Test-Retest Reliability and Responsiveness of the Shortened Disability Arm Shoulder Hand (QuickDASH) Questionnaire: A Pilot Study.

Kate Polson

A dissertation submitted to
Auckland University of Technology
in partial fulfilment of the requirements for the degree
of Master of Health Science

2007

School of Physiotherapy

Primary Supervisor: Duncan Reid

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ATTESTATION OF AUTHORSHIP

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning

Signed……. Date……..

Name…….

ACKNOWLEDGEMENTS

I would like to acknowledge the physiotherapists at the following Physiotherapy Clinics for generously assisting in this study. The Principals were kind enough to permit the study to be conducted in their clinics and all took the time to collect the data from their patients.

- Louise Johnson - Pakaranga Physiotherapy
- Lisa Hansen - Lisa Hansen Physiotherapy Clinic
- Graeme White - Unisportsmed Physiotherapy
- Pauline Newton - AUT Musculoskeletal Physiotherapy Clinic
- Robyn Lindstrom - Physio Plus Te Kuiti
- Mike Lovell - Sports Med Physiotherapy

I owe a special gratitude to Robyn Lindstrom and the fellow staff at Physio Plus Te Kuiti who have always encouraged my academic pursuits, maintained a lovely closeness, which I treasure, and supported me even when I was clearly preoccupied. Thank you to my 'study buddy' Shelley Johnson and Lynley Bradnam-Roberts for their friendship and editorial advice.

Duncan Reid and Peter Larmer have been exceptional supervisors, encouraging me at a time when giving up would have been an option. Without their ongoing support, guidance and patience when the demands of research were in danger of getting the better of me, this project would never have reached fruition.

I would also like to thank Peter McNair for his invaluable assistance with the statistical analysis and data presentation.
Ethics approval was granted by Auckland University of Technology Ethics Committee on 12/10/06 AUTEC reference number 06/173.

Finally, a special thank you needs to go to my family, specifically my husband John, children Oliver and Paddy and my sister Louise Johnson. Your unconditional love and encouragement throughout this project has been wonderful. It has given me the will to carry on during rough and troubled times. You have put up with my selfishness, distractedness and boring conversation. You have all continued to believe in me, taught me to stay calm under fire and provided many laughs which lightened my load.

ABSTRACT

Objective

The aim of this dissertation was to determine the test-retest reliability and responsiveness of the shortened Disability Arm Shoulder Hand (QuickDASH) questionnaire as a functional outcome measure in a primary health care setting.

Background

The QuickDASH Questionnaire was developed to reduce the burden of administration to both the clinician and patient when assessing functional impairment of the upper extremity. An analysis of current literature on the QuickDASH has provided limited evidence that it is a reliable and responsive instrument in a primary health care setting.

Study Design

A literature review on the key features of functional outcome measures plus instruments available to physiotherapists used in assessing functional status of the upper extremity was conducted. The QuickDASH questionnaire was identified as requiring further psychometric analysis. The pilot study was a questionnaire based, cross sectional, longitudinal study of patients with upper extremity injuries.

Methods

Participants presenting for treatment for upper extremity injuries (n=35) were recruited from private physiotherapy clinics. Participants completed the QuickDASH questionnaire on three occasions, the first prior to treatment, the second 24 -48 hours later and the third at discharge or at six weeks, which ever event occurred first. At this time the participant and therapist were asked to complete a Global Discharge Questionnaire, which rated the current overall condition of the upper extremity and the perceived level of change that had occurred.

Reliability was analysed by calculating an Intraclass Correlation Coefficient and Bland and Altman plots and limits of agreement. Responsiveness was analysed by calculating the Effect Size and Standard Response Mean. The Minimal Clinical Importance Difference (MCID) to assess responsiveness was calculated using an anchor approach. The correlation between the physiotherapist and patient perceived level of overall

condition of upper extremity disability was also analysed by an Intraclass Correlation Coefficient.

Results

Results demonstrated inconclusive test-retest reliability. Responsiveness was demonstrated to be very strong (ES = 1.02, SRM = 1.1). The MCID was calculated to be 19 points. There was a strong correlation between patient and therapist perception of overall current condition of upper extremity disability with an ICC of 0.88.

Conclusion

These results support that the QuickDASH is a responsive instrument at measuring patient perceived functional change over time in a primary health setting. Reliability was not conclusively established. Additionally, support was found that therapists are in agreement with patients as to the level of functional impairment over a six week timeframe.

Chapter 1 Introduction

Musculoskeletal injuries to the upper extremity can leave patients functionally impaired. Functional impairment can be defined as a person's inability to perform activities of daily living (ADL) such as personal hygiene, dressing and housework plus work and recreational activities (Giang, 2006). Disability can arise from symptoms such as pain, numbness, tingling, weakness or stiffness to any part of the upper limb. Measuring the degree of functional incapacity at any given point in time over a patient's recovery may be assisted by the use of functional outcome measures. Functional outcome measures may involve the use of self-reported questionnaires that rely on the patient's perception of their functional ability. Effective functional outcome measures must also be valid and reliable instruments and be adept at assessing change over time, specifically change that is considered important to the patient (Laing, Lew, Stucki, Fortin, & Daltroy, 2002; Schmitt & Fabio, 2004; Terwee, Dekker, Wiersinga, Priummel, & Bossuyt, 2003). Assessing function is in keeping with the promotion of the World Health Organisation (WHO) core values where Health Related Quality of Life (HRQOL) assessment of injury or disease consequence, as they affect the person, is applied to rehabilitation principles in the health sector, inclusive of physiotherapy practice (Cieza & Stucki, 2005; Horner & Larmer, 2006; Jette, 1993).

The economic strain to individuals and society as a result of upper extremity injuries is well documented globally (Driver, 2006; Giang, 2006; Keogh, Nuwayhid, Gordon, & Gucer, 2000). In 2002 workplace musculoskeletal injuries accounted for costs of $53.4 billion dollars in the United States with spinal and upper extremity injuries making up the majority of workers compensation claims involving time away from work (Driver, 2006; Stover, 2004). Upper extremity injuries are also common in New Zealand. The New Zealand Accident Compensation Commission (ACC) injury statistics, for the period of June 2005 - July 2006, quote a total of 27 % of all entitlement claims derive from upper extremity injuries (ACC, 2006). These injuries occurred across all age groups, in the home, work place, sporting and motor vehicle accidents. Specifically, of this group injuries to the shoulder were the highest (30%).

Physiotherapists are ideally placed to manage and optimise recovery from upper extremity injuries and minimise functional impairment. Confirmation of this is viewed with physiotherapy management included in many evidence based treatment guidelines developed for upper extremity conditions, for example shoulder injury management (Green, Buchbinder, & Hetrick, 2005), elbow conditions (Vicenzino, Collins, Benson, & Wright, 1998) and in the domain of hand therapy (Michlovitz, Harris, & Watkins, 2004). Funding agencies such as ACC support evidence-based practice (EBP). It is the result of research, specifically randomised control trials that are measured by functional outcomes, that lead to recommendations for and against treatment guidelines (ACC, 2003). However, the implementation of functional outcome measures in physiotherapy is low leading to poor collection of data on treatment effectiveness (Abrams et al., 2006; MacDermid & Stratford, 2004).

Evaluating effectiveness of treatment is challenging for physiotherapists for several reasons. Firstly, a patient functional outcome may vary from traditional objective measures. For example a patient with a shoulder injury may not be able to raise their arm high above their head, however some patients may not find this a dysfunction with activities such as dressing or even household chores such as hanging out the washing, nor may it disturb their sleep. Additionally, patients with finger injuries may not manage fine finger dexterity or have the same grip strength prior to the injury. They may train themselves to perform all tasks with the other hand or manage the necessary hands tasks in an adapted fashion and thus consider themselves functionally able. Secondly, using functional outcome measures can be viewed as a burden to both the physiotherapist and patient (Abrams et al., 2006). Abrams et al., (2006) identified that the main barriers for clinical application of functional outcome measures in a physiotherapy setting were; time to administer, lack of accessibility, training and familiarity with the instruments, and an inability to interpret scores and change of scores.

The ability to incorporate self-rated change in condition by the patient needs to become an important component of physiotherapy practice. The knowledge gained from obtaining this information can be used to develop patient targeted treatment strategies. Using functional outcome measures that are psychometrically robust and user friendly to both the physiotherapist and patient reduces the burden of administration and can raise the level of implementation. The Disability Arm Shoulder Hand (DASH)

questionnaire is a well recognised functional outcome measure of the upper limb (Beaton, Katz, Fossel, Wright, & Tarasuk, 2001)**.** However one of the limitations of the DASH is the length of time it takes to complete (Beaton, Wright, & Katz, 2005). The QuickDASH has been developed as a shortened version of the DASH in order to reduce the time of administration. Whilst the DASH has had robust psychometric scrutiny, the QuickDASH has had limited analysis to date, specifically in a primary health setting (Bot et al., 2004; Gummesson, Ward, & Atroshi, 2006; Imaeda et al., 2006; Jester, Harth, Wind, Germann, & Sauerbier, 2005; MacDermid & Stratford, 2004; Paul et al., 2004; SooHoo, McDonald, Seiler, & McGillivary, 2002).

The construct validity of the QuickDASH was established by Beaton et al., (2005) and again by Stover (2004). Reliability and responsiveness have been demonstrated in pre and post surgical patient population's only (Gummesson et al., 2006; Imaeda et al., 2006). Of particular importance to clinical management is assessment of the ability for the QuickDASH to reliably measure the change that has occurred that is considered important or significant to the patient. This is known as the Minimal Clinical Important Difference (MCID) and is the change in score over time that reflects a true meaningful change in health status has occurred as perceived by the patient (Laing et al., 2002; Schmitt & Fabio, 2004). Knowledge of this scale will allow the physiotherapist to become clinically familiar with a time efficient instrument that will measure true change in individual patients perceived health status.

1.1 Purpose

The purpose of this dissertation is to examine the literature in order to review functional outcome measures used in the upper extremity and to undertake a pilot study to test the reliability and responsiveness of the QuickDASH questionnaire.

Chapter 2 Literature Review

This chapter contains a review of the current literature concerning the use of outcome measures for objective assessment of the impact of functional limitations resulting from upper extremity injuries. Firstly the methodology is outlined followed by a review of the characteristics of outcome measures and finally a critique of selected upper limb measures in use is presented.

2.1 Methodology

A literature search using Pubmed, Cinhal and Multisearch database was performed using the key words: upper extremity, upper limb, outcome measures and functional outcome measures. A further search included key words: validity, reliability and responsiveness. The years were restricted to 1980-2006. An inclusion criteria was structured to include instruments that were whole limb measures and had had studies published on their validity, reliability and responsiveness. Exclusion criteria included instruments that were not readily accessible via search engines such as Google Scholar or free websites. The rationale for this was to ensure that the functional outcome measures selected were available to all physiotherapists, not only those who had access to university library or interloan systems and also to avoid breach of copyright laws (Davidson, 2004). The instrument specifically had to be self completed in order to interpret the patient's perception of their functional capacity, therefore any instrument that included a physical examination as a critical component was also excluded.

2.2 Results

Three functional outcome measures met the inclusion criteria. The three functional outcome measures are: The Upper Extremity Functional Index (UEFI) (Stratford, Binkley, & Stratford, 2001), the Disability Arm Shoulder Hand (DASH) (Beaton et al., 2001) and more recently the shortened version of the Disability Arm Shoulder Hand (QuickDASH) (Beaton et al., 2005).

2.3 Discussion

2.3.1 Assessing Upper Extremity Injuries

2.3.1.1 Objective assessment

Physiotherapists customarily assess upper extremity injuries in a clinical environment by using both subjective and objective examinations of the patient, which leads to formation of a diagnosis and the development of a treatment plan (Kirkley & Griffin, 2003; MacDermid & Stratford, 2004). Examples of objective measures include joint range of movement (ROM) measured by a goniometer, muscle strength and testing for neural involvement (Kirkley & Griffin, 2003). These are performed to determine a diagnosis and the resulting physiological impact of the disorder. These assessment techniques have an important place in establishing the nature and severity of the injury, form the basis for the development of treatment planning and act as indicators of a response to treatment (Beaton et al., 2001). There is a justification for treatment options being based on evidence-based medicine, especially those developed from research that have examined the effectiveness of treatment using objective outcome measures (Green et al., 2005). An example of evidence-based treatment profiles is the recent publication of a review on shoulder injuries by the ACC, 'The Diagnosis and Management of Soft Tissue Shoulder Injuries and Related Disorders' (ACC, 2004). These guidelines document objective assessment techniques for making a diagnosis plus objective measures on which clinicians can base a management profile in order to achieve an optimal outcome within normal timeframes.

However, this document fails to acknowledge how to utilise functional outcome measures in the treatment of shoulder injuries. Thus in the realm of evidence based medicine; how does the physiotherapist determine if the treatment profile implemented also has had an effect on a patients ability to function? For example, if restoration of shoulder function is an indicator to the patient of treatment success, then the specific range of shoulder joint movement measured by a goniometer may be of less importance to them than whether or not they can reach a high shelf or return to work. While the two measures may be related, a change in goniometer reading may have little functional meaning to the patient who is seeking a pain free return to their ADL's (Cieza & Stucki, 2005; Jette, 1993).

Further, in relation to the upper extremity, evidence suggests that a clinical objective test and patient function are poorly correlated (Hopkins, 2000; Jost, Zumstein,

Pfirrmann, Zanetti, & Gerber, 2005). There are many factors besides tissue damage that contribute to work disability such as fear avoidance behaviour or loss of job prospects, that are not fully explained by symptoms or by biological, and physiological variables (Stover, 2004).

2.3.1.2 Functional outcome measures

Whilst objective measures still have a place in the clinical setting, recent debate has surfaced challenging the sole use of objective measures as outcome measurements (Cieza & Stucki, 2005; Jette, 1993; Kirkley & Griffin, 2003; MacDermid & Stratford, 2004). There is a current trend towards evaluating the efficacy of treatment by measuring functional outcomes that are relevant to the patient. Further, it has been suggested that measuring functional status discriminates severity and predicts subsequent disability better than physical examination or laboratory measures (Jester, Harth, Wind et al., 2005; Stover, 2004). In clinical practice, using subjective measures, such as questionnaires or functional outcome tools that assess the ability to function in daily life plus psychological aspects of health ensures the focus of treatment is on the patient rather than the disease. This is considered relevant in a clinical physiotherapy setting (Higginson & Carr, 2001). Examples of where physiotherapists may enhance patient management by the use of functional outcome measures are; working in a multidisciplinary team requiring effective and consistent communication between health professionals, planning and setting patient orientated goals and predicting time frames for recovery such as return to work planning (Jette, 1993). Moving practitioners away from treating problems and towards treating people, without losing evidence-based treatment practice, is a conceptual framework of functional assessment that can be employed by physiotherapists in the clinical setting (Horner & Larmer, 2006; Jette, 1993). As the goals of health care change the available means of health outcome measures must adapt as well.

