Physiotherapy for pain: a meta-epidemiological study of randomised trials

Elisabeth Ginnerup-Nielsen,[1] Robin Christensen,[1] Kristian Thorborg,[2] Simon Tarp,[1] Marius Henriksen[1,3]

[1]The Parker Institute, Copenhagen University Hospital Bispebjerg & Frederiksberg, Copenhagen, Denmark
[2]Sports Orthopedic Research Center Copenhagen (SORC-C), Arthroscopic Centre Amager, Copenhagen University Hospital, Copenhagen, Denmark
[3]Department of Physio- & Occupational Therapy, Copenhagen University Hospital Bispebjerg & Frederiksberg, Copenhagen, Denmark

Correspondence to
Professor Marius Henriksen, The Parker Institute & Department of Physio- & Occupational Therapy, Copenhagen University Hospital Bispebjerg & Frederiksberg, Nordre Fasanvej 57, Copenhagen F 2000, Denmark; marius.henriksen@regionh.dk

ABSTRACT

Objectives To empirically assess the clinical effects of physiotherapy on pain in adults.

Design Using meta-epidemiology, we report on the effects of a 'physiotherapy' intervention on self-reported pain in adults. For each trial, the group difference in the outcome 'pain intensity' was assessed as standardised mean differences (SMD) with 95% CIs. Stratified analyses were conducted according to patient population (International Classification of Diseases-10 classes), type of physiotherapy intervention, their interaction, as well as type of comparator group and risks of bias. The quality of the body of evidence was assessed based on GRADE methodology.

Data sources Systematic searches were carried out in MEDLINE and PEDro from 1 January 2004–31 December 2013. 174 trials (224 comparisons) met the inclusion criteria for the meta-analysis.

Eligibility criteria for selecting studies Randomised trials using 'no intervention' or of a sham-controlled design were selected. Only articles written in English were eligible.

Results An overall moderate effect of physiotherapy on pain corresponding to 0.65 SD-units (95% CI 0.57 to 0.73) was found based on a moderate inconsistency (I^2=51%). Stratified exploration showed that therapeutic exercise for musculoskeletal diseases tended to be more beneficial than multimodal interventions (difference 0.30 95% CI 0.03 to 0.57; p=0.03). Trials with a 'no intervention' comparator tended to have a higher overall effect size than trials with a sham comparator (difference 0.25; 95% CI 0.09 to 0.41; p=0.004). In general, our confidence in the estimates was low, mainly due to high risk of performance biases and between-study heterogeneity.

Conclusions Physiotherapy reduces pain in adults, but standardisation of interventions and focus on trial research with low risks of bias and reproducible treatment modalities are needed.

Trial registration number CRD42014008754.

INTRODUCTION

Pain and pain-related deficiencies affect an enormous part of the world's adult population.[1][2] Pain is complex, and has many treatment options, of which pharmacological treatments are the most common.[3] However, over recent decades, pain management has changed and non-pharmacological interventions have gained focus and interest.[3] Among these, physiotherapy is common.[4][5] The increased use and belief in the efficiency of physiotherapy as a pain treatment with no or limited safety concerns has led to physiotherapies being included in several international recommendations on management of major painful diseases such as low back pain and osteoarthritis.[6][7]

In contrast to many specialties of medicine, physiotherapy lacks an identifiable and clearly defined organ system to convene around and 'physiotherapy' is therefore not easily defined. The 'World Confederation of Physical Therapists' (WCPT) describes physiotherapy as "… services provided in circumstances where movement and function are threatened by ageing, injury and pain",[8] which can result in the broad definition: 'physiotherapy is what physical therapists do'. By consequence, a large range of physiotherapy treatment options exists, but scientific and practical information that can guide treatment choices for individual patients is often elusive causing variations in clinical practice.

In 2008, Consolidated Standards of Reporting Trials (CONSORT) had already proposed an extension of the CONSORT Statement for trials of non-pharmacological treatments,[9] and Helmhout et al[10] addressed methodological issues of non-pharmacological studies and, among other things, proposed: "Specification of a theoretical framework for a therapy when designing the intervention, identification of subgroups of patients most likely to benefit from treatment, and categorisation of interventions, in terms of concept, mode, intensity, duration, frequency, and length."

Still, variability in physiotherapy is reflected in clinical research where 'routine physiotherapy' is often used as experimental treatment or comparator, presumably to reflect daily clinical practice. For example, in a study of acupressure for low back pain, routine physiotherapy was the comparator and included 'pelvic manual traction, spinal manipulation, thermotherapy, infrared light therapy, electrical stimulation and exercise therapy, as decided by the physical therapist'.[11] In a study of physiotherapy for hip osteoarthritis, the provided physiotherapy comprised '…core components plus optional techniques and exercises depending on assessment findings'.[12] These examples highlight the variability in physiotherapy observed across countries but also between (and within) institutions in the same country.

In clinical trial research, this generates validity issues because the mediating factor rather than any specific modality becomes the individual therapist.[13] In daily clinical practice, these possible variations in the provided services complicate the estimation of what benefit patients may expect regarding their painful condition. While standardisation of physiotherapy for painful conditions is

pending, the provisional question, 'is physiotherapy generally effective in management of pain?' is reasonable to ask.

