Article details: 2020-0222	
Title	The effectiveness of physical activity interventions in older adults living with frailty or pre-frailty: A systematic review and meta-analysis
Authors	Megan Racey PhD, Muhammad Usman Ali MD MSc, Diana Sherifali RN PhD, Donna Fitzpatrick-Lewis MSW, Ruth Lewis BA, Milos Jovkovic MSc; Danielle R. Bouchard PhD, Anik Giguere PhD, Jayna Holroyd-Leduc MD, Ada Tang PhD, Leah Gramlich MD, Heather Keller RD PhD, Jeanette Prorok PhD, Perry Kim PhD, Amanda Lorbergs PhD, John Muscedere MD; for the Canadian Frailty Network
Reviewer 1	Lindsay Wallace
Institution	Medicine, Faculty of Medicine, Dalhousie University, Halifax, NS
General comments (author response in bold)	1) A justification or explanation of the link between the many outcomes examined would be useful. As many of these items are components of frailty or linked to frailty, a description or explanation of their relationship and whether they were examined in lieu of or in parallel to frailty would improve the rationale. **We wanted to be as broad as possible in terms of outcomes as there is no established consensus as to what the most important outcomes are in this patient population. We briefly highlighted the process that we went through to select the outcomes chosen in the manuscript and have provided more details in our response here. A more fulsome list of outcomes was developed by an interdisciplinary steering committee based on clinical relevance, previous research, and expert opinions on being related to frailty and direct components of frailty measures. This thorough and detailed list was subsequently voted upon anonymously by members of this group. The rankings were than averaged and all outcomes ranked as critical were extracted and analyzed. The novelty of this review is the use of our random effects multi-level meta-analytic approach which allows us to combine multiple measurements for one outcome while accounting for dependency between effect sizes i.e. the correlation between effect sizes due to multiple measures or sub-measures of same outcome with-in a study or comparison of multiple interventions to a single control group. The outcomes in this review (with the associated measures) were all voted as critical. A limitation of this review and gap in the literature is the lack of primary studies which measured frailty as an outcome. Where possible, we did extract and analyze frailty as an outcome, but this was only found in 4 of the included studies. Therefore, these outcomes are examined in parallel to frailty since frailty was rarely reported on as an outcome.** (p. 7) 2) Consider including the forest plot for the main outcome (frailty) in the body of the manuscript (rather than as supplementary material) to help readers visualize the results more readily. **We did not identify a main outcome for our review as all outcomes analyzed were ranked as critical by the committee. We have kept the forest plots in the appendices and highlighted all the results in our manuscript.** (N/A) 3) Consider organizing results section by outcome rather than exposure to be more informative and improve readability. From a reader perspective, particularly in CMAJ, many will be interested in certain outcomes and it would be helpful to be able to select an outcome and find in that paragraph the best evidence for physical activity interventions. **As this review supports Clinical Practice Guidelines that have been written by exposure/PA type, we have structured the review results to match these**

	requests. However, given the readability of outcome-specific information, we have re-structured and revised the discussion of the manuscript to consider PA interventions by outcomes. (pp. 13-15) 4) Consider revising sentence "titles and abstracts were reviewed in duplicate" to say titles and abstracts were reviewed by two independent reviewers or some such (it is a bit confusing due to **We have made this change to be clearer in our language.** (p. 7) 5) More detail on the 'development, pilot, and deployment of standard forms for data extraction' would be appreciated. **With the limited word count restrictions, we have added a few clarifying details of these data extraction forms.** (p. 7) 6) The frailty phenotype was originally created in the Cardiovascular Health Study, some explanation of the difference in tools cited as CHS vs. Frailty Phenotype would be useful. **These measures can be combined. They are similar, but were described as such in papers (some said Fried vs others said CHS) and were thus captured this way for transparent reporting in the manuscript. Fried's criteria were derived from the Cardiovascular health study and such can be considered the same. We have updated the manuscript and corresponding tables.** (p. 10, Table 1) 7) A comment on the use of non-validated frailty tools (e.g. inability to descend stairs step over step without holding the railing) would add to the discussion. This commentary has been added to the discussion. (p. 15) 8) A comment on how confounders were handled would be useful; i.e. were all models from the original papers unadjusted effect estimates, and if not, how might this influence the results? **Our data extraction and analysis was always from the study's raw data with unadjusted mean and SD. We do not take any adjusted estimates from studies (such as SMDs or Cohen's d).** (p. 7) 9) In cases where only one study was included on the forest plot, it would be appropriate to remove the diamond indicating a combined effect. **While we appreciate the comment that it may be redundant to have the diamond indicating combined effect in a forest plot where only one study was included, it is the reporting template standard both methodologically and in the software we used to present the forest plots as such. This avoids confusion for those reading and reviewing the figures and provides them the ability to quickly identify what statistical results (diamond) were considered in the overall pooled estimate.** (N/A)
Reviewer 2	Jennifer Jones
Institution	Princess Margaret Cancer Centre, Cancer Survivorship, Toronto, Ont.
General comments (author response in bold)	1) Overall the paper is well written and the methods have no major flaws (see specific methods questions below). My only concern with this paper is that there is already appears to be good quality evidence including systematic reviews and meta-analyses to evaluate various physical activity interventions and their components related to frailty prevention, progression, and reversal. In the