2.3.2 Health Outcome Measurement Frameworks

Established health outcome measures involve two main frameworks: Health Related Quality of Life (HRQOL) and the International Classification of Functioning and Disability (ICF). Measuring the subjective health status of patients involves the use of psychometric properties, that is the science of psychological testing (Hicks, 1999). The measurement of HRQOL, where the physical, social and psychological health of the

patient is assessed, is an example of a psychometric measure used commonly in the health sector, inclusive of physiotherapy practice (Jette, 1993). Whilst HRQOL does not specifically measure physical impairment it does describe and evaluate the patient's perception of how their impairment impacts on their functioning and health status (Cieza & Stucki, 2005; Higginson & Carr, 2001; Jette, 1993; Kirkley & Griffin, 2003). Thus HRQOL assesses the consequence or the burden of the disease or impairment as they affect the person not the specific assessment of the disease at the organ or body system level.

The International Classification of Functioning, Disability and Health (ICF) (World Health Organisation, 2001) have classified this psychosocial framework further. The ICF is not an assessment tool itself but is a framework that can be applied to rehabilitation, funding policy, outcome measurement and research (Cieza & Stucki, 2005). The ICF categories are divided up into areas representing body structures and functions including physiological and psychological function, the activities that people do, life situation and social roles that people participate in, and environmental factors (Cieza & Stucki, 2005; Horner & Larmer, 2006). All these factors may be directly or indirectly affected by injury or illness. Thus incorporating these concepts into rehabilitation planning, setting patient goals and being able to measure the outcome is important to physiotherapy management.
Using psychometric measures or tools, the physiotherapist can evaluate the functional status of, or outcome that a treatment intervention has had on a patient as well as the objective change in impairment such as ROM or muscle strength.

2.3.3 What Are Functional Outcome Measures?

Whilst there is an evolving classification of functional outcome measures, four main categories emerge from the literature: generic health profiles (Cieza & Stucki, 2005), disease or condition specific instruments (Kirkley & Griffin, 2003), patient generated formatted instruments (Hoving, O'Leary, Niere, Green, & Buchbinder, 2003) and, more recently, whole limb outcome measures (Beaton et al., 2001).

2.3.3.1 Generic health profiles

Generic health profiles are instruments that focus on activities and participation relevant to all health conditions and patient populations (Cieza & Stucki, 2005; Higginson &

Carr, 2001). Examples of generic health profiles are the Nottingham Health Profile (NHP) (Hunt, McEwan, & McKenna, 1985) which measures illness behaviour, the Sickness Impact Profile (Bergner, Bobbit, Carter, & Gilson, 1981) and the Medical Outcomes Study Short Form (SF-36) (Ware & Sherbourne, 1992). These are single instruments that have several categories representing different aspects of health that can be aggregated into one overall score. For example the SF-36 contains scales that cover the dimensions of physical health, mental health, social functioning, role functioning, general health, pain and vitality. Two summary scales can be obtained: the Physical Component Summary Score (PCS) and the Mental Component Summary Score (MCS) (Ware & Sherbourne, 1992). Generic health status instruments however, have limitations that may affect physiotherapy practice. They are costly to access and are time consuming to complete. Further, generic instruments are limited in that they are very broad and may not focus enough on aspects of quality of life that are of relevance to a particular disease or treatment. For example there is not one question in the SF-36 that is specific to shoulder conditions. This inadequate focus may neglect a small but clinically significant change in the quality of life to the patient with the disease (Kirkley & Griffin, 2003). Cieza & Stucki (2005) recently reviewed a range of commonly used generic health profiles against the ICF classification framework. They identified that not one individual generic health profile met all the concepts of the ICF and that when selecting health status outcome measures additional information may be required specific to a condition or patient population.

2.3.3.2 Disease specific instruments

Disease specific instruments have been designed to overcome the limitations of generic health profiles. They are considered to be more sensitive to the disorder under consideration and are therefore more likely to reflect clinical changes (Hudak, Amadio, Bombardier, & Group, 1996; Kirkley & Griffin, 2003). Specific conditions can reflect a number of areas, for example: diseases (instability, arthritis, whiplash, carpal tunnel), anatomical areas (shoulder, knee, neck) or patient population (children, older adult) (Hudak et al., 1996; Roach, Budiman-Mak, Songsiridej, & Lertratanakul, 1991; Stratford et al., 2001). Disease specific instruments, particularly in the musculoskeletal field, have been developed further with instruments being available for very specific conditions. Examples of this are prevalent in the upper extremity with instruments designed for instability (Kirkley, Griffin, Mclintock, & Ng, 1998), rotator cuff disease (Holtby & Razmjou, 2005), and the hand (Chung, 1998). Disease specific instruments

that are narrowed to an exact condition have a place in the field of research, particularly if evaluation is needed for a precise treatment intervention such as a surgical technique (Baysal, Balyk, Otto, Luciak-Corea, & Beaupre, 2005). Whilst disease specific instruments are widely used they may also not reflect all the HRQOL or ICF domains. For this reason it is not uncommon to use a combination of outcome measures such as a disease specific plus a generic health profile instrument. Despite this being a common practice in research, consideration needs to be given to the practicality of using this method to obtain outcome measures in a clinical setting (Bot et al., 2004; Horner & Larmer, 2006).

A further limitation of the degree of specificity in disease specific instruments is that one instrument may not be applicable to another condition, requiring a different instrument for every diagnosis. In the upper extremity some conditions may occur concurrently, making it difficult for the clinician to decide which instrument is best to use. This also highlights the burden of determining the depth of psychometric analysis performed on individual outcome measures (MacDermid & Stratford, 2004). In other words; is the instrument reliable, valid and responsive and thus potentially useful in a clinic?

2.3.3.3 Whole limb outcome measures

An extension of the concept of disease specific instruments and to address the aforementioned limitations is the development of extremity outcome measures. This is where the function of the limb is measured as a regional unit versus a specific area or condition involving part of the limb. This has been viewed in both the upper extremity (Hudak et al., 1996; Stratford et al., 2001) and lower extremity (Blinkley, Stratford, Lott, Riddle, & Research, 1999).

The design purpose is to specifically conceptualise the extremity as a single functioning unit with the analysis reflecting how well the patient is functioning as a whole. Thus for the upper extremity, the instrument would be used as an outcome measure for any upper extremity pathology, musculoskeletal area, arm dominance and across all population groups (Hudak et al., 1996). The practical aspect of this concept is that in a clinical setting one outcome measure can be used for single or multiple disorders. This may have a particular benefit in encouraging familiarisation and utilisation of outcome measures by clinicians (MacDermid & Stratford, 2004). However, upper extremity

outcome measures are still problematic in that they may lose the specificity required for a particular pathology plus the fixed items may not reflect every patient's functional task requirements (MacDermid & Tottenham, 2004).

2.3.3.4 Patient generated instruments

Patient generated instruments are a format that allows the patient to identify the functional activities they are experiencing difficulty with. The key point to this type of outcome measure is that the patient provides the items of importance not the therapist, thus focusing on the patient's perceived functional rating. Hoving et al., (2003) compared the validity of neck disability questionnaires with a patient preference questionnaire (the problem elicitation technique [PET]). Their research revealed that the existing fixed item questionnaires did not cover all the problems that whiplash patients judged important to them. However, administering PET questionnaires requires a trained interviewer; whilst this is not prohibitive to clinical administration it would also not be readily available to most clinical practices (Hoving et al., 2003).

2.3.4 Components of a Functional Outcome Measure

In order for a clinician to interpret questionnaires an insight into the following characteristics is important. The key components are how the validity, reliability and responsiveness of an instrument are established. The significance of these components to the clinician is that they indicate that an instrument has had the psychometric properties evaluated for a specific population and for a specific purpose. For an outcome measurement instrument to be successful it needs to be established that the burden to the clinician and patient is minimal and that the instrument has been deemed valid and reliable in a clinical setting (Kirkley & Griffin, 2003; MacDermid & Stratford, 2004). The next section will review these key characteristics of outcome measures.

2.3.4.1 Validity

Validity indicates the extent an instrument measures the construct or theory that it is intended to measure (Jette, 1993). Validity does not belong to the instrument directly but is a property of an instrument when it is administered to a specific sample under specific circumstances (Hicks, 1999; Horner & Larmer, 2006; Kirkley & Griffin, 2003). The most recognised method of statistical analysis for validation is the correlation coefficient, which is a numerical value between -1.0 and +1.0 and indicates the degree

and nature of the association between sets of data. Correlation statistics include Pearsons Product Moment Correlation Coefficient (r) and the Spearmans Rank Coefficient (r_s) Guidelines to the correlation scale can be viewed in Table 1 (pg 14). The forms of validation identified are: content validity, criterion (includes concurrent and predictive) validity and construct (includes divergent and convergent) validity.

Content validity refers to the overall appearance and content of the instrument judged by those who are going to use it. It examines the extent to which the domain of interest is comprehensively sampled by the items in the questionnaire. Therefore for upper extremity questionnaires, items on the questionnaire must reflect functional areas that are important to patients with upper extremity issues (Bot et al., 2004; Hicks, 1999; Horner & Larmer, 2006; Kirkley & Griffin, 2003). There is no statistical analysis known to measure this so a panel of experts is considered a judge of content validity. Improved content validity is supported when patients have been part of the panel of experts and have had an input into the questionnaire design and pilot study (Bot et al., 2004).

Criterion validity is the measure of the degree to which an instrument correlates with an external criterion variable that is currently available and already known to be valid. If the criterion exists in the present this is known as concurrent validity. Predictive validity applies to an external criterion that is to be measured in the future. Ideally the external criterion comparison should be considered a gold standard. However, for attributes measured in clinical practice such as pain, functional status and HRQOL there are no known gold standards, therefore the research must identify why a particular external criterion was selected (Kirkley & Griffin, 2003; MacDermid & Stratford, 2004). If there is an absence of a gold standard external criterion then the validation of a measure draws heavily on construct validity (MacDermid & Stratford, 2004).

Construct validity should validate the hypothesis or theory under investigation. This refers to the extent to which scores on a particular instrument relate to other measures consistent with the theoretically derived hypothesis concerning the constructs that are being measured (Bot et al., 2004; Horner & Larmer, 2006). Techniques to determine construct validity include convergent and divergent validity, which are the measure of two scales looking for a likeness or difference of constructs respectively. Strong correlations between subscales purporting to measure the same construct are thought to

be evidence of convergent validity. Weak correlations between subscales are thought to measure different constructs and are viewed as evidence of divergent validity (Bot et al., 2004; Horner & Larmer, 2006).

There is discussion among researchers regarding the best method of validity testing. Primarily, competing goals exist between measurement precision and low response burden to both the patient and the clinician. The validity of a measure is enhanced as the number and diversity of comparison standards supporting its application grows (MacDermid & Stratford, 2004). This reflects that establishing the validity of questionnaires is an ongoing process with confidence in validity developing the more an instrument is tested and researched. The New Zealand population has a considerable cultural diversity and work type activities. This may have implications on measures, as it is important to assess validity across differing populations such as gender, age, ethnic and cultural groups.

In summary, there may be several approaches taken to measure validity, which need to be reported in the research. Importantly, validity of an instrument is enhanced with exposure to a wide study population.

2.3.4.2 Reliability

Reliability refers to the ability of an instrument to be repeatedly administered to stable subjects over time and will produce a stable result that has no measurement error (Horner & Larmer, 2006; Kirkley & Griffin, 2003). There are several synonyms with reliability including reproducibility, repeatability and precision. Most HRQOL instruments are self-administered therefore variance related to interviewers (intra and inter-rater reliability) becomes irrelevant. However if a rater is involved the most common method of quantifying reliability is to use the Intraclass Correlation Coefficient (ICC) (Horner & Larmer, 2006; Kirkley & Griffin, 2003; MacDermid & Stratford, 2004). For guidelines to ICC scales refer to Table 1.

Another analytical technique used to measure limits of agreement between repeated measures is Bland and Altman distribution plots (Bland & Altman, 1986). The graph is an X-axis with range of two standard deviation above and below (positive and negative) the axis indicating a range of error. Of the results being recorded, 95% of the data should fall within this range hence creating a 95% limit of agreement (Rankin & Stokes,

1998). The graphs are easy to visually inspect, making data interpretation of reliability simple and convenient. An appreciation can be gained of where the error is across the whole data set, for example if a systematic error is present between tests. The size and range of measurements can be appreciated at a glance. Evidence of bias or outliers is present plus the relationship between the size of differences and the size of the mean. Reliability from the Bland and Altman plots can indicate a reliability that is suitable for clinical application (Rankin & Stokes, 1998).

Of more relevance however is the internal consistency and test-retest reliability. Internal consistencies refer to how well individual items in an instrument are homogenous or "hang" together. This is important when instruments with multiple items are summed to form a total score as it reflects that patients have answered in a consistent way to all items (MacDermid & Stratford, 2004). Internal consistency is measured using Cronbachs alpha or a coefficient alpha (Horner & Larmer, 2006; MacDermid & Stratford, 2004). For guidelines to Chronbachs alpha scales see Table 1.