The objective of this study was to assess the reported effects of physiotherapy on populations of patients with pain. Using meta-epidemiology, we report on the association of estimates of treatment benefits according to different stratifications from the available evidence.

METHODS

The methods of this study's search strategy, inclusion criteria and data analysis were prespecified in a protocol (see online supplementary file 1) that was registered before any study-related activities were started (PROSPERO 2014: CRD42014008754).

Data sources and searches

The bibliographic databases MEDLINE and PEDro were searched from inception to the 2 of March 2014.

During full-text reviewing, it became clear that the reporting quality of many older trials was poor and the practised interventions were outdated. It was therefore decided to focus on trials published after implementation of the CONSORT statement in most medical societies.[14] Assuming a reasonable reporting standard and recently updated treatment modalities after implementation of the CONSORT statement, we focused our review on articles published after 1 January 2004.

Search terms included free-text and MeSH terms related to physical therapy, physiotherapy, physiotherapy modalities and pain. The MEDLINE search was combined with the 'Cochrane highly sensitive search strategy for identifying randomised trials in MEDLINE: sensitivity-maximising and precision-maximising version (2008 revision); based on the Ovid form',[15] modified with the search terms 'sham' and 'quasi' to make the search strategy more sensitive to trials that were quasi randomised or included sham treatments. (Full search strategy is available in online supplementary file 2).

Study selection

We included randomised controlled trials assigning adults (≥18 years of age) with pain to a physiotherapy intervention compared to a control group receiving either sham treatments (placebo) or no intervention. We predefined physiotherapy as 'any treatment modality delivered by one or more physical therapists (or physiotherapists) or explicitly defined as physiotherapy (or physical therapy) by the authors of the article'. We only included randomised controlled trials reporting pain intensity as an outcome. Only articles written in English were eligible.

Selection was made by two independent reviewers (EGN and MH), who screened titles and abstracts and reviewed full text of potentially eligible articles. Any disagreements were resolved through discussion. Study selection, assessment of eligibility criteria, data extraction and statistical analyses were performed according to the 'Cochrane Collaboration Methodological Expectations for Cochrane Intervention Reviews' (MECIR) guidelines.[16]

Data extraction and quality assessment

All data were extracted on pretested forms and entered into an Excel spreadsheet, and included: First author, year of publication, journal, country of origin, number of participants, source of financial support, mean age, gender, condition treated, type of intervention and control intervention, pain assessment scale, study duration from baseline to assessment of primary pain outcome at follow-up and study-attrition. One reviewer (EGN) performed the extraction, with a second reviewer (MH) checking the data for accuracy.

Risk of bias as either 'low', 'unclear', or 'high' was assessed for the pain outcome per study using 'Cochrane's Collaboration's tool for assessing risk of bias in randomised trials'.[17] The domains 'Sequence generation', 'Allocation concealment', 'Blinding of participants', 'Blinding of therapists', 'Incomplete outcome data reporting' and 'Selective outcome reporting' were assessed. Considering the importance of minimising performance bias when the outcome is subjective,[18] blinding was evaluated based on 'blinding of participants' and 'blinding of treating therapists' instead of the default 'Blinding of participants and personnel' and 'Blinding of outcome assessors'.

Using the: 'Grading of Recommendations Assessment, Development and Evaluation' (GRADE) system for grading evidence, the quality of the total body of evidence concerning physiotherapy treatment for pain was assessed through: study limitations (risk of bias),[19] consistency of results (eg, I^2 statistics),[20] indirectness[21] and imprecision (95% CI).[22]

One reviewer (EGN) assessed risk of bias and graded the quality of the body of evidence. In case of doubt, a second reviewer (MH) was consulted and disagreements were resolved by consensus.

Data synthesis and analysis

Pain intensity was the prespecified outcome of interest and had to be assessed on a recognised scale at baseline and follow-up. Assuming that the first follow-up assessment would show the most pronounced improvement on pain, this time point was chosen as default. If several pain scale instruments had been used, the one referred to as the primary pain outcome or the first reported pain-outcome, was chosen.

Study outcomes were expressed as standardised mean differences (SMDs) calculated from each trial based on means and SDs, preferably based on the change from baseline to follow-up. If change data were not stated, follow-up values for each group were applied, assuming that there was no relevant difference between follow-up and change data SMDs.[23] Before performing meta-analysis, the (Cohen'SD) SMD values were converted into Hedges' g value (ie, referred to as bias-corrected SMD)[24] with positive values favouring physiotherapy. Using generic inverse variance analysis, a pooled effect size was calculated using a random effects model. The primary advantage of SMD is the fact that it provides a common metric and is an intuitive index for many people.[25] SMD values of 0.20, 0.50 and 0.80 correspond to small, medium and large effects in the social sciences.[26]

Including several physiotherapy modalities in our review, we anticipated considerable between-study heterogeneity. Therefore, the estimation of the pooled treatment outcomes was given using four meta-analytic strategies to test the robustness of the estimates.[27]