	Introduction of the manuscript, the authors have stated that it still unclear the best interventions to support older adults with frailty and that there is also a need for a comprehensive and systematic literature search with more homogeneous participants in the included studies. To address this, the meta-analysis examined of the impact of all physical activity interventions together and also included sub-group analyses based on the intervention category (aerobic, muscle-strengthening, mobilization/rehabilitation, and mixed). However, this review did not 1) compare the effectiveness of different types or categories of physical activity; 2) only included immediate post-intervention outcomes and did not analyze interventions based on frequency, intensity, or duration; and 3) did not look at long term outcomes. In the end, the interpretation was made largely on all physical activity interventions together and the authors were challenged by the fact that frailty itself was infrequently measured at both baseline and postintervention (n=4). Other limitations that the authors have highlighted include a large number of exclusions due to the frailty criteria and the variety of tools and definitions used to describe participants still made for diverse study population. **The authors recognize that while there are other systematic reviews which focus on older adults and physical activity/exercise and may comment on frailty, this review is unique in that (to our knowledge) it is the only one focused solely on persons living with frailty/pre-frailty using our clear inclusion criteria of such a population. We feel that a clearly defined population is important when creating guidelines from evidence that will support clinical practice. Additionally, a head-to-head comparison requires knowledge of what intervention is gold standard or optimal. Currently, in this population, the most effect PA strategy or type is unknown making it useful to compare to a control group. Lastly, our analysis is unique and maximizes the inclusion of outcome measures for a more robust presentation of the results. By putting studies together, we are able to present more precise synthesis and add to the existing literature in this field. Consequently, due to this specific focus on frail/pre-frail populations, there was insufficient evidence to conduct sub-group analyses such as those mentioned by the reviewer. There were also very few studies that looked at effectiveness beyond post-intervention, limiting our ability to investigate long-term outcomes. Our inability to compare the types of PA or speak to long-term outcomes were already mentioned in our discussion and limitation sections of our manuscript. Where possible, we have tried to add explicit content about the novelty of our review.** (pp. 5, 14, 15, 16) 2) An interesting finding in this paper was the number of adverse effects. Of the 26 included studies, 7 reported adverse effects or harms that were directly related to the intervention. It would be helpful to see some discussion of this and if possible to include recommendations on how these might be avoided or how future studies could screen for risk. **Thank you for this comment and highlighting an interesting finding of our review. We have added some more information in the results section and further described this finding in the discussion in context with other research.** (pp. 10, 14) 3) Finally, given this was a meta analyses, I would have liked to see a bit more in the discussion on next steps based on the results- for example highlight the weaknesses of previous studies and provide recommendations on how to improve the design of future trials and gaps that need to be addressed in order to move the

	science forward. **Thank you for this suggestion. We have added some language about next steps to the manuscript including: There was only one included study specific to aerobic activity, so further research in this area and frailty is warranted. Overall, interventions could have been better described (ex. lack of specificity around type of exercise/dose) to improve generalizability. Additionally, outcome measures were variable and future studies would benefit from standardization of outcomes so that interventions could be compared across studies. However, this was also a strength in one sense as the positive outcomes from the variation in interventions allows for a diversity of things to be applied in clinical and real-world setting that would be of benefit.** (p. 15) 4) Given the potential for publication bias, what strategies did the researchers use to identify trials that are unpublished? **While there is a potential for publication bias and many of our included studies are small RCTS, our effect estimates are within the funnel plot estimates and there was no significant deviation from pooled effect estimates. We did not observe any significant asymmetry across overall PA funnel plots for outcomes.** (p. 16 and Appendix 5) 5) The authors have stated that studies were also included if a sub-analysis was conducted on a portion of the participants who were pre-frail or frail. In the published PROSPERO protocol it was stated that only studies with a mixed (non-frail and frail) group with 80% of the population being pre-frail or frail would be included. Was this criteria followed? **These outline two different inclusion criteria to follow. If the study was a mixed population, 80% or more had to be pre-frail or frail. If the study population was primarily high functioning older adults, but there was a subgroup analysis on those that were pre-frail/frail, then that subgroup analysis and the associated results (being 100% pre-frail/frail) would be considered. Both criteria were followed to identify studies that were in pre-frail/frail older adults AND those that had subgroups which could be included.** (N/A) 6) Two or more authors abstracted the information from studies independently. Please indicated if the reviewers were blinded to the authors and institution of the studies undergoing review and confirm that the results from the data abstraction were compared after completing the review of the articles. Agreement among the reviewers should be reported. **We followed to Cochrane Handbook for Systematic Reviews of Interventions Version 6, 2019 for methods. During independent title and abstract screening, reviewers were blinded to institution of the authors, but not the authors themselves. During independent full-text screening and data abstraction, reviewers were not blinded to authors or institutions of the studies as this level of screening and extraction required the need to view the PDF in its entirety. During screening and extraction, authors and institutions are not reviewed in detail as they are not part of our inclusion process. Our reviewers are not content experts and therefore the inclusion of study authors names is not considered a source of potential bias. Results of screening and abstraction were compared as agreement was needed for inclusion or exclusion and all data was also verified independently by the**

	statistician. This is outlined in the methods of our manuscript. There were no conflicts of consequence in our screening. (N/A) 7) While forest plots were included is it does not appear that a sensitivity analysis was done to evaluate consistency of results and factoring in quality from the GRADE approach. Also, please add a comment of the clinical heterogeneity and methodological heterogeneity of the studies and the researcher's confidence that outcomes from individual studies should have been pooled. **Based on visual inspection of forest plots, most of the observed statistical heterogeneity in our review is likely due to small vs large study effects observed across pooled studies. We are still confident that outcomes from individual studies can and should be pooled based on the methods used in multi-level statistical analysis. We also used GRADE to inform our confidence of the evidence, which mostly ranged from very low to moderate. The results of GRADE were downgraded to inform effect estimates and this indicates that we have a low confidence in the effect estimates from our review. Further, our sub-group analysis based on the type of PA intervention noted some differences in how the effect was reported and highlights some of the factors influencing our results.** (N/A)