Test-retest reliability provides information about the stability of person's responses over time in persons who have truly remained unchanged. This is achieved by performing a repeat measure from the baseline measure after a time period where a patient will not remember their original responses but also for their condition not to have changed. The appropriate statistical test is the ICC, as this not only measures the association between repeated measures but the agreement (see Table 1). A measure of agreement is important to quantify measurement error and detect systematic difference between two measurements (Bot et al., 2004; Horner & Larmer, 2006; Kirkley & Griffin, 2003; MacDermid & Stratford, 2004)

Table 1: *Selection of Common Statistical Analysis Methods for Psychometric Testing*

Property	**Statistical Test**	**Guidelines to Interpretation**
Validity	Correlation Statistics: • Peasons Product Moment-Correlation Coefficient (r) • Spearmans Rank-Correlation Coefficient (r_s)	.10 = small .30 = medium .50 = large (Hicks, 1999)
Reliability	Intraclass Correlation Coefficient (ICC)	< .40 = poor < .70 = inadequate $\geq$ .70 = good $\geq$ .80 = excellent (Hicks, 1999)
	Bland and Altman	Graph depicting 95% limits of agreement (Hicks, 1999)
	Chronbachs Alpha (α)	< 70 = inadequate $\geq$ .70 = good $\geq$ .80 = excellent (Hicks, 1999)
Responsiveness	Group Level Statistics: • Effect Size (ES)	.20 = small size .50 = moderate size .80 = large size (Hicks, 1999)
	• Standard Response Mean (SRM) (Cohens effect size)	.01 = small .06 = moderate .14 = large (Hicks, 1999)
	Individual Level Statistics: • Standard Error of Measurement (SEM)	SEM = (sample standard deviation) $\sqrt{}$ 1 – reliability coefficient Associated 90% & 95% confidence levels should be defined. (Deyo et al., 1991)
	• Reciever Operating Characteristics, area under the curve ROC (auc)	$\leq$ .50 = inadequate discrimination > .60 = adequate discrimination $\geq$ .80 = good discrimination (Deyo et al., 1991)

.

2.3.4.3 Responsiveness

Responsiveness is defined as the ability of an instrument to detect small but important clinical change in a patient's status (Deyo, Diehr, & Patrick, 1991). Synonyms found in the literature include sensitivity to change and interpretability. There is discord among researchers regarding the use of the word 'sensitivity' as in research it has other clinical and epidemiological meaning (Deyo et al., 1991; Horner & Larmer, 2006). A change in patient status is clinically relevant as this is the fundamental way a clinician will develop an individualised treatment plan. Objective knowledge of a change in patients' status can also be important for wider use in the health arena such as following an evidence based trend in rehabilitation time lines and for funding reasons. The ability of an instrument to be responsive to change therefore is representative of the most relevant area of an instrument's use in a clinical setting. Thus the clinician needs to have an understanding of how responsiveness is interpreted for individual instruments. **Sensitivity** is interpreted as a minimal statistical or numerical change detected by the instrument that is not error or minimal change detected (MCD). Whilst this ability to quantify change is desirable it may not have any clinical meaning nor may it be interpreted as a significant enough change by either the patient or the clinician. Therefore the **responsiveness** of an instrument is the ability to detect a clinically meaningful or minimal clinical important difference (MCID) in patient health status; that is, the patient can subjectively distinguish that a change has occurred in *their* health status.

Interpretability is the amount the score changes (in either direction) to indicate this MCID and is the method by which clinician interprets that a change in patient status has occurred. This gives a meaning to the patients score with the desired goal being that the clinician is able to assess or judge the importance of the findings. In other words, a qualitative meaning to a quantitative score (Bot et al., 2004; Deyo et al., 1991; Horner & Larmer, 2006; Kirkley & Griffin, 2003; Schmitt & Fabio, 2004). This has been defined further by Schmitt & Fabio (2004), who report that MCID is when the *patient* perceives that there has been an important change in health status.

There is no agreement among researchers as to how to measure and report responsiveness to clinical change (Horner & Larmer, 2006; Kirkley & Griffin, 2003; MacDermid & Stratford, 2004; Schmitt & Fabio, 2004; Terwee et al., 2003). Various statistics have been used and often combinations of statistics are reported in a single

study. The most commonly used statistics in the literature are: effect size (ES), standardised response mean (SRM), the standard error of measurement (SEM), and the receiver operating characteristic (ROC) area under the curve (ROCauc). For a summary of responsiveness statistics refer to Table 1. The effect size is calculated as the ratio of the mean score change divided by the baseline score for a group of patients who have changed over time, this method is known as Kazis effect size (Kazis, Anderson, & Meenanm, 1989). For interpretation of effect sizes refer to Table 1. The SRM also uses data from subjects deemed to have improved and also qualify as group-level statistics. It uses the standard deviation of the change in scores for the denominator instead of the baseline scores (Deyo et al., 1991; Schmitt & Fabio, 2004). A further effect size measure is Guyatt's effect size or the responsiveness statistic (also known as the responsiveness ratio). This statistic is calculated with data from both patients who have improved and those who have not (Guyatt, Walter, & Norman, 1987).

The effect size measure is often used in research to determine an outcome for particular treatment techniques over larger populations or in randomised clinical trials. This has important implications when a clinician is attempting to interpret research results as to whether or not an instrument of choice is deemed to be responsive for individual patients. That is, can the same outcome measure be responsive at tracking an individual patient through a course of treatment? Schmitt & Fabio (2004) argue that a statistically significant change at a group level may not be significant at the individual level. A statistical method reported to be more reliable in detecting individual-level of change is the standard error of measurement SEM (this is not to be confused with the standard error of mean which has the same abbreviation). The SEM reflects the amount of error associated with an individual subject assessment and not the error associated with the instrument, thus calculates statistically reliable change. The resulting numerical figure can be interpreted as to the degree of confidence, using 90% or 95% confidence intervals (CI), that the score is an individual or patient's true score.

The ability to detect change over time can also be quantified by construction of ROC curves. ROC curves synthesise information on sensitivity and specificity; discriminating between patients reporting clinical improvement and patients reporting clinical stability. ROC analysis is performed by plotting sensitivity to change on the y-axis and 1-specificity on the x-axis, against the patients' global assessment. The area under the curve (ROCauc) can be interpreted as the probability of correctly identifying an

improved patient from randomly selected pairs of improved and stable patients (Deyo et al., 1991). Given the debate and lack of resulting consensus of what statistical analysis is best suited for determining responsiveness, the use of multiple analytical techniques in individual studies remains the norm. Indeed, a study where more than one measure of responsiveness has been used is believed to provide a more comprehensive representation for the clinician.

The use of a concurrent patient global disability rating scale is commonly used in research as the external criterion to determine a change in patient perceived status (Kirkley & Griffin, 2003). This often consists of a single question about the effect of injury on the subject's daily function over the past week and is administered at each test occasion. It is usually worded along the lines of "have you experienced: improvement, no change or deterioration"? It is commonly used as the criterion or 'anchor' that determines if the patients have improved or are considered stable. This baseline anchor permits a comparison of statistical values to be made in order to calculate the effect size and ROC statistics. It also provides the patients current global functional assessment without requiring extensive recall (Paul et al., 2004; Schmitt & Fabio, 2004).

Further difficulty with measuring responsiveness is documented by researchers Terwee et al., (2003) who studied the currently used statistical calculations of responsiveness and argue that the MCID may vary for some conditions and for differing levels of functioning, thus it is not a fixed property of an instrument. Furthermore they report that defining a MCID for a specific study requires a judgement of what important change is, thus if the purpose of measuring HRQOL is from the patient's perspective then only a patient can judge what an important change is. This line of argument therefore indicates what is judged as MCID will vary with data obtained from different population groups and levels of disability (Terwee et al., 2003). Thus extensive studies on responsiveness need to be performed in order to make comparisons between studies possible.

Interpretation of the MCID scores by clinicians needs to occur in order to incorporate the use of outcome measures into clinical practice, particularly to answer the 'why bother' question. In other words, has the clinician secured enough information that has made the exercise of obtaining the information worthwhile and will it assist with the rehabilitation planning for that patient? Additionally the question can be asked, is this information gathering process beneficial for the patient from the patient's perspective?

It raises the issue of how much the patient is engaged in the development of their rehabilitation planning and their understanding of the role of functional outcome measures and how they may be used advantageously for them. However, this topic is outside the scope of this dissertation.

In summary, responsiveness is the ability of an instrument to measure change over time and allows the clinician to interpret if a meaningful change in patient status has occurred. This is a pivotal use of an instrument in a clinical setting in order to determine treatment strategies or supply quantitative results for interested parties. Clinicians need to be aware of how to interpret the MCID in instruments of choice and have confidence that these indices have been established following rigorous research. Furthermore, clinicians also need to be aware that the MCID needs to be established in order for the instrument to have some clinical value but that the parameters for these indices are far from clear due to research difficulties.

2.3.5 Specific Whole Limb Upper Extremity Functional Outcome Measures

Outcome measures that are disease or area specific to the upper extremity are numerous (Beaton et al., 2001). In a systematic review of the literature by Bot et al (2004) specifically evaluating only shoulder disability questionnaires no less than 16 questionnaires were identified that met the inclusion criteria of having evidence of evaluation of psychometric scrutiny. The authors revealed that no one questionnaire demonstrated satisfactory results for all recommended psychometric properties. In particular they identified that not all questionnaires presented research involving responsiveness and determining the MCID index. This example highlights the dilemma a clinician has in choosing an outcome measure for the upper extremity that has had psychometric testing and user-friendly clinical application. In a recent survey monitoring change in the implementation of outcome measures by physiotherapists it was documented that several barriers existed to using validated outcome measures in a private practice clinical setting. The key barriers to use were: the time required to administer the tests, lack of training and familiarity with the tests and measures, in particular the inability to interpret the scores and change in scores and finally the accessibility to appropriate outcome measure instruments themselves (Abrams et al., 2006). To address these particular limitations of outcome measure use for clinical

settings, the ‘whole limb’ approach has been successfully adopted. An advantage of this concept is not having to be familiar with numerous condition specific outcome measures for the whole upper extremity range of possible pathologies, some of which may not have evidence of psychometric testing. Accessibility is also an issue for many clinicians where copyright laws and cost may prohibit ready access for the clinician to not only the questionnaires but instructions on interpreting scores (Abrams et al., 2006; Horner & Larmer, 2006). Three upper extremity questionnaires are documented as having had psychometric testing for interpretability and meet the criteria of accessibility: the Upper Extremity Functional Index (UEFI) (Stratford et al., 2001), the Disability Arm Shoulder Hand (DASH) (Beaton et al., 2001) and more recently the shortened version of the Disability Arm Shoulder Hand (QuickDASH) (Beaton et al., 2005). For a summary of the psychometric analysis of the whole limb upper extremity functional outcome measures refer to Table 2 (pg 25).

2.3.6 Upper Extremity Functional Index (UFEI)

The UEFI is a 20-item instrument looking exclusively at ADL function using a Likert scale. The scores are totalled with a possible score being between 0-80. A lower score indicates poor function. An important goal of the designers was for the instrument to be completed in less than five minutes by the patient and be scored quickly by the therapist without computer or calculator aides. The MCID is calculated at 6 points. To date this instrument has been examined in one pilot study only (Davidson, 2004; Stratford et al., 2001). There were limitations with the study in that there were no reports on the actual pathologies being assessed, thus it is not certain if a range of upper limb pathologies were being examined.

2.3.6.1 Validity

The validity of the UEFI was examined by measuring it against the Upper Extremity Functional Scale (UEFS) (Stratford et al., 2001). The authors measured 47 patients with upper limb conditions at a physiotherapy outpatient setting at baseline, 48 hours and at a three week timeframe. With respect to cross-sectional convergent validity the coefficient between the UEFI and the UEFS plus two pain scales was examined. The results revealed a correlation of > 0.6 suggesting that the UEFI has a moderate to good relationship with the UEFS. However a major limitation with this study was that there was no documentation that the UEFS is a valid measure to make a comparison with.

2.3.6.2 Reliability

The test-retest reliability results indicated evidence of reliability with an ICC of > 0.8, however this is the only data on reliability to date (Stratford et al., 2001).

2.3.6.3 Responsiveness

The responsiveness was calculated by the SEM 95% CI. The results demonstrated a SEM (95% CI) of 3.9 (3.2 to 5.0) with the resulting MCID of a patients score being 6 points (Stratford et al., 2001).

In summary, the UEFI may be easily obtainable from a web site and be a very simple format to use in the clinical setting. However, care should be taken in interpreting the scores in the clinical setting until more research has established its psychometric properties. For concise summary refer to Table 2.

2.3.7 Disability Arm Shoulder Hand Index (DASH)

The DASH is a 30-item questionnaire using a Likert scale with domains that cover functional activity, symptoms (including pain) and social function. There is an optional high performance sport/music and work section, a total of 4 items on each section. This section has been designed to identify the specific difficulties that may be experienced in a sport or work place but not impact on the ADL's of the patient and thus goes undetected in the 30-item section.

At least 27 of the 30 items must be completed for the score to be calculated. The assigned values for all completed responses are summed and averaged, producing a score out of five. This value is then transformed to a score out of 100 by subtracting one and multiplying by 25. A higher score indicates greater disability. It is rated as a moderately difficult scoring system, defined as needing a calculator to determine the score (Bot et al., 2004). The MCID recommended is a change of 15 points (MacDermid & Stratford, 2004). This parameter was calculated by comparison to normative data and across groups of disabilities (Beaton et al., 2001; Skutek, Fremerey, Zeichen, & Bosch, 2000).

2.3.7.1 Validity

Beaton and co-workers (2001) examined the validity of the DASH with specific reference to the upper limb acting as a single functional unit by evaluating patients with

a variety of upper limb pathologies over a three-month period. The patients that had shoulder conditions were compared with the Shoulder Pain and Disability Index (SPADI) and those that had hand conditions were compared with the Brigham questionnaire (disease specific to hand conditions). The DASH was found to correlate well with the other questionnaires (r = >0.69) and to discriminate well, for example between patients who were working and those who were not (Beaton et al., 2001). Additional support for the construct validity was demonstrated by evaluating the DASH with both the physical (PCS) and mental (MCS) components of the SF-36 in patients with upper limb pathologies over a four month period (SooHoo et al., 2002). Using Pearson correlation coefficients the DASH revealed moderate correlation (-0.3 to -0.6) to several of the subscales in the SF-36, irrespective of anatomical area of complaint, supporting the construct validity of the DASH as a measure of health status. This result also encourages the broader use of function versus focusing on the anatomical area of the shoulder or hand. Floor and ceiling effects were only viewed with comparison to the MCS subscales. However there were a large number of participant dropouts during this study, which potentially affects results.

An extensive number of studies document validity of the DASH involving a wide range of sample populations, including postoperative rotator cuff rehabilitation (Skutek et al., 2000), humeral fracture (Robinson & Page, 2003), hand and elbow conditions (MacDermid & Tottenham, 2004), different cultures (Imaeda et al., 2006), non-clinical populations (Jester, Harth, & Germann, 2005) and prognostic evaluations (Jester, Harth, Wind et al., 2005).

