A comparison of individual patient analysis versus pooled study meta-analysis

methodologies of exercise training trials in heart failure p0atients

Neil A. Smart [1] and Michael Steele [2, 3]

[1]University of New England
[2]School of Business, Bond University
[3]Griffith Graduate Research School, Griffith University

Abstract:

Background: A fixed effects meta-analysis of ten exercise training in trials heart failure patients was conducted. The aim of this current work was to compare different approaches to meta-analysis using the same dataset from the previous work on ten exercise training trials in heart failure patients.
Methods: The following different meta-analysis techniques were used to analyse the data and compared the effects of exercise training on BNP, NT-pro-BNP and peak VO2 before and after exercise training:

(1)   Trial level (traditional) level MA
 i) Follow up (post-exercise training intervention) outcome only.
 ii) Baseline-follow up difference
(2)   Patient level MA by Post-Stage ANCOVA
i) naive model does not take into account trial level
ii)   Single Stage
iii)   Two Stage
(3)   Post outcome only
i) Single stage
ii) Pre-post outcome difference Single stage

Results: The Individual patient data (IPD) analyses produced smaller effect sizes and 95% confidence intervals compared to conventional meta-analysis. The advantage of the one-stage model is that it allows sub-group analyses, while the two-stage model is considered more robust but limited for sub-analyses.
Conclusions: Our recommendation is to use one-stage or two-stage ANCOVA analysis, the former allows sub-group analysis, while the latter is considered to be more technically robust.

Key words: Individual patient data meta-analysis, heart failure, exercise training, brain natriuretic peptide.

1. Background

Meta-analyses are based on systematic reviews and are regarded as the gold standard for the practise of evidence-based medicine. However, such analyses are typically based on study level group data that may be incomplete across all included studies. A potential solution is a meta-analysis of individual patient data (IPD) that has become increasingly common. IPD has an advantage in that one can reanalyse the individual patient data for each study thus combining them in a more consistent way. Of particular importance are time to event data where one may analyse data using Cox regression models and continuous outcomes where one may analyse data adjusted for baseline values i.e. ANCOVA. However the obvious potential limitation of IPD is obtaining raw data from authors and permission to use it.   Gaining this permission typically involves setting up an IPD trial collaboration. A notable example in exercise training for heart failure patients is ExTraMATCH [1].

Here we present several methods of conducting a meta-analysis, the examples given have examined the same dataset. The dataset was generated from 10 studies measuring the impact of exercise training on brain natriuretic peptide (BNP), the N-terminal portion of BNP (NT-Pro-BNP) and peak V02 where study level meta-analysis and IPD data have been previously published.   The first meta-analysis used pooled post-intervention data from individual studies that has inherent limitations, possibly the most important being that baseline differences are not accounted for [2]. This first meta-analysis therefore did not utilize baseline data and in some of the included 10 studies, some of the outcome measures were not matched at baseline and therefore using post-exercise training data only may have affected the findings.

The second meta-analysis used individual patient data (IPD) from the same 10 studies and combined them into one dataset [3]. This approach is often con-sidered more methodologicall robust. In addition individual characteristics of the patient E.g. age and disease severity can be linked to the magnitude of improvement, as such investigators can predict which patients may benefit most.

2. Methods

This work has emanated from two previous meta-analyses. Following the ac- ceptance for publication of a meta-analysis evaluating the effects of exercise train- ing on BNP and NT-pro-BNP2, a collaborative group was formed and prospective data collection was agreed and a common dataset of collected variables was pro- vided by all included study authors. This collaborative group

produced a second publication which was an individual patient analysis from the same 10 studies[3]. Details of the search strategy and inclusion and exclusion criteria, data extraction and analysis can be found in these works [2][3].

For the purpose of this work, several different meta-analysis approaches were taken to analyse data from the original 10 studies. We compared the effects of exercise training on BNP, NT-pro-BNP and peak VO2 before and after exercise training:

Trial level (traditional) level MA
i) Follow up (post-exercise training intervention) outcome only.
ii) Baseline-follow up difference
Patient level MA by Post-Stage ANCOVA
iii) Naive model does not take into account trial level
iv) Single Stage
v) Two Stage Post outcome only
vi) Single stage
vii) Pre-post outcome difference Single stage

Analyses were all conducted with random effects models, because heterogeneity appeared across all 3 outcome measures. ANCOVA models i.e. exercise- control comparison of baseline-follow up difference adjusted for baseline values. We first ran the one-stage hierarchical model that took into account both the ANCOVA analysis but the study level structure of data (as a random effects). We then ran the meta-analysis as a two-stage approach i.e. 1st estimated ANCOVA effect size in each individual trial and then pooled the aggregated effect size across trials using random effects meta-analysis (allows Forest plots to be produced).