The first strategy was a random effects meta-analysis of all trials. The second strategy was inference restricted to the 25% of largest trials (number of participants). The third strategy included the SE of each study estimate as a covariate. The fourth strategy was predicting an adjusted pooled estimate of effect from the regression line, the pooled effect size for an ideal study of infinite size (hence with zero SE).[28] A priori, we defined a relevant study level covariate as one that would decrease the between-study variance, as a consequence of inclusion in the (mixed-effects) statistical model.[29]

Stratified analyses were performed according to type of physiotherapy approaches and disease populations (8

populations defined by International Classification of Diseases Tenth Edition (ICD-10) from the WHO[30] and 6 major physiotherapy approaches adopted from the World Confederation for Physical Therapy (WCPT).[8] Populations were: musculoskeletal diseases, nervous system, neoplasms, genitourinary system, mental disorders, pregnancy/childbirth, external causes and other populations. Interventions were: therapeutic exercise, passive therapy, mechanical modalities, electrotherapeutic modalities, education and other interventions. The stratified analyses were conducted using the meta-analytic strategy associated with the lowest between study heterogeneity (based on the I^2 statistics and Cochran's Q test.[31] In addition, the interaction between physiotherapy approaches and disease populations was analysed and stratifications were given on each risk of bias item, dividing these in 3 levels: low, unclear and high risk.

All statistical models were based on Restricted Maximum Likelihood (REML).[32] Analyses were performed using SAS software (V.9.3, by SAS Institute Inc, Cary, North Carolina, USA) and Review Manager (V.5.2).

RESULTS
Literature search and study selection
The search identified 5159 unique potentially eligible references. The initial screening revealed 1355 full text articles of which 770 articles were published from 1 January 2004 to 31 December 2013. The final selection of randomised controlled trials consisted of 174 trials of which 43 trials included more than one comparison. All in all, 224 comparisons with 14 687 participants were included. The search is visualised in the flow diagram (figure 1).
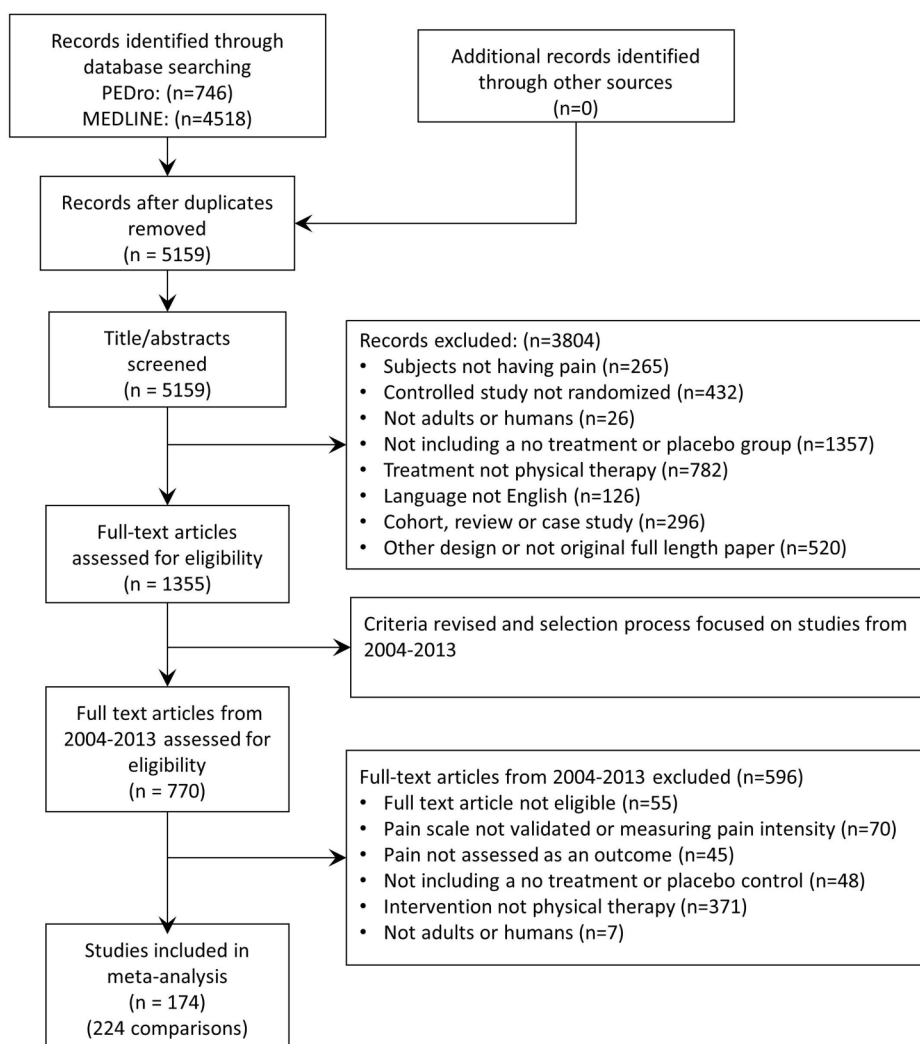
During the data extraction process, no trials investigating mental disorders or genitourinary diseases were found. Twenty-three studies assessed combined interventions including at least two of the following: education, therapeutic exercise, passive therapy or electrotherapeutic treatment in one intervention. An intervention factor named 'multimodal modalities' was therefore included in the analysis.

