Challenges in Applying Best Evidence to Physiotherapy

Mark Jones, BS (Psych), MAAppSc (Manipulative Therapy)  
Karen Grimmer, PhD  
Ian Edwards, PhD  
Joy Higgs PhD  
Franziska Trede, PhD

1. Program Director, Senior Lecturer, Postgraduate Coursework Masters Programs, School of Health Sciences, University of South Australia  
2. Director, Centre for Allied Health Evidence, at the University of South Australia  
3. Lecturer, School of Health Sciences, University of South Australia  
4. Professor, University of Sydney  
5. Diversity Health Coordinator at The Prince of Wales Hospital, Sydney

Abstract

Purpose: As the first of two papers addressing challenges in applying best evidence to physiotherapy practice, this paper highlights the role of qualitative research and challenges researchers face in producing sound evidence for practice. The challenge of what constitutes acceptable evidence to inform evidence-based practice is addressed in this paper with critique of the current over-reliance on quantitative methodologies that excludes a substantial body of valuable qualitative evidence to support sound practice. The current trend of relying on questionnaires to obtain psychosocial data is challenged on the basis of limitations of research design. Challenges to researchers from the quantitative and qualitative worlds to break down the political barriers separating these two groups are put forward. Lastly, we consider challenges that clinicians face in maintaining best practice based on when evidence is still largely not available or is compromised by limitations to research design with respect to population homogeneity, diagnostic inclusion criteria, intervention details, outcome measures, and critical appraisal tools. Conclusion: Quantitative research alone is insufficient to understand patients' pain and disability experiences. Researchers are challenged to improve their reporting of research with greater detail provided regarding populations, therapeutic environment, and interventions used if clinicians are to be able to apply research findings in practice.

Introduction

This is the first of a pair of papers that address the challenges that clinicians face in applying evidence-based practice and researchers and clinicians face in producing relevant and sound evidence for practice. This first paper addresses the core question of the nature of evidence and takes a critical look at how such evidence can be understood and critiqued.

The Importance of Evidence-based Practice

Without doubt, the evidence-based practice drive has had a significant positive impact on the physiotherapy profession. While the search for physiotherapy evidence is not something new, the quality of physiotherapy research has never been scrutinised to the extent it is today, with the positive effect of a steady increase in the quality of quantitative and qualitative physiotherapy research. Clinicians are wise to embrace evidence-based practice because without the critical appraisal evidence-based practice demands, misconceptions in clinical theory and practice can continue unchecked. Research and reports of clinical practice on humans are subject to unrecognised bias, taken-for-granted assumptions, and errors of reasoning that necessitate audits of practice as encouraged by the evidence-based movement. In fact, Thomas Kuhn, a science historian, highlights how the majority of misconceptions through the history of science, including such things as the function of the heart as the organ of thought, can almost universally be attributed to a lack of critical appraisal of contemporary theory. Evidence-based practice is essential in order to avoid prolonged misconceptions and to understand how best to work with patients in their health management.
However, as essential as evidence-based practice is to the growth of physiotherapy research and clinical practice, a somewhat narrow conception of what constitutes evidence-based practice and what constitutes acceptable evidence is creating challenges to clinicians wanting to apply best evidence to their physiotherapy practice. In this paper we draw attention to challenges both researchers and clinicians face in determining what constitutes acceptable evidence.

Challenge of What Constitutes Acceptable Evidence
The dominant research paradigm underpinning the evidence-based practice movement is without doubt quantitative theory, with the pre-eminent primary research approach valued in evidence-based practice being the randomised controlled trial (RCT) used to investigate the effectiveness of therapeutic interventions under tightly controlled conditions. This is evident in the currently available 'Levels of Evidence' hierarchies for ranking the quality of information available to guide practice (See Table 1).4, 5, 6

Table 1. Levels of evidence (Sackett et al., 2000, pp. 17-21)

<table>
<thead>
<tr>
<th>Level</th>
<th>Type of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Evidence from at least one systematic review (with homogeneity) of multiple randomised controlled trials</td>
</tr>
<tr>
<td>II</td>
<td>Evidence from systematic review of cohort studies (including at least one randomised controlled trial)</td>
</tr>
<tr>
<td>III</td>
<td>Evidence from systematic review (with homogeneity) of case-control studies or individual case-control studies</td>
</tr>
<tr>
<td>IV</td>
<td>Evidence from well designed case series (and poor-quality cohort and case-control studies)</td>
</tr>
<tr>
<td>V</td>
<td>Expert opinion without explicit critical appraisal or based on physiology, bench research or 'first principles'</td>
</tr>
</tbody>
</table>

