in task failure); (k) which muscle/muscle groups were tested after the fatiguing protocol; (l) the outcome measure used to

---

\(^3\) The term ‘power’ is commonly misused in sport, exercise, and physical activity literature considering its specific definition in Newtonian mechanics; instead it would be more accurate to speak of the ability generate impulse in most applications where it is used [see 21]. Indeed, power is defined as the rate of performing work. However, with respect to muscular performance, ‘power’ is often thought of as the product of force and velocity which as Winter et al. [21] note in fact refers to the impulse as derived from the impulse-momentum relationship. However, given its wide colloquial use to refer to such applications, we continue to use the term ‘power’ here for ease of communication given the primary intention of this article is not to debate terminology.
measure performance; (m) the time after the fatiguing protocol that the measurement was taken; (n) the means and standard deviations for outcome measures including both control and fatigue conditions (either pre- and post-, or just post-fatiguing protocol), or pre- and post-fatiguing protocols for studies without control conditions; and (o) sample sizes. Of note, we included data for all performance outcomes reported in studies and for all time points for which they were measured; thus, if a study reported multiple performance measures these were all included, and if a study reported measures taken more than once post-fatiguing protocol these were all included (the dependency between effect sizes was handled using a multi-level robust variance estimation approach – see statistical analysis). Some studies only reported relative outcomes (i.e., performance as a percentage normalised to a pre-fatiguing protocol measurement), and further, some studies only reported data graphically, did not report outcomes in a manner conducive to extraction for our analysis, or despite investigating NLMF effects did not report outcomes relevant for our analysis as they were secondary in those studies. In all of these cases, authors were contacted for either appropriate summary statistics or raw data for absolute measures of performance (if collected) to facilitate inclusion in our analysis. Some authors were unable to share data due to no longer having access to it. We followed up with authors a maximum of twice if we received no response. For those studies reporting only graphical data and that we did not receive responses for, WebPlotDigitizer (v4.3, Ankit Rohatgi; https://apps.automeris.io/wpd/) was used to extract data for inclusion in our analysis.

2.4 Synthesis and analysis

The meta-analysis was performed using the ‘metafor’ [22] package in R (v 4.0.2; R Core Team, https://www.r-project.org/). All analysis code utilised is presented in the supplementary materials (https://osf.io/6x4ct/).

Studies were grouped by their design for appropriate calculation of standardised effects sizes using the ‘escalc’ function in ‘metafor’. For between-participant study designs (i.e., fatigue and control conditions conducted on separate groups of participants) with post-fatigue protocol measures, only the standardised mean difference was calculated between groups. For between-participant study designs with both pre- and post-fatigue protocol measures, the standardised mean change using the raw values was calculated within each condition using the post-fatigue standard deviations for each condition, and then the difference between these (i.e., the standardised mean change for fatigue conditions minus the standardised mean change for control conditions) was calculated in addition to the variance (i.e., the sum of variances from both conditions). Similarly, for within-participant study designs (i.e., fatigue and control conditions
conducted on the same group of participants) with post-fatigue protocol measures, only the standardised mean change using the raw values was calculated between the conditions. For within-participant study designs with both pre- and post-fatigue protocol measures the standardised mean change using the raw values was calculated within each condition using the post-fatigue standard deviations for each condition, and then the difference between these was calculated in addition to the variance. Lastly, for within-participant study designs with only a single condition (i.e., fatigue) and pre- and post-fatigue protocol measures the standardised mean change using the raw values within that condition was calculated. The magnitude of standardised effect sizes was interpreted with reference to Cohen’s [23] thresholds: trivial (<0.2), small (0.2 to <0.5), moderate (0.5 to <0.8), and large (>0.8). Standardised effects were calculated in such a manner that negative effect size values indicated the presence of fatigue (i.e., a drop in the performance measure post-fatiguing protocol), whereas positive effect size values indicated an improvement in performance.