2.3.7.2 Reliability

Beaton et al., (2001) established a test-retest ICC of 0.96 in a study population of 86 patients, a subgroup of a study population of 172 patients with upper extremity dysfunction. Further support for the test-retest reliability has come from Schmitt & Fabio (2004) with an ICC rating of 0.91 and MacDermid & Stratford (2004) where the ICC ratings exceed 0.9. This provides strong evidence that the DASH is a reliable instrument in a variety of sample populations.

2.3.7.3 Responsiveness

The responsiveness of the DASH was examined by researchers with the results (SRM 0.74 – 0.80, ROCauc –15 or –20 correctly rated 68% and 72% accuracy respectively)

indicating that the DASH could measure change at a group and individual level analysis (Beaton et al., 2001).

Additional evaluation of responsiveness by Schmitt & Fabio (2004) was undertaken using a range of statistical methods with particular reference to the indices that would reflect change in individuals. The authors compared the DASH to other shoulder questionnaires with the results documenting that the DASH SEM ranged from 5.22 to 5.86 and was lower than the disease specific shoulder questionnaires. This may be reflective of the DASH compromising responsiveness in the shoulder for being an instrument that can measure a wider range of conditions. Evidence for the DASH being a responsive instrument is viewed in a range of study populations; in hand injury patients (MacDermid & Stratford, 2004), following rotator cuff and carpal tunnel surgery (Gummesson, Atroshi, & Ekdahl, 2003; Kotsis, Chung, & Arbor, 2005) and thumb surgery (Smet, 2004), again signalling a depth of psychometric analysis.

In summary, the DASH is a widely used outcome measure for the upper limb. It has evidence of validity, reliability and responsiveness in upper extremity conditions. For a concise summary refer to Table 2. The DASH has been assessed across a variety of sample populations with a degree of consistency among the results. Its limitations are that it has a moderately complex scoring system, as previously defined and it is considered a lengthy questionnaire. This has lead to the development of the shortened version, the QuickDASH.

2.3.8 Shortened Disability Arm Shoulder Hand (QuickDASH)

The QuickDASH comprises of an eleven-item version chosen from the full length DASH, plus the additional sections on work and sport/recreation section. Beaton et al (2005) compared three different but independently validated methods of item-reduction to create the eleven-item version of the DASH with the following items being selected: six from the ADL domain (items 1, 7, 10, 14, 16, & 18), two from the social domain (22 & 23) and three from the symptom related domain (24, 26 & 29), (see Appendix A). The scoring system remains the same as the DASH with a tolerance of only one missing item.

2.3.8.1 Validity

Research to date has determined validity of the QuickDASH as an instrument to assess upper extremity musculoskeletal disorders (Beaton et al., 2005; Gummesson et al., 2006; Imaeda et al., 2006). Beaton et al., (2005), as part of the item retention research on the DASH, established construct validity with further validity being substantiated by Gummesson and co-workers (2006) using a similar technique of analysing the correlation coefficients between the DASH and the QuickDASH in a large post operative study population. Additional validation however has been documented by comparing the QuickDASH responses with the SF-36 subscales demonstrating construct validity of the QuickDASH in both a cross-cultural population (Japanese) (Imaeda et al., 2006) and in a large working population (n = 559) with both non symptomatic and varied upper extremity disorders populations in the study group (Stover, 2004). There is however a scarcity of research that analyses the validity of the work and sport/music sections of both the DASH and QuickDASH. Stover (2004) assessed the validity of the QuickDASH as a measure of screening upper extremity disorders in the workplace with the population including workers from industrial and sedentary environments plus a range of experience, repetition and forces present in the job descriptions. The work module score was assessed separately and compared to the SF-12 physical and mental component scores. The results confirmed that the work module could discriminate well between groups based on diagnosis and symptom severity status. Interestingly, workers reported less effect from upper extremity disorders on their work, as addressed in the work module, than on household chores and other ADL's as addressed by the QuickDASH, for all but the most severe symptoms. The author proposed that this may be due to workers being reluctant to reveal the degree or impact their disorder may have on their work to their employers (Stover, 2004).

2.3.8.2 Reliability

Whilst there is a paucity of research on the QuickDASH, the initial findings reveal that it has a strong reliability in the population groups studied. Beaton et al., (2005) reported an ICC ≥ 0.94 and Cronbachs alpha ≥ 0.92, which is supported by Gummesson et al., (2006) with an ICC of > 0.9. Recently, the QuickDASH was researched in a Japanese society, to review the psychometric properties following translation, with a slightly lower reliability being established (ICC 0.82 and Chronbach's alpha 0.88) (Imaeda et al., 2006).

2.3.8.3 Responsiveness

In the literature to date there have been two studies that have specifically examined the responsiveness of the QuickDASH, both measuring change in a pre and post surgical patient population (Gummesson et al., 2006; Imaeda et al., 2006). However, in both studies the QuickDASH responses were extracted from the full-length DASH. The results indicate that the QuickDASH is sensitive to change over time in the population groups studied. Gummesson et al., (2006) specifically analysed both group and individual level responsiveness in a large (n = 109) Swedish population awaiting elective surgery for upper extremity disorders (using the Swedish version of the DASH). The effect size and standardised response mean revealed a moderate result (ES 0.50, SRM 0.63) with the ROCauc signalling a stronger responsiveness by being able to discriminate among groups that differed in self rating improvement. Similar group level responsiveness results are recorded in the Japanese population receiving surgery for carpal tunnel syndrome (ES 0.37, SRM 0.54) (Imaeda et al., 2006). Consideration needs to be given to the participants being pre and post surgical intervention, which possibly results in large score change compared to other upper extremity disorders that are exposed to other treatment interventions (Terwee et al., 2003).

In summary, the QuickDASH is reported to be a more efficient version of the DASH while retaining the psychometric properties. For a concise summary refer to Table 2. If further research supports this concept then the QuickDASH is likely to have more clinical benefits as an upper limb outcome measure.

Table 2: *Summary of Psychometric Analysis of Whole Limb Upper Extremity Functional Outcome Measures*

Functional Outcome Measure	Validity	Reliability	Responsiveness	Advantages	Limitations
Upper Extremity Functional Index (Stratford et al., 2001)	Moderate validity established	ICC established	Established MCID 6 points	Quick and easy to use. Simple scoring system	Very limited research to date
Disability Arm Shoulder Hand Index (DASH) (Hudak et al., 1996)	Repeated evidence of validity established	Repeated evidence of reliability established	Repeated evidence of responsiveness established. MCID 15pts	Extensive body of supporting research	Lengthy questionnaire. Moderately complex to score.
Shortened DASH QuickDASH (Beaton et al., 2005)	Moderate validity established	Moderate reliability established	Limited responsiveness established	Quick to complete. Evolving body of supportive research	Limited research to date

2.3.9 Clinical Limitations of the QuickDASH

All studies on the QuickDASH in a patient population have been computed from the full-length DASH responses. Research needs to be conducted to determine if the patient responses to the 11 items would differ if only the QuickDASH was administered. Reducing the burden to using functional outcome measures to both the physiotherapist and patient is desirable and may be more achievable with a shortened instrument, which takes a quicker time to complete and minimizes the risk of missing data.

Research available to date has indicated that the QuickDASH has potential to be a valid, reliable and responsive instrument for measuring upper extremity musculoskeletal disorders (Beaton et al., 2005; Gummesson et al., 2006). However, there is still a lack of research that documents vigour of psychometric strength in the ability for the QuickDASH to measure health related quality of life parameters. It is yet to be established if the QuickDASH is reliable and responsive in a New Zealand primary care population. Further, to ascertain psychometric properties, an instrument needs to be tested over a diverse study population, especially the population on which its use is intended. Current research looking at the responsiveness of the QuickDASH in a patient

population has examined only pre and post surgical status where large changes in symptoms are expected.

2.4 Summary

Upper extremity injuries are common and costly to society, inclusive of New Zealand. They may result in a functional deficit that can remain a long-term burden to the patient. Traditionally physiotherapists have assessed upper extremity injuries objectively to measure outcome from treatment. However, there is a growing need to measure the functional deficit clinically in order to ensure that treatment protocols are restoring the patient's functional capacity. Incorporating psychosocial assessment into physiotherapy practice permits an evaluation of the functional status resulting from a treatment intervention. Functional outcome measures need to demonstrate validity, reliability and responsiveness. Responsiveness is of particular importance to the clinician as it identifies if a meaningful change has occurred in a patient's status. Key barriers to the use of functional outcome measures in clinics have been identified. The 'whole limb' concept addresses most of these barriers, in particular time of administration and familiarity with the forms. The DASH is a recognised upper extremity functional outcome measure with the recently developed shortened version (QuickDASH) showing a greater potential for clinical use. However the QuickDASH still has limited psychometric analysis especially in the primary health care setting.

The purpose of this pilot study is to examine the QuickDASH in a clinical physiotherapy setting, as there is no research to date that has examined the QuickDASH in this population group. It needs to be established if the QuickDASH is sensitive enough to measure responsiveness where a range of upper extremity disorders and associated treatment strategies occur and therefore where a change in health status may be more difficult to discriminate.

Chapter 3 Pilot Study

This chapter outlines a pilot study that investigated the test-retest reliability and responsiveness of the QuickDASH functional outcome measure. The chapter is divided into three sections; firstly the aim of the pilot study is outlined followed by an account of the methodology and the mode of data analysis is summarised.

3.1 Aim

The aim of this study was to analyse the test-retest reliability and responsiveness of the shortened version of the Disability Arm Shoulder Hand Questionnaire (DASH), the QuickDASH, in New Zealand private physiotherapy clinics on participants with upper extremity musculoskeletal injuries. While the QuickDASH includes optional modules on sport/performing arts and work related tasks, an analysis of these components was not examined and was considered outside the scope of this dissertation.

3.2 Methodology

3.2.1 Recruitment

In accordance with the Auckland University of Technology Ethics Committee (AUTEC 06/173) (Appendix B) approval, private physiotherapy practices of convenience were contacted by the investigator (KP) and invited to participate in the pilot study. The practices were selected from both rural and urban locations and from a variety of areas of practice, for example upper limb and sports medicine physiotherapy clinics. Six practices were recruited between November 2006 and February 2007. They were informed of the study protocol, inclusion and exclusion criteria. Signed consent from the practice principal was gained in order to conduct research in their clinic and to collect data from patients (Appendix C). Any physiotherapist working in these clinics was able to recruit patients. Participants that presented for treatment for an upper extremity musculoskeletal injury and met the inclusion criteria were invited by the treating physiotherapist to participate. Those participants who agreed and met the inclusion criteria were given a participation information sheet (Appendix D) and signed a consent form (Appendix E).

3.2.2 Inclusion criteria

Participants aged 20 and over who suffered from a musculoskeletal condition to the upper extremity that required physiotherapy treatment were considered eligible for the research.

3.2.3 Exclusion criteria

Participants were excluded from the research if they suffered from systemic or metabolic diseases. They were also excluded if there were symptoms arising from the cervical spine including radiculopathy or in the first four months following major reconstructive surgery. Subjects were excluded if they were unable to read or were not fluent in written English.

3.2.4 Procedure

The participants filled the QuickDASH questionnaires on three occasions (Appendix A). The first was filled out at the commencement of treatment (QuickDASH 1), with the second (QuickDASH 2) being completed 24 - 48 hours later either at the second appointment or being returned by the patient if the next treatment did not occur within that time frame. The third questionnaire (QuickDASH 3) along with the Patient Discharge Global Assessment Questionnaire (Appendix F) was completed at discharge or at six weeks after the treatment, whichever event occurred first. At this point the participants were asked to complete the Discharge Global Assessment Questionnaire to rate their a) current overall condition of their upper extremity injury and b) the perceived change in status of their condition. These provided the anchors for calculation of the MCID. At the same time the treating physiotherapist also completed a Physiotherapist discharge Global Assessment Questionnaire to record their perceived current overall condition of the patient's upper extremity injury (Appendix G). It was requested of the treating physiotherapist to complete their Discharge Global Questionnaire independent of and without discussion with the patient. At this time the physiotherapist also completed a patient demographic information form (Appendix H).

3.3 Data Analysis

The QuickDASH questionnaire scores were standardised to a 0 -100 scale across all three questionnaires. The assessment of reliability involved the calculation of the following statistics. Short-term reliability was measured by comparing the results of QuickDASH 1 and QuickDASH 2, which was collected 48 hours later.

An Intraclass Correlation Coefficient (ICC) was calculated using a two-way mixed model with the mode of assessment as the fixed variable and the subjects as a random variable (Deyo et al., 1991). Additionally, Bland and Altman plots were used to demonstrate the distribution error between QuickDASH 1 and QuickDASH 2, in addition to the calculation of bias and limits of agreement (Bland & Altman, 1986).

The assessment of responsiveness was achieved by calculating the: effect size (ES) and the standard response mean (SRM). The effect size was calculated by taking the difference in the means between QuickDASH 1 and QuickDASH 3, and dividing this difference by the standard deviation of QuickDASH 1 results (Deyo et al., 1991). The SRM was calculated by the mean change score of improved subjects divided by the standard deviation of the change score in improved subjects (Deyo et al., 1991).
The examination of correlation between patient and physiotherapist discharge global assessment on the rated current overall condition of the injury was performed with an ICC analysis.

Data was analysed using the Statistical Package for Social Sciences (SPSS for Window, Rel 14, Chicago, USA; SPSS Inc) and Graph Pad Prism (Graphpad Software Inc, San Diego, Ca 92130, USA). The alpha level was set to 0.05.

Chapter 4 Results

This chapter is divided into four parts. Firstly the participants demographic results and secondly the responses to the three QuickDASH and Global Discharge responses are reported. Thirdly the results from the reliability analysis and finally the responsiveness analysis are reported.

4.1 Demographics

The demographic data is presented in Table 3. There were a total of thirty-five participants who completed all three QuickDASH questionnaires plus the global discharge questionnaire. There were twenty-one males and fourteen females, with the mean age being 48 years. The average time between date of injury and commencement in the study (first treatment) was nine weeks, with the range being one to forty-five weeks. This data was only available from thirty participants as in five cases the date of injury was recorded as a month and year only preventing the calculation of time from date of injury to commencement in the study from being conducted. In 18 subjects the affected arm was the left arm and for 17 the affected arm was the right arm, with just over half the population having the problem in the dominant arm (n=19).