For those researchers, clinicians, and academics who have campaigned for a broader definition of what constitutes acceptable evidence, there is some satisfaction to be found in Sackett et al.’s most recent definition of evidence-based medicine as "the integration of best research evidence with clinical expertise and patient values." 4-7,14 However, while this research group acknowledges patient values and the need for qualitative research to inform evidence-based practice and help clarify patients’ experiences and understandings of disease, disability, assessment, and management, the contribution made by qualitative research to understanding patients’ perspectives and clinical expertise is excluded from all current hierarchies of evidence. Scientific evidence obtained through measurement of observable phenomenon is not only given ascendency over other forms of evidence, it is literally the only form of evidence considered worthwhile.9 This raises the critical issues of what constitutes the best research design(s) to answer different research questions.

A broader definition of what constitutes acceptable evidence for specific research questions is urgently required, supported by systems to critically scrutinise all forms of evidence using standard notions of research rigour. The exclusion of qualitative research evidence from current hierarchies of evidence creates a significant challenge for clinicians and researchers as it excludes a substantial body of available evidence to which clinicians could turn, and it discourages researchers from using qualitative methodologies even when they may be the most appropriate paradigm for the question(s) being asked. If we continue down the current path of relying almost entirely on quantitative research to understand our patients and evaluate our practice, exclusively using narrow levels of evidence, the result will be suboptimal decisions not fully informed by the full scope of relevant, valid and available information,7, 9, 15, 11, 13, 16 Experienced clinicians use all of these forms of evidence to optimise their patient-centred, evidence-based clinical decision making. The fundamental question for physiotherapy practice is what should constitute evidence that is mindful of the clinical reality?

The Role of Qualitative Research
Quantitative research emphasises measurement and analysis of causal relationships between variables. Qualitative research aims to understand human phenomena working under the premise that humans and the human body cannot be separated from social experiences that contribute to shaping their individual illness, pain, or disability experiences.

The scientific process and quantitative research paradigm have clearly provided an important means of generating knowledge that has contributed significantly to medical and physiotherapy healthcare. A good example from the medical arena is the breast cancer prevention drug tamoxifen that was investigated in the 1990s through a double-blind, randomised controlled trial of 13,388 women with identified risk factors for breast cancer.12 However, the trial was terminated early on the basis of its “successful” results to date despite the lack of any research of the drug’s social, emotional, or long-term consequences. While the trial demonstrated successful empirical results, some of the human/life results were lost in the generalised findings. Miller and Crabtree...
highlighted these lost results in brief stories of three
women from the trial. Camille, 50 years of age, is now
a single-breasted trial failure who is struggling financially,
frightened of her spreading cancer, and sees her life as
shredded. Gloria was also in the study and while she did
not develop breast cancer, she didn’t like taking the daily
tablet, found herself worrying about everything since
joining the study, and has since elected to discontinue
the treatment. Alice was also considered a success in
that she did not develop breast cancer. However Alice
did develop a drug related blood clot that left her with
chronic back and leg pain and significant disability. Such
stories remind us that health interventions are never 100
percent effective and too often research reports fail to
give the full picture, instead reducing the illness
experience and treatment effect to quantifiable measures
where individual suffering and social consequences are
either neglected all together or reduced to homogeneous
variables oversimplified through quantitative
assessment. A close look at many of our own
physiotherapy randomised controlled trials reveals an
analogous incomplete picture where the impact the
health problem and the physiotherapy intervention had
on the research subjects’ lives is not fully elucidated,
leaving those patients who failed to respond as mere
statistics, outliers, and unheard voices. We need to
recognise that psychosocial factors cannot be separated
from biomedical factors and as such psychosocial effects
must be considered along side biomedical outcomes —
at the very least they are important in terms of the
client’s informed consent to the proposed treatment
regime. Should physiotherapy provide universal and
monocultural treatments to biophysical body parts or
individualised and appropriate treatments to patients
inclusive of their diverse perspectives?

Psychosocial data collection within the quantitative
paradigm – Is it adequate?
Numerous variables are now recognised as potentially
contributing to our health, either as genuine risk factors
to the development of a health condition or as factors
influencing how we cope with injury or illness. Traditionally, variables such as beliefs, feelings, and
coping strategies were completely omitted from the data
collected in quantitative research and even now are only
superficially obtained through the plethora of
questionnaires available to measure such things. Subjective views of health status, quality of life, and
patients’ perspectives on their pain, disability, and future
prospects, traditionally gleaned from the subjective
examination, are increasingly being measured in
quantitative studies in recognition that assessment of
physical function or impairment alone is insufficient to
provide a complete health profile for describing patient
populations and detecting change. While psychometric
testing ensures questionnaire test-retest reliability, and
internal consistency and validity are established against
other accepted measures (criterion validity) and
population-expected trends (construct validity),
significant limitations exist with the variability of
meanings of questions and depth of information that can
be obtained from this format of data collection. Despite convincing evidence from psychology and
sociology that processes involved in interpreting
questions and selecting answers are complex and non-
uniform, questionnaire research has yet to investigate
the issue of respondent variability in interpretations.

