Watson, JA, Ryan, CG, Atkinson, G, Williamson, P, Ellington, D, Whittle, R, Dixon, J and Martin, DJ

Inter-Individual Differences in the Responses to Pain Neuroscience Education in Adults With Chronic Musculoskeletal Pain: A Systematic Review and Meta-Analysis of Randomized Controlled Trials

Title Page

Inter-individual differences in the responses to pain neuroscience education in adults with chronic musculoskeletal pain: A systematic review and meta-analysis of randomised controlled trials.

Short title: Individual differences in response to pain neuroscience education

Authors:

James A Watson BSc (Hons) [1]
J.A.Watson@tees.ac.uk

Cormac G Ryan BSc, MSc, PhD[1]
C.Ryan@tees.ac.uk

Greg Atkinson BSc (Hons), PhD[1]
Greg.Atkinson@tees.ac.uk

Philip Williamson, BSc (Hons), MSc, PhD[3]
phil.williamson@york.ac.uk

Dominic Ellington BSc (Hons) [2]
D.Ellington@tees.ac.uk

Robbie Whittle BSc (Hons) [2]

robbie.whittle@nhs.net

John Dixon BSc, PhD[1]

John.Dixon@tees.ac.uk

Denis J Martin BSc. MSc. DPhil[1]

D.Martin@tees.ac.uk

[1]School of Health and Life Sciences, Teesside University, Middlesbrough, Tees Valley, TS1 3BX, United Kingdom.

[2]North Tees and Hartlepool NHS Foundation Trust, University Hospital of North Tees, Hardwick Road, Stockton on Tees, Cleveland, TS19 8PE, United Kingdom.

[3]York Trials Unit, Department of Health Sciences, University of York, Heslington, York, YO10 5DD, United Kingdom.

Corresponding author: James Watson, J.A.Watson@tees.ac.uk

James Watson

School of Health and Life Sciences

Teesside University

Campus Heart

Southfield Rd

Middlesbrough

TS1 3BX

Telephone number and fax number not available.

Disclosures

This research was funded by Teesside University.

The authors have no conflict of interest to declare.

Highlights:

- Pain neuroscience education (PNE) is a pain management intervention.
- Little evidence of true individual differences in response to PNE for disability.
- Findings should be interpreted cautiously due to very wide prediction intervals.
- Estimating individual differences should be applied to other pain interventions.

Abstract

Pain neuroscience education (PNE) is an approach used in the management of chronic musculoskeletal pain (CMP). Previous reviews on PNE and other pain interventions, have focussed on mean treatment effects, but in the context of "precision medicine", any inter-individual differences in treatment response are also important to quantify. If inter-individual differences are present, and predictors identified, PNE could be tailored to certain people for optimising effectiveness. Such heterogeneity can be quantified using recently-formulated approaches for comparing the response variance between the treatment and control groups. Therefore, we

conducted a systematic review and meta-analysis on the extracted standard deviations of baseline-to-follow up change to quantify the inter-individual variation in pain, disability and psychosocial outcomes in response to PNE. Electronic databases were searched between 01/01/2002 and 14/06/2018. The review included five randomised controlled trials (n=428) in which disability outcomes were reported. Using a random effects meta-analysis, the pooled SD (95% CI) for control group-adjusted response heterogeneity to PNE was 7.36 units /100 (95% CI: -3.93 to 11.12). The 95% prediction interval for this response heterogeneity SD was wide (-10.20 to 14.57 units /100). The control group-adjusted proportion of “responders” in the population who would be estimated to exceed a clinically important change of 10/100 ranged from 18-45%. Therefore, when baseline-to-follow up random variability in disability is taken into account (informed by the control arm), there is currently insufficient evidence for the notion of clinically important inter-individual differences in disability responses to PNE in people with CMP. The protocol was published on PROSPERO (CRD42017068436).

Perspective

We bring a novel method to pain science for calculating inter-individual differences in response to a treatment. This is conducted within the context of a systematic review and meta-analysis on PNE. We highlight how using erroneous methods for calculating inter-individual differences can drastically change conclusions when compared to appropriate methods.

Key words

Pain, neuroscience, education, Individual response variance

Introduction

Pain neuroscience education (PNE) is an educational approach used in the management of chronic pain. PNE aims to reconceptualise an individuals' understanding of their pain as less threatening to facilitate rehabilitation[23]. Since its inception PNE has become increasingly popular in clinical practice[24]. Our group recently published a mixed-methods systematic review and meta-analysis on the effectiveness of PNE for adults with chronic musculoskeletal pain (CMP)[39]. Quantitatively we found no evidence to indicate that PNE results in clinically important changes over control for pain or disability. In contrast we found moderate quality evidence that PNE produces small clinically important changes over control for pain catastrophising and kinesiophobia. Qualitatively we found that achieving some degree of pain reconceptualisation following PNE can enhance peoples' ability to cope with their condition.

One question that arose during our previous research work was whether PNE may be effective for some types of people, implying that there may be some individual differences in response to PNE[39]. The quantitative component of our review focused on the mean intervention/treatment effect. This focus on mean intervention effect whilst common in research on pain interventions[5,15,30] could have obscured important inter-individual differences in response to PNE[16,41]. Such response heterogeneity is particularly important within the context of precision medicine, an increasingly popular field which encompasses 'tailor-made' therapies based on the person's individual response to a given intervention[31]. This individualised approach to medicine aims to improve the quality of care and reduce costs[33]. The potential

importance of a tailored approach has been highlighted by some of our previous qualitative work on PNE. The relevance of PNE to the individual (i.e. how tailored the material is to that individual) appears to be an important factor in the success of PNE[17,18,29,39]. Where PNE was reported to be relevant, people reported greater perceived benefit. The opposite was found where PNE was deemed not relevant[17,18,29].

Some researchers[27] have attempted to complement the quantification of mean treatment effects with a quantification of how many people in each intervention group change above or below a pre-set threshold, termed sample responder counts. Crucially, this approach does not provide any information about response heterogeneity to a given intervention in the context of precision medicine. In fact, these responder counts lack statistical power and may merely reflect within-subject random variation between timepoints and/or group differences in mean change. Furthermore, the dichotomisation (responder or non-responder) also creates problems adjusting for baseline differences between study groups (comprehensive reviews are available [2,32]). These sample responder counts tell us little about whether different people respond to different degrees to the same intervention, which is one of the fundamental questions in precision medicine. Should any inter-individual differences be falsely identified using the above-mentioned methods, any follow-up analysis to explore potential moderators of the intervention effect to explain the individual differences in response are therefore unwarranted[1,2]. Subsequent follow-up studies on the same participants is a waste of resources, and potentially unethical, if no true inter-individual differences in response exist to explain.