Table 3: *Participant Demographics*

Characteristic	n
Gender	
Male	21 (60%)
Female	14 (40%)
Age (years)	
Mean (SD)	48.7 years (15.7)
Minimum	19 years
Maximum	78 years
Time between injury and participation in study (weeks) n = 30	
Mean (SD)	9.77 weeks (11.50)
Median	6
Minimum	1 week
Maximum	45 weeks
Arm Affected	
Left	18
Right	17
Arm Dominance	
Left	4
Right	31
Dominant arm affected	19

SD = Standard Deviation

Table 4 presents the range of upper extremity pathologies that subjects presented with, the most common being rotator cuff injuries (n = 17) followed by tennis elbow pathology (n = 5).

Table 4: *Range of Upper Extremity Pathologies Presented*

Pathology	n
Rotator Cuff Injury	17
Tennis Elbow	5
De Quervains	3
Fracture Radius	2
Shoulder Instability	2
Wrist Strain	2
Fracture Clavicle	1
Fracture Humerus	1
Thumb Strain	1
Shoulder Capsulitis	1

4.2 QuickDASH Responses

The maximum possible score on the QuickDASH is 100, indicating the greatest level of patient perceived disability, the lowest score being zero, indicating no patient perceived disability. Table 5 presents the mean, standard deviation, median and range of scores for all three time intervals.

Table 5: *Descriptive Statistics for QuickDASH 1, 2 and 3*

Questionnaire	Mean	St Deviation	Median	Range
QuickDASH 1	37.7	18.6	36.3	11 - 73
QuickDASH 2	33.3	19.2	29.5	2 - 70
QuickDASH 3	18.5	15.9	15.9	0 - 55

The Patient Discharge Global Assessment Questionnaire (PGA) had two scales: one rating the overall condition of the upper extremity, the second scale measured change in overall status (discussed later in this section). The Physiotherapist Discharge Global Assessment Questionnaire (PhGA) had one scale rating the overall condition of the upper extremity at the present time, identical to the first scale on the PGA. The overall condition scale was a ten anchor scale.

This scale was grouped according to anchors: 1 – 2 measuring significant limitations that affect activities of daily living, 3 – 5 moderate limitations that affect activities of daily living, 6 – 8 some limitation of daily living and 9 – 10 no limitation to daily living. The results are presented in Table 6.

Table 6: *Comparison of the Two Global Discharge Scores Rating the Current Overall Condition of Upper Extremity*

Range	**Patient Raw Score**	**Percentage**	**Therapist Raw Score**	**Percentage**
1 – 2	0	0%	0	0%
3 – 5	10	28%	8	23%
6 - 8	14	40%	16	46%
9 - 10	11	31%	11	31%
Total	**35**		**35**	
Mean score (SD)	6.98 (2.18)		7.11 (2.21)	

SD = Standard Deviation

The correlation between the patient's perception of the current overall condition of their upper extremity and the therapist's perception of the patient's level of current disability was high with a Pearsons Correlation of 0.88 (Table 7).

Table 7: *Correlation Between Patients and Therapists Perception of Current Overall Condition of Their Upper Extremity*

		Therapist	**Patient**
Therapist	Pearsons Correlation	1	.88*
	Sig. (2-tailed)		.00
	N	35	35
Patient	Pearsons Correlation	.88*	1
	Sig. (2-tailed)	.00	
	N	35	35

* Correlation is significant at the 0.01 level 2 tailed

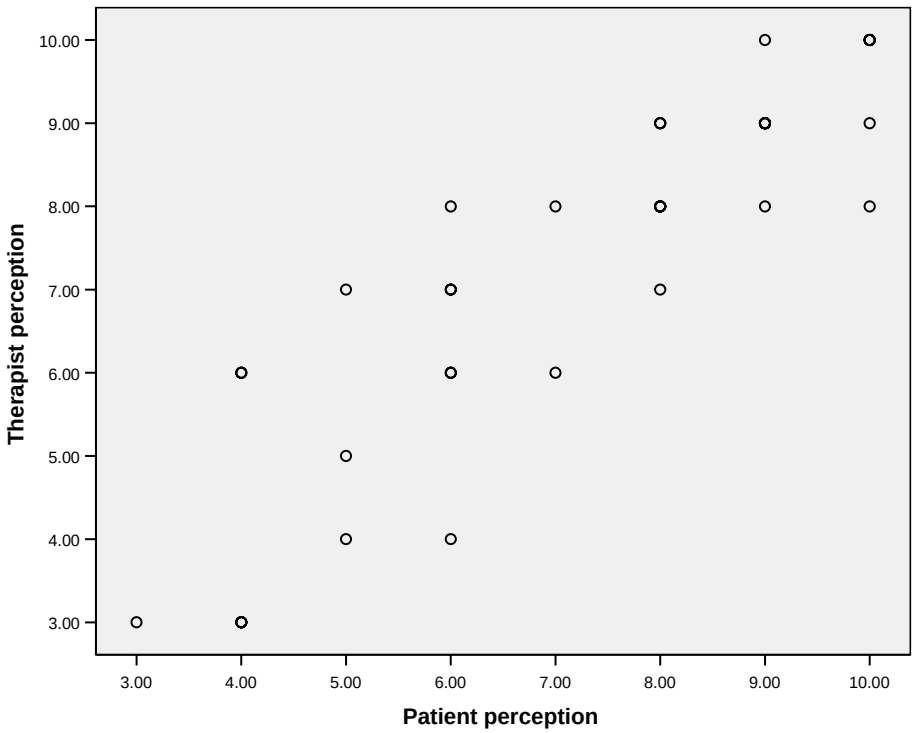
An internal consistency between the patient and therapist perception of level of disability was analysed with a Cronbach Alpha statistic of 0.93. There was a high level of agreement with the Intraclass Correlation Coefficient of 0.88, refer to Table 8.

Table 8: *Intraclass Correlation Coefficient Between Patient and Therapist Perception of Current Overall Condition of Their Upper Extremity*

	95% Confidence Interval			F test with True Value 0			
	Intraclass Correlation	**Lower Bound**	**Upper Bound**	**Value**	**df1**	**df2**	**Sig**
Single Measures	.88	.78	.94	15.65	33	33	.00
Average Measures	.94	.87	.97	15.65	33	33	.00

See Figure 1 for illustration of correlation graph where the data points demonstrate a consistent pattern

Figure 1. Correlation graph between patient and therapist perception of current overall condition of the upper extremity



4.3 Test-Retest Reliability

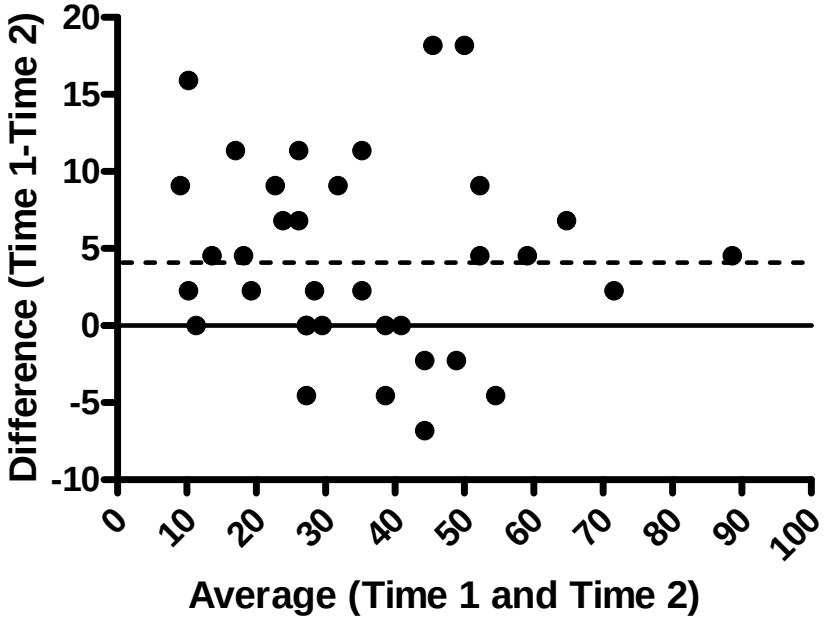
The ICC between QuickDASH 1 and QuickDASH 2 was 0.94 (>0.8), see Table 9

Table 9: *Intraclass Correlation Coefficient for QuickDASH 1 and 2*

	Intraclass Correlation	95% Confidence Interval		F Test with True Value 0			
		Lower	Upper	Value	df1	df2	sig
Single measures	.94	.89	.97	34.98	34.0	34	.00
Average Measures	.97	.94	.99	34.98	34.0	34	.00

Bland and Altman plots revealed a bias of 4.3, with a standard deviation of 6.3 and 95% limits of agreement ranging from -8.03 to 16.73. The spread of data can be viewed in Figure 2. The Bland and Altman graph shows a large distribution of error across the range of scores that may not be random, while the moderate bias score of 4.3 (>1.0) may indicate a systematic error occurring across days

Figure 2. Bland and Altman data showing the differences in scores across QuickDASH 1 and QuickDASH 2 versus the average of the questionnaires



Difference (Time 1 – Time 2) = between days difference between measurements, Average (Time1 – Time 2) = mean measurements between QuickDASH1 and QuickDASH 2

4.4 Responsiveness

Responsiveness measures were taken from two time measures QuickDASH 1 and QuickDASH 3 and using the Patient Discharge Assessment Questionnaire scale to record patient perceived change in overall status.

4.4.1 Effect Size

Group descriptive statistics between QuickDASH 1 and QuickDASH 3 are outlined in Table 10.

Table 10: *Group Descriptive Statistics of Mean Change in Score Between QuickDASH 1 and 3*

	n	Mean	Std Deviation
QuickDASH 1	35	37.7	18.6
QuickDASH 3	35	18.5	15.9
Difference		-19.1	17.2

The Kazi's effect size was calculated to be 1.02 (difference QuickDASH 1 + 3 / SD QuickDASH 1 score) and the Standard Response Mean (SRM) (difference QuickDASH 1and 3 / SD of the change in scores) was calculated to 1.1, both are considered to be a strong effect size (< 0.8).

4.4.2 Minimal Clinical Importance Difference

On the Patient Discharge Global Assessment Questionnaire, the change in overall status had six anchors; zero measuring very much improved, one much improved, two minimally improved, three no change, four minimally worse, five much worse and six very much worse. There were no patients who perceived their change to have deteriorated (Table 11).

Table 11: *Patient Responses to Overall Change in Upper Extremity Change*

Anchors	n	percentage
0 = very much improved	10	28%
1 = much improved	15	43%
2 = minimally improved	8	23%
3 = no change	2	6%
4 = minimally worse	0	
5 = very much worse	0	

Descriptive statistics for the four groups is viewed in table 12. An improved score is from a high score to a low score thus the change in score is represented as a negative

Table 12: *Descriptive Statistics for Patient Perceived Change*

	Very much improved (n=10)	Much improved (n=15)	Minimally improved (n=8)	No change (n=2)
Mean change of score (SD)	-30.45 (20.33)	-18.78 (12.63)	-13.07 (9.55)	10.22 (14.46)
95% CI Lower	-45.00	-25.79	-21.05	-119.72
95% CI Upper	-15.90	-11.79	-5.09	140.17
Max	-65.91	-47.73	-31.82	0.00
Min	-6.82	-2.27	-2.27	20.45
SEM	6.43	3.26	3.37	10.22

SD= Standard Deviation, CI = Confidence Interval, Max= Maximum score, Min= Minimum score, SEM= Standard error of mean

The anchor chosen for calculating the MCID was from the much improved measure with the mean change in score being -18.78, therefore the MCID for the QuickDASH is deemed to be 19 points.

As the minimal clinical change was of interest, a T-Test was performed using two groups: group one was the patients who considered their change as much improved and group two was the combined group of those patients who had minimal improvement or no change. The comparison did not reach statistical significance ($p = 0.07$). The results are outlined in Table 13.

Table 13: *Group Statistics Comparing Minimum Change and No Change with Much Improved Patient Perceived Changes*

Group	n	Mean	SD	SEM
Diff group 1	15	-18.78	12.63	3.26
Diff group 2	10	-8.40	13.80	4.36

Group 1 = much improved, Group 2 = minimally improved + no change,
Diff = difference, SD = Standard Deviation, SEM = standard error of mean

Chapter 5 Discussion

This chapter is divided into eight sections. Firstly the results from the test-retest reliability analysis will be discussed followed by discussion on responsiveness analysis. Thirdly the relationship between patient and physiotherapist's subjective analysis of ability at discharge will be covered followed by discussion on the mean scores. In the fifth section the demographic findings will be discussed. Limitations to the present study will be covered in the sixth section followed by implications for further research. The eighth section will outline the conclusion from this pilot study.

5.1 Test-Retest Reliability

The test-retest reliability results indicate that the QuickDASH is a reliable instrument with a high ICC (0.94) and a Bland and Altman moderate bias (4.3 SD: 6.3). Test-retest reliability is measured by the presence of a stable population to determine instrument error. What determines a stable population is a patient global rating questionnaire, thus it can be argued that in the current study the test-retest reliability was not conclusively established. This was due to the absence of a patient global rating questionnaire being completed concurrently with the QuickDASH 2 questionnaire in order to identify patients that may have changed in health status (Bot et al., 2004). Therefore, it could not be conclusively determined that the test-retest of the QuickDASH was measuring instrument error versus patients that may have had a change in health status. This is particularly relevant in physiotherapy management as patients often receive treatment on the day of the initial assessment, (when the baseline QuickDASH 1 was completed) creating an environment where change could have occurred between completing the two reliability questionnaires.

Despite not establishing a stable population via a global rating questionnaire at the time of completing QuickDASH 2, some confidence has to be gained in that there was only a 4-point difference between the mean scores of QuickDASH 1 and QuickDASH 2. Additionally, the average time from injury to presentation for treatment (baseline measure) was nine weeks. At nine weeks post injury the healing time frame would categorise this as sub-acute, with one week considered acute, 2-12 weeks sub-acute and 12 weeks on chronic (Hunter, 1994). Given that many upper extremity pathologies are

likely to be in the sub-acute phase of healing it could be estimated that there would be no dramatic changes in health status after 1-2 days which was the time duration between completing QuickDASH 1 and 2 (ACC, 2004; Belvins, Djurasovic, Flatlow, & Vogel, 1997)..