Mallinson investigated patients’ understanding of
questions and response options to the SF-36, one of the
most widely used health-related quality of life
questionnaires aimed at portraying generic health status
and well-being regardless of age, condition or treatment
group. Particular difficulties faced by the respondents
included the use of two-part questions, respondents’
limited perspectives of distance (e.g. ability to walk x
yards), and limited context provided in the question (e.g.
variability in weather influencing outdoor activities). The
complexities that comprise an individual’s health status
and quality of life in such tick box responses are reduced
to a numerical score and unique individual differences
and their interpretations are not captured.

When arguably even more complex patient perceptions
such as those assessed in the Fear-Avoidance Beliefs
Questionnaire are judged on the basis of categorical
questions presented out of context, it should be evident
that taking a totally quantitative approach to the
assessment of complex phenomena such as disability,
beliefs, and quality of life is insufficient on its own.
Questionnaire research is moving toward shorter and
shorter forms to reduce time for completion such that
one version of a Pain Beliefs and Perceptions Inventory
has reduced the number of questions to four. When
patients’ personal perspectives are reduced to numerical
ratings based on such limited information, the underlying
uniqueness of the human story and the sensitivity to
detect variation in treatment effects across these
qualitative variables will inevitably be lost.

In contrast to the quantitative research paradigm,
qualitative research operates under the philosophical
assumption that truth and reality are not absolute.
Instead the qualitative or interpretive paradigm sees
reality as constructed by individuals through their life
experiences, resulting in unique and contextually framed
experiences. For instance, experiences such as pain,
disability, and diminished self-efficacy are interpreted
differently by people living in different circumstances.
Where quantitative research takes a phenomenon apart
to uncover the effect of different variables, qualitative
research seeks to examine the selected lived
phenomenon in context. Where quantitative research in
the health sciences endeavours to establish cause and
effect and measure effectiveness, qualitative research
focuses more on understanding the meaning people
have constructed, or how they make sense of their world
and the experiences they have in their world. The value
qualitative research offers evidence-based practice is
supported by increasing recognition of the difficulties
Challenges in Applying Best Evidence to Physiotherapy

Qualitative research can provide the context currently through a more sophisticated multimethod approach. Some of these questions have been and are being asked by quite philosophically separated researchers rather than researchers. Based on these arguments we need to recognise that evidence in evidence-based practice must include knowledge derived from qualitative research as well as quantitative research.

Challenges to Researchers
Traditionally different foci of quantitative and qualitative research meant that the two research worlds were kept separate with different aims, methods of data collection and analysis, and typically very different minded researchers. However, recognition of the strengths and limitations of both paradigms has led to a call to breakdown the political barriers separating these two groups with clinical questions being the common ground on which to unite them. Qualitative research provides a different form of evidence that needs to be recognised alongside the RCT. Miller and Crabtree have conceptualised a multimethod RCT with qualitative methods addressing issues of context and meaning and quantitative methods providing measurement and analysis toward generalisation. While the two paradigms have different aims, by linking them through clinical questions they can inform each other to produce a more complete understanding. Consider the range of questions that could be asked concerning the multifactorial entity chronic pain using integrated qualitative and quantitative research.

- What are the qualitative and quantitative patient outcomes of specific physical, cognitive-behavioural (or educative) and combined management interventions?
- What is the lived experience and meaning chronic pain has for patients: physically, behaviourally, emotionally, cognitively, culturally, spiritually?
- How does chronic pain and chronic pain management affect patients' families and social structures?
- What are patients' self concept, hopes, despair, fears and insecurities concerning chronic pain?
- What power and political influences exist that support or hinder the development and management of chronic pain?

Some of these questions have been and are being researched but typically as separate questions, asked by quite philosophically separated researchers rather than through a more sophisticated multimethod approach. Qualitative research can provide the context currently lacking in the traditional quantitative approach that dominates evidence-based practice. Achieving this integrated balance of research paradigms requires cross-disciplinary and multi-paradigm relationships between qualitative and quantitative researchers and utilisation of a range of consultants including psychologists, sociologists, epidemiologists, statisticians and health economists as required. There are multiple ways in which qualitative and quantitative research might be combined with a clear need for studies to be performed longitudinally in series toward a larger research agenda.