Because of the nested structure of the effect sizes calculated from the studies included (i.e. effects nested within groups nested within studies), multilevel mixed-effects meta-analyses with both study and intra-study groups were included as random effects in the model were performed to explore the effect of fatiguing protocols on performance measures (i.e., whether or not they induced NLMF). Cluster robust point estimates and precision of those estimates using 95% compatibility (confidence) intervals (CIs) were produced [24], weighted by inverse sampling variance to account for the within- and between-study variance (tau-squared). Restricted maximal likelihood estimation was used in all models. A main model was produced including all effects sizes.

Several exploratory subgroup comparisons and meta-regression analyses of moderator’s analyses were conducted. Subgroup analyses included comparison of study designs (between participant vs. within-participant designs), homologous vs. heterologous effects, upper vs. lower body effects, type of outcome measure (grouped as ‘strength’, ‘endurance’, or ‘power’), and trained vs. recreationally active participants. For each comparison, multilevel models with robust estimates were produced for each subgroup, and fixed-effects with moderator’s model used to compare the models. Moderators examined using meta-regression included the time post-fatigue protocol that measures were taken (seconds; though this was limited to effects measured within a 24 hours maximum time frame), the time post-fatigue protocol that measures were taken but limited to only the first measure taken (to separate the possible impact of multiple post-fatigue protocol tests), the ‘severity’ of the fatiguing protocol used (this was calculated from those studies where data was available as the ‘relative load’ x ‘sets/bouts’ x ‘time under load’ to give a measure
of the total ‘workload’ performed)\(^4\) [25], the percentage males in the sample, and mean age of the sample.

For all models, we opted to avoid dichotomizing the existence of an effect for the main results and therefore did not employ traditional null hypothesis significance testing, which has been extensively criticised [26, 27]. Instead, we consider the implications of all results compatible with these data, from the lower limit to the upper limit of the interval estimates, with the greatest interpretive emphasis placed on the point estimate.

The risk of small study bias was examined visually through contour-enhanced funnel plots. \(Q\) and \(I^2\) statistics were also produced and reported [28]. A significant \(Q\) statistic is typically considered indicative of effects likely not being drawn from a common population. \(I^2\) values indicate the degree of heterogeneity in the effects: 0-40% were not important, 30-60% moderate heterogeneity, 50-90% substantial heterogeneity, and 75-100% considerable heterogeneity [29].

For within-participant effects, pre-post correlations for measures were rarely reported; thus, we assumed a range of values for correlation coefficients (\(r = 0.5, 0.7, \text{ and } 0.9\)) and explored the sensitivity of results to each of these. As overall findings were relatively insensitive to this range, we report the results for \(r = 0.7\) here and include the results for inclusion of the other assumed correlation coefficients in the supplementary materials (https://osf.io/6n283/ & https://osf.io/88bj/). Heterogeneity of individual effect estimates included was also examined through production of ‘GOSH’ plots\(^5\) (graphical display of study heterogeneity; [30].

### 3.0 Results

#### 3.1 Included studies

After initial searches and screening, 39 studies were identified that met the inclusion criteria. Additional search approaches identified a further 13 studies that met inclusion criteria. Thus, the final number of studies included was 52 [31-53, 118-129: citations for studies not otherwise referenced in the manuscript but included in the data extraction table]. Details of the search and inclusion process are shown in the flow chart (figure 1). The pooled number of participants in the studies included was 303 across 57 groups within studies and with sample sizes ranging from 6 to 45 participants (median = 11). Full details of all included studies can be seen in the data extraction table (https://osf.io/4g5jw/).

\(^4\) This in essence was similar to the manner in which resistance training dose is often operationalised as ‘volume-load’ relative to one-repetition maximum (1RM) i.e. sets x repetitions x %1RM [25].

\(^5\) Given that the number of possible combinations (\(2.1469889656237000007512297816 \times 10^{27}\)) based on the number of effect sizes extracted prohibited the computation of all possible overall effect estimates, we limited this to a random sample of 1000000 possible combinations.
Figure 1. Flow chart illustrating different phases of the search and study selection.