Inter-individual differences in response can be quantified by comparing the SDs of the baseline-to-follow-up changes between the experimental and control groups[1,4]. The difference between these SDs represents the SD for individual responses (SDir) which quantifies the individual variability in treatment response *per se*. The SD of the mean change score solely for the intervention group comprises treatment response variance *in addition to* the random variability in measurements between the baseline and follow-up timepoints. The SD of the changes in the control group represents this random variability in measurements between baseline and follow up – the random within-subjects variance component and measurement error.

Our qualitative analysis highlighted that PNE may be effective for some people but not for others implying that true inter-individual differences in response to PNE may exist which could be explored to facilitate appropriate targeting of PNE to those most likely to benefit[39]. However, clinically relevant inter-individual response variation should first be conducted using appropriate methodology[1,2,13,40,41] to confirm the presence of such inter-individual responses. If individual differences are observed, and predictors of individual response are identified, then PNE could be tailored to the individual optimising its effect[41].

To date, there has been no investigation of 'true' individual response variation of the effect of PNE, or indeed any pain management intervention. Therefore, we aimed to conduct a systematic review and meta-analysis of the available research to quantify the 'true' inter-individual variation in pain, disability and psychosocial outcomes in response to PNE in adults with CMP.

Methods

The protocol for the systematic review was published on PROSPERO (CRD42017068436). The analysis of inter-individual differences is presented here in detail to ensure the background and rationale for this novel method within the field of pain is adequately reported. A detailed account of the full review-methods has been published elsewhere[39] but a brief summary is provided below.

Inclusion and Exclusion Criteria

Inclusion criteria

- Studies including adults (≥18 years) who have CMP consistent with the British Pain Society definition (chronic pain, that lasts beyond the time that tissue healing would normally be expected to have occurred, often taken as ≥3 months)[35].
- RCTs that (i) compared the intervention with no treatment (true control) or usual care (ii) concomitant studies where PNE was delivered in addition to another intervention where that other intervention was received by both groups and (iii) head-to-head studies where PNE was compared to another active intervention.
- Studies reporting either pain and/or disability and/or psychosocial wellbeing.
- The SD of the changes for the intervention and control groups must have been included within the publication, have been available from the author upon request, or could be calculated from other information given such as the

standard error. This is an additional criterion that was not included in the registered protocol.

Exclusion criteria

- Studies that included participants with non-musculoskeletal pain such as cancer pain, visceral pain or post stroke pain.

Search Strategy

Pre-identified keywords (Pain AND (Physiology OR Neurophysiology OR Neuroscience OR Biology) AND Education) and index terms were searched across all included databases (The Cochrane Library, AMED, CINAHL Complete, MEDLINE, PsycINFO, PEDro, Scopus, EMBASE, Education Resources Information Centre (ERIC), Web of Science, clinicaltrials.gov, dissertations indexed with ProQuest Dissertations and Theses Global and EThOS) from 2002-25 July 2017, and updated on 14 June 2018.

After removing duplicates, the title and abstracts were screened by two authors and disagreements were resolved through discussion or a 3rd reviewer. The full-text was obtained for all records that could potentially fit the criteria. Upon reading the full-texts those deemed not to meet the inclusion criteria were rejected. See Supplementary Digital Content 1 for a list of excluded publications and reasons for exclusion.

226

227

228 Deviation from protocol

229

230 In our previous review[39] when the SD of change was not reported, and could not be
231 obtained by contacting the authors, it was either calculated from other information
232 given such as standard error, or estimated from the baseline and follow up SDs,
233 according to methods described in the Cochrane handbook[10]. Where there was
234 uncertainty regarding the validity of baseline, follow up and change score SDs from
235 included studies we opted not to use this data to inform our calculations to estimate
236 the SD of change scores. Instead, we used a robust data set of individuals with CMP
237 where we were confident in the validity of the baseline, follow up and change score
238 SDs. However, for the current review, given that to calculate the true inter-individual
239 differences in response to an intervention the SD of the mean change score is of
240 central importance[1], it would be inappropriate to estimate the SD of the change or
241 use a robust data set. Thus, an additional criterion for inclusion was created for the
242 current review where the SD of the changes for the intervention and control groups
243 must have been published in the article, available upon request by the author, or
244 could be calculated from other information given, such as the standard error.

245

246 Assessment of methodological quality and data extraction

247

248 Articles selected for critical appraisal were independently assessed by two reviewers
249 using the Cochrane tool for assessing risk of bias[9]. Two reviewers independently
250 extracted the data using JBI-SUMARI[36] including details about the interventions,

populations, study methods and outcomes of relevance to the review question/objectives. The Grades of Recommendation, Assessment, Development and Evaluation (GRADE) approach[7] was used to rate the overall quality of quantitative evidence for each outcome. A summary of findings table created using GradePro is presented (Table 1 and 2).

Meta-analysis

To contextualise the results for individual response variance we conducted a random-effects meta-analysis for the mean difference in disability across the included studies using a restricted maximum likelihood (REML) model combined with the Knapp-Hartung method. This method uses quantiles of the t distribution to calculate a confidence interval for the average effect instead of the standard normal distribution in the more conventional methods[37] . The Knapp-Hartung method has been shown to be superior to the DerSimonian-Laird method where there is a small number of studies (<20) and heterogeneity is present[11]. We then extracted the standard deviation of the changes in disability for both control (C) and PNE (I) groups. The true individual response variance (intervention minus control) was then calculated by $\sqrt{(SD_{I2}-SD_{C2})}$[13]. The standard error (SE) for this variance was then calculated using the equation: $SE = \sqrt{[2(SD_I^4/DF_I + SD_C^4/DF_C)]}$, where DF_I and DF_C are the degrees of freedom of the standard deviation in the PNE group and the control groups[13]. A negative value for the individual response variance for the confidence intervals or prediction intervals implies greater variability in the changes in disability in the control versus PNE group.

The individual response variances and their SEs were meta-analysed using an REML model combined with Knapp-Hartung method. It's important to highlight that the variances are unbiased, whereas the SD is not, and deriving a SE for the SD for individual responses is also problematic. Thus, we synthesised the individual response variances instead of the SDs for individual responses. The point estimate for the pooled individual response variance were derived together with a 95% CI to express its uncertainty. The point estimate and CIs were then square rooted to convert to an SD metric. If the lower limit was negative, the sign was ignored, the square root taken, and the sign re-applied. This approach is consistent with the 'no bound' option in SAS/STAT® software, which permits negative variances (SAS Institute Inc. 2017. SAS/STAT 14.3 User's Guide. Cary, NC: SAS Institute Inc.).

Using the methods of Swinton et al.[34] the proportion of responders in the population of interest within each included RCT was estimated. To estimate this, the observed mean change score and true individual response variance are needed for each RCT. Normal variance is assumed. The total area of any probability distribution is equal to one, thus the estimate of the proportion of response can be obtained by calculating the area of the derived normal distribution that lies beyond the minimally clinically important difference (MCID). An MCID of 10% was used in recent NICE guidelines for back and radicular pain[25]. The calculation estimating the proportion of response was performed via an online calculator[28]. The proportion of response was estimated for the intervention and control groups for all RCTs and has been used to demonstrate the difference in results, and thus conclusion that could be made if researchers erroneously ignored the control group data.