3. Results

3.1 Study characteristics

Ten randomized, controlled studies (8 discrete datasets, overlapping data was not duplicated) met our eligibility criteria, with an aggregate number of 565 subjects (313 exercise participants and 252 controls) were included. Five studies measured BNP, six measured NT-pro-BNP and 1 study measured NT-pro-BNP and BNP. BNP data from 230 patients, NT-pro-BNP data in 466 patients and both BNP and NT-pro-BNP in 133 patients was available. The overall mean exercise-control estimates can be seen in table 1.

Table 1: Comaprsion of Effect Sizes and 95% Confidence Intervals for Change in Outcome Measures for Different Analysis Methods. BNP

	Peak VO$_2$ (ml/kg/min) N=6 trials, N=511 patients	BNP (pg/ml) N=4 trials, N=233 patients	NT-pro-BNP (pg/ml) N=5 trials, N=466 patients
Trial data level analysis Post outcome only* Pre-post outcome difference*^	2.74 (1.85 to 3.63) P<0.0001, I^2 = 55% 2.16 (0.96 to 3.36) P<0.0001, 1^2=82%	-76 (-172 to 20) P=0.12, I^2 = 59% -69 (-19 to -120) P=0.007, I^2 = 46%	-564 (-695 to -433), P<0.0001, I^2=0% -404 (-571 to -236), P<0.0001, I^2=67%
IPD analysis Post outcome ANCOVA *Naive analysis+* *Single stage* *Two stage* Post outcome only *Single stage* Pre-post outcome difference *Single stage*	2.24 (1.89 to 2.61), P<0.0001, I^2=NA 2.36 (2.01 to 2.70), P<0.0001, I^2=NA 2.17 (1.16 to 3.17), P<0.0001, I^2=90% 2.45 (1.83 to 3.08), P<0.0001, I^2=NA 1.58 (0.92 to 2.23), P<0.0001, I^2=NA	-74 (-39 to -109), P<0.001, I^2=NA -77 (-44 to -110), P<0.001, I^2=NA -71 (-16 to -126), P=0.01, I^2=59% -91 (-38 to -144), P=0.001, I^2=NA -72 (-36 to -108), P<0.0001, I^2=NA	-438 (-570 to -306), P<0.0001, I^2=NA -467 (-597 to -337), P<0.0001, I^2=NA -434 (-278 to -589), P<0.0001, I^2=64% -607 (-811 to -404), P<0.0001, I^2=NA -413 (-554 to -275), P<0.0001, I^2=NA

3.2 Major findings from the analyses

Logistic regression has several characteristics that are suitable for implement- ing the inference model as listed below:

- The inference of all models in this example is the same i.e. highly significant statistics in favour of the exercise group for all outcome measures.

- The inference of one-stage ANCOVA models is that the method permits the examination of subgroup and moderation effects with more power. Mean estimates for both ANCOVA models appear to fall between the group and single database analyses, despite this the 95% confidence intervals are wider.

- The trial data values are different to those in the two published meta- analyses because these trial level aggregate analyses were calculated di- rectly from the IPD, whilst previously published analyses used the trial level results reported originally.

4. Discussion

Our work compared several meta-analysis methods including traditional trial level analyses using both follow up outcome only, and also baseline-follow up difference, we also used one and two step ANCOVA methods. As can be seen the overall mean estimates for all primary outcomes are similar and significantly improved in the exercise versus control groups for all models utilized. The statistical significance probably reflects the large effects sizes and that these data were highly significant and therefore unlikely to yield results that were statistically dif- ferent, irrespective of the method of analysis. It should however, be highlighted that mean estimates   for both ANCOVA models appear   to fall between the group and single database analyses, despite this the 95% confidence intervals are wider. This suggests that in this case, the ANCOVA   models are more conservative and hence more robust.   We contend that using the actual individual patient data gives a more fair comparison of methods.