The test-retest reliability results from this study compare favourably with the two previous studies that have examined the QuickDASH. Gummesson et al., (2006) investigated test-retest reliability on a subgroup of 30 patients on two occasions prior to surgery with a resulting ICC of 0.9; however it was not clear if this patient group was stable via an anchoring questionnaire. Imaeda et al., (2006) also examined reliability on a subgroup of 38 patients who received no treatment (surgery or medication) and were retested 1- 2 weeks later with a resulting ICC of 0.82. Again there was no conclusive evidence in the research to determine if this had been performed on a stable population via a global rating questionnaire. Additionally, neither study performed any concurrent data analysis of reliability such as Bland and Altman plots or typical error (Hopkins, 2000; Rankin & Stokes, 1998). Rankin & Stokes, (1998) argue that singular reliability studies such as an ICC provide insufficient information and recommend that both ICC and Bland and Altman plots are used concurrently to measure reliability**.** The ICC for this study was 0.94. In clinical situations an ICC of 0.9 is regarded as the minimal measure of consistency and 0.95 is desirable (Bland & Altman, 1986).The Bland and Altman graph (figure 2) demonstrated a moderate bias score of 4.3 (>1.0).The relevance of this is that this study cannot indicate if the bias is a result of instrument error or if in fact measurement has included patients who have had a change in health status as presented by a change in QuickDASH 2 scores. The 95% limits of agreement are also large, ranging from -8.03 to 16.73 points. This range reports that 95% of the time, an individuals difference scores will fall within this range and indicates to the clinician a normal variation (Rankin & Stokes, 1998). It is further argued that limits of agreement are affected by sample size (Hopkins, 2000). Given that the reliability studies to date on the QuickDASH involve small but similar sample sizes (n = 30-38), false representation of reliability may have occurred.

5.2 Responsiveness

5.2.1 Effect Size and Standard Response Mean

The findings of this study indicate that the QuickDASH is a responsive questionnaire in a clinical setting as determined by the comparable and strong effect size (ES) (1.02) and standard response mean (SRM) (1.1). The ES and SRM are similar conceptually with the values being interpreted as large if > 0.8, moderate if between 0.5-0.8 and small for 0.2-0.5 (Horner & Larmer, 2006). In comparison to other research, this study had a large ES and SRM. Gummesson et al., (2006) and Imaeda et al., (2006) both reported moderate to low ES and SRM (ES 0.5, SRM 0.63: ES 0.3, SRM 0.5 respectively). The difference in results may be due to the methodology of computing the QuickDASH from the full DASH questionnaire. A higher but still moderate SRM (0.7) was reported by Beaton et al., (2005) when designing the QuickDASH. As this is the first study where the QuickDASH was presented in its independent format, the resulting ES and SRM support the QuickDASH as being a responsive instrument and may indicate that it is a more sensitive to change when used in this particular format. This view is reinforced with the knowledge that the population group studied by both Gummesson et al., (2006) and Imaeda et al., (2006) were measured pre and post surgical intervention and measured over a longer period of time (3 to 12 months) where a larger degree of patient perceived change may have been expected. Additionally, this study has demonstrated that the QuickDASH questionnaire is responsive to intervention involving a six week period of physiotherapy treatment, making it clinically relevant for use in a physiotherapy clinic.

The analytical technique for ES and SRM assume that all patients change in the same direction and imprecision is introduced if patients who do not improve are included in the summary statistic (Laing et al., 2002). In this study no patient reported that their condition had deteriorated during the six week treatment period giving support to the findings reported.

5.2.2 Minimal Clinical Importance Difference

The MCID was calculated to be 19 points based on the anchor approach. In this study the patients who rated themselves as having no change or minimally improved were used as an indicator of measurement noise whereas values for patients who reported they were much improved indicated the point of separation between instrument noise

and meaningful results. The patients in the much improved group had a mean score change from QuickDASH 1 to QuickDASH 3 of 18.78 (SD:12.6) points.

A secondary analysis was performed to determine if the QuickDASH was sensitive to change in patient's perceived change by comparing those in the no change and minimally changed group with the much improved group. The result did not reach statistical significance ($p = 0.07$). This result however may be because the sample size for these groups was very low (n=2 in the no change group, n=8 in the minimally change group and n=15 in the much improved group). A larger sample size may have produced a significant result.

When comparing the QuickDASH MCID to the full DASH questionnaire, the findings of this study rate the MCID as slightly larger. MacDermid & Stratford, (2004) summarised the DASH psychometric findings and report an MCID of 15 as being an acceptable measure of true patient perceived change in a physiotherapy clinical setting (MacDermid & Stratford, 2004). It is difficult to know given the limited research on the QuickDASH if this difference of 4 points between the QuickDASH and DASH is of significance or is in fact a comparable finding.

The clinical implications of the results of this study are that physiotherapists using the QuickDASH for upper limb pathologies can be confident that when a score change occurs from a baseline measure by 19 points then a true change has occurred in the patients perceived health status.

5.3 Patient and Physiotherapist Discharge Global Assessment

The discharge global assessment comprised of two scales. In the first instance patients were asked to rate the overall condition of their upper limb at discharge or at the six week mark; this was replicated in the physiotherapist assessment. As the scale was only used at discharge it does not therefore measure perceived change in health status but rates the perceived level of disability as recorded by the patient and (separately) by the physiotherapist. Thus this scale documents the disease impact at the time of completing the questionnaire, not the treatment benefit. There was a strong correlation and level of agreement between the determined level of disability recorded by the patient and physiotherapist (ICC of 0.88). The significance of this finding is that physiotherapists

appear to accurately interpret the impact the disease has on the patient. Whilst this finding is outside the scope of this dissertation comment needs to be made on the relevance of this result. There is paucity in the literature specifically looking at the correlation between differing subjective assessments of disease impact. Schmitt and Fabio, (2005) examined the validity of prospective and retrospective global change criterion measures and assessed the agreement of the patient's current global function between clinical therapist and patients at baseline. The study was undertaken on patients with upper limb injuries and compared with functional outcome measures for the upper limb. The results documented a moderate agreement between therapist and patient global disability rating with an ICC of 0.64 for patients with proximal pathologies and 0.53 for patients with distal pathologies. In the current study the agreement between therapist and patient was considerably higher. However there were several methodological differences which made comparison difficult, such as the use of a seven point global rating scale instead of a ten point anchor plus therapists completed a global rating scale at baseline and at three months, not just at six weeks or at discharge as in this study.

5.4 Mean Scores

As this study looked at the QuickDASH in its independent format, a comparison of the mean scores to those computed from the full DASH is indicated. The clinical relevance of understanding mean scores is evident in questionnaires where standardised scoring has been established. The DASH has been able to establish some baseline data to indicate a standardised score, for example an average 55 year old male in the general population will have a mean DASH score of 12 and a mean preoperative score for those patients with rotator cuff tears awaiting surgery is around 49 (MacDermid & Stratford, 2004; Skutek et al., 2000). Beaton et al., (2001) reported a mean DASH score of 50.7 for persons unable to work because of their upper extremity problem. The clinical implication of this is that at this early stage of psychometric analysis on the QuickDASH, the clinician needs to exercise caution when interpreting a score as a indicator of disability. More research needs to be conducted involving a variety of pathologies and intervention to determine predicted and normative mean score data.

The mean QuickDASH baseline score was 37.7. This was similar to Gummesson et al., (2006) where a baseline QuickDASH score of 39 was recorded in patients waiting for

elective upper extremity surgery. Both these scores however, are slightly higher than the full DASH score of 35, calculated by Gummesson et al., (2006). In contrast there has been a large range reported by two other studies. Imaeda et al., (2006) found a baseline QuickDASH score of 28, also calculated from the full DASH and Beaton et al (2005) when developing the QuickDASH had a range of 37.9 to 45 in the three item reducing techniques used. It may be implied from these results that the QuickDASH baseline score is variable depending on the study population being investigated and may represent potentially better precision in detecting differing degrees of disability.
The QuickDASH 2 score incurred a 4-point difference of 33.3. There have been two other studies using a test-retest methodology on a sub group of the study population, neither revealed the actual mean score of the QuickDASH 2, thus no comparison can be made (Gummesson et al., 2006; Imaeda et al., 2006).

Finally, QuickDASH 3 had a mean score of 18.5, a 19.2 difference from the baseline QuickDASH. Gummesson et al., (2006) revealed a mean score of 27 in the follow up QuickDASH, presented 12 months after surgery, a difference of 12 points. It cannot be determined if the difference in these post intervention mean scores are because of the presenting pathologies and intervention, the time line involved or the methodology of computing the QuickDASH from the full DASH.

5.5 Demographics

Demographic data indicates that 60% were male and 40% were female and with a combined average age of 48 years (Table 3). This may represent a manual working population (Keogh et al., 2000; MacDermid & Stratford, 2004). The most common presentation was rotator cuff pathology (48%) (Table 4). This is also in keeping with the New Zealand ACC Statistics where shoulder injuries are the fourth most common presentation and of those, rotator cuff injuries are the most common pathology (ACC, 2006). The second most common presentation was termed ‘tennis elbow’ by the physiotherapists. Tennis elbow is not a diagnosis and it has been concluded that the diagnosis is likely to be lateral common extensor tendinopathy of the elbow (Khan, Cook, Bonar, Harcourt, & Astrom, 1999). The average time from injury to participation in the study was 9.7 weeks with the minimum presentation being one week. This is of relevance to this study in that the dominant pathology presented were rotator cuff injuries, which have a known long time frame for recovery. The Diagnosis and

Management of Soft Tissue Shoulder Injuries and Related Disorders Guidelines refer to a partial tear taking up to six weeks to show signs of recovery and full thickness tears taking in excess of 12 weeks to recover (ACC, 2004).

5.6 Study Limitations

A key limitation to this study was the methodology of analysing reliability. The methodological flaw was the absence of securing a population of patients who were deemed to be stable (no change in health status had occurred) before completing QuickDASH 2. At the time of administering the second QuickDASH a concurrent patient global rating of change questionnaire would have been of benefit (Schmitt & Fabio, 2004). This prevented an accurate reliability data analysis being conducted resulting in inconclusive findings. The reliability results undertaken in this study were performed as an exercise only to see if there was some comparability with documented results from previous studies and should not be interpreted as evidence of reliability. In addition, the ability to not accurately determine reliability (in particular the ICC) placed a restriction on what responsiveness statistical analysis could be performed. The Standard Error of Measurement and ROCauc both require the calculation of an ICC from a stable population (Deyo et al., 1991; Schmitt & Fabio, 2004). The result of this is that the group effect of the QuickDASH responsiveness could only be determined as opposed to calculating the responsiveness involving tracking an individual patient through a course of treatment (Schmitt & Fabio, 2004)

A further limitation was the small sample size. Injuries to the upper extremity can be varied and range significantly in severity thus the sample size may not be representative of the range of injuries that typically occur (Giang, 2006). However as the current study is a pilot study only, a small sample size is to be expected.

The data collection for this study involved only private practice physiotherapy clinics. No data was collected from hospital physiotherapy outpatient departments. Bias of data collection involving patient selection, patient availability and injury type as a result cannot be conclusively ruled out. As this was a pilot study, data collection was specifically defined to private clinics only in the methodology design. This however, is still valuable in the primary care setting.

The time frame for collection of data is also considered a limitation of this study. Although 94% of study population improved with treatment with the criteria for completion of the study being at six weeks or discharge, whichever event occurred first, this may be considered a tight timeframe for assessing treatment benefits in a clinical setting. There are some pathologies where a change in health status may not be evident within such a limited timeframe, namely rotator cuff injuries which made up the larger proportion of recorded injuries. The second time frame identified as a limitation was the collection of the second QuickDASH. This was meant to be administered 24-48 hours after the QuickDASH 1. However many clinicians involved in the data collection see their patients on a weekly basis making administering QuickDASH 2 problematic. In some circumstances it was reported that the patient took QuickDASH 2 home and returned it on the next appointment in order to get around this problem. This creates uncertainty that the patient is the person completing the questionnaire and also when exactly it was completed.

The additional results detected involving the correlation between the patient and physiotherapist subjective rating of disease impact may have included potential crosstalk. It was intended that the patient and physiotherapist fill out these components of the study independently of each other, (instructions to do so were printed on top of the physiotherapists section). However as this data was collected in various physiotherapy clinics without direct supervision, therfore it cannot be guaranteed that this procedure occurred.

Finally, this current study is largely compared to two other similar studies where one of the aims was to validate and provide evidence of test-retest reliability and responsiveness in a translated form of the QuickDASH to Swedish (Gummesson et al., 2006) and Japanese (Imaeda et al., 2006). As this is the first study of its kind looking at the non-translated English format some findings may not be considered comparable due to the process of translation.

5.7 Area For Future Research

The key area for future research is to determine the reliability of the QuickDASH as a reliable instrument for clinical use in assessing functional outcomes in upper extremity injuries. This needs to occur with an improved methodology, specifically involving a sample of the study population who are deemed to be stable at the time of completing a second questionnaire. Further areas of research need to involve repeat studies looking at larger sample size to strengthen the findings. This should include a longer time period for data collection plus a larger number of physiotherapy clinics and hospitals ensuring that the results were more representative of a wide variety of upper extremity injuries. Furthermore prospective longitudinal studies could be used to explore the use of the QuickDASH as a prognostic indicator of upper extremity injuries. This knowledge would allow clinicians to better determine treatment management.

5.8 Conclusion

The QuickDASH was developed to reduce the burden of administration in assessing the functional outcome from upper extremity injuries (Beaton et al., 2005). Previous investigations have indicated that the QuickDASH is a reliable and responsive instrument to use in a primary health setting (Gummesson et al., 2006; Imaeda et al., 2006). Test-retest reliability of the QuickDASH was not established as a result of this study due to methodological error thus unable to support the existing findings of reliability. This current study reinforces that the QuickDASH is a responsive tool capable of measuring functional disability in subjects with upper extremity injuries. Evidence of high levels of responsiveness was established giving clinicians confidence that the QuickDASH can measure relevant patient perceived changes in health status in a primary health setting.