Variations in the descriptions of similar physiotherapy approaches were observed and intervention lengths varied from a few days to 1 year within the modalities education, therapeutic exercise and electrotherapeutic modalities, and from a few days to 3 months within mechanical modalities and passive therapy. Multimodal modalities, the most diverse approach, varied from 2 weeks to 1.5 years. (see characteristics of the included studies in online supplementary file 3).

Data synthesis and stratified analyses
The first meta-analysis strategy (random effects model) showed an effect on pain in favour of physiotherapy compared to sham treatments or no intervention, with an overall statistically significant SMD of 0.58 (95% CI 0.49 to 0.68; p<0.0001), that is, a 'moderate'—and clinically relevant—effect size. The relatively

Figure 1 PRISMA flow diagram.

narrow CI indicates a fairly precise effect estimate, but the analysis also showed substantial between-study heterogeneity ($\tau^2=0.41$) with an I^2 of 80% (table 1). When focusing the random effects model on the 25% largest trials (strategy 2) the effect size was reduced to 0.45 SMD (95% CI 0.26 to 0.64; p<0.0001) with substantial heterogeneity across studies ($\tau^2=0.48$) corresponding to an I^2 of 82%. This strategy supports the notion of small study effects, that is, a tendency of small trials to report larger benefits than larger trials (typically due to publication biases, less methodological rigour and/or individual therapist effects in smaller trials compared to larger trials).

The random effects model with SE applied as a covariate (strategy 3) showed a higher numerical effect size (SMD=0.65; 95% CI 0.57 to 0.73; p<0.0001) in favour of physiotherapy. The statistical model fit of this strategy was better than the first strategy as indicated by the change in τ^2 from 0.41 to 0.26 and lower between-study heterogeneity (I^2: 51%; table 1).

The fourth strategy repeated the random effects model with SE as covariate but estimated the effect from an imaginary trial of infinite size (SE=0) and revealed no clinical value of physiotherapy (ie, model predicts an effect on pain in favour of the control group; −0.51 SMD (95% CI −0.74 to −0.28; p<0.0001); table 1), further supporting the notion of a small study effect (ee figure 2 for association between each study SMD and SE of SMD).

The meta-analytical strategy adding individual study estimate SE as a covariate (strategy 3) was associated with the lowest between-study heterogeneity and a better statistical model fit (table 1). Accordingly, stratifications and interaction analyses were carried out using this model.

Comparing outcomes between trials with high, unclear and low risk of bias

More than 50% of the comparisons had unclear or high risk of bias within one or more domains: high or unclear risks of bias due to lack of blinding of participants and treating therapists was present in 117 comparisons (52%), while inadequate or unclear sequence generation and allocation concealment was present in 121 comparisons (54%). Unclear outcome data reporting was observed in 148 comparisons (66%). Six comparisons had no domains with high risk of bias, and 30 comparisons had low risk of bias in the domains sequence generation, allocation concealment and outcome data reporting (see online supplementary file 3). Within the domain 'blinding of participants', effect sizes tended to be less beneficial in the 40 comparisons with low risk of bias compared with the 137 comparisons with high risk, corresponding to a difference in effect of: SMD 0.41 (95% CI 0.19 to 0.63; p<0.0001). Low risk of bias within outcome data reporting did not have significant impact on effect sizes (table 1).
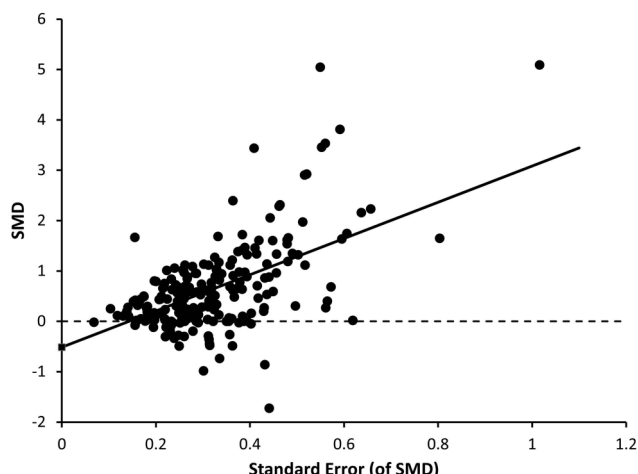
Table 1 Meta-analytic strategies and stratifications