3.2 Main model – All effects

The main model including all effects sizes (278 across 50 clusters [median = 4, range = 1 to 18 effects per cluster) revealed a trivial point estimate with high precision for the interval estimate (-0.02 [95%CIs = -0.14 to 0.09]), yet with substantial heterogeneity ($Q_{277} = 642.3$, $p < 0.01$), $I^2 = 67.4\%$). Figure 2 presents all effect sizes and interval estimates in an ordered caterpillar plot, figure 3 presents the funnel plot for all studies, and the GOSH plot, which showed generally high levels of between-study heterogeneity, can be viewed in the supplementary materials (https://osf.io/sbmze/).
Figure 2. Ordered caterpillar plot of all effects.
3.3 Subgroup and meta-regression analyses

3.4 Between-participant vs. within-participant study design

Subgroup models revealed trivial yet opposite sign effects for both between participant study designs (0.17 [95%CIs = 0.01 to 0.34]; 21 effects across 2 clusters [median = 10.5, range = 5 to 16 effects per cluster; $I^2 = 0.0\%$]) and for within participant designs (-0.04 [95%CIs = -0.15 to 0.08]; 257 effects across 48 clusters [median = 4, range = 1 to 18 effects per cluster; $I^2 = 69.6\%$]), which significantly differed ($z = -3.43, p < 0.001$).

3.5 Homologous vs. heterologous effects

Subgroup models revealed trivial yet opposite sign effects for both homologous (0.06 [95%CIs = -0.09 to 0.22]; 131 effects across 27 clusters [median = 3, range = 1 to 18 effects per cluster; $I^2 = 59.1\%$]) and heterologous effects (-0.09 [95%CIs = -0.26 to 0.07]; 147 effects across 28
clusters [median = 3.5, range = 1 to 16 effects per cluster; $I^2 = 74.1\%$], which did not significantly differ ($z = 1.43, p = 0.152$).

3.6 Upper body vs. lower body effects

Subgroup models revealed trivial yet opposite sign effects for both upper body (0.07 [95%CIs = -0.07 to 0.22]; 110 effects across 22 clusters [median = 4, range = 1 to 12 effects per cluster; $I^2 = 58.9\%$) and lower body effects (-0.11 [95%CIs = -0.27 to 0.04]; 165 effects across 37 clusters [median = 3, range = 1 to 18 effects per cluster; $I^2 = 73.9\%$], which did not significantly differ ($z = 1.813, p = 0.07$).

3.7 Type of outcome measure

Subgroup models revealed trivial effects for both strength (0.11 [95%CIs = 0.01 to 0.21]; 174 effects across 35 clusters [median = 4, range = 1 to 18 effects per cluster; $I^2 = 39.5\%$) and power outcomes (-0.01 [95%CIs = -0.24 to 0.22]; 43 effects across 8 clusters [median = 4, range = 2 to 12 effects per cluster; $I^2 = 45.9\%$), yet moderate effects for endurance outcomes (-0.54 [95%CIs = -0.95 to -0.14]; 61 effects across 19 clusters [median = 2, range = 1 to 16 effects per cluster; $I^2 = 94.5\%$]) which differed significantly from both strength ($z = 3.29, p = 0.001$) and power outcomes ($z = 2.49, p = 0.013$).

3.8 Participants’ training status

Subgroup models revealed trivial effects for both trained (-0.01 [95%CIs = -0.16 to 0.13]; 121 effects across 19 clusters [median = 3, range = 1 to 18 effects per cluster; $I^2 = 50.8\%$) and for recreationally active participants (-0.04 [95%CIs = -0.21 to 0.13]; 157 effects across 31 clusters [median = 4, range = 1 to 18 effects per cluster; $I^2 = 75.6\%$], which did not significantly differ ($z = 0.289, p = 0.772$).