References

Abrams, D., Davidson, M., Harrick, J., Harcourt, P., Zyliniski, M., & Clancy, J. (2006). Monitoring the change: Current trends in outcome measure usage in physiotherapy. *Manual Therapy, 11*, 46-53.

ACC. (2003). *New Zealand Acute Low back Pain Guide*. Wellington: New Zealand Accident Compensation Commission.

ACC. (2004). *The Diagnosis and Management of Soft Tissue Shoulder Injuries and Related Disorders*. Wellington: New Zealand Accident Compensation Commission.

ACC. (2006). ACC Injury Statistics 2006. Retrieved 19.12.2006, 2006, from http://www.acc.co.nz

Baysal, D., Balyk, R., Otto, D., Luciak-Corea, C., & Beaupre, L. (2005). Functional outcome and health-related quality of life after surgical repair of full-thickness rotator cuff tear using a mini-open technique. *American Journal of Sports Medicine, 33*(9), 1346-1355.

Beaton, D., Katz, J., Fossel, A., Wright, J., & Tarasuk, V. (2001). Measuring the whole or the parts? Validity, reliability and responsiveness of the disabilities of the arm, shoulder and hand outcome measure in different regions of the upper extremity. *Journal of Hand Therapy, 14*, 128-146.

Beaton, D., Wright, J., & Katz, J. (2005). Development of the QuickDASH: Comparison of three item-reduction approaches. *Journal of Bone and Joint Surgery (British), 87-A*(5), 1038-1046.

Belvins, F., Djurasovic, M., Flatlow, E., & Vogel, K. (1997). Biology of the rotator cuff. *Orthopedic Clinics of North America, 28*(1), 1-16.

Bergner, M., Bobbit, R., Carter, W., & Gilson, B. (1981). The sickness impact profile: Development and final revision of a health status measurement. *Medical Care, 19*, 787-905.

Bland, J., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet, 1*, 307-310.

Blinkley, J., Stratford, P., Lott, S., Riddle, D., & Research, N. A. O. R. (1999). The lower extremity functional scale (LEFS): Scale development, measurement properties and clinical application. *Physical Therapy, 79*(4), 371-383.

Bot, S., Terwee, C., van der Windt, D., Bouter, L., Dekker, J., & de Vet, H. (2004). Clinimetric evaluation of shoulder disability questionnaires: a systematic review. *Annals of Rheumatoid Diseases, 63*, 335-341.

Chung, K. (1998). Reliability and validity testing of the Michigan Hand Outcomes Questionnaire. *Journal of Hand Surgery (Am), 23*, 575-587.

Cieza, A., & Stucki, G. (2005). Content comparison of health-related quality of life (HRQOL) instruments based on the international classification of functioning, disability and health (ICF). *Quality of Life Research, 14*, 1225-1237.

Davidson, M. (2004). Guideline to selecting and interpreting standardised outcome measures: Orthopaedic conditions. Retrieved 13.04.2006, 2006, from http://www.tac.vic.govt/au/jsp/corporate/sitemap/physicaltherapyresources

Deyo, R., Diehr, P., & Patrick, D. (1991). Statistics and strategies for evaluation. *Controlled Clinical Trials, 12*, 142S- 158S.

Driver, D. (2006). Occupational and physical therapy for work-related upper extremity disorders: How can we influence outcomes. *Clinics in Occupational and Environmental Medicine, 5*(2), 471-482.

Giang, G. (2006). Epidemiology of work-related upper extremity disorders: understanding prevalence and outcomes to impact provider performances using a practice management reporting tool. *Clinics in Occupational & Environmental Medicine, 5*(2), 267-283.

Green, S., Buchbinder, R., & Hetrick, S. (2005). Physiotherapy interventions for shoulder pain. *Cochrane Database of Systematic Reviews, 2*.

Gummesson, C., Atroshi, I., & Ekdahl, C. (2003). The disabilities of the arm, shoulder and hand (DASH) outcome questionnaire: longitudinal construct validity and measuring self-rated health change after surgery. *Biomedcentral Musculoskeletal Disorders, 4*(11).

Gummesson, C., Ward, M., & Atroshi, I. (2006). The shortened disabilities of the arm, shoulder and hand questionnaire (QuickDASH): validity and reliability based on responses within the full-length DASH. *Biomedcentral Musculoskeletal Disorders, 7*(44).

Guyatt, G., Walter, S., & Norman, G. (1987). Measuring change over time: assessing the usefulness of evaluative instruments. *Journal of Chronic Disease, 40*, 171-178.

Hicks, C. (1999). *Research Methods For Clinical Therapists* (3rd ed.). Edinburgh: Churchill Livingston.

Higginson, U., & Carr, A. (2001). Measuring quality of life: Using quality of life measures in the clinical setting. *British Medical Journal, 322*, 1297-1300.

Holtby, R., & Razmjou, H. (2005). Measurement properties of the Western Ontario rotator cuff outcome measure: A preliminary report. *Journal of Shoulder and Elbow Surgery, 14*, 506-510.

Hopkins, W. (2000). Measures of reliability in sports medicine and science. *Sports Medicine, 30*(1), 1-15.

Horner, D., & Larmer, P. (2006). Health outcome measures. *New Zealand Journal of Physiotherapy, 34*(1), 17-24.

Hoving, J., O'Leary, E., Niere, K., Green, S., & Buchbinder, R. (2003). Validity of the neck disability index, Northwick Park neck pain questionnaire, and problem elicitation technique for measuring disability associated with whiplash-associated disorders. *Pain, 102*, 273-281.

Hudak, P., Amadio, P., Bombardier, C., & Group, U. E. C. (1996). Development of an upper extremity outcome measure: The DASH (Disabilities of the arm, shoulder and hand). *American Journal of Industrial Medicine, 29*, 602-608.

Hunt, S., McEwan, J., & McKenna, S. (1985). Measuring health status: a new tool for clinicians and epidemiologists. *Journal of the Royal College of General Practitioners, 35*(273), 185-188.

Hunter, G. (1994). Specific soft tissue mobilisation in the treatment of soft tissue lesions. *Physiotherapy, 80*(1), 15-21.

Imaeda, T., Toh, S., Nakao, Y., Nishida, J., Hirata, H., Ijichi, M., et al. (2006). Validation of the Japanese society for surgery of the hand version of the Disability of the Arm, Shoulder and Hand questionnaire. *Journal of Orthopaedic Science, 10*, 353-359.

Jester, A., Harth, A., & Germann, G. (2005). Measuring levels of upper-extremity disability in employed adults using the DASH questionnaire. *Journal of Hand Surgery, 30A*(5), 1074-1082.

Jester, A., Harth, A., Wind, G., Germann, G., & Sauerbier, M. (2005). Disabilities of the arm, shoulder and hand (DASH) questionnaire: Determining functional activity profiles in patients with upper extremity disorders. *Journal of Hand Surgery (Br), 30B*(1), 23-28.

Jette, A. (1993). Using health-related quality of life measures in physical therapy outcomes research. *Physical Therapy, 73*(8), 528-538.

Jost, B., Zumstein, M., Pfirrmann, C., Zanetti, M., & Gerber, C. (2005). MRI findings in throwing shoulders: abnormalities in professional handball players. *Clinical Orthopeadics and Related Research, 434*, 130-137.

Kazis, L., Anderson, J., & Meenanm, R. (1989). Effect sizes for interpreting changes in health care. *Mediacl Care, 27*(3), s178-s189.

Keogh, J., Nuwayhid, I., Gordon, J., & Gucer, P. (2000). The impact of occupational injury on injured worker and family: outcomes of upper extremity cumulative trauma disorders in Maryland workers. *American Journal of Industrial Medicine, Nov 38*, 498-506.

Khan, K., Cook, J., Bonar, F., Harcourt, P., & Astrom, M. (1999). Histopathology of common tendinopathies. *Sports Medicine, 27*(6), 393-408.

Kirkley, A., & Griffin, S. (2003). Development of disease-specific quality of life measurement tools. *Journal of Arthroscopic and Related Surgery, 19*(10), 1121-1128.

Kirkley, A., Griffin, S., Mclintock, H., & Ng, L. (1998). The development and evaluation of a disease-specific quality of life measurement tool for shoulder instability. *American Journal of Sports Medicine, 26*(6), 764-772.

Kotsis, S., Chung, K., & Arbor, A. (2005). Responsiveness of the Michigan Hand Outcomes Questionnaire and the Disabilities of the Arm, Shoulder and Hand Questionnaire in carpal tunnel surgery. *Journal of Hand Surgery, 30A*(1), 81-86.

Laing, M., Lew, R., Stucki, G., Fortin, P., & Daltroy, L. (2002). Measuring clinically important chnages with patient-orientated questionnaires. *Medical Care, 40*(4), 45-51.

MacDermid, J., & Stratford, P. (2004). Applying evidence on outcome measures to hand therapy practice. *Journal of Hand Therapy, 17*(2), 165-173.

MacDermid, J., & Tottenham, V. (2004). Responsiveness of the Disability of the Arm, Shoulder and Hand (DASH) and patient related Wrist/Hand Evaluation (PRWHE) in evaluating change after hand therapy. *Journal of Hand Therapy, 17*, 18-23.

Michlovitz, S., Harris, B., & Watkins, M. (2004). Therapy interventions for improving joint range of motion: a systematic review. *Journal of Hand Therapy, 17*(2), 118-131.

Paul, A., Lewis, M., Shadforth, M., Croft, P., van der Windt, D., & Hay, E. (2004). A comparison of four shoulder-specific questionnaires in primary care. *Annals of Rheumatoid Diseases, 63*, 1293-1299.

Rankin, G., & Stokes, M. (1998). Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. *Clinical Rehabilitation, 12*, 187-199.

Roach, K., Budiman-Mak, E., Songsiridej, N., & Lertratanakul, Y. (1991). Development of a shoulder pain and disability index. *Arthritis Care & Research, 4*(4), 143-149.

Robinson, C., & Page, R. (2003). Severely impacted valgus proximal humeral fractures. Results of operative treatment. *Journal of Bone and Joint Surgery (AM), 85*, 1647-1655.

Schmitt, J., & Fabio, R. (2004). Reliable change and minimum important difference (MID) proportions facilitated group responsiveness comparisons using individual threshold criteria. *Journal of Clinical Epidemiology, 57*, 1008-1018.

Skutek, M., Fremerey, R., Zeichen, J., & Bosch, U. (2000). Outcome analysis following open rotator cuff repair. Early effectiveness validated using four different outcome assessment scales. *Archives Orthopaedic Trauma Surgery, 120*, 432-436.

Smet, L. (2004). Responiveness of the DASH score in surgically treated basal joint arthritis of the thumb: preliminary results. *Clinical Rheumatology, 23*, 223-224.

SooHoo, N., McDonald, A., Seiler, J., & McGillivary, G. (2002). Evaluation of the construct validity of the DASH questionnaire by correlation to the SF-36. *Journal of Hand Surgery (Am), 27A,* 537-541.

Stover, B. (2004). *Validation and evaluation of a disability measure for upper extremity musculoskeletal disorder screening in the workplace and prognostic factors of long term disability.* University of Washington, Washington.

Stratford, P., Binkley, J., & Stratford, D. (2001). Development and initial validation of upper extremity functional index. *Physiotherapy Canada, 53*(4), 259-267.

Terwee, C., Dekker, F., Wiersinga, W., Priummel, M., & Bossuyt, P. (2003). On assessing responsiveness of health-related quality of life instruments: Guidelines for instrument evaluation. *Quality of Life Research, 12*, 349-362.

Vicenzino, B., Collins, D., Benson, H., & Wright, A. (1998). An investigation of the interrelationship between manipulative therapy-induced hypoalgesia and sympathoexcitation. *Journal of Manipulative and Physiological Therapeutics, 21*(7), 448-453.

Ware, J., & Sherbourne, C. (1992). The MOS 36 item short form health survey (SF-36). A conceptual framework and item selection. *Medical Care, 30*, 473-483.

World Health Organisation. (2001). International Classification of Functioning, Disability and Health. Retrieved April 30, 2006, from http:/www3.who.int/icf/icftemplate.cfm

Time 1
Participant ID No

THE QuickDASH OUTCOME MEASURE

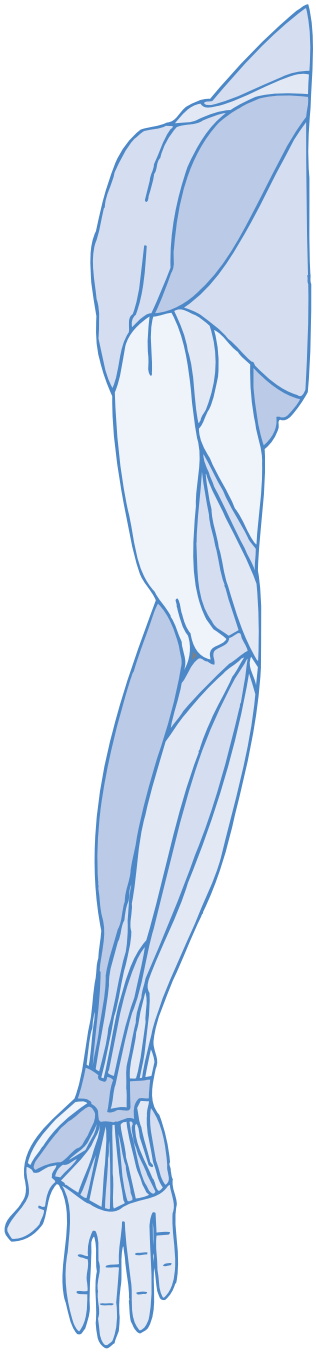
INSTRUCTIONS

This questionnaire asks about your symptoms as well as your ability to perform certain activities.

Please answer *every question*, based on your condition in the last week, by circling the appropriate number.

If you did not have the opportunity to perform an activity in the past week, please make your *best estimate* of which response would be the most accurate.

It doesn't matter which hand or arm you use to perform the activity; please answer based on your ability regardless of how you perform the task.



*Quick*DASH

Please rate your ability to do the following activities in the last week by circling the number below the appropriate response.