Stratified analyses	Number of comparisons	SMD (95% CI)	p Value	τ^2	I^2*	p Value for interaction
Random effects model (strategy 1)	224	0.58 (0.49 to 0.68)	0.0000	0.41	80%	–
25% largest trials (strategy 2)	57	0.45 (0.26 to 0.64)	0.0000	0.48	82%	–
SE as covariate (strategy 3)	224	0.65 (0.57 to 0.73)	0.0001	0.26	(51%)†	–
Infinite trial size (SE=0) (strategy 4)	224	−0.51 (−0.74 to −0.28)	0.0000	0.26	(51%)†	–
Comparator				0.24		0.004
Sham	76	0.48 (0.35 to 0.61)	0.0000			
No intervention	148	0.73 (0.64 to 0.83)	0.0000			
Sequence generation				0.26		0.43
Low risk	161	0.67 (0.58 to 0.77)	0.0000			
Unclear	55	0.60 (0.43 to 0.76)	0.0000			
High risk	8	0.42 (−0.01 to 0.85)	0.0551			
Allocation concealment				0.26		0.37
Low risk	114	0.70 (0.59 to 0.82)	0.0000			
Unclear	28	0.58 (0.35 to 0.81)	0.0000			
High risk	82	0.59 (0.46 to 0.72)	0.0000			
Blinding participants				0.23		0.00
Low risk	40	0.34 (0.16 to 0.52)	0.0002			
Unclear	47	0.60 (0.43 to 0.77)	0.0000			
High risk	137	0.75 (0.66 to 0.85)	0.0000			
Blinding personnel				0.25		0.06
Low risk	12	0.31 (−0.03 to 0.64)	0.0728			
Unclear	12	0.86 (0.53 to 1.19)	0.0000			
High risk	200	0.65 (0.57 to 0.74)	0.0000			
Incomplete outcome data				0.26		0.86
Low risk	80	0.63 (0.50 to 0.77)	0.0000			
Unclear	50	0.69 (0.52 to 0.86)	0.0000			
High risk	94	0.64 (0.51 to 0.76)	0.0000			
Selective outcome reporting				0.26		0.27
Low risk	26	0.78 (0.55 to 1.00)	0.0000			
Unclear	171	0.65 (0.56 to 0.74)	0.0000			
High risk	27	0.51 (0.28 to 0.74)	0.0000			

*Calculated in Review Manager.
†I^2 Derived from the overall I^2 and τ^2. All other values were calculated in SAS.
SMD, standardised mean differences; τ^2 (τ^2), estimated between study variance.

Figure 2 Interaction between study SMD and SE of SMD. The line represents the linear solution of the random effects model used in the analytic strategy 4 (meta-regression line; corresponding to R^2=34%). SMD, standardised mean differences.

Estimating outcomes according to physiotherapy approach and population

The interaction between different physiotherapy approaches and populations is illustrated in table 2. Analyses showed that therapeutic exercise for musculoskeletal diseases (70 comparisons) tended to be more beneficial than multimodal interventions (20 comparisons) (SMD difference 0.30; (95% CI 0.03 to 0.57); p=0.03). Furthermore, trials with a 'no intervention' control group (148 comparisons) tended to have a higher overall effect size than trials using sham treatment (76 comparisons) SMD difference 0.25; (95% CI 0.09 to 0.41; p=0.004). This suggests a likely exaggeration of clinical effect of 50% in trials using 'no intervention' control groups.

Quality of the body of evidence

In the stratified analyses, moderate inconsistency was found with a between-study variance (τ^2) of 0.23–0.26 for each stratum and an overall I^2 of 51% (table 1). Estimates were generally imprecise with broad 95% CIs in 24 strata. Only two analyses, therapeutic exercise for musculoskeletal diseases and electrotherapeutic modalities for musculoskeletal diseases, had 95% CIs where clinical actions would not differ if either the upper or lower CI boundary represented the truth (table 2). No analyses using indirect comparisons were carried out.

The quality of the body of evidence was judged as 'low', meaning that further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate of effect of physiotherapy on pain.

DISCUSSION

Our meta-analysis revealed an overall moderate positive effect of physiotherapy although variations occurred depending on the meta-analytic strategy chosen. Stratified analyses showed large, statistically significant effect sizes for, among others, therapeutic exercise and mechanical modalities for musculoskeletal conditions. Our infinite study size analysis and the results from 25% of largest trials indicate that a small study effect is present in the scientific literature concerning physiotherapy for pain. While the results of the different analytic strategies support overall beneficial effects of physiotherapy for pain, the results also suggest that some physiotherapy modalities are less (or not) efficient for certain populations.

Table 2 Interaction* between modalities and populations

	Education (95% CI)	Electrotherapeutic modalities (95% CI)	Mechanical modalities (95% CI)	Multimodal modalities (95% CI)	Other modalities (95% CI)	Passive therapy (95% CI)	Therapeutic exercise (95% CI)	Overall†
External causes	NA	0.20 (−0.34 to 0.74) (4 comparisons)	NA	1.08 (0.36 to 1.8) (2 comparisons)	NA	0.62 (−0.48 to 1.72) (1 comparison)	0.86 (0.27 to 1.45) (3 comparisons)	0.62 (0.26 to 0.98) (10 comparisons)
Musculoskeletal diseases	2.30 (1.33 to 3.27) (1 comparison)	0.55 (0.39 to 0.71) (52 comparisons)	0.89 (0.49 to 1.28) (10 comparisons)	0.49 (0.24 to 0.73) (20 comparisons)	0.84 (−0.25 to 1.91) (1 comparison)	0.39 (0.16 to 0.62) (25 comparisons)	0.79 (0.65 to 0.92) (70 comparisons)	0.64 (0.56 to 0.74) (179 comparisons)
Neoplasms	NA	NA	NA	NA	NA	NA	1.05 (0.41 to 1.69) (3 comparisons)	1.05 (0.37 to 1.74) (3 comparisons)
Nervous system	NA	0.97 (0.41 to 1.53) (4 comparisons)	1.06 (0.05 to 2.06) (1 comparison)	NA	NA	0.48 (−0.18 to 1.14) (3 comparisons)	0.34 (−0.3 to 0.98) (3 comparisons)	0.67 (0.31 to 1.03) (11 comparisons)
Other	2.24 (1.08 to 3.39) (1 comparison)	0.18 (−0.42 to 0.78) (4 comparisons)	0.76 (−0.28 to 1.8) (1 comparison)	NA	NA	0.65 (0.46 to 1.75) (1 comparison)	0.20 (−0.29 to 0.69) (5 comparisons)	0.46 (0.11 to 0.81) (12 comparisons)
Pregnancy/childbirth	0.59 (−0.39 to 1.57) (1 comparison)	2.18 (0.8 to 3.55) (1 comparison)	NA	0.23 (−0.81 to 1.27) (1 comparison)	NA	1.15 (0.06 to 2.25) (1 comparison)	0.68 (0.22 to 1.14) (5 comparisons)	0.75 (0.37 to 1.13) (9 comparisons)
Overall†	1.67 (1.05 to 2.28) (3 comparisons)	0.55 (0.41 to 0.70) (65 comparisons)	0.89 (0.54 to 1.24) (12 comparisons)	0.53 (0.30 to 0.77) (23 comparisons)	0.84 (−0.27 to 1.94) (1 comparison)	0.44 (0.23 to 0.64) (31 comparisons)	0.75 (0.63 to 0.87) (89 comparisons)	0.65 (0.57 to 0.73) (224 comparisons)

Estimates presented as standardised mean differences (SMD) with SE as covariate and 95% CIs.
*Including both main effects and interaction.
†Only main effects estimates.
NA, not available (no comparisons available).

Our results related to the beneficial effect of therapeutic exercise for painful musculoskeletal conditions are supported by similar meta-analyses concerning knee osteoarthritis that have found effect sizes of therapeutic exercise and passive treatment modalities from 0.2 to 0.9 SMD.[33] [34] Other reviews have found similar effect sizes of exercise for fibromyalgia and low back pain.[35] [36] Similar to this study, these reviews found methodological issues within the reviewed trials especially due to risk of bias. Thus, other studies support our findings and our data support therapeutic exercise as particularly relevant in treatment for pain secondary to musculoskeletal conditions.

Combining all available evidence into one estimate can ease interpretation, but the estimate of a meta-analysis depends largely on the ways the data are analysed.[37] [38] Our analytic strategies revealed effect estimates ranging from physiotherapy being harmful to having moderate positive effect on pain. We cannot know which effect is closest to the truth but the result of the infinite study model seems unrealistic as this model is based on the assumption of a linear relationship between study and effect size. Considering pain, a linear relationship between study size and outcome is unlikely, because all pain scales have defined minimum and maximum boundaries. Nevertheless, the infinite study strategy supports the notion that larger trials generally report lower estimates of effect,[39] and indicates that efforts are needed to optimise physiotherapy for pain in larger populations. Further, larger trials tend to have more than one therapist engaged, and lower effect sizes with larger samples, supporting the notion of therapist dependency in daily practice.[40] This highlights that once physiotherapy is compared across therapists, institutions or countries, standardisation is required and the population-wide beneficial effect is likely to be reduced. Small study effects in meta-analyses can also be due publication bias.[41]

In 2009, Chalmers and Glaziou[42] estimated the extent of research investments being wasted to about 85%. In 2015, Yordanov et al[43] estimated the amount of avoidable study waste in clinical trials included in Cochrane reviews between April 2012 and 2013 to 42%. Avoidable waste was defined as: 'trials with at least one domain at high risk of bias for which easy adjustments with no or minor cost could change all domains to low risk'. In our review, issues around sequence generation/allocation concealment and outcome data reporting (present in 54% and 66% of the trials, respectively) could be easily avoided, while issues around blinding of participants and therapists are difficult or even impossible to implement in trials of therapeutic exercise and passive therapies, but more easily implemented in trials with electrotherapeutic modalities.

Variability in otherwise similar interventions is likely a main cause of the between-study heterogeneity observed in this study. We observed heterogeneity within similar physiotherapy interventions with regard to length and delivery complicating estimation of effect of a specific physiotherapy treatment. These results indicate that it is of utmost importance for the physiotherapy community to aspire for standardisation of the research interventions in order for physiotherapy research to be relevant and applicable by every therapist across local, regional, national and cultural borders.

A systematic review and meta-analysis share the limitations of the studies included. As the general quality of evidence is low, partly due to a sometimes inevitable bias caused by lack of blinding, the present results should be interpreted with caution and it is likely that further high-quality research will change our confidence in the estimates.

An important limitation to this review is the restriction on both date and language. We focused on studies published in English from 2004 to 2013. Therefore, a large amount of literature representing the use of physiotherapy to manage pain was excluded before any methodological scrutiny, and thus this review possibly represents a 'limited' view of physiotherapy applied worldwide.

In the absence of a clear definition of physiotherapy, we defined physiotherapy as: 'Any treatment modality performed by physiotherapists or explicitly defined as physiotherapy'. This led to filtration of interventions possibly carried out by physiotherapists, but not specifically defined as such. Thus, it is not reasonable to conclude anything about the effects of specific treatment modalities as no complete search covering all trials testing one modality, has been undertaken. Furthermore, we have only compared physiotherapy with no intervention or sham treatment, and cannot conclude on the comparative effectiveness of physiotherapy and, for example, pharmacological interventions. Even if the large number of trials in this analysis leading us to a robust overall estimate of physiotherapy is considered a strength, the very broad and arbitrary interpretability is a limitation.

Physiotherapy can reduce pain but standardisation and optimisation are needed to harmonise physiotherapy for painful conditions, and to increase the quality and external validity of research in physiotherapy. There is a need for physiotherapy research with minimal risks of bias and focus on prespecified, well-defined and reproducible treatment modalities. Considering the significant heterogeneity in populations with pain, identification of strata within each population would facilitate the process of defining the right treatment modality for each patient and make the patient-specific adjustments to a physiotherapy modality rest on a scientific background and less on therapist skills and discretion.

What is already known on this topic

▶ Physiotherapy is increasingly used as a pain treatment and is included in several international recommendations on management of major painful diseases.
▶ A large range of physiotherapy options exists, but scientific and practical information guiding treatment choices for individual patients is often elusive causing large variation in clinical practice.
▶ This variation complicates the estimation of what benefit a patient may expect from physiotherapy.

What this study adds

▶ Physiotherapy per se has an overall moderate and clinically relevant effect on pain.
▶ However, the quality of evidence is generally low.
▶ We observed large variations in the descriptions of similar physical therapy approaches and intervention lengths varied from a few days to more than 1 year.
▶ There is a need to increase the quality of research in physiotherapy, with focus on standardisation of prespecified, well-defined and reproducible treatment modalities for well-defined populations.

In conclusion, physiotherapy, per se, has a moderate effect on pain with the most consolidated effect seen within musculoskeletal diseases. Greater effect sizes are generally revealed in smaller trials and in trials with high risk of performance or attrition biases. Study inconsistency is considerable, probably due to the clinically heterogeneous groups of patients, as well as intervention variability. Thus, for the overarching question of whether physiotherapy reduces pain, the quality of the evidence is generally low. To increase the value of future research in physiotherapy, focus should be on reducing avoidable research waste by limiting bias, and exploring effects of specific, well-defined and reproducible physiotherapies on well-defined patient populations.