3.9 Time of post-fatigue protocol measurement

Meta-regression suggested that NLMF effects were not moderated by the time that post-fatigue protocol measurements were taken whether looking at all effects within a 24 hour period post-fatigue protocol ($\beta = -0.00 [95\%\text{CIs} = -0.0001 to 0.00]$; 265 effects across 50 clusters [median = 4, range = 1 to 18 effects per cluster), or when limited to only the first post-fatigue protocol measurement ($\beta = -0.0006 [95\%\text{CIs} = -0.0016 to 0.0005]$; 129 effects across 47 clusters [median = 2, range = 1 to 10 effects per cluster). Figures in the supplementary materials ([https://osf.io/gd9cm/](https://osf.io/gd9cm/) and [https://osf.io/3jhb/](https://osf.io/3jhb/)) show meta-analytic scatterplots for these effects.
3.10 ‘Severity’ of fatigue-protocol

Meta-regression suggested that NLMF effects were not moderated by severity of the fatigue protocol used ($\beta = -0.00$ [95%CI = -0.00 to 0.00]; 251 effects across 39 clusters [median = 5, range = 1 to 18 effects per cluster]. The figure showing the meta-analytic scatterplots for these effects can be found in the supplementary materials (https://osf.io/ajngp/).

3.11 Percentage of males in sample

Meta-regression suggested that NLMF effects were not moderated by the percentage of males in the sample ($\beta = -0.0021$ [95%CI = -0.0001 to 0.0042]; 268 effects across 47 clusters [median = 4, range = 1 to 18 effects per cluster]. The figure showing the meta-analytic scatterplots for these effects can be found in the supplementary materials (https://osf.io/rdc45/).

3.12 Mean age of sample

Meta-regression suggested that NLMF effects were not moderated by the mean age of participants in the sample ($\beta = -0.0008$ [95%CI = -0.0078 to 0.0062]; 278 effects across 47 clusters [median = 4, range = 1 to 18 effects per cluster]. The figure showing the meta-analytic scatterplots for these effects can be found in the supplementary materials (https://osf.io/7fhy8/).

4.0 Discussion

The major findings of this meta-analysis were that when all performance measures indicative of neuromuscular fatigue were examined (muscle strength, power, and endurance measures combined), the effect estimate was close to zero and at best compatible with only trivial effects. Thus, there does not appear to be clear evidence of a general NLMF effect, though between-study heterogeneity was high even across various iterations of effect sizes included (see GOSH plot https://osf.io/sbmze/). Further, NLMF effects were similarly absent and not moderated by study design (between vs. within-participant), homologous vs. heterologous effects, upper or lower body effects, participant training status, sex, age, the time of post-fatigue protocol measurement, or the severity of the fatigue protocol. However, when examining types of performance measures separately, there did appear to be some evidence for a moderate NLMF effect upon endurance-based outcomes (i.e. TTF); though the interval estimate was relatively imprecise showing compatibility with a range from possibly trivial to large effects and again there was considerable between-study heterogeneity ($I^2 = 94.5\%$).

Prior fatiguing actions of another muscle induced trivial effects upon subsequent single discrete maximal strength and power contractions of the tested muscles. The finding of trivial
effects in our meta-analytic model does not imply that all studies demonstrated trivial NLMF results with single testing contractions. There are some studies that have reported NLMF with single discrete maximal contractions [8, 13, 15, 54, 55]. However, this sample of studies illustrating single maximal contraction NLMF impairments is counterbalanced by an even greater number of studies showing no significant changes; thus, the multilevel meta-analysis models estimates suggest that the mean of the distribution of the true effect sizes is approximately zero and at best only trivial.

In contrast, NLMF produced moderate magnitude impairments of a previously non-exercised, non-local muscle when tested with an endurance test. In addition, we conducted a post-hoc exploratory analysis of longer duration (i.e. where the endurance tests involved bouts to task failure that lasted >75-s) endurance outcomes as these have previously been considered as ‘true’ endurance by others [56, 57]. These longer duration endurance outcomes demonstrated a large, albeit still imprecise and ranging from small to large, effect estimate (-0.85 [95%CI = -1.46 to -0.24]), which still showed a high between-study heterogeneity ($I^2 = 89.3\%$). It should be noted that there are comparatively fewer studies having examined endurance-based outcomes, which likely explains the imprecision of this estimate; further research is therefore encouraged. However, the results from this meta-analysis are in accord with the conclusion of Halperin et al. [7] in their narrative review, where they also reported that muscle endurance-based outcomes provided clearer evidence of NLMF vs. strength or power-based outcomes. They speculate on a range of possible mechanisms for this effect which we briefly discuss here in the context of our findings.