The tau statistic (τ) was used to quantify between-study heterogeneity – a SD that describes the typical variability of the mean effect between studies[3,8]. A 95% prediction interval was calculated using the tau and the SE for the pooled mean effect to quantify the expected range of true effects in future similar studies[12]. Stata (StataCorp. 2019. Stata Statistical Software: Release 16. College Ststion, TX: StataCorp LLC.) was used to conduct all statistical analysis.

Results

Following removal of duplicates, 12,136 publications were identified (Fig. 1). Fifty-seven full text articles were screened. Forty-nine articles were excluded at this stage. See document, supplementary digital content 1 for a list of excluded publications and reasons for exclusion. Thus, six publications reporting five RCTs were included[6,19,20,21,26,38]. The included studies encompassed a total of 428 participants (I = 212, C = 216). Table 3 provides further details regarding the studies.

Methodological quality

Quality scores ranged from 1-6 out of 7 (Table 4). There was a high risk of performance bias due to lack of blinding of participants and personnel (Fig. 2 and 3 produced by using RevMan software (Review Manager. Version 5.3. Copenhagen: The Nordic Cochrane Centre. The Cochrane Collaboration, 2014).

Study outcomes

Jackson and Turner[14] recommend only pooling data where the number of studies is ≥5 to ensure adequate statistical precision. Disability was the only outcome measured consistently in all five included studies, thus our analysis focused solely on this outcome.

The pooled mean group difference in pre/post changes in disability (intervention minus control) was -2.26 units /100 (95% CI: -6.49 to 1.97). See Fig. 4. Between study heterogeneity in mean treatment effect was observed (τ = 2.49; 95% CI: 0.48 to 4.51). The prediction interval revealed that, were investigators to undertake a future trial, the 95% plausible range for mean disability change versus control would be -11.56 to 7.04 units /100.

The pooled point estimate for the inter-individual variability in disability change in response to PNE (SD_{IR}) was 7.36 units /100 (95% CI: -3.93 to 11.12). Substantial between-study heterogeneity was observed (τ = 6.55). The 95% prediction interval for true inter-individual responses was -10.20 to 14.57. Appendix 1 provides a step by step guide for the calculations here.

Using the methods of Swinton et al.[34] we estimated the proportion of responders in the population of interest within each included RCT (Table 5). The threshold reduction in disability for clinical relevance was set at -10/100, in keeping with recent NICE guidelines for back and radicular pain[25]. These proportions were adjusted for

the apparent proportions exceeding this threshold in the comparator groups that were estimated to be due wholly to random variability in the pre to post measurements of disability. It can be seen that these proportions are generally lower than the proportion of participants who exceed the threshold in the intervention groups *per se*.

Discussion

We conducted a systematic review and meta-analysis of the literature in order to quantify the control-group adjusted inter-individual variation in pain, disability and psychosocial outcomes in response to PNE in adults with CMP. Several potential studies did not report the SD of the mean change, and this information could not to be obtained upon request meaning our analysis was restricted to disability.

The inter-individual difference in disability change in response to PNE, as indicated by our SDir of 7.36 /100 units, did not reach our criterion for clinical significance (10 /100 units). Therefore, there is insufficient evidence at present for the existence of inter-individual differences in people's response to PNE over and above random within-subjects variability between baseline and follow-up observations. Although this finding, seems at odds with previous qualitative study findings from our group[17,18,29], that qualitative work focused upon patient experience rather than attempting to objectively quantify inter-individual differences. Considering the upper 95% CI (11.12 /100 units) and wide 95% prediction interval -10.20 to 14.57 of the SDir, any inferences regarding "true" inter-individual responses are unclear. Given the small number of included studies, the wide prediction intervals are unsurprising and this

illustrates the importance of statistical power in any analysis of response heterogeneity[1,2].

Therefore, it is apparent that more high quality RCTs are needed that sufficiently report relevant data. We encourage researchers and reviewers of academic journals to ensure that the means and standard deviations of the change scores in all treatment groups are reported. This will provide the information required to include the study within meta-analyses of both individual responses and mean effect of treatment.

It is worth highlighting that the very common act of simply looking at the intervention group responses (Table 4) would have falsely led a researcher to think that substantial response heterogeneity was present. This may have led to follow-up analyses to explore potential moderators which may be unwarranted and a waste of resources. Furthermore, any follow-up studies on the same participants may be unethical if there are no true individual differences in response present to explain[1].

This is the first systematic review and meta-analysis to employ the method of calculating true inter-individual differences in response to an intervention within the pain sciences[34]. Given the huge global burden of chronic pain, and the limited efficacy of current treatment options for matching peoples' individual responses to treatments, appropriate methodology needs to be applied across the pain field. This will hopefully lead to improved quality of care, reduced costs[33] and ultimately improve the quality of life of people with pain.

Limitations

Only five studies were eligible for this review which meant that we could only analyse disability data and the inter-individual differences in response to PNE for other outcomes are unknown. Six studies that were otherwise eligible, were excluded because they did not report the appropriate data needed to conduct an inter-individual differences meta-analysis and this data was not available upon email request. We have no reason to believe that authors would withhold this data and thus assume these studies are missing at random. Only studies published in English were eligible for inclusion as no facility for translation was available. Thus, important data from non-English studies may have been missed.

The nature of the comparison group will influence the calculation of the inter-individual difference. In the case of usual care comparisons and other intervention comparisons, if these have inherent variability in response within them, beyond random variability (noise) of a true no intervention control, this may mask the degree of interindividual variability seen within the PNE (intervention of interest) group. Thus, this could have influenced the findings. Nevertheless, in the case of intervention vs usual care, if there are true individual differences in the responses to the novel component(s) of the intervention under study, then this should, in theory, manifest itself in a larger change variance in the intervention group vs the usual care group.

Conclusion

This is the first study to investigate “true” inter-individual differences in response within the field of pain. By this, we mean a quantification of response heterogeneity that takes into account the individual differences in baseline to follow-up change that can be observed in the comparator groups, and are attributable to random fluctuation in pain scores over time. Our findings provide little evidence at present of “true” variation in peoples’ response to PNE regarding disability, but the evidence is very uncertain. Furthermore, given the wide 95% confidence and prediction intervals any inferences made regarding true individual variation in peoples’ response to PNE are unclear. Moreover, given the small number of studies included in the analysis further work is warranted before firm conclusions can be drawn. Therefore, the data currently available does not allow us to clearly identify if individual differences in disability occur for people with CMP following PNE. We would recommend against studies exploring which factors may explain which people will benefit from PNE until such time as the existence of inter-individual differences has been confirmed using appropriate methodology and we would extend this recommendation to all pain interventions.

Acknowledgements

This research was funded through Teesside University.

The authors have no conflict of interests to declare.