REFERENCES

1 Breivik H, Collett B, Ventafridda V, et al. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. Eur J Pain 2006;10:287–333.
2 Breivik H, Eisenberg E, O'Brien T. The individual and societal burden of chronic pain in Europe: the case for strategic prioritisation and action to improve knowledge and availability of appropriate care. BMC Public Health 2013;13:1229.
3 Kumar N. WHO normative guidelines on pain management. Geneva: World Health Organization, 2007.
4 Fransen M, McConnell S, Bell M. Exercise for osteoarthritis of the hip or knee. Cochrane Database Syst Rev 2003;(3):CD004286.
5 Rueda JR, Solà I, Pascual A, et al. Non-invasive interventions for improving well-being and quality of life in patients with lung cancer. Cochrane Database Syst Rev 2011;(9):CD004282.
6 Savigny P, Kuntze S, Watson P, et al. Low back pain: early management of persistent non-specific low back pain. London: National Collaborating Centre for Primary Care and Royal College of General Practitioners, 2009:14.
7 McAlindon TE, Bannuru RR, Sullivan MC, et al. OARSI guidelines for the non-surgical management of knee osteoarthritis. Osteoarthritis and Cartilage, 2014.
8 World Confederation for Physical therapy. Policy statement—description of physical therapy. 2011. Available at http://www.wcpt.org/sites/wcpt.org/files/files/PS_Description_PT_Sept2011_FORMATTED_edit2013.pdf (accessed Nov 2015).
9 Boutron I, Moher D, Altman DG, et al. Methods and processes of the CONSORT Group: example of an extension for trials assessing nonpharmacologic treatments. Ann Intern Med 2008;148:W60–6.
10 Helmhout PH, Staal JB, Maher CG, et al. Exercise therapy and low back pain: insights and proposals to improve the design, conduct, and reporting of clinical trials. Spine (Phila Pa 1976) 2008;33:1782–8.
11 Hsieh LL, Kuo CH, Lee LH, et al. Treatment of low back pain by acupressure and physical therapy: randomised controlled trial. BMJ 2006;332:696–700.
12 Bennell KL, Egerton T, Martin J, et al. Effect of physical therapy on pain and function in patients with hip osteoarthritis: a randomized clinical trial. JAMA 2014;311:1987–97.
13 Glasziou PP, Chalmers I, Green S, et al. Intervention synthesis: a missing link between a systematic review and practical treatment(s). PLoS Med 2014;11:e1001690.
14 Moher D, Schulz KF, Altman DG, CONSORT. The CONSORT statement: revised recommendations for improving the quality of reports of parallel group randomized trials. BMC Med Res Methodol 2001;1:2.
15 Higgins JPT, Green S. Highly sensitive search strategies for identifying reports of randomized controlled trials in MEDLINE. Cochrane Handbook for Systematic Reviews of Interventions 4.2. 5 [updated May 2005]; Appendix 5b. In: The Cochrane Library. Issue 3. 2005.
16 Chandler J, Churchill R, Higgins J, et al. Methodological standards for the conduct of Cochrane Intervention Reviews (version 2.3, 02 December 2013). http://editorial-unit.cochrane.org/sites/editorial-unit.cochrane.org/files/uploads/MECIR_conduct_standards%202.3%2002122013.pdf (accessed Nov 2015).
17 Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928.
18 Hróbjartsson A, Kaptchuk TJ, Miller FG. Placebo effect studies are susceptible to response bias and to other types of biases. J Clin Epidemiol 2011;64:1223–9.
19 Guyatt GH, Oxman AD, Vist G, et al. GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias). J Clin Epidemiol 2011;64:407–15.
20 Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 7. Rating the quality of evidence—inconsistency. J Clin Epidemiol 2011;64:1294–302.
21 Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 8. Rating the quality of evidence—indirectness. J Clin Epidemiol 2011;64:1303–10.
22 Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines 6. Rating the quality of evidence—imprecision. J Clin Epidemiol 2011;64:1283–93.
23 da Costa BR, Nüesch E, Rutjes AW, et al. Combining follow-up and change data is valid in meta-analyses of continuous outcomes: a meta-epidemiological study. J Clin Epidemiol 2013;66:847–55.
24 Hedges LV. Distribution theory for Glass's estimator of effect size and related estimators. J Educ Behavl Stat 1981;6:107–28.
25 Ghogomu EA, Maxwell LJ, Buchbinder R, et al. Updated method guidelines for Cochrane musculoskeletal group systematic reviews and metaanalyses. J Rheumatol 2014;41:194–205.
26 Cohen J. Statistical power analysis. Curr Dir Psychol Sci 1992;1:98–101.
27 Dechartres A, Altman DG, Trinquart L, et al. Association between analytic strategy and estimates of treatment outcomes in meta-analyses. JAMA 2014;312:623–30.
28 Moreno SG, Sutton AJ, Thompson JR, et al. A generalized weighting regression-derived meta-analysis estimator robust to small-study effects and heterogeneity. Stat Med 2012;31:1407–17.
29 Eriksen P, Bartels EM, Altman RD, et al. Risk of bias and brand explain the observed inconsistency in trials on glucosamine for symptomatic relief of osteoarthritis: a meta-analysis of placebo-controlled trials. Arthritis Care Res (Hoboken) 2014;66:1844–55.
30 WHO. International Statistical Classification of Diseases and Related Health Problems 10th Revision. 2010. Available at https://www.google.dk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&ved=0ahUKEwiyo4z3_87LAhUsEpoKHeYzCLUQFggoMAE&url=http%3A%2F%2Fwww.who.int%2Fclassifications%2Ficd%2FICD10Volume2_en_2010.pdf&usg=AFQjCNFV_goNOWkNtO23OOx-XpKbtPKSbg (accessed Nov 2015).
31 Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. BMJ 2003;327:557–60.
32 Thompson SG, Higgins JP. How should meta-regression analyses be undertaken and interpreted? Stat Med 2002;21:1559–73.
33 Corbett MS, Rice SJ, Madurasinghe V, et al. Acupuncture and other physical treatments for the relief of pain due to osteoarthritis of the knee: network meta-analysis. Osteoarthritis Cartilage 2013;21:1290–8.
34 Juhl C, Christensen R, Roos EM, et al. Impact of exercise type and dose on pain and disability in knee osteoarthritis: a systematic review and meta-regression analysis of randomized controlled trials. Arthritis Rheumatol 2014;66:622–36.
35 Busch AJ, Barber KA, Overend TJ, et al. Exercise for treating fibromyalgia syndrome. Cochrane Database Syst Rev 2007;(4):CD003786.
36 Pennick V, Liddle SD. Interventions for preventing and treating pelvic and back pain in pregnancy. Cochrane Database Syst Rev 2013;8:CD001139.
37 Gøtzsche PC, Hróbjartsson A, Maric K, et al. Data extraction errors in meta-analyses that use standardized mean differences. JAMA 2007;298:430–7.
38 Tendal B, Higgins JP, Juni P, et al. Disagreements in meta-analyses using outcomes measured on continuous or rating scales: observer agreement study. BMJ 2009;339:b3128.
39 Nuesch E, Trelle S, Reichenbach S, et al. Small study effects in meta-analyses of osteoarthritis trials: meta-epidemiological study. BMJ 2010;341:c3515.
40 Herbert RD, Bø K. Analysis of quality of interventions in systematic reviews. BMJ 2005;331:507–9.
41 Egger M, Davey Smith G, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:629–34.
42 Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. Lancet 2009;374:86–9.
43 Yordanov Y, Dechartres A, Porcher R, et al. Avoidable waste of research related to inadequate methods in clinical trials. BMJ 2015;350:h809.