Accumulation of metabolites and increased acidosis in the working (exercised) muscle(s) may be distributed globally by the cardiovascular system having non-local effects on muscle disrupting contractile kinetics, enzymatic functioning, and action potential propagation [58-61]. Increases in potassium [62, 63], hydrogen [62, 64], and blood lactate [14, 62, 64, 65] have been observed in the non-exercised muscles following contralateral exercise. While the impact of blood lactate and hydrogen ions on muscle fatigue is disputed [66, 67], they have been reported to decrease the force per cross-bridge [68, 69], and reduce myofibrillar Ca\(^{2+}\) sensitivity [68, 70]. Increased potassium gradients with repeated contractions [71], can diminish force [72] and reduce muscle excitability contributing to fatigue [63]. Exercise-induced heat shock proteins [73] are evident in non-exercised systems [74] and are reported to impact the recovery of force-producing capacity [75]. Fatigue-induced changes to the metabolic environment can also activate group III and IV muscle afferents [76-78] inhibiting the central nervous system, attenuating non-local, or global muscle performance [55]. However, the lack of effects upon other outcomes (strength and power) question the influence of factors such as group III and IV afferents effects on central motor
drive. Further questioning the impact of centrally mediated neural mechanisms, studies that have examined electromyographic activity (EMG) of the non-exercised muscle are conflicting in whether there appears to be evidence of neural inhibition [8, 10, 15, 79-81]. While there is evidence suggestive of reduced cerebral oxygenation possibly impacting NLMF [82, 83], evidence from studies of transcranial magnetic stimulation (TMS) regarding specific levels neural influence (i.e. supraspinal [cortical or cerebral], spinal, and peripheral) is varying; studies have reported inhibitory [84-88], excitatory [89-93] and no significant NLMF effects upon corticospinal excitability [94, 95].

The lack of general NLMF effects, or comparative differences in both homologous and heterologous, and upper and lower body, outcomes question the role of systemic metabolic mechanisms. Further, findings regarding possible neural mechanisms are highly variable and, accompanied by the general lack of NLMF effect upon strength and power outcomes (which suggests that a general reduction in force-producing capacity does not explain reduced endurance performance), it appears they may play a little role; a similar conclusion to that reached with a meta-analysis by Miller et al. [18], who reported trivial reductions in non-local heterologous spinal and supraspinal excitability. Thus, it is not clear whether physiological mechanisms might explain the possible NLMF effects upon endurance performance. However, a further possible explanation that has been speculated upon is psychological.

Maintaining muscular actions to the point of task failure is uncomfortable, sometimes painful, and necessitates focus and concentration (cognitive demand) to maintain task demands in both traditional aerobic endurance activities [96] and resistance exercise tasks [97]. Indeed, even cognitively fatiguing tasks performed alone have been suggested to impede subsequent physical performance, especially with endurance-based tasks [98-100]; though this effect has also been questioned [96]. Nevertheless, a deficit in cognitive capacity may influence phenomenological experience more globally and potentially explain NLMF especially during endurance-based tasks [7]. Cognitively fatiguing tasks may lead individuals to perceive a subsequent task to be more effortful, resulting in an earlier cessation of the activity [98-100]; though again this finding is not consistent and may be small [101]. Actual fatigue (and indeed effort) is distinct from the perception of fatigue (and perception of effort) [102]. Steele [102] has most recently defined perception of effort in relation to perception of fatigue as “…the perception of that which must be done in attempting to achieve a particular demand, or set of demands, and which is determined by the perception of current task demands relative to the perception of capacity to meet those demands…”. Given the separation of actual- and perception of-, and the link between perception of capacity (which is synonymous with, albeit the opposite sign of, perception of fatigue) and perception of effort, the NLMF effects seen for endurance tasks may result from prior fatiguing non-local tasks