References:

1. Atkinson G, Batterham AM: True and false interindividual differences in the physiological response to an intervention. *Exp Physiol 100*:577-588, 2015

2. Atkinson G, Williamson P, Batterham AM: Issues in the determination of "responders" and "non-responders" in physiological research. *Physiol* 2019
3. Borenstein M, Higgins JP, Hedges LV, Rothstein HR: Basics of meta-analysis: I2 is not an absolute measure of heterogeneity. Res Synth Methods. 8(1):5-18, 2017
4. Cortés Martínez J, González Alastrué JA, Medina MN, Vogler M, Vilaró Pacheco M, Elmore M, Senn SJ, Campbell MJ, Cobo Valeri E: Does evidence support the high expectations placed in precision medicine? A bibliographic review. *F1000Res 7*:1-11, 2018
5. Furlan AD, Sandoval JA, Mailis-Gagnon A, Tunks E: Opioids for chronic noncancer pain: a meta-analysis of effectiveness and side effects. *CMAJ 174*:1589-1594, 2006
6. Gallagher L, McAuley J, Moseley, GL: A randomized-controlled trial of using a book of metaphors to reconceptualize pain and decrease catastrophizing in people with chronic pain. Clin J Pain 29:20-5, 2013
7. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ: GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 336:924-926, 2008
8. Higgins JPT: Commentary: heterogeneity in meta-analysis should be expected and appropriately quantified. *Int J Epidemiol* 37:1158–1160, 2008
9. Higgins JP, Altman DG, Sterne JA: Chapter 8: Assessing risk of bias in included studies. In: Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org. [Accessed 26/06/2018].

10. Higgins JPT, Deeks JJ, Altman DG (editors). Chapter 16: Special topics in statistics. In: Higgins JPT, Green S (editors), Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from www.handbook.cochrane.org. [Accessed 26/06/2018].
11. IntHout J, Ioannidis JP, Borm GF: The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Med. Res. Methodol. 14*(1):25, 2014
12. IntHout J, Ioannidids JP, Rovers MM, Goeman JJ: Plea for routinely presenting prediction intervals in meta-analyses. *BMJ Open* 6:e010247, 2016
13. Hopkins W: Individual responses made easy. *J Appl Physiol* 118:1444–1446, 2015
14. Jackson D, Turner R: Power analysis for random-effects meta-analysis. *Res Synth Methods* 8:290-302, 2017
15. Kamper SJ, Apeldoorn AT, Chiarotto A, Smeets RJEM, Ostelo RWJG, Guzman J, Van Tulder MW: Multidisciplinary biopsychosocial rehabilitation for chronic low back pain: Cochrane systematic review and meta-analysis. *BMJ 350*;p.h444, 2015
16. King NA, Hopkins M, Caudwell P, Stubbs RJ, Blundell JE: Individual variability following 12 weeks of supervised exercise: identification and characterization of compensation for exercise induced weight loss. *Int J Obes* 32: 177–184, 2008
17. King R, Robinson V, Ryan CG, Martin DJ: An exploration of the extent and nature of reconceptualisation of pain following pain neurophysiology

education: A qualitative study of experiences of people with chronic musculoskeletal pain. *Patient Educ Couns* 99:1389-93, 2016

18. King R, Robinson V, Elliot-Button HL, Watson JA, Ryan CG, Martin DJ: Pain reconceptualisation after Pain Neurophysiology Education in Adults with Chronic Low Back Pain: A Qualitative Study. *Pain Res Treat* 1-10, 2018
19. Louw A, Diener I, Landers MR, Puentedura EJ: Preoperative pain neuroscience education for lumbar radiculopathy: a multicenter randomized controlled trial with 1-year follow-up. *Spine* 39:1449-57, 2014
20. Louw A, Diener I, Landers MR, Zimney K, Puentedura EJ: Three-year follow-up of a randomized controlled trial comparing preoperative neuroscience education for patients undergoing surgery for lumbar radiculopathy. J Spine Surg 2:289, 2016
21. Malfliet A, Kregel J, Meeus M, Roussel N, Danneels L, Cagnie B, Dolphens M, Nijs J: Blended-Learning Pain Neuroscience Education for People With Chronic Spinal Pain: Randomized Controlled Multicenter Trial. *Phys Ther* 98:357-68, 2018
22. Moher D, Liberati A, Tetzlaff J, Altman DG: Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med 151*:264-269, 2009
23. Moseley GL. Reconceptualising pain according to modern pain science. *Phys Ther Rev 12*:169-178, 2007
24. Moseley GL, Butler DS. Fifteen years of explaining pain: the past, present, and future. *J Pain* 16:807-13, 2015
25. National Institute for Health and Clinical Excellence (NICE). Low back pain and sciatica in over 16s: Assessment and management draft. (NICE guideline

NG59). Available at: www.nice.org.uk/guidance/ng59/documents/draftguideline. Accessed October 1, 2018

26. Pires D, Cruz EB, Caeiro C. Aquatic exercise and pain neurophysiology education versus aquatic exercise alone for patients with chronic low back pain: a randomized controlled trial. *Clin Rehabil* 29:538-47, 2015
27. Pires D, Caeiro C, Cruz EB. Individual patient responder analysis of the effectiveness of a pain neuroscience education programme in chronic low back pain. In *IFOMPT Conference* 2016.
28. Rice University: *Online Statistics Education: A Multimedia Course of Study.* 2019. Available at: http://onlinestatbook.com/2/calculators/normal_dist.html (Accessed: 02/09/2019).
29. Robinson V, King R, Ryan CG, Martin DJ: A qualitative exploration of people's experiences of pain neurophysiological education for chronic pain: The importance of relevance for the individual. *Man Ther* 22:56-61, 2016
30. Searle A, Spink M, Ho A, Chuter V: Exercise interventions for the treatment of chronic low back pain: a systematic review and meta-analysis of randomised controlled trials. *Clin Rehabil 29*:1155-1167, 2015
31. Senn S, Rolfe K, Julious SA: Investigating variability in patient response to treatment—a case study from a replicate cross-over study. *Stat Methods Med Res* 20:657–666, 2011
32. Snapinn SM, Jiang Q: Responder analyses and the assessment of a clinically relevant treatment effect. *Trials 8*:31-37, 2007
33. Spear BB, Heath-Chiozzi M, Huff J: Clinical applications of pharmacogenetics. *Trends Mol Med* 7:201–204, 2001

34. Swinton PA, Hemingway BS, Saunders B, Gualano B, Dolan E: A statistical framework to interpret individual response to intervention: paving the way for personalised nutrition and exercise prescription. *Frontiers in nutrition 5*:41, 2018;

35. The British Pain Society. Guidelines for Pain Management Programmes for adults - An evidence-based review prepared on behalf of the British Pain Society. 2nd Ed. The British Pain Society. London: The British Pain Society; 2013. https://www.britishpainsociety.org/static/uploads/resources/files/pmp2013_main_FINAL_v6.pdf. Accessed October 1, 2018.

36. The Joanna Briggs Institute (JBI). The System for the Unified Management, Assessment and Review of Information (SUMARI) is the Joanna Briggs Institute's premier software for the systematic review of literature. JBI SUMARI. https://www.jbisumari.org. 2017. Accessed October 1, 2018.

37. van Aert RC, Jackson D: A new justification of the Hartung-Knapp method for random-effects meta-analysis based on weighted least squares regression. *Res Synth Methods.* 2019

38. van Ittersum MW, Wilgen CP, Schans CP, Lambrecht L, Groothoff JW, Nijs J: Written pain neuroscience education in fibromyalgia: A multicenter randomized controlled trial. *Pain Pract* 14:689-700, 2014

39. Watson JA, Ryan CG, Cooper L, Ellington D, Whittle R, Lavender M, Dixon J, Atkinson G, Cooper K, Martin DJ: Pain neuroscience education for adults with chronic musculoskeletal pain: a mixed-methods systematic review and meta-analysis. *J Pain* 2019

40. Williamson PJ, Atkinson G, Batterham AM: Inter-individual responses of maximal oxygen uptake to exercise training: a critical review. *Sports Med 47*:1501-1513, 2017
41. Williamson PJ, Atkinson G, Batterham AM: Inter-individual differences in weight change following exercise interventions: a systematic review and meta-analysis of randomized controlled trials. *Obes Rev 19*:960-975, 2018
42. Wolfe F, Smythe HA, Yunus MB, Bennett RM, Bombardier C, Goldenberg DL, Tugwell P, Campbell SM, Abeles M, Clark P, Fam AG: The American College of Rheumatology 1990 criteria for the classification of fibromyalgia: report of the Multicenter Criteria Committee. *Arthritis Rheum* 33:160-72, 1990

Table 1 Summary of findings:

PNE compared to control for treatment of adults with chronic musculoskeletal pain

Patient or population: treatment of adults with chronic musculoskeletal pain
Setting:
Intervention: PNE
Comparison: control

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with control	**Risk with PNE**				
Change in disability score in the short term. (ST Disability) assessed with: Validated measure of disability converted to percentage Scale from: 0 to 100 (worse)	The mean change in disability score in the short term. was **-8.63** units	mean **2.26 units lower** (6.49 lower to 1.97 higher)	-	428 (5 RCTs)	⊕◯◯◯ VERY LOW a,b,c,d,e,f,g,h	PNE may reduce/have little to no effect on change in disability score in the short term. but the evidence is very uncertain.

***The risk in the intervention group** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval

GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Explanations

a. A large proportion of the weight came from a study where there was concern over selection bias, performance bias, attrition bias, reporting bias and other bias. There was concern with most studies over performance bias which whilst normal of these types of studies may still impact the results.
b. Some variation is size of the effect, however the difference between studies does not reach a clinically meaningful difference
c. Good overlap of the confidence intervals.
d. I-Squated above 50%
e. Tau-Squared higher than point estimate.
f. Sample of chronic musculoskeletal pain comparing PNE against control using an appropriate outcome measure.
g. Has over 400 participants but imprecise due to prediction interval including null effect and clinically important benefit.
h. A comprehensive search was conducted on electronic databases and trials registries. References lists and citing articles of included studies were searched to identify any further articles.

Table 1 Legend: Summary of findings, PNE compared to control for treatment of adults with chronic musculoskeletal pain

Table 2 Summary of findings:

Do inter-individual differences in disability change in response to PNE exist in adults with chronic musculoskeletal pain?

Patient or population: treatment of adults with chronic musculoskeletal pain
Setting:
Intervention: PNE
Comparison: control

Outcomes	**Estimated absolute inter-individual difference in response** (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
Inter-individual variability in disability change in the short term. SD_{IR} assessed with: Validated measure of disability converted to percentage Scale from: 0 to 100 (worse)	mean **7.36 units** (3.93 lower to 11.12 higher)	428 (5 RCTs)	⨁◯◯◯ VERY LOW [a,b,c,d,e,f,g]	Little evidence of "true" variation in peoples' response to PNE for disability, but the evidence is very uncertain.

CI: Confidence interval

GRADE Working Group grades of evidence
High certainty: We are very confident that the true difference in response lies close to that of the estimate of the difference in response
Moderate certainty: We are moderately confident in the difference in response estimate: The true difference in response is likely to be close to the estimate of the difference in response, but there is a possibility that it is substantially different
Low certainty: Our confidence in the difference in response estimate is limited: The true difference in response may be substantially different from the estimate of the difference in response
Very low certainty: We have very little confidence in the difference in response estimate: The true difference in response is likely to be substantially different from the estimate of difference in response

Explanations

a. A large proportion of the weight came from a study where there was concern over selection bias, performance bias, attrition bias, reporting bias and other bias. There was concern with most studies over performance bias which whilst normal of these types of studies may still impact the results.
b. Some variation in size of the effect, however the difference between studies does not reach a clinically meaningful difference
c. Good overlap of the confidence intervals.
d. Tau-Squared higher than point estimate.
e. Sample of chronic musculoskeletal pain comparing PNE against control using an appropriate outcome measure.
f. While the analysis includes over 400 participants this lack precision due to the very wide prediction interval including both a clinically important positive effect and clinically important negative effect.
g. No evidence of publication bias. Sample sizes ranged from 62-120. A comprehensive search was conducted on electronic databases and trials registries. References lists and citing articles of included studies were searched to identify any further articles.

Table 2 Legend: Summary of findings, Do inter-individual differences in disability change in response to PNE exist in adults with chronic musculoskeletal pain?

628 Table 3 Characteristics of included studies
629

Study	Methods	Sample size (baseline)/ gender/ mean age in years	Participants	Intervention(s)	Duration of educational intervention	Control	Authors conclusions/notes	Setting/country
van Ittersum et al. 2013[38]	RCT	N = 105 7% M 46.7	Fibromyalgia diagnosed according to The American College of Rheumatology 1990 criteria[42]. 18-65 years of age. Baseline pain as mean % = 71.5% Duration of pain in mean months = unknown	Written PNE + 1 phone call for motivation/questions +/- 2x phone calls/emails for further clarification/questions	Unknown	Written Relaxation exercises + 1 phone call for motivation/questions +/- 2x phone calls/emails for further clarification/questions	Written PNE alone is not effective for changing the impact of the illness on daily life, pain catastrophising, or illness perceptions in fibromyalgia patients.	Specialised centres for chronic pain and chronic fatigue. Belgium.
Gallagher, McAuley	RCT	N = 79 39% M	18-75 years of age with pain that had	80-page booklet divided into 11 sections -	Unknown	80-page booklet divided into 11	Written material using metaphors to explain key	Unknown Unknown