In addition, research methods can be integrated more intimately in a single study. For example, Borkan et al. used an epidemiological cross-sectional design to investigate hospitalised elderly patients with hip fractures in an attempt to better understand the relationship between hip fracture, disability, and patient pain/disability perspectives. Traditional rehabilitation outcome measures (e.g. fracture type) were used as the dependent variables while multiple biomedical indicators (e.g. ambulation status) were measured as independent variables. Unique to this study was their inclusion of in-depth interviews with each patient regarding the meaning their hip fracture had in their lives. The qualitative analysis identified distinguishable narratives or plots to patients' stories that were then entered as another independent variable in the statistical outcome modelling. Interestingly they found narrative type was the most powerful predictor of rehabilitation outcome, better than the traditional biomedical indicators. That is, patients' illness experience and the meaning they attributed to their hip fracture was the best predictor of their outcome. The question this raises is what outcomes are we measuring: outcomes as measurable by external variables or outcomes as perceived by patients?

Challenges from insufficient reporting of research details and results
The evidence-based practice movement has had the positive effect of promoting greater rigour in physiotherapy research. Moseley et al. highlight how the methodological rigour of experimental studies archived on the Physiotherapy Evidence Database (PEDro) has increased steadily over the past five years. However a number of problems still exist with physiotherapy effectiveness studies. These include:

- high rates of drop-out or loss to follow-up,
- lack of blinding (patient, therapist, measurer),
- lack of random and concealed allocation to treatment arms,
- lack of adequate identification of population subgroups, and
- artificial isolation of treatment interventions in determining their effect and lack of evidence of sustainable outcomes.
To address these deficits, a new wave of physiotherapy research has commenced. Patients with conditions such as whiplash associated disorder are no longer considered homogenous, but rather, researchers recognise subgroups demonstrating differences in measures of motor system impairment, sensory hypersensitivity and psychological effects. This suggests differences not only in clinical presentation following whiplash but also differences in central processing and psychological effects. This more sophisticated research approach will clearly assist our subsequent refinement of management strategies for the various subgroups discovered.

Patients’ perspectives that potentially influence their health condition are also increasingly being assessed. Nevertheless, our current body of research is either incomplete, or incomplete in its reporting, to adequately guide therapists in their recognition and management of the multitude of patient problems we face. Therein lies another challenge to researchers – greater reporting of population, environment, and intervention details and greater care with conclusions reached in the absence of good research.

Confronted with a growing number of systematic reviews, based largely on effectiveness research of potentially limited methodology, practicing clinicians face the daunting challenge of maintaining best practice based on best evidence when the evidence is still largely not available or is incomplete. Even when a systematic review or a high hierarchy primary research study (such as an RCT) can be located that investigates the condition of interest, numerous issues must be considered for the clinician to have confidence in the applicability of the findings to their patients.

**Homogeneity of the sample and external generalisability of the study**

Fundamental to establishing the relevance of effectiveness studies to clinical practice are two key issues:
- the homogeneity of subjects in the study (do the subjects all have the same problem and/ or the same potential for responding to the intervention?), and;
- the external generalisability of study (can the study findings be applied to patients in clinical practice)?

Diagnostic/inclusion classification of patients for inclusion/exclusion study purposes is thus clearly needed in order to establish clinical relevance, and is in need of continual refinement in line with advances in clinical reasoning, diagnostic tests, and understanding of symptomatology. Coupled with this is the growing recognition of the influence of patients’ health perspectives on their health condition, necessitating ongoing investigation of the relationship between these variables with patients’ pain perception and disability and our therapeutic interventions. A good recent example of high quality use of multiple measures of outcome reflecting financial and opportunity costs, patient preference and adherence to rehabilitation programs, and functional outcome are the studies by Fairbank et al. which examine the costs, outcomes and risk analysis of surgery versus intensive rehabilitation for patients with chronic low back pain.

**Diagnostic/inclusion criteria**

Regarding diagnostic/inclusion criteria, lack of capacity to compare findings between studies usually relates to researchers’ assumptions regarding sample homogeneity. Inclusion criteria for chronic conditions usually only stipulate a minimum amount of time the condition has been suffered. It is rare that researchers classify the presenting condition in any other way (i.e. by employment and family stresses, or by previous treatment approaches or migration status). Individual perspectives or their behavioural adaptations to their health condition are not taken into account except by generalised patient responses to standard surveys. Thus opportunities to understand how an intervention changes individual perspectives or behaviour is rarely considered. As a result, when no significant change between the pre- and post-test group standard survey instrument scores is the result (as it often is because variability in individual response will be attenuated within the group response analysis), the conclusion will be that the intervention is not effective (for anyone). Confidence intervals provide a group estimate of improvements and poor response; however, important information on individual responses cannot be provided to identify individuals, or subgroups of patients, who actually did significantly better with the intervention than any other subjects.

