

List of Revisions
NRES-D-08-00075

Title: Meta-analysis of quality-of-life outcomes from physical activity interventions

We appreciate the careful review of our manuscript and the several positive comments about the research methods and the paper. The list below contains verbatim reviewer comments (reviewer number in parentheses) and associated manuscript revisions. Citations in the list of revisions are to sources in the manuscript reference list.

Comment: Title Page -- Supply running head of less than 50 characters (no abbreviations).(Checklist for style)

Response: A running head was added to the title page: quality-of-life and physical activity meta-analysis.

Comment: Some of the writing style is difficult to read. For example research question 1 "What are the overall effects of interventions to increase PA on QOL outcomes after interventions?" This could be stated in a much better and easier to understand fashion.(#1)

Response: The original first research question was divided into two questions to increase clarity. The remainder of the manuscript was evaluated for difficult writing and appropriate revisions were made.

Comment: The introduction is also confusing.(#1)

Response: The introduction was revised.

Comment: Please avoid the use of first person in the text.
(Checklist for style)

Response: The manuscript was revised to remove first person wording.

Comment: Many comments refer back to previous work. I don't want to read the previous work to understand what was done. Specifically what MESH words were searched? What diverse strategies were used for the searches? The selection criteria could be presented in a table form and why 1970? That is a rather broad time frame. Was there a reason for it? (#1)

Response: To preserve full disclosure, citations to the previous work (which includes any of the primary studies in this manuscript while addressing different outcomes) were retained.

The methods section was expanded to address these issues. A reference librarian performed computerized searches in 11 databases (MEDLINE, Cochrane Central Register of Controlled Trials, Dissertation Abstracts International, PsychInfo, SportDiscus, HealthStar, Clinical Evidence, Scopus, DARE, ABI/Inform, Cumulative Index to Nursing & Allied Health Literature). Broad search terms were used for intervention (adherence, behavior therapy,

clinical trial, compliance, counseling, evaluation, evaluation study, evidence-based medicine, health care evaluation, health behavior, health education, health promotion, intervention, outcome & process assessment, patient education, program, program development, program evaluation, self care, treatment outcome, validation study) and physical activity (exercise, physical activity, physical fitness, exertion, exercise therapy, physical education & training, walking) (Conn, Isamaralai et al., 2003).

The diverse search strategies are listed in the methods section (computerized searches of 11 databases, ancestry searches, computerized author searches, journal hand searches, conference abstract searches).

As suggested, a new table was created for the inclusion criteria. The sample section of the manuscript narrative was shortened since inclusion criteria were moved to the new Table 1.

Although many aspect of chronic illness health care have changed since 1970, there is little reason to believe quality of life (QOL) outcomes of physical activity (PA) interventions would change over the broad time span. Previous meta-analyses found little evidence of an association between publication year and PA outcomes.

Comment: The search of published literature was supplemented with a search for unpublished studies, a strength of the project. It would be helpful to the readers to know (a) the proportion of these 66 studies that were unpublished, (b) how they approached the investigators to obtain the results of unpublished studies, and (c) the response rate.(#2)

Response: The methods section was expanded to address this issue. Several strategies were used to search for unpublished studies. Dissertation Abstracts International was thoroughly searched (author searches were conducted on dissertations located in Dissertation Abstracts International to determine if a publication reported dissertation findings). Conference abstracts were evaluated for eligible studies and subsequent author searches were conducted to locate published studies. If published reports were not available, abstract authors were contacted for additional information. Project staff emailed all of the first authors on studies included in the meta-analysis to solicit additional published or unpublished studies. These comprehensive search strategies yielded ten unpublished research reports.

This was a very large project (the QOL outcomes are a small subset of the larger project focused on PA and health outcomes) where over 12,000 research reports and papers were initially reviewed. Contacts soliciting unpublished manuscripts were not tracked due to the size of the project, thus response rate information is not available. Recording response rate, in a more manageable project, would be interesting.

Comment: Several points could be explained more thoroughly. First, "common language effect size" is discussed in the narrative as a proportion (53% (58%) of the subjects would have a better quality of life value) but the derivation of this is not clear in Table 1. (#2)

Response: The description of Common Language Effect Size (CLES) in the statistical analysis table was expanded. CLES measures the difference between two populations as the probability that a score sampled from one population would be higher than a score from the other population.

Two-group: $CLES = Pr(Z < d/SQRT(2))$

Pre-post: $CLES = Pr(Z < d/SQRT[2(1 - \rho)])$

The CLES indicates the probability that a random treatment subject would attain better a QOL score than a random control subject, or that a subject would score higher after than before treatment. An effect size (ES) of 0 (no effect) would have a corresponding CLES of .50.

Comment: The "assumptions of association" discussed in the narrative should be explained further. (#2)

Response: The explanation in the statistical analysis table was expanded. Pre- and post-intervention scores are probably correlated to some extent in single-group design studies. Some measure of the correlation is needed for computing pre-post ES. No studies provided data regarding this association. Lacking empirical evidence, the analyses were conducted under assumptions of no (linear) correlation ($\rho_{12} = 0$) and high correlation ($\rho_{12} = .8$). A comment was added to the text referring readers to the explanation in the statistics table.