and Moseley 2013[6]		43.5	been sufficient to disrupt their activities of daily living for more than the previous 3 months. Baseline pain as mean % = 65% Duration of pain in mean (SD) months = 28 (19.5)	Metaphors and stories to help understand the biology of pain		sections - Advice about managing pain (The back book and Manage your pain)	biological concepts increased knowledge of pain biology and decreased catastrophic thought processes about pain and injury when compared to material that presented biopsychosocial advice for pain management.	
Pires, Cruz and Caeiro, 2015[26]	RCT	N = 62 35% M 51	Low back pain >3 months duration +/- leg pain. 18-65 years of age. Baseline pain as mean % = 42.9% Duration of pain in mean (SD) months = unknown	2x 1.5h Group PNE. 12 sessions of aquatic exercise over 6 weeks. 30-50m each session.	PNE 3h Control 3h	12 sessions of aquatic exercise over 6 weeks. 30-50m each session.	PNE is a clinically effective addition to aquatic exercise. The addition of PNE resulted in statistically significant reduction in pain intensity at 3-month follow up. No statistically significant differences were found for pain intensity at 6 weeks follow up or functional disability at either follow up.	Outpatient clinic. Portugal
Louw et al. 2014/16[19,20]	RCT	N = 67 46% M	Patients with lumbar	0.5h individual PNE.	PNE 0.5h	Lumbar surgery alone + usual care	Providing a single PNE session to patients prior to lumbar	7 Clinical sites in the US.

		49.6	radiculopathy, scheduled for lumbar surgery. 18-65 years of age. Baseline pain as mean % = 48.4% Duration of pain in mean (SD) months = 3 (7.5)	PNE booklet "your nerves are having back surgery" & Lumbar surgery + usual care	Control 0		surgery results in significant reduction in healthcare costs 3-years after LS.	
Malfliet et al. 2018[21]	RCT	N = 120 39.2% M 39.8	Non-specific chronic spinal pain (neck and lower back) at least 3 days a week for at least 3 months since the first symptoms. 18-65 years of age Baseline pain as mean % = 50.65	3 PNE sessions 1. 0.5-1h group (maximum of 6 patients). Information booklet provided at the end. 2. ~0.63h home-based online e-learning module containing 3 explanatory videos and	PNE 1.88h Control 1.88h	3 biomedical education sessions 1. 0.5-1h group (maximum of 6 patients). Information booklet provided at the end. 2. ~0.63h Home-based online e-learning module containing 3 explanatory videos	PNE, and not neck/back school education, is able to improve kinesiophobia, beliefs regarding the negative impact of the illness on quality of life and functional capacity, and beliefs regarding the chronicity of pain and the time scale of illness symptoms. However, none of the educational programs of this study were able to decrease the participants perceived disability due to pain. Nevertheless, as kinesiophobia is generally	University hospitals in Ghent and Brussels, Belgium.

			Duration of pain in mean (SD) months = 82 (143.25)	questions about pain. 3. 0.5 Individual education. Focus on patients' personal needs following difficulties with session 2. Focus on the application of knowledge to participants life.		3. 0.5 Individual. Focus on patients' personal needs following difficulties with session 2. Focus on the application of knowledge to participants life.	considered to be a strong predictor and mediator of chronic pain, PNE is preferred as the educational approach for people with non-specific chronic spinal pain.	

Table 3 Legend: Randomized controlled trial, RCT. Male,

Table 3 Legend: Characteristics of included studies. PNE, Pain neuroscience education. SD, Standard deviation. RCT, Randomised controlled trial

Table 4 Critical appraisal of quantitative studies

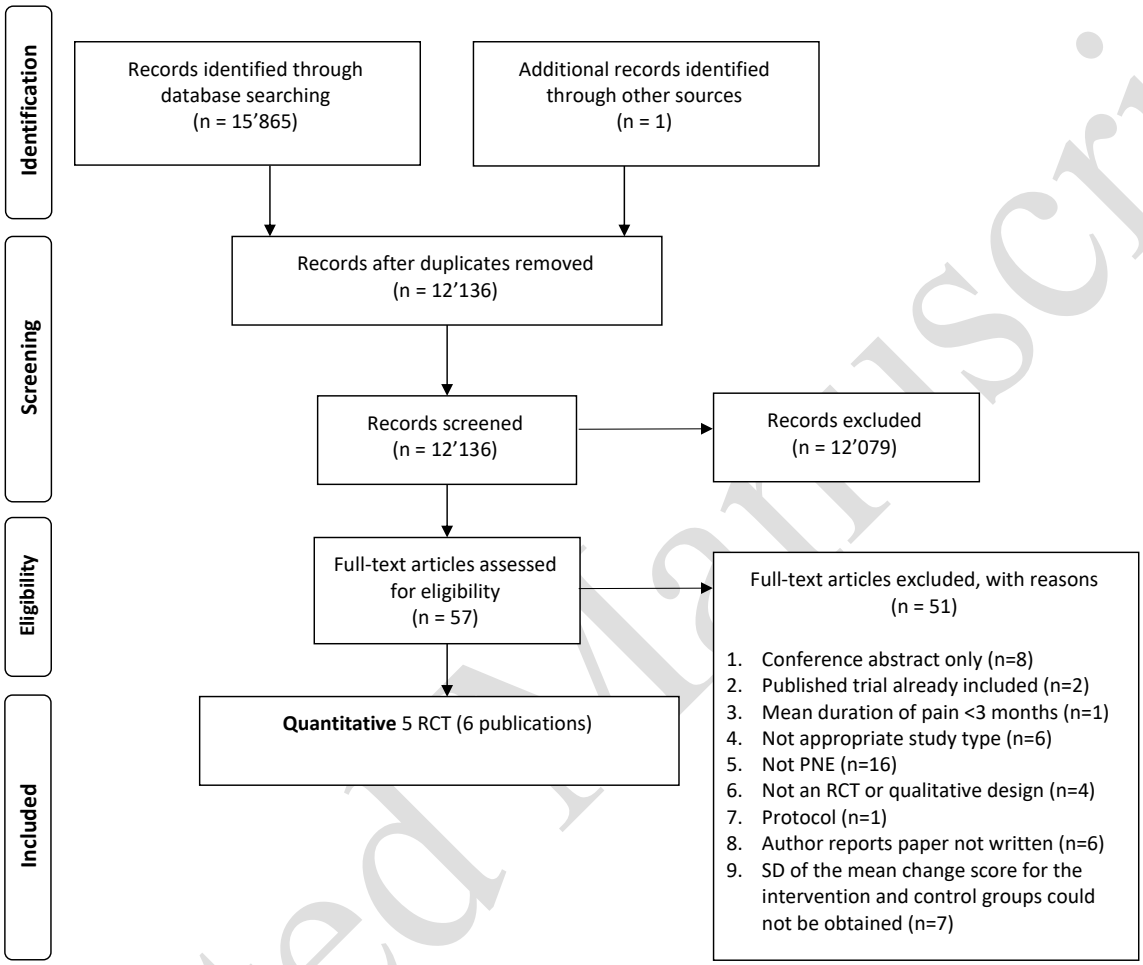
Study	Score /7	Percentage
Gallager 2013[6]	5	71%
Louw 2014/16[19,20]	3	43%
Malfliet 2018[21]	6	86%
Pires 2015[26]	3	43%
van Ittersum 2013[38]	1	14%

Figure 4 Legend: Forest plot of PNE versus control in the short term; primary outcome disability mean difference.