	NO DIFFICULTY	MILD DIFFICULTY	MODERATE DIFFICULTY	SEVERE DIFFICULTY	UNABLE
1. Open a tight or new jar.	1	2	3	4	5
2. Do heavy household chores (e.g., wash walls, floors).	1	2	3	4	5
3. Carry a shopping bag or briefcase.	1	2	3	4	5
4. Wash your back.	1	2	3	4	5
5. Use a knife to cut food.	1	2	3	4	5
6. Recreational activities in which you take some force or impact through your arm, shoulder or hand (e.g., golf, hammering, tennis, etc.).	1	2	3	4	5

	NOT AT ALL	SLIGHTLY	MODERATELY	QUITE A BIT	EXTREMELY
7. During the past week, *to what extent* has your arm, shoulder or hand problem interfered with your normal social activities with family, friends, neighbours or groups?	1	2	3	4	5

	NOT LIMITED AT ALL	SLIGHTLY LIMITED	MODERATELY LIMITED	VERY LIMITED	UNABLE
8. During the past week, were you limited in your work or other regular daily activities as a result of your arm, shoulder or hand problem?	1	2	3	4	5

Please rate the severity of the following symptoms in the last week. *(circle number)*	NONE	MILD	MODERATE	SEVERE	EXTREME
9. Arm, shoulder or hand pain.	1	2	3	4	5
10. Tingling (pins and needles) in your arm, shoulder or hand.	1	2	3	4	5

	NO DIFFICULTY	MILD DIFFICULTY	MODERATE DIFFICULTY	SEVERE DIFFICULTY	SO MUCH DIFFICULTY THAT I CAN'T SLEEP
11. During the past week, how much difficulty have you had sleeping because of the pain in your arm, shoulder or hand? *(circle number)*	1	2	3	4	5

*Quick*DASH DISABILITY/SYMPTOM SCORE = $\left(\left[\frac{\text{(sum of n responses)}}{n}\right] - 1\right) \times 25$, where n is equal to the number of completed responses.

A *Quick*DASH score may <u>not</u> be calculated if there is greater than 1 missing item.

WORK MODULE (OPTIONAL)

The following questions ask about the impact of your arm, shoulder or hand problem on your ability to work (including homemaking if that is your main work role).

Please indicate what your job/work is:_____

❒ I do not work. (You may skip this section.)

Please circle the number that best describes your physical ability in the past week.

Did you have any difficulty:	NO DIFFICULTY	MILD DIFFICULTY	MODERATE DIFFICULTY	SEVERE DIFFICULTY	UNABLE
1. using your usual technique for your work?	1	2	3	4	5
2. doing your usual work because of arm, shoulder or hand pain?	1	2	3	4	5
3. doing your work as well as you would like?	1	2	3	4	5
4. spending your usual amount of time doing your work?	1	2	3	4	5

SPORTS/PERFORMING ARTS MODULE (OPTIONAL)

The following questions relate to the impact of your arm, shoulder or hand problem on playing *your musical instrument or sport or both*. If you play more than one sport or instrument (or play both), please answer with respect to that activity which is most important to you.

Please indicate the sport or instrument which is most important to you:_____

❒ I do not play a sport or an instrument. (You may skip this section.)

Please circle the number that best describes your physical ability in the past week.

Did you have any difficulty:	NO DIFFICULTY	MILD DIFFICULTY	MODERATE DIFFICULTY	SEVERE DIFFICULTY	UNABLE
1. using your usual technique for playing your instrument or sport?	1	2	3	4	5
2. playing your musical instrument or sport because of arm, shoulder or hand pain?	1	2	3	4	5
3. playing your musical instrument or sport as well as you would like?	1	2	3	4	5
4. spending your usual amount of time practising or playing your instrument or sport?	1	2	3	4	5

SCORING THE OPTIONAL MODULES: Add up assigned values for each response; divide by 4 (number of items); subtract 1; multiply by 25.
An optional module score may <u>not</u> be calculated if there are any missing items.

A p p e n d i x B

M E M O R A N D U M

To: Duncan Reid
From: **Madeline Banda** Executive Secretary, AUTEC
Date: 12 October 2006
Subject: Ethics Application Number 06/173 **Test-retest reliability and responsiveness of the shortened DASH (QuickDASH) questionnaire: a pilot study.**

Dear Duncan

Thank you for providing written evidence as requested. I am pleased to advise that it satisfies the points raised by the Auckland University of Technology Ethics Committee (AUTEC) at their meeting on 11 September 2006 and as the Executive Secretary of AUTEC I have approved your ethics application. This delegated approval is made in accordance with section 5.3.2.3 of AUTEC's *Applying for Ethics Approval: Guidelines and Procedures* and is subject to endorsement at AUTEC's meeting on 13 November 2006.

Your ethics application is approved for a period of three years until 11 October 2009.

I advise that as part of the ethics approval process, you are required to submit to AUTEC the following:

- A brief annual progress report indicating compliance with the ethical approval given using form EA2, which is available online through *http://www.aut.ac.nz/research/ethics*, including when necessary a request for extension of the approval one month prior to its expiry on 11 October 2009;
- A brief report on the status of the project using form EA3, which is available online through *http://www.aut.ac.nz/research/ethics*. This report is to be submitted either when the approval expires on 11 October 2009 or on completion of the project, whichever comes sooner;

It is also a condition of approval that AUTEC is notified of any adverse events or if the research does not commence and that AUTEC approval is sought for any alteration to the research, including any alteration of or addition to the participant documents involved.

You are reminded that, as applicant, you are responsible for ensuring that any research undertaken under this approval is carried out within the parameters approved for your application. Any change to the research outside the parameters of this approval must be submitted to AUTEC for approval before that change is implemented.

Please note that AUTEC grants ethical approval only. If you require management approval from an institution or organisation for your research, then you will need to make the arrangements necessary to obtain this.

To enable us to provide you with efficient service, we ask that you use the application number and study title in all written and verbal correspondence with us. Should you have any further enquiries regarding this matter, you are welcome to contact Charles Grinter, Ethics Coordinator, by email at charles.grinter@aut.ac.nz or by telephone on 921 9999 at extension 8860.

On behalf of the Committee and myself, I wish you success with your research and look forward to reading about it in your reports.

Yours sincerely

Madeline Banda
Executive Secretary
Auckland University of Technology Ethics Committee
Cc: Kathryn Anne Polson jkpolson@xnet.co.nz

Appendix C

Consent Form

For release of Patient Information

Project title: Test- retest reliability and responsiveness of the shortened DASH (QuickDASH) questionnaire; a pilot study.

Project Supervisor: Duncan Reid

Researcher: Kate Polson

- ○ I have read and understood the information provided about this research project.
- ○ I have had an opportunity to ask questions and to have them answered.
- ○ I understand that I may withdraw from the research or have any information that I have provided for this project at any time prior to completion of data collection withdrawn, without being disadvantaged in any way.
- ○ I agree to take part in this research.
- ○ I wish to receive a copy of the report from the research (please tick one): Yes○ No○

Principal's signature:

.....……
……

Principal's name:

.....……
……

Practice Contact Details

……..
……..
……..
……..

Date:

Approved by the Auckland University of Technology Ethics Committee on 12/10/06
AUTEC Reference number 06/173.

Appendix D

Participant Information Sheet

Date Information Sheet Produced:

08 August 2006

Project Title

Test-retest reliability and responsiveness of the shortened DASH (QuickDASH) questionnaire; A pilot study.

An Invitation

As a patient of …… (name of clinic) Physiotherapy Clinic you have been invited to participate in a research that is evaluating the use of a questionnaire for upper limb injuries. The research is being conducted by Kate Polson (registered physiotherapist) and is a dissertation which is a part requirement for a Masters of Health Science (Physiotherapy). The research involves the completion of the Quick Disability Arm Shoulder Hand (QuickDASH) questionnaire plus another disability questionnaire. Your participation in the research is entirely voluntary and you are welcome to withdraw at any stage without penalty or consequence to your physiotherapy management.

What is the purpose of this research?

The purpose of this research is to investigate if the QuickDASH questionnaire can measure functional outcomes from the patient's perspective, in a New Zealand population. Physiotherapists can use this information, if it is valid and reliable, to assist in designing physiotherapy treatment plans. The research completes the final part of a Masters in Health Science degree. On completion of this research it will be published in a physiotherapy related journal.

How was I chosen for this invitation?

You have been selected for this research by your physiotherapist if you suffer from an upper limb injury or disease and are aged 20 years or over. As it is a fixed format questionnaire written in English, you must also have English language skills.

What will happen in this research?

You will complete three QuickDASH questionnaires and one global disability questionnaire:

- The first one at your initial treatment
- The second one 24 hours later or at the next appointment

- The third one at six weeks or at discharge, whichever event occurs first. At this time you will also be asked to complete a second questionnaire assessing your overall perceived recovery.

Your physiotherapist will complete a demographic sheet at the beginning and also complete an overall disability questionnaire at the completion.
It is important to note that your physiotherapy treatment will not be affected nor influenced by this research.

What are the discomforts and risks?

The main risk is the disclosing of personal information by completing the questionnaire. This information refers to your ability to perform every day upper limb tasks and questions about your pain. A second barrier to this research is the time needed to complete the questionnaire which should take on average 5 minutes.

How will these discomforts and risks be alleviated?

You will not be identified by the researcher; the questionnaires are identified by a number, therefore ensuring that their identity remains confidential to the physiotherapy clinic. The QuickDASH is designed to reduce the time taken to complete questionnaires in order to make collecting information like this more user friendly to both the patient and therapist.

What are the benefits?

It is important when designing a physiotherapy treatment plan that the outcomes, from a patient's perspective, are achieved and are measurable. One method for a therapist to gain insight into where the patient experiences their dysfunction is by using questionnaires. However, some questionnaires like the QuickDASH have problems associated with their use, unless they are well tested on the people you intend to use them on. The benefit of this research is that the QuickDASH questionnaire may provide physiotherapists with a tool for measuring functional outcomes in their patients. The benefits to the patients therefore are that a quick and easy questionnaire could assist the physiotherapist in developing their best treatment plan to make sure their functional goals are achieved.

How will my privacy be protected?

Once you have consented to the project, you will be entered into the research data base and referred to by number only. In that way whenever anything is published and reported from the research you will remain anonymous. The data concerning the project is stored in locked cabinets at AUT or password protected computers. Only the principal researcher and the supervisor have access to this information

What are the costs of participating in this research?

Time is the only cost to your participation. It is estimated that the questionnaire takes on average 3-5 minutes to complete.

What opportunity do I have to consider this invitation?

Once your physiotherapist has considered you may be appropriate for the study you will have chance to read this information sheet. If you have any questions they will be addressed at that time.

How do I agree to participate in this research?

If you agree to participate you will need to sign a consent form on the day of the first treatment you receive from the physiotherapist

Will I receive feedback on the results of this research?

You will be able to receive a written summary of the findings of the study on request to the principal researcher

What do I do if I have concerns about this research?

Any concerns regarding the nature of this project should be notified in the first instance to the Project Supervisor:

Duncan Reid, duncan.reid@aut.ac.nz , PH 09 921 9999 ext 7806

Concerns regarding the conduct of the research should be notified to the Executive Secretary, AUTEC, Madeline Banda, *madeline.banda@aut.ac.nz* , 921 9999 ext 8044.

Whom do I contact for further information about this research?

Researcher Contact Details:
Kate Polson
P.O. Box 227
Te Kuiti
Project Supervisor Contact Details:
Duncan Reid
Head of Division
Division of Rehabilitation and Occupational Studies
Auckland University of Technology
Ph 09 921 9999 ext 7806
duncan.reid@aut.ac.nz

Approved by the Auckland University of Technology Ethics Committee on *12/10/06*, AUTEC Reference number *06/173.*

Appendix E

Patient Consent Form

Project title: Test- retest reliability and responsiveness of the shortened DASH (QuickDASH) questionnaire; a pilot study.

Project Supervisor: Duncan Reid

Researcher: Kate Polson

- ○ I have read and understood the information provided about this research project.
- ○ I have had an opportunity to ask questions and to have them answered.
- ○ I understand that I may withdraw from the research or have any information that I have provided for this project at any time prior to completion of data collection withdrawn, without being disadvantaged in any way.
- ○ I agree to take part in this research.
- ○ I wish to receive a copy of the report from the research (please tick one): Yes○ No○

Patients signature:

.....……
……

Patients name:

.....……
……

Practice name :

……..
……..

Date:

Approved by the Auckland University of Technology Ethics Committee on 12/10/06
AUTEC Reference number 06/173.

Appendix F

AUCKLAND UNIVERSITY OF TECHNOLOGY
TE WĀNANGA ARONUI O TAMAKI MAKAU RAU

Patient Discharge

Global Assessment Questionnaire

Physiotherapist to complete

Clinic ID:_____ Patient Number:_____ Date:_____

Patient to complete

Rate the **overall condition** of your upper limb/arm at the present time

Please circle one number below.

1	2	3	4	5	6	7	8	9	10

Scale	Description
2	I have significant limitations that affect activities of daily living
4	I have moderate limitations that affect activities of daily living, e.g. no sports possible
6	I have some limitations e.g. with sports or housework, but I can participate; I compensate
10	I am able to do whatever I wish with no problems

Overall Status:

Since I started treatment my overall status is:

Please circle one number below.

0	1	2	3	4	5	6

Scale	Description
0	Very much improved
1	Much improved
2	Minimally improved
3	No change
4	Minimally worse
5	Much worse
6	Very much worse

Appendix G

AUCKLAND UNIVERSITY OF TECHNOLOGY
TE WĀNANGA ARONUI O TAMAKI MAKAU RAU

Physiotherapist <u>Discharge</u>

Global Assessment Questionnaire

Clinic ID:_____ Patient Number:_____

Date:_____

When answering this question please reflect on your experience of treating upper limb pathologies and from the current patient's history.

<u>Do not ask the patient this question directly</u>

<u>Patients severity</u>

When comparing this upper limb pathology to other similar pathologies you have seen, how do you rate the overall condition of this pathology at the present time?

Please circle one number below.

1	2	3	4	5	6	7	8	9	10

Scale	Description
2	They have significant limitations that affect activities of daily living.
4	They have moderate limitations that affect activities of daily living, e.g. no sports possible.
6	They have some limitations e.g. with sports over housework, but they can participate; they compensate.
10	They are able to do whatever they wish with no problems.

Appendix H

Patient Demographic Form

Date:

Clinic ID:

Patient Number:

Date of Injury

Date of first treatment

Patient Age: **Gender: M F** **Ethnicity:**

Occupation :

Diagnosis:

-
-
-

Arm Affected: L R **Arm Dominance: L R**