DOI: 10.31236/osf.io/w7x3v

SportRxiv is free to access, but not to run. Please consider donating at www.storkinesiology.org/annual

15
impacting perception of fatigue, and thus perception of effort experienced during subsequent task performance, as opposed to any ‘direct’ physiological mechanism. Indeed, Greenhouse-Tucknott et al. [103] show some evidence for this; they found that no actual NLMF fatigue (operationalised as knee extensor neuromuscular function), yet prior handgrip exercise increased perception of fatigue and subsequently perception of effort (and affect) impacting upon endurance performance. Thus, increased perception of fatigue may have global repercussions increasing perception of effort during subsequent task performance. However, given the varying evidence for performance deficits from prior cognitive tasks inducing fatigue [57, 104], further studies should examine NLMF effects in endurance tasks and examine these possible mechanisms using appropriate approaches to mediation analysis.

Biomechanical alterations have also been discussed as possible factors influencing the apparent presence of NLMF [7]; that muscle groups not ‘directly’ involved in the fatiguing task still contribute to the performance of working muscle groups using factors such as stabilization and mechanical energy transference and thus have the potential to influence NLMF effects [105-108]. For example, a 20% reduction of power production during 30 seconds of lower body cycling was observed when the lower body cycling was performed without gripping the handlebar for stability [108]; thus, a reduction in grip strength could contribute to reduced performance through this mechanism. Similarly, high levels of activation in the trunk muscles (abdominal and lower back) have been reported during upper [109] and lower body [105] movements due to the trunk muscles acting in a stabilizing role to allow for efficient proximal to distal transfer of mechanical energy (i.e. kinetic chain) [110]. Thus, fatigue in these muscles might impact subsequent tasks and would seem unlikely to differentially impact homologous over heterologous muscle contractions. However, it seems likely that such biomechanical alterations would be highly specific to the exact task performed in both the fatiguing protocol and subsequently for performance testing. Thus, it may be fruitful for future studies to more specifically focus on circumstances where NLMF may appear to be present and whether or not this is explained by biomechanical alterations not typically examined.

A review of the mechanisms that might explain a possible NLMF effect in endurance tasks suggests that neither metabolic nor central neuromuscular activation deficits, are likely to be responsible. Possible mechanisms however include a global impact upon perception of fatigue and increased perception of effort during subsequent task performance, in addition to possible biomechanical alterations induced by fatigue in other muscle groups not ‘directly’ involved in the fatiguing task. However, further research is needed to establish the extent of an NLMF effect upon endurance tasks, and whether these mechanisms mediate this.
The lack of moderating effects from other studies, participant, and protocol characteristics is interesting and suggests that considering the consistently high between-study heterogeneity, other characteristics may explain the variation in effects seen across individual studies. Halperin et al. [7] previously suggested that NLMF effects may be more pronounced in the lower body, for male participants, and that training background may have some influence; indeed these characteristics have been less well explored and so the lack of moderating effects may merely be due to the relative lack of data.

The lack of impact of sex in the present meta-analysis may be related to the much lower proportion of women participating in these studies. The possibility of sex differences is difficult to identify as only 13 studies reported including female participants and only three of these directly compared men and women. Martin and Rattey [8] and Ye et al., [50] were two of only three studies to directly compare sex-related NLMF effects, finding greater NLMF with males vs. females. Doix et al. [80] contrastingly discovered that males experienced greater fatigue in the exercised leg but there was no significant NLMF with either sex, although vastus lateralis EMG activity demonstrated a greater magnitude deficit during the post-test with females. Females tend to exhibit greater muscle endurance [111] and indeed experience less local muscular fatigue [112], which has been attributed to lower absolute muscle forces with the same relative work, contributing to a lower muscle oxygen demand and vasculature compression [113]. In contrast, males are more reliant on glycolytic pathways and present a greater neuromuscular activation impairment after fatigue [111, 113]. If the aforementioned metabolite dispersion impacts NLMF, the higher vascular compression and glycolytic-based metabolite release might promote the greater male NLMF effects. Nonetheless, with such few female participants and comparative studies, future studies must include and compare both sexes.

There was also little difference between trained and untrained participants in our analysis. However, the training background in studies tended to be dichotomised into either recreationally active participants or trained participants. Only Triscott et al. [114] compared three training backgrounds: healthy, strength-endurance, and resistance-trained individuals. Following unilateral elbow flexors fatigue, the contralateral elbow flexors’ single MVC did not significantly change in any group regardless of training history. As observed with other NLMF studies, a contralateral post-fatigue endurance test (submaximal task to failure) was more sensitive to NLMF with both the healthy and resistance-trained participants showing significant decrements, whilst the strength-endurance trained subjects were not significantly affected. Similar to the sex differences rationale, strength-endurance trained participants could have produced less metabolic by-products due to their lower reliance on glycolytic pathways resulting in reduced metabolite
distribution by the cardiovascular system and diminished inhibitory metaboreceptor afferent input [115]. Though it is not clear whether this could be a result of their prior training history, or due to self-selection bias (i.e. those with such characteristics were more likely to be strength-endurance athletes). Thus further investigations should examine this issue including examining whether NLMF can be induced or prevented through the implementation of longitudinal training interventions.

A final area with little research is the possible moderating role of age in NLMF effects. None of the studies identified had directly compared younger (children and adolescents) and older cohorts, with the majority using younger adults (typically of university age). Yet, it has been suggested that both children and older adults exhibit greater endurance capacity than younger adults, though interestingly children tend to display large reductions in capacity (i.e. fatigue) comparable to that seen in endurance athletes, whereas older adults may exhibit less [116, 117]. Considering this, there remain intriguing possibilities to examine whether or not age may impact the absence or presence of NLMF effects.

5.0 Conclusions

The results of this comprehensive meta-analysis of findings from the NLMF literature suggest that, following a fatiguing intervention, previously non-exercised do not typically appear to exhibit a general NLMF effect when considering muscle strength and power-based performance outcomes. However, there is some evidence suggestive of NLMF for muscle endurance-based outcomes, though relatively fewer studies have explored this and the possible mechanisms for this specific effect are unclear. General NLMF effects were similarly absent and not moderated by study design (between vs. within-participant), homologous vs. heterologous effects, upper or lower body effects, participant training status, sex, the time of post-fatigue protocol measurement, or the severity of the fatigue protocol. There is however a paucity of studies including female participants and conducting direct sex comparison, in addition to studies of younger (children and adolescents) and older populations, and of different training backgrounds or indeed intervention studies examining the effects of implementing differing prior training. Thus, future studies are needed to clarify the possible NLMF effect upon endurance outcomes in addition to the potential mechanisms, in addition to studies examining population differences (i.e. sex, age, training background). Considering the lack of evidence for a general NLMF effect, we suggest that future research on NLMF effects should be redirected towards these still relatively unexplored areas.
**Funding information**
Partial financial support for graduate students (ER, MMIM, JW) was received from the Natural Science and Engineering Research Council of Canada.

**Data and Supplementary Material Accessibility**
All materials, data, and code are available on the Open Science Framework project page for this study [https://osf.io/cs9e2/](https://osf.io/cs9e2/)

**REFERENCES**
1. Tipton CM. History of Exercise Physiology. Windsor, Ontario, Canada: Human Kinetics Publisher Inc.; 2014; pp. 16-64.

DOI: 10.31236/osf.io/w7x3v

SportRxiv is free to access, but not to run. Please consider donating at www.storkinesiology.org/annual


Salveson FS. The effect of maximal vs. submaximal contractions on crossover fatigue between limbs. NTNU Open. 2020.

DOI: 10.31236/osf.io/w7x3v

SportRxiv is free to access, but not to run. Please consider donating at www.storkinesiology.org/annual

25


