Published in final edited form as:

Arch Phys Med Rehabil. 2014 January ; 95(1): 153–162. doi:10.1016/j.apmr.2013.07.022.

Patient-Reported Shoulder Outcome Measures Utilized in Breast Cancer Survivors: A Systematic Review

Shana Harrington, PT, PhD, SCS, MTC^{1,2}, Lori A Michener, PT, PhD, ATC³, Tiffany Kendig, PT, DPT, MPH⁴, Susan Miale, PT, DPT, PCS⁵, and Steven Z George, PT, PhD⁶

¹Department of Clinical and Applied Movement Sciences, University of North Florida, Jacksonville, Florida

²K12 Scholar, Department of Physical Therapy, University of Florida, Gainesville, FL

³Department of Physical Therapy, Virginia Commonwealth University – MCV Campus, Richmond, Virginia

⁴Rehabilitation Medicine Service, Department of Neurology, Memorial Sloan-Kettering Cancer Center, New York, New York

⁵Department of Physical Therapy, School of Health Technology and Management, Stony Brook University, Stony Brook, New York

⁶Department of Physical Therapy, University of Florida, Gainesville, Florida

Abstract

Objective—1) To identify English Language published patient-reported upper extremity outcome measures used in breast cancer research and 2) To examine construct validity and responsiveness in patient-reported upper extremity outcome measures used in breast cancer research.

Data Sources—PubMed, CINAHL and ProQuest MEDLINE® databases were searched up to February 5, 2013.

Study Selection—Studies were included if a patient-reported upper extremity outcome measure was administered, the participants were diagnosed with breast cancer, and published in English.

Data Extraction—Eight hundred and sixty-five articles were screened. Fifty-nine full text articles were assessed for eligibility. A total of 46 articles met the initial eligibility criteria for aim 1. Eleven of these articles reported mean and standard deviations for the outcome scores, and included a comparison group analysis for aim 2.

Data Synthesis—Construct validity was evaluated by calculating effect sizes for known group differences in 6 studies using the Disabilities of Arm, Shoulder and Hand (DASH), Penn Shoulder Score, Shoulder Disability Questionnaire-Dutch, and 10 Questions by Wingate (Wingate). Responsiveness was analyzed comparing a treatment and control group by calculating the coefficient of responsiveness in 5 studies for the DASH and Wingate.

Conclusions—Eight different patient-reported upper extremity outcome measures have been reported in the peer-review literature for women with breast cancer, some (n=3) were specifically developed for breast cancer survivors and others that were not (n=5). Based on the current evidence we recommend administering the DASH to assess patient-reported upper extremity function in breast cancer survivors because the DASH had most consistently large effects sizes for construct validity and responsiveness. Future large studies are needed for more definitive recommendations.

Keywords

Breast neoplasms; Upper extremity; Outcome assessment; Rehabilitation

Breast cancer is the most common cancer affecting women in the United States, excluding cancers of the skin.¹ It is estimated that in 2013, approximately 232,000 new cases of breast cancer will be diagnosed in American women.¹ Currently there are approximately 2.9 million American women who are surviving with breast cancer, commonly referred to as breast cancer survivors.² While survival rates continue to improve the focus on survivorship issues and quality of life has gained more attention.³ One survivorship issue well documented in the literature is upper extremity function in breast cancer survivor quality of life.^{4–15}

The ability to use the upper extremity during daily activities in order to participate in desired family, social and occupational roles^{16,17} can be quantified with patient-reported outcome measures.¹⁸ Third-party payers increasingly require health care providers to measure clinical outcomes, and patient-reported outcome measures are one way to do so. Patient-report outcome measures can be used to assess the difficulty of a variety of upper extremity activities during participation in daily roles. Rehabilitation professionals can use the information from patient-reported outcomes to assist clinical decision-making and to assess the impact of the effects of rehabilitation treatment on what is ultimately most important to the patient – the daily use of their upper extremity.^{18–20}

Understanding the usefulness of patient-reported outcome measures in terms of construct validity and responsiveness is essential to help guide clinical decision-making.²¹ The most commonly used patient-reported outcome measures fall into one of two broad categories, disease/condition specific or region specific.²¹ Currently, there are over 30 patient-reported disease or region-specific upper extremity measures with documented psychometric properties.²¹ Some of the most commonly used upper extremity patient-reported outcome measures include the American Shoulder and Elbow Surgeons (ASES), Disabilities of Arm, Shoulder and Hand (DASH), Shoulder Disability Questionnaire (SDQ), and Shoulder Pain and Disability Index (SPADI). Many of these questionnaires were created so the outcome score could be interpreted for a variety of individuals regardless of the upper extremity pathology or diagnosis.²¹ The development and validation of these measures occurred in individuals with diagnoses such as humeral fractures, glenohumeral arthritis, rotator cuff tendinitis, impingement syndrome, status-post shoulder surgery, glenohumeral instability/dislocation, rotator cuff syndrome, adhesive capsulitis, hemiarthroplasty, shoulder weakness, and soft tissue disorders.^{22–24} Because many of the current patient-reported outcome

measures were originally designed and measurement properties characterized for use predominantly in general musculoskeletal disorders of the shoulder or upper limb, it is unclear if the psychometric properties of these outcome tools have been adequately studied for use in the breast cancer population.

The authors are aware of three published breast cancer specific upper extremity patient-reported outcome measures, Kwan's Arm Problem Scale (KAPS), Upper Limb Disability Questionnaire (ULDQ) and Wingate, although these all lack reported psychometric properties. Additionally, we are not aware of any studies that have identified which outcome measures have been administered to the breast cancer population and systematically assessed the construct validity and responsiveness of these measures when administered to women with breast cancer. Therefore, the purpose of this systematic review was to: 1) identify English Language published patient-reported upper extremity outcome measures used in breast cancer research and to 2) examine the construct validity and responsiveness of these patient-reported measures. This information will provide an improved understanding of which patient-reported outcome measures have been administered to breast cancer survivors and reveal select psychometric properties specific to the breast cancer population. Rehabilitation professionals can use this information to guide their decisions when determining which measure is most appropriate to administer when examining upper extremity function in breast cancer survivors, in planning future psychometric studies, and for calculating sample sizes based on data reported for existing measures.

Methods

The Evaluation Database to Guide Effectiveness (EDGE) task force was developed by the Research Section of the American Physical Therapy Association. The goal of the EDGE is to establish a framework to facilitate the evaluation of outcome measures. The Oncology Section of the American Physical Therapy Association developed a taskforce, whose goals are to facilitate identification of valid and reliable tests and measures that reflect clinically important outcomes for standard use across selected patient groups in the area of oncology rehabilitation. Breast cancer was chosen as the initial target practice area and this systematic review was performed in conjunction with the goals of this taskforce. Inclusion criteria for article retrieval for both aims in this review included women diagnosed with breast cancer, methodology indicated the use of a patient-reported upper extremity outcome measure, and published in English. Studies were excluded for both aims if the patient-reported upper extremity outcome measure was designed for a specific patient condition (i.e. rheumatoid arthritis, pulmonary dysfunction), the measure was not standardized, if the measure contained clinical-rated or performance based measures as part of the total score of the measure (i.e. Constant Shoulder Scale, 6 minute walk test), the outcome measure was modified from the original format, or if the outcome measure was dichotomous yes or no (i.e. Simple Shoulder test). Additionally, for Aim 2, studies were excluded if there was no comparison group reported and/or if there was no reported outcome total score mean or standard deviation.

Data Sources and Searches

A primary systematic search using PubMed was performed from its inception up to February 5, 2013 and resulted in the retrieval of 716 publications. The search strategy began with the filters (((“upper extremity” [MeSH Terms] OR shoulder)) AND ((disability) or function)) AND (breast neoplasm [MeSH Terms]). A second systematic search strategy using CINAHL was performed up to February 5, 2013 and yielded 24 publications using the following search terms MM “Breast Neoplasms” and MM ((shoulder or axilla or upper extremity) AND ((function or disability))). A third systematic search strategy using ProQuest MEDLINE® up to February 5, 2013 and yielded 115 publications using the following search terms (breast neoplasms) AND (upper extremity OR shoulder) AND (disability OR function).

Study Selection

Three of the authors examined reference lists from all of the selected publications to verify that no pertinent publications were missed during the above-described electronic searches. Retrieved abstracts were reviewed for possible inclusion. When warranted, full articles were obtained for review. Many ($n = 684$) of the publications did not meet the following inclusion criteria: women diagnosed with breast cancer, methodology indicated the use of a patient-reported upper extremity outcome measure, and published in English. After abstracts were reviewed and duplicates were identified, 59 full text articles remained to assess for eligibility.

Data Synthesis

One method of construct validation is the analysis of “known groups” differences.²⁵ In this method, participants are grouped on the basis of an attribute – for example, lymphedema or no lymphedema. Using this method, we calculated effect sizes for the differences between two groups using the means and standard deviations when reported for the patient-reported measures. Comparisons were made within each study that defined groups at a single point in time for between group effect sizes to be calculated. The comparison groups for construct validity were: 1) lymphedema versus no lymphedema, 2) diagnosis of breast cancer on a dominant extremity versus a diagnosis of breast cancer on a non-dominant extremity, 3) breast cancer diagnosis versus a healthy (no breast cancer diagnosis) control, 4) a normal or abnormal ultrasound versus pain or no pain in the affected shoulder, and 5) axillary lymph node dissection (ALND) versus sentinel lymph node dissection (SLND). These groups were determined after the literature search was conducted. Construct validity effect sizes were calculated as Cohen’s d ; $(M_1 - M_2)/\text{pooled SD}$, where M is the mean for each group and the pooled standard deviation is the root mean square of the two standard deviations from each group.²⁶ Standardized effect sizes were calculated for individual studies and not pooled due to sample heterogeneity. Responsiveness to change was assessed for intervention studies when means and standard deviations were reported between two groups for the patient-reported outcome measure over two or more points in time. Baseline and post-intervention reported data were used to calculate effect sizes for responsiveness using the coefficient of responsiveness equation: $\text{mean}(\text{test}_1 - \text{test}_2)_{\text{treatment}} - \text{mean}(\text{test}_1 - \text{test}_2)_{\text{control}} / \text{SDtest}_{\text{1pooled}}$.²⁷ Confidence intervals (CI) for construct validity and responsiveness were

calculated as $95\% \text{ CI} = \bar{x} \pm (1.96) s_{\bar{x}}$. Effect sizes for construct validity and responsiveness were categorized as small (>0.20), moderate (>0.50), or large (>0.80).²⁸

RESULTS

Study Selection

After full text review of the 59 articles, 13 were removed due to the following exclusion criteria: 1) the outcome measure was designed for use in a specific condition not related to breast cancer (i.e. rheumatoid arthritis), 2) the outcome measure was not standardized, 3) a performance based component existed in the outcome measure, 4) modifications were made to the outcome measure, and 5) a dichotomous measure was used; leaving a total of 46 articles that met the initial eligibility criteria for aim 1 (Figure 1). Thirty five of these articles were further excluded due to 1) no reported total score, mean, and/or standard deviation or 2) no comparison group. A total of 11 studies were included in the review for aim 2.

Study Characteristics

A total of 8 different patient-reported upper extremity outcome measures were administered to the breast cancer survivor population with the DASH used in 46% of the studies (Table 1). Of these, the KAPS, ULDQ and Wingate were specifically designed to be administered to women who received breast cancer treatment. The 11 final studies assessed in this review included 886 breast cancer survivor participants. The time of assessment varied across these studies from prior to breast cancer surgery to 7.8 years after surgery.^{15,29–35} Although, 2 of the 11 studies^{32,36} did not report the average duration since breast cancer surgery and one study³⁷ was vague stating participants “received surgical treatment for breast cancer within the past 5 years”. For studies that involved an intervention, the duration of the interventions were variable from 1 month, 2 months, 3 months, and 6 months.

Synthesis of Results

Construct validity was calculated from data assessed in a total of 6 studies, including 1 study that administered two different patient-reported outcome measures.¹⁵ Of these 6 studies, 3 different patient-reported outcome measures were used: DASH, Penn, and SDQ-NL. Effect sizes for construct validity ranged from 0.27 – 1.63. (Table 2) The effect sizes for construct validity varied, dependent on the patient-reported outcome measure used and the comparison group, which included healthy individuals and breast cancer survivors, breast cancer survivors with and without lymphedema, breast cancer survivors who had no pain and a normal ultrasound of the shoulder, breast cancer survivors who had pain and an abnormal ultrasound of the shoulder and breast cancer survivors who had pain and a normal ultrasound of the shoulder. The confidence intervals were variable, however in those studies with a sample size over 100^{29,30,33} the confidence limits were narrower. Construct validity could not be examined for the KAPS, SPADI, SRQ, ULDQ and Wingate.

Responsiveness was analyzