Evidence for physiotherapy practice: A survey of the Physiotherapy Evidence Database (PEDro)

Anne M Moseley, Robert D Herbert, Catherine Sherrington and Christopher G Maher
Centre for Evidence-Based Physiotherapy, The University of Sydney

Evidence-based practice involves the use of evidence from systematic reviews and randomised controlled trials, but the extent of this evidence in physiotherapy has not previously been surveyed. The aim of this survey is to describe the quantity and quality of randomised controlled trials and the quantity of systematic reviews relevant to physiotherapy. The Physiotherapy Evidence Database (PEDro) was searched. The quality of trials was assessed with the PEDro scale. The search identified a total of 2,376 randomised controlled trials and 332 systematic reviews. The first trial was published in 1955 and the first review was published in 1982. Since that time, the number of trials and reviews has grown exponentially. The mean PEDro quality score has increased from 2.8 in trials published between 1955 and 1959 to 5.0 for trials published between 1995 and 1999. There is a substantial body of evidence about the effects of physiotherapy. However, there remains scope for improvements in the quality of the conduct and reporting of clinical trials. [Moseley AM, Herbert RD, Sherrington C and Maher CG (2002): Evidence for physiotherapy practice: A survey of the Physiotherapy Evidence Database (PEDro). Australian Journal of Physiotherapy 48: 43-49]

Key words: Evidence-Based Medicine; Meta-Analysis; Physical Therapy; Randomized Controlled Trials

Introduction

Over the past decade, physiotherapists have been encouraged to take an evidence-based approach to the teaching and practice of physiotherapy (eg MacIntyre et al 1999, Research Committee (Victorian Branch) of the Australian Physiotherapy Association and contributors 1999). Evidence-based practice has been defined by Sackett et al (2000, p. 246) as “the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients”. This involves “integrating individual clinical expertise with the best available external clinical evidence from systematic research”. The best available evidence of the benefits and harms of therapy is provided by systematic reviews of randomised controlled trials (Level I) and well-designed randomised (Level II) or pseudo-randomised (Level III-1) controlled trials (National Health and Medical Research Council 2000). Many physiotherapists have only limited access to this high level evidence (due partly to restricted access to databases that archive clinical trials and reviews or even an awareness of these databases), which has led to the belief that there is little evidence about the effects of physiotherapy interventions (Bithell 2000).

The Physiotherapy Evidence Database (PEDro) was launched in October 1999 to support an evidence-based approach to the teaching and practice of physiotherapy (Sherrington et al 2000). It is a free, Internet-based resource (http://ptwww.cchs.usyd.edu.au/pedro) developed and maintained by the Centre for Evidence-Based Physiotherapy. PEDro contains bibliographic details and author abstracts of systematic reviews and randomised controlled trials in physiotherapy.

To be included on PEDro, randomised controlled trials must fulfil the following criteria:

- the trial compares at least two interventions (ie at least one intervention compared with a control or sham, or a comparison of two or more interventions);
- at least one of the interventions is currently part of physiotherapy practice, or could become part of physiotherapy practice;
- the interventions are applied to human subjects who are representative of those to whom the intervention might be applied in the course of physiotherapy practice (ie people with or at risk of developing a health condition or disability);
- there is random allocation or intended-to-be-random allocation of subjects to interventions; and
- the trial is published as a full paper in a peer-reviewed journal.

The second criterion has been interpreted broadly in order to include the range of treatments that may be used by physiotherapists internationally. Systematic reviews are included on PEDro if they are published in a peer reviewed journal, contain a methods section, and review at least one trial that satisfies the above criteria.

To assist users of PEDro to interpret the results of research, randomised controlled trials on the database are rated for methodological quality by trained PEDro staff or volunteer physiotherapists (all raters complete a training package and pass a rating accuracy test) using the PEDro scale (see Appendix 1). This scale is based on the Delphi list developed by Verhagen et al (1998), a nine-item list...
established by expert consensus (items: eligibility criteria specified, subjects randomly allocated to groups, concealed allocation, groups similar at baseline, blinding of subjects, therapists and assessors, intention-to-treat analysis, point measures and measures of variability reported). Two additional items not on the Delphi list have been included in the PEDro scale (ie outcome measures obtained from more than 85% of subjects and reporting of results of between-group statistical comparisons). Four PEDro scale items have been validated empirically in the medical literature: randomisation, concealed allocation, blinding, and adequacy of follow-up (Moher et al 1999). The other items have face validity but are yet to be empirically validated. The PEDro scale has been shown to have acceptable inter-rater reliability (Moseley et al 1999). Observed agreement for individual PEDro items ranged from 70% (groups similar at baseline) to 96% (blinding of therapists), and the intraclass correlation coefficient for the total PEDro score (ie the score derived from adding all PEDro scale items except the specification of eligibility criteria) was 0.54. To increase the accuracy of quality ratings on the PEDro database, each trial is independently rated by two reviewers, with a third rater arbitrating on items where consensus cannot be reached.