Comment: Explain how the funnel plot (effect size by variance) is a publication bias assessment. (#2)

Response: Content was added to the statistical methods table to explain funnel plots. As sample size increases, sampling error should decrease. Thus smaller sample studies should scatter more broadly around the mean ES at one end of the plot because they have larger sampling error, while larger sample studies should cluster more tightly at the other end of the plot since they have smaller sampling error. A symmetrical funnel-shaped plot suggests absence of publication bias. An asymmetrical plot may suggest that some studies with small (or negative) ES were not published. Funnel plots are difficult to assess when few primary studies are available.

Comment: The results are described concisely and clearly, although information from the tables could be explained a bit further in the narrative. For example, clearly stating how many of the 66 studies were unpublished, including the median as well as the range of sample size, attrition, and mentioning the distribution of time and types of interventions included. (#2)

Response: The number of unpublished studies was added to the results. The descriptive statistics table (new Table 3) contains the median, minimum, and maximum for sample size, attrition, and intervention time. Additional details from table 3 were placed in the text. More details about the interventions were added as a new second paragraph in the results section. Details found in Tables 5 and 6 were judiciously added to the text (e.g. indicating $\hat{\mu}_s$ and $\hat{\beta}_1$ in the text) due to concerns about adding length to the manuscript as well as a different reviewer comment that the manuscript is number dense.

Comment: I realize that there are too many studies for you to present Forest plots... but the manuscript is really heavy number-wise and not very "reader-friendly"; nurses need to become more comfortable reading meta-analyses and I wonder if there is any way for you to make the content here more accessible. (#3)

Response: As noted by the reviewer, there too many studies to present Forest plots. A stem-and-leaf figure was added as a visual display of the two-group post-test findings.

Further details about the analyses were added in the statistics table. Four citations to excellent introductory texts regarding meta-analysis methods were added. The meta-analysis methods, and findings, may move beyond the research expertise of some readers. Nurse researchers need to become more familiar with meta-analyses. The report may stimulate further reading regarding meta-analysis methods.

We agree we report many findings in this dense manuscript. We removed the fixed-effect findings from the table of effect size point estimates and tests. We attempted to balance this comment about the manuscript being 'heavy number-wise' with another reviewer's suggestion to add more details from the tables in the text.

An alternative approach to reduce the 'heavy number-wise' burden in the manuscript would be to separate the findings into two manuscripts. This approach was not selected because the analysis of QOL outcomes is one project and should be presented in one manuscript. Another strategy is to present less data. The approach was used in the manuscript by not reporting considerable additional findings (e.g. fixed-effects analyses). The authors believe the manuscript presents the minimum amount of data to reasonably describe the findings. The authors would thoughtfully consider further specific *Nursing Research* recommendations about deleting some data to make the manuscript less 'heavy number-wise'.

Comment: Further explanation is needed here for the CLES and a context for the minimal difference in probability of better

quality of life scores between control (51%) and treatment (53%) subjects. (#2)

Response: The first paragraph of the discussion section includes content which addresses ES magnitude and CLES scores. The ES magnitude, as calculated and as depicted by CLES scores, seems modest. It is unclear what ES would represent a clinically meaningful improvement in QOL among chronically ill adults. People with major chronic illnesses experience many reasons for declining QOL, including the disease itself and physical or psychosocial sequelae of the disease, and onerous or distressing treatments. Studies may recruit chronically ill study subjects from specialty medical practice settings where the subjects may already be receiving optimal medical care. Even a small change in QOL may be important since QOL is a complex phenomenon likely affected by diverse factors. It is important to note that these findings were heterogeneous, as expected, though less so for two-group than pre-post comparisons and less so for QOL outcomes than for other health and PA outcomes reported in these same primary studies (Conn, Hafdahl, Brown et al., 2008; Conn, Hafdahl, Mehr et al., 2007; Conn, Hafdahl, Minor et al., 2008; Conn, Hafdahl, Moore et al. 2008).

Comment: Another potential point to consider in the discussion of the larger effect size from unpublished studies is the possibility that the investigators in those studies may have been more likely to share the results for the meta-analysis if their results were clearly positive. (#2)

Response: This suggestion was incorporated in new content in the discussion section.

Comment: I would think that the interventions also would have varied considerably in terms of intensity, duration, and approach; assessment of physical activity and QOL must also have varied. All of this leaves the quality, hence, validity of the studies in question. Can you comment on this in manuscript? (#3)

Response: The interventions and measures were heterogeneous, as expected. Interventions varied dramatically from brief motivational sessions to extended supervised exercise programs. Diverse measures were used to assess QOL and PA. No gold standards exist for interventions or measures of QOL and PA. Other important factors that may affect validity of primary studies' findings, which are infrequently reported in primary studies, could not be assessed (e.g. treatment fidelity). As more primary research accumulates, future meta-analyses may be able to determine if ES are related to research methods.

Comment: Update "in press" references. (Checklist for style)

Response: The arthritis paper citation was updated because the article was published. The cardiac article has been published online, but not in print. The cardiac paper citation was changed to indicate online access by providing a DOI. The journal has not indicated a print publication date for the cardiac article.