The primary purpose of this work was to examine if any of the methods em- ployed produced outlying effects, this was not the case.   However   each of the methods   employed have their   own inherent strengths and weaknesses. The orig- inal method used was to analyze post-data only in a group data analysis. As all three primary outcomes produced continuous data the main problem with analysis was calculating pre-post intervention change in standard deviation.   An established method for overcoming this is to analyze only using post-intervention data 4 and this is inherently flawed as no adjustment is made for baseline dif- ferences between groups. One method to overcome this is to use both pre- and post-intervention data and impute the change in standard deviations using p val- ues 5. While this method adjusts for baseline differences between groups, the technique itself has inherent flaws. The potential for error is highest when calcu- lating change in standard deviation from studies not reporting precise p values. An example of this can be seen if the actual p value is say 0.015 but the p value reported is p¡0.05, in which case a default value of p=0.049 can be used which is likely to make a large difference to the calculation of change in standard deviation. The third

method we employed combined the individual patient data provided from all study authors into one database. There are several advantages of this technique, it adjusts for baseline differences between groups, allows pre-post- intervention calculation in standard deviation change, it also allows sub-group analyses and the results in this case were similar to the other methods employed, although it should be noted that the use of this technique is not widely endorsed.

We now move onto the ANCOVA models, the two-stage model takes into account of the full variance between trials not only in the baseline-control differ- ence but also baseline values, in that sense, this is probably the more technically robust result. However, the inference of one-stage and two-stage models in this analysis is the same i.e. highly significant stats difference in favour of exercise for both BNP and NT-pro-BNP. The strength of the one-stage model is that it gives a structure allows subgroups to be tested at individual patient level. To this end the methods allows the analysis of subgroups, both as categorical and con- tinuous variables to quantify treatment effect interactions which are technically more appropriate than stratified comparison of subgroups within the intervention group.

5. Recommendations

Based upon our experiences with these different methods of conducting meta- analysis we make the following recommendations:

- If it is possible to get the individual data sets then an individual patient data (IPD) analysis is preferred to a trial level analysis, although the IPD data set should be analyzed so that one can generate the trial level results for each study and then use these in the comparisons. The alternative of course is to use actual trial results reported by the authors, but we would contend that using the actual IPD data gives a more fair comparison of methods.

- Use an ANCOVA model, the one-stage model allows subgroups to be tested at the individual patient level, both as categorical and continuous variables. The two stage ANCOVA model is considered more robust.

- Run the subgroups as treatment effect interactions (technically more appro- priate than stratified comparison of subgroup in the intervention groups).

6. Conclusions

Several methods are available to conduct meta-analyses on the same datasets, each method has inherent strengths and weaknesses. The preference is to use one-stage or two-stage ANCOVA analysis, the former allows sub-group analysis, while the latter is considered to be more technically robust.

7. List of Abbreviations

BNP Brain Natriuretic Peptide IPD - Individual Patient Data NT-Pro-BNP- N-terminal portion of BNP

8. Competing Interests

None

9.   Author's Contributions

Neil Smart - wrote the protocol and collected the data and wrote the draft paper Michael Steele -edited the protocol, conducted   the analyses and edited the draft paper

10. Acknowledgements

References

O'Connor CM, Whellan DJ and Lee KL (2009).   Efficacy and safety of exercise training in patients with chronic heart failure: HF-ACTION randomized controlled trial. In JAMA, 301(14):1439-1450

.Smart NA and Steele M. (2010). Systematic review of the effect of aerobic and resistance exercise training on systemic brain natriuretic

peptide (BNP) and N-terminal BNP expression in heart failure patients. In *Int J Cardiol*, **140**(3):260-265.

Smart N MT, Butterfield J, Faddy S, Passino C, Malfatto G, Jonsdottir S, Sarullo F, Wisloff U, Vigorito C and Giallauria F. (2011). Individual patient meta-analysis of exercise training effects on systemic brain natriuretic peptide expression in heart failure. Eur J Cardiovasc Prev Rehabil.2011.

Higgins JPT GS. (2012). Cochran Handbook for Systematic Reviews of Interventions. 2011. http://www//Cochrane-handbook.org.

DerSimonian R, Laird N. (1986). Meta-analysis in clinical trials. *Control Clin Trials.* **7**(3):177-188

Associate Professor Neil A. Smart
University of New England
School of Science and Technology
Armidale, NSW 2351, Australia
Tel: +61-2-6773-4076
Fax: +61-2-6773-5011
E-mail: nsmart2@une.edu.au

Assistant Professor Michael Steele
School of Business, Bond University
Gold Coast QLD 4229, Australia