**Interventions**

Regarding interventions which have been tested in primary research studies (or collated into systematic reviews), adoption of research findings by clinicians can hinge on whether the intervention is replicable in clinical practice. It is increasingly common in experimental studies to test a number of interventions. Very few studies provide sufficient detail and justification of the treatments (how decisions were negotiated, what was done, how was it done, who did it, how often, for how long, in what setting, in what order of implementation of different treatments, rationales for differences such as number of occasions for different treatments, or the content of education programs) to enable clinicians to adopt any of the treatment approaches, irrespective of their effectiveness.

The lack of publication detail on interventions may well be due to lack of rigour in study design and management; however, it may also be due to journal publication restrictions (e.g. word limit). Although the researchers could be contacted by clinicians for additional information, this lack of published information poses a constraint on consumers of research readily adopting the research findings. There is an encouraging
move to address this issue by a number of the electronic journals, which are currently moving to publish the complete methodologies of trials as separate papers.

Authors should disclose research bias and assumptions that underpin their research design so that the reader can make their own informed decisions of the credibility and quality of key indicators for treatment decisions. Sufficient detail should be provided in research publications to enable others to understand the key indicators on which treatment decisions were based with sufficient contextual information regarding the therapeutic environment and manner in which interventions were delivered (educatively, behaviourally and humanistically) to give readers a sense of being observers of the therapy itself. Such “thick descriptions” of context are characteristic of research reporting from the qualitative paradigm and anything less than this is not a valid account of events and does not do justice to the real-life intricacies of clinical practice. This is illustrated in an example provided by Linton where the way in which an early intervention for acute back pain was administered became the difference between a highly successful intervention and one that may have actually increased the problem. This would not be surprising to practicing clinicians who know well that the manner in which a therapy is provided is critical to its success. The nature and quality of the physiotherapist-patient relationship can contribute in a significant way to effective treatment outcomes. The human factor, the engagement, the rapport between practitioner and patient should not be marginalised or excluded from the evidence used as a basis for practice.

Outcome measures
The clinical relevance of studies is often constrained by the limited and non-standard nature of the chosen outcome measures. Many outcome measures may be used for research into specific conditions; thus, studies on, for example, neck pain may well measure various ranges of movement, pain (intensity, duration, medication used to deal with pain, effect of lifestyle), use of disability aids such as collars, recent disability using any number of standard indices (SF36, Neck Disability Index, Patient Specific Scale), depression, cost of treatment etc. It is possible that 10 primary research reports on neck pain may well report 10 different outcome measures, even if the same intervention has been tested. Capacity to compare between studies is therefore limited, and opportunities to identify the strongest evidence for an intervention are reduced to general descriptions of the direction of change of outcome measures which were significantly influenced by the intervention. Lack of standard approaches to reporting change in outcome measures constrains capacity to compare between research reports.

Critical Appraisal
Few of the critical appraisal tools developed for effectiveness studies assign a score to patient homogeneity, clarity of intervention, or consistency or breadth of outcome measures. Thus, to assist clinicians to evaluate clinical relevance (one of five key evidence dimensions), an increased focus is required on developing additional appraisal criteria. These could include a more critical description of the individuals who consent to participate in a study (for instance their pain experience perspectives) as well as specific information on the interventions tested (to enable them to be repeated clinically). Moreover, discussions should be held between researchers and clinicians about the most useful outcome measures for specific conditions, to ensure that all stakeholder perspectives are considered, and that the range of ways in which the intervention may be effective have been measured.

Conclusion
The physiotherapy profession is faced with a number of challenges with respect to applying best evidence to practice. Quantitative research alone is insufficient to understand patients’ pain and disability experiences, and relying on questionnaires alone to obtain data regarding patients’ health status, quality of life, beliefs and fears is arguably invalid. Researchers are challenged to improve their reporting of research with greater detail provided regarding populations, therapeutic environment, and interventions used if clinicians are to be able to apply research findings in practice. In the case of systematic reviews, care is needed not to confuse “no evidence of effect” with “evidence of no effect” and conclusions made on the basis of methodologically poor studies need to be limited to the former.

References