The extent of the external evidence relevant to neurological physiotherapy archived on PEDro has recently been surveyed (Moseley et al 2000). There were 238 randomised controlled trials (54% were categorised as being of moderate to high quality, rating five or more out of 10 on the PEDro scale) and 27 systematic reviews relevant to neurological physiotherapy. However, the quantity and quality of the external evidence on therapy relevant to other subdisciplines of physiotherapy or physiotherapy in general has not been surveyed. Information on the quantity and quality of the external evidence on therapy is important to the physiotherapy profession for a number of reasons. This information can assist the profession in identifying areas of practice that have not been well evaluated. Most importantly, this information could guide the profession in deciding what services should be part of contemporary physiotherapy practice. An evidence-based practice perspective would suggest that the profession should discontinue using treatments that are clearly ineffective and should consider take-up of new treatments when there is evidence that they are effective.

This survey aims to describe the quality and quantity of randomised controlled trials and the quantity of systematic reviews indexed on PEDro. As PEDro probably is the most complete database of trials and reviews in physiotherapy, this survey will provide some insights into the scope of evidence that can be used to guide decisions about the effects of therapy. The PEDro database is unique in that it provides a measure of the quality of the randomised controlled trials it indexes. This permits statements to be made about the quality of randomised controlled trials in physiotherapy.

**Method**

Data were extracted from PEDro on April 17, 2001. The authors, title, source, year of publication, publication type (randomised controlled trial or systematic review) and quality scores for the records were analysed. Three analyses were performed. The first focused on the amount of external evidence available each year. Cumulative totals of randomised controlled trials and systematic reviews were calculated. The relationship between the quantity of reviews and trials and the year of publication was described using least-squares non-linear curve fitting.
The second analysis investigated the methodological quality of the randomised controlled trials. The analysis of methodological quality was performed on the sub-set of trials that had been quality rated using the PEDro scale (the remaining trials had bibliographic information entered on the PEDro database but had not yet been rated). The mean and standard deviation of the total PEDro score (ie the score derived by adding all scale items except for Item 1, the specification of eligibility criteria) for five-year epochs and the frequency distribution of the total PEDro score for all rated trials were calculated. In addition, the percentage of trials satisfying each of the items on the PEDro scale was determined.

The last analysis examined the evidence available for each subdiscipline of physiotherapy. All PEDro records were coded for up to three areas of physiotherapy (cardiothoracics, continence and women's health, ergonomics and occupational health, gerontology, musculoskeletal, neurology, orthopaedics, paediatrics, and sports – see Appendix 2). The number of randomised controlled trials and systematic reviews were calculated for each area of physiotherapy.

Results

There were 2,708 records indexed on PEDro. Of these, 2,376 were randomised controlled trials and 332 were systematic reviews (that is, one systematic review for every 7.2 trials)(a). Most records (93%) were in English. The first archived trials in physiotherapy were published in 1955 (a), about seven years after the first modern randomised controlled trial in medicine (Medical Research Council 1948). These trials investigated manipulation for low back pain (Coyer and Curwen 1955) and wax baths for rheumatoid arthritis (Harris and Millard 1955). The number of randomised controlled trials has increased rapidly and approximately exponentially since then (Figure 1). An exponential growth curve fits the data well (number of trials = $e^{0.18x}$ years since 1955; $r^2 = 0.94$). An average of 155 trials were published each year from 1995 to 1999.

The first archived systematic review relevant to physiotherapy was published in 1982 and investigated the effects of weight loss treatments (including exercise) on hypertension (Hovell 1982). Since then, there has been an exponential increase in the number of systematic reviews, with 59 published in 1999. Again, an exponential growth curve fits the data very well (number of reviews = $e^{0.33x}$ years since 1982; $r^2 = 0.99$).

Of the 2,376 randomised controlled trials archived on PEDro at April 17, 2001, 2,297 (or 97%) had been rated for methodological quality at least once and 1,235 (52%) had ratings confirmed by a second or third rater (the remaining 79 trials had bibliographic information entered on the PEDro database but had not yet been rated). Trial quality ranged from 0/10 to 10/10, with most rating 4 or 5 (Figure 2). Fifty-two per cent of trials were nominally of moderate to high quality, rating 5 or more on the PEDro scale.

The quality of randomised controlled trials in physiotherapy has increased gradually since 1955 (Figure 3). The mean total PEDro score has increased from 2.8 for trials published between 1955 and 1959, to 5.0 for trials published between 1995 and 1999 (the mean for trials published in 2000 or 2001 is 5.7).

The percentage of rated randomised controlled trials that satisfy each item of the PEDro scale is illustrated in Figure 4. Most trials (94%) fulfilled the randomisation item of the PEDro scale (ie subjects were randomly allocated to experimental conditions). This is not surprising given that...
trials are only eligible for PEDro if they use a randomised or intended-to-be-randomised method of allocation. Other high-rating PEDro scale items were the reporting of between-group statistical comparisons (89% of trials) and the provision of point estimates and measures of variability (82% of trials). The trials showed lower compliance for three features that have been shown to predict bias in clinical trials (concealed allocation to groups, blinding, and adequacy of follow-up; Moher et al 1999). Only 16% clearly used concealed allocation and 59% of trials clearly had a dropout rate of less than 15%. The prevalence of blinding was low. Very few trials used blinded therapists (5% of trials) or subjects (9% of trials), but about one-third (34% of trials) blinded the assessors. Only 12% of trials reported using an intention-to-treat analysis.

The systematic reviews and randomised controlled trials covered most areas of physiotherapy (Figure 5). Note that the columns in this figure do not tally with the total number of records on PEDro because trials and reviews can be relevant to more than one area of physiotherapy. Trials and reviews can be coded for up to three areas of physiotherapy. The ‘other’ category in Figure 5 is used when the record cannot be classified using the subdiscipline definitions (see Appendix 2). The largest number of trials and reviews (39% of trials and 36% of reviews) were in the area of musculoskeletal physiotherapy.

**Discussion**

Contrary to popular belief, there is a significant body of evidence to guide the practice and teaching of physiotherapy. Much of this evidence is archived on PEDro which, in April 2001, contained 2,376 randomised controlled trials and 332 systematic reviews. This may be an underestimate of the amount of evidence, although it is difficult to estimate how many trials and reviews are not on the PEDro database. Trials published in languages other than English are probably significantly under-represented because the majority of PEDro staff and volunteers are fluent in English only.

There were randomised controlled trials and systematic reviews relevant to all subdisciplines of physiotherapy. Not surprisingly, the largest number of trials and reviews were in the area of musculoskeletal physiotherapy. The evidence for neurological physiotherapy, which had been surveyed previously (Moseley et al 2000), ranked fourth. Paediatrics and ergonomics were the areas of physiotherapy with the smallest number of trials and reviews. These areas could be priority areas for future research funding.

While the PEDro database contains 332 systematic reviews, it is likely that many clinical questions in physiotherapy are yet to be addressed by these reviews. It is unclear from this survey whether this is primarily due to a lack of randomised controlled trials in particular areas or a lack of systematic reviews summarising completed trials.

A substantial number of the randomised controlled trials in physiotherapy are nominally of moderate to high quality.
trials that evaluate exercise, manual therapy and education, but it should be possible in trials that evaluate electrotherapies such as laser and ultrasound.

While it is desirable to blind subjects, this is difficult for many physiotherapy interventions. Two recent papers have shown some interesting approaches to this difficult problem. Streitberger and Kleinheinz (1998) reported a protocol for sham acupuncture that patients could not distinguish from real acupuncture. The protocol uses a ‘sham’ needle that retracts into the handle when it touches the skin. The patients feel a pricking sensation and see the shortening of the needle as would occur with real acupuncture. Van Tulder et al (2000) have advocated that where blinding is not feasible, the use of a sham treatment that was demonstrated to be equally credible and acceptable to the patient should be regarded as passing the criterion for subject blinding if the trial formally assesses the equal credibility of treatment. In contrast, Vickers and de Craen (2000) have argued that sham or placebo-controlled trials sacrifice external validity for internal validity and do not inform clinical decision making as well as trials in which the control group is not administered a sham therapy.

Like randomised controlled trials, the quality of systematic reviews vary. While the methodological quality of systematic reviews have not been evaluated on the PEDro database, physiotherapists could use a number of resources to identify reviews that are likely to be valid. First, the quality of Cochrane systematic reviews is likely to be high as they are performed according to stringent guidelines. Second, the Database of Abstracts of Reviews of Effectiveness (DARE; University of York 2000) provides structured abstracts and quality assessments for many of the systematic reviews archived on PEDro. About 50% of the systematic reviews relevant to neurological physiotherapy have structured abstracts on DARE (Moseley et al 2000). Third, readers can assess individual systematic reviews for the presence of features known to introduce bias. The filters suggested by the National Health and Medical Research Council (2000) include the adequacy of the search strategy used, the use of appropriate inclusion criteria that were applied in an unbiased way, quality assessment of included studies, the appropriate summarising of the characteristics and results of the individual studies, the use of appropriate methods for pooling the data, and exploration of sources of heterogeneity.

The amount of external evidence about therapy relevant to physiotherapy is significant and growing at an exponential rate. This poses a challenge to even the most dedicated physiotherapist who, based on current publication rates, would have to read approximately 155 trials and 59 reviews each year to keep abreast of the physiotherapy literature. Some solutions to this problem of information overload are to restrict reading to the area of current practice, read only high quality randomised controlled trials and systematic reviews, and make use of distilled literature such as the Critically Appraised Papers section of this journal.

Conclusion

This survey of the Physiotherapy Evidence Database (PEDro) revealed a significant body of high level external evidence (both randomised controlled trials and systematic reviews) that can be used to support decision-making about therapy for all subdisciplines of physiotherapy. The amount of evidence is expanding at an exponential rate and the quality of trials is increasing. While there is a relatively large amount of high quality evidence available, there still remains scope for improvements in the quality of the conduct and reporting of clinical trials.

Footnote  Since this paper was accepted for publication, the PEDro database has continued to expand. In January 2002 PEDro contained 2,712 randomised controlled trials and 411 systematic reviews. We have also identified some trials that were published before the first trials in physiotherapy cited in this paper. Two trials were published before 1955, both in the area of cardiothoracic physiotherapy. They are Palmer and Sellick (1952) and Palmer and Sellick (1953).

Acknowledgements Financial support for PEDro has been provided by the Motor Accidents Authority of New South Wales, the New South Wales Physiotherapists Registration Board, and the Australian Physiotherapy Association. The support of the Rehabilitation and Related Therapies Field of the Cochrane Collaboration, the School of Physiotherapy at The University of Sydney, the Rehabilitation Studies Unit at The University of Sydney, PEDro staff, and the volunteer physiotherapists who have assisted with literature searching and the quality rating of trials is also acknowledged.

Authors Anne Moseley, Rehabilitation Studies Unit, Department of Medicine, The University of Sydney, Royal Rehabilitation Centre, Post Office Box 6, Ryde, New South Wales 1680. E-mail: amoseley@mail.usyd.edu.au (for correspondence). Robert Herbert, School of Physiotherapy, The University of Sydney, Post Office Box 170, Lidcombe, New South Wales 1825. Catherine Sherrington, Prince of Wales Medical Research Institute, Barker Street, Randwick, New South Wales 2231. Christopher Maher, School of Physiotherapy, The University of Sydney, Post Office Box 170, Lidcombe, New South Wales 1825.

References

Appendix 1: Criteria for the PEDro scale

1. eligibility criteria were specified
2. subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received)
3. allocation was concealed
4. the groups were similar at baseline regarding the most important prognostic indicators
5. there was blinding of all subjects
6. there was blinding of all therapists who administered the therapy
7. there was blinding of all assessors who measured at least one key outcome
8. measures of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups
9. all subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention to treat”
10. the results of between-group statistical comparisons are reported for at least one key outcome
11. the study provides both point measures and measures of variability for at least one key outcome

Notes on administration of the PEDro scale:

All criteria Points are only awarded when a criterion is clearly satisfied. If on a literal reading of the trial report it is possible that a criterion was not satisfied, a point should not be awarded for that criterion.

Criterion 1 This criterion is satisfied if the report describes the source of subjects and a list of criteria used to determine who was eligible to participate in the study.

Criterion 2 A study is considered to have used random allocation if the report states that allocation was random. The precise method of randomisation need not be specified. Procedures such as coin-tossing and dice-rolling should be considered random. Quasi-randomisation allocation procedures such as allocation by hospital record number or birth date, or alternation, do not satisfy this criterion.

Criterion 3 Concealed allocation means that the person who determined if a subject was eligible for inclusion in the trial was unaware, when this decision was made, of which group the subject would be allocated to. A point is awarded for this criterion, even if it is not stated that allocation was concealed, when the report states that allocation was by sealed opaque envelopes or that allocation involved contacting the holder of the allocation schedule who was “off-site”.

Criterion 4 At a minimum, in studies of therapeutic


interventions, the report must describe at least one measure of the severity of the condition being treated and at least one (different) key outcome measure at baseline. The rater must be satisfied that the groups’ outcomes would not be expected to differ, on the basis of baseline differences in prognostic variables alone, by a clinically significant amount. This criterion is satisfied even if only baseline data of study completers are presented.

Criteria 4, 7-11 Key outcomes are those outcomes which provide the primary measure of the effectiveness (or lack of effectiveness) of the therapy. In most studies, more than one variable is used as an outcome measure.

Criterion 5-7 Blinding means the person in question (subject, therapist or assessor) did not know which group the subject had been allocated to. In addition, subjects and therapists are only considered to be “blind” if it could be expected that they would have been unable to distinguish between the treatments applied to different groups. In trials in which key outcomes are self-reported (eg visual analogue scale, pain diary), the assessor is considered to be blind if the subject was blind.

Criterion 8 This criterion is only satisfied if the report explicitly states both the number of subjects initially allocated to groups and the number of subjects from whom key outcome measures were obtained. In trials in which outcomes are measured at several points in time, a key outcome must have been measured in more than 85% of subjects at one of those points in time.

Criterion 9 An intention to treat analysis means that, where subjects did not receive treatment (or the control condition) as allocated, and where measures of outcomes were available, the analysis was performed as if subjects received the treatment (or control condition) they were allocated to. This criterion is satisfied, even if there is no mention of analysis by intention to treat, if the report explicitly states that all subjects received treatment or control conditions as allocated.

Criterion 10 A between-group statistical comparison involves statistical comparison of one group with another. Depending on the design of the study, this may involve comparison of two or more treatments, or comparison of treatment with a control condition. The analysis may be a simple comparison of outcomes measured after the treatment was administered, or a comparison of the change in one group with the change in another (when a factorial analysis of variance has been used to analyse the data, the latter is often reported as a group × time interaction). The comparison may be in the form hypothesis testing (which provides a p value, describing the probability that the groups differed only by chance) or in the form of an estimate (for example, the mean or median difference, or a difference in proportions, or number needed to treat, or a relative risk or hazard ratio) and its confidence interval.

Criterion 11 A point measure is a measure of the size of the treatment effect. The treatment effect may be described as a difference in group outcomes, or as the outcome in (each of) all groups. Measures of variability include standard deviations, standard errors, confidence intervals, interquartile ranges (or other quantile ranges) and ranges. Point measures and/or measures of variability may be provided graphically (for example, SDs may be given as error bars in a Figure) as long as it is clear what is being graphed (for example, as long as it is clear whether error bars represent SDs or SEs). Where outcomes are categorical, this criterion is considered to have been met if the number of subjects in each category is given for each group.

Appendix 2: PEDro definitions for each sub-discipline of physiotherapy

Cardiothoracics includes, but is not restricted to, papers evaluating acute and rehabilitation cardiothoracic interventions or fitness training on those with conditions affecting the cardiothoracic system. This subdiscipline does not include studies of general fitness training among patient populations. Studies of general fitness training for healthy populations are not indexed on PEDro.

Continence and women’s health includes, but is not restricted to, male and female incontinence and pre- and post-natal interventions for the mother.

Ergonomics and occupational health includes, but is not restricted to, interventions based at workplaces or on workers for work-related conditions.

Gerontology includes papers where the average age of the study sample is over 60, and papers on conditions which commonly affect older people (eg arthritis).

Musculoskeletal includes, but is not restricted to, low back pain, rheumatoid disease, entrapment syndromes and neuralgia.

Neurology includes, but is not restricted to, lesions of the central and peripheral nervous systems excluding those whose primary presentation is pain or paraesthesia such as carpal tunnel syndrome, neuralgia or sciatica.

Orthopaedics includes only fractures and intervention before or after orthopaedic surgery (eg knee replacements, ligament repairs).

Paediatrics includes papers where the average age of the study sample is under 16, and papers on conditions which commonly affect children (eg cystic fibrosis).

Sports includes papers which specifically mention sports injuries as well as conditions which commonly affect sports people (eg ligament repairs).

If the area for a particular paper does not fit under any of the above categories, the code “no appropriate value in this field” is used.