Table 5: Proportions of responders.

Study	Mean Change (PNE)	SD (PNE)	% responders (PNE)	Mean change (Con)	SD (Con)	% Responders (Con)	Mean treatment effect (PNE-Con)	SD for true Ind diffs	% Responders based on SDir[34]
van Ittersum et al. 2013[38]	0.7	4.2	0	0.3	2.9	0	0.4	3.0	0
Pires, Cruz and Caeiro, 2015[26]	-11.1	15.8	53	-7.7	10.6	41	-3.4	11.7	29
Louw et al. 2014/16[19,20]	-12.0	18.5	54	-11.1	13.8	53	-0.9	12.3	23
Malfliet et al. 2018[21]	-1.1	13.8	26	1.6	11.2	15	-2.7	8.1	18
Gallagher, McAuley and Moseley 2013[6]	-36	17	94	-27.0	15.0	87	-9.0	8	45

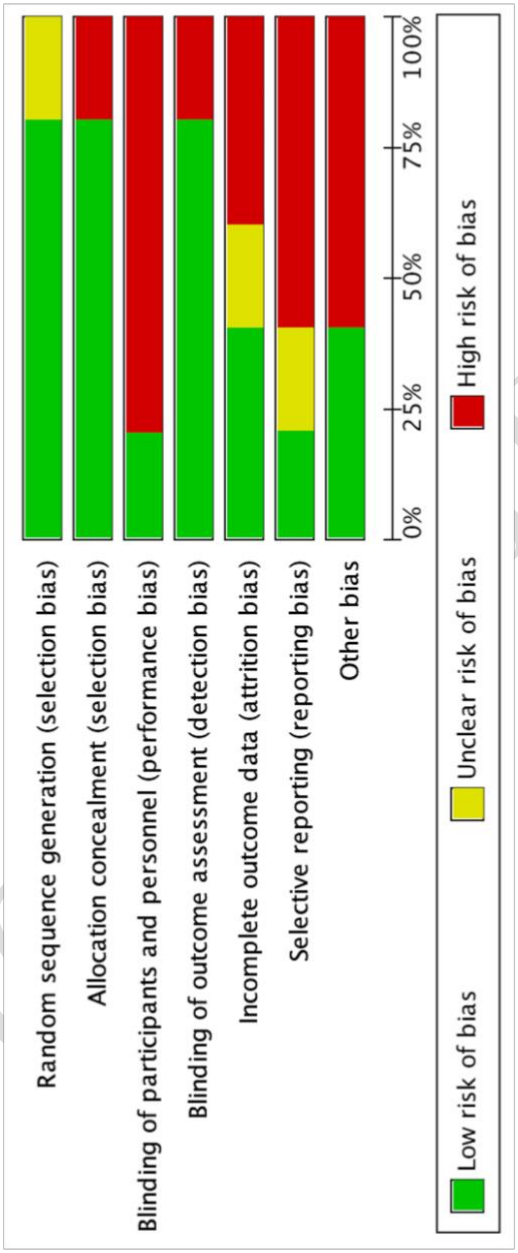
Table 5 Legend: Proportions of responders. PNE, Pain neuroscience education. Con, Control. SD, Standard deviation. SDir, Standard deviation for individual responses.



646

647 *Figure 1 Legend: PRISMA flow diagram of search and study selection process.*

648 *(Adapted from Moher et al.[22]).*

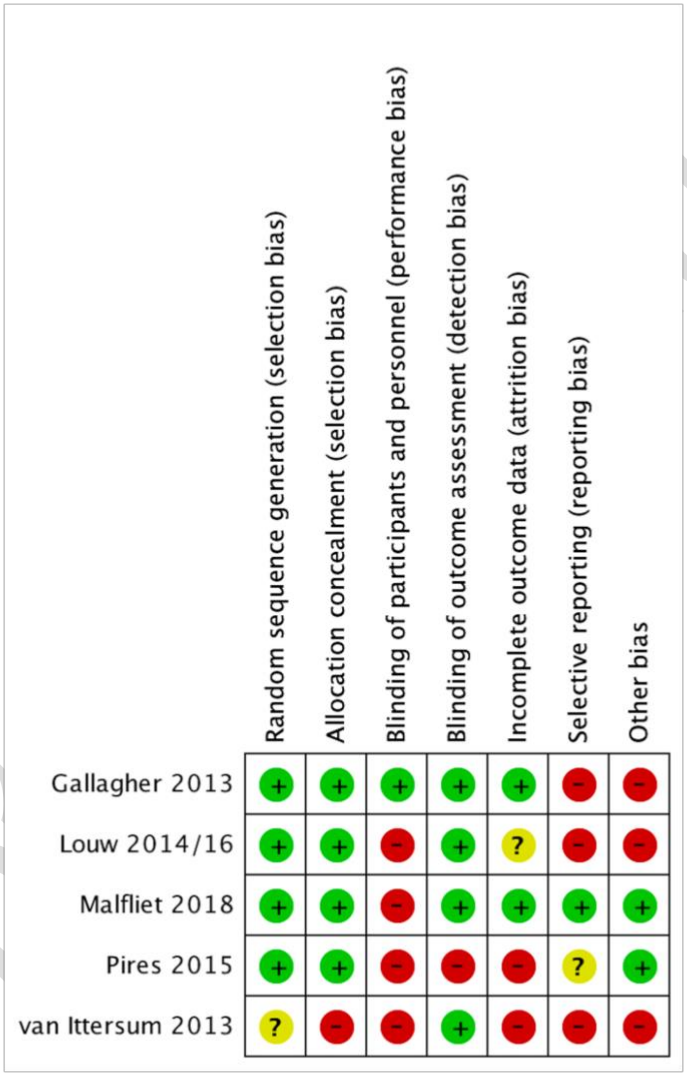


649

650 *Figure 2 Legend: Risk of bias graph.*

651

652



653

654 *Figure 3 Legend: Risk of bias summary*

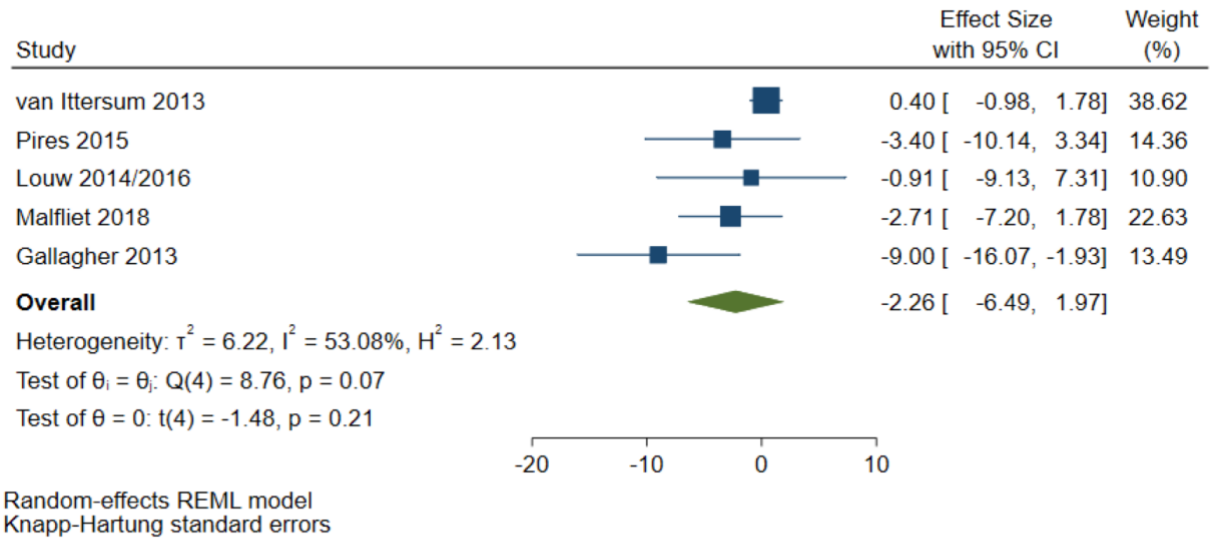


Figure 4 Legend: Forest plot of PNE versus control in the short term; mean difference of disability between groups.

Supplementary Appendix 1 - Calculations for inter-individual differences meta-analysis

Step 1

We extracted the standard deviation (SD) of the changes in disability for both control (C) and PNE (I) groups.

Study	SDC	Mean I
van Ittersum	2.9	0.7
Gallagher	15	-36
Pires	10.6	-11.1
Louw	13.79	-12
Malfliet	11.15	-1.1

Step 2

The true individual response variance (intervention minus control) was then calculated by $\sqrt{(SD_I^2-SD_C^2)}$ (Hopkins, 2015).

Study	IR_Variance	SDI	SDC
van Ittersum	9.23	4.2	2.9
Gallagher	64	17	15
Pires	137.28	15.8	10.6
Louw	152.09	18.5	13.79
Malfliet	65.84	13.79	11.15

Step 3

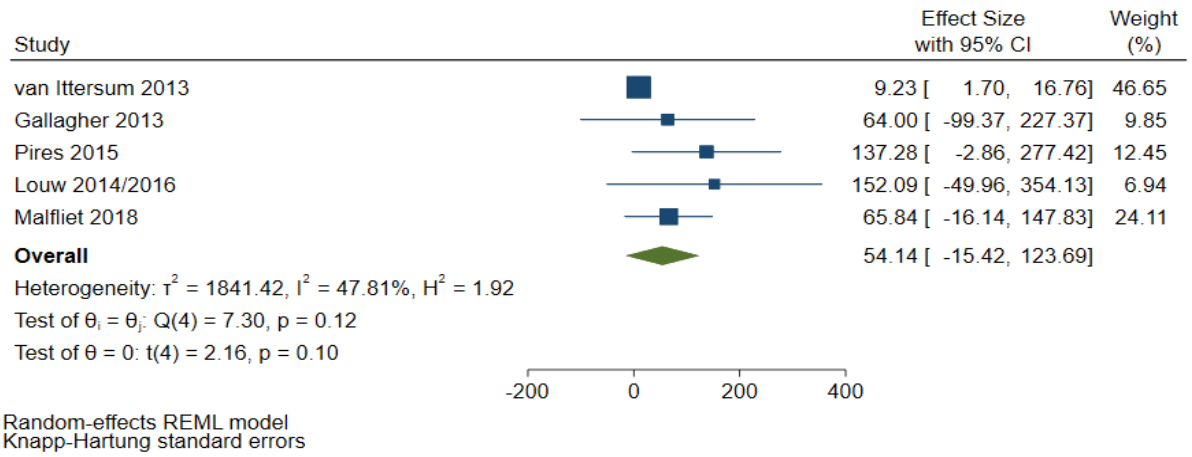
The standard error (SE) for this variance was then calculated using the equation: $SE = \sqrt{[2(SD_I4/DF_I + SD_C4/DF_C)]}$, where DF_I and DF_C are the degrees of freedom of the standard deviation in the PNE group and the control groups (Hopkins, 2015).

Study	IR_Variance	SE	SDI	SDC	n I	n C
van Ittersum	9.23	3.83949378	4.2	2.9	53	52
Gallagher	64	83.3522758	17	15	40	39
Pires	137.28	71.5013373	15.8	10.6	30	32
Louw	152.09	103.087047	18.5	13.79	29	33
Malfliet	65.84	41.8303551	13.79	11.15	60	60

Step 4

The individual response variances and their SEs were meta-analysed using an REML model combined with Knapp-Hartung method. It's important to highlight that the variances are unbiased, whereas the SD is not, and deriving a SE for the SD for individual responses is also problematic. Thus, we synthesised the individual response variances instead of the SDs for individual responses. The point estimate for the pooled individual response variance were derived together with a 95% CI to express its uncertainty.

Forest plot of **Variance** Meta-analysis for estimating individual differences in response:



Step 5

The point estimate and CIs were then square rooted to convert to an SD metric. If the lower limit was negative, the sign was ignored, the square root taken, and the sign re-applied. This approach is consistent with the 'no bound' option in SAS/STAT® software, which permits negative variances (SAS Institute Inc. 2017. SAS/STAT 14.3 User's Guide. Cary, NC: SAS Institute Inc.).

	As variance	SD without sign	As SD with sign re-applied
Total point estimate	54.14		7.35798886
Lower CI -	15.42	3.92683078	-3.9268308
Upper CI	123.69		11.1216006

Steps to calculate the prediction interval for the inter-individual differences point estimate

PI = pooled estimate +/- $t_{(n-2)}$ x SQRT(SE^2 + tau^2)

Pooled Est 54.14

3.182 is the two-tailed t value for n-2 degrees of freedom = 3 degrees of freedom, and P=0.05. See: http://www.ttable.org/student-t-value-calculator.html

$t_{(n-2)}$ = 3.182
SE= 25.0508232
SE^2 = 627.543743 (From STATA)
tau^2 = 1841.4235

(SE^2 + tau^2) = 2468.96724
SQRT(SE^2 + tau^2) = 49.6887034

PI =	Pooled est	+/-	t(n-2)	x	SQRT(SE2 + tau2)
PI =	54.14	+/-	3.182	x	49.6887034
PI =	3.182	+/-	158.109454		

PI Upper = 212.249454
PI Lower = -103.96945

Square root the above values to convert from variance to SD to get to the PI for the SDir:

PI Upper = 14.5687836
PI Lower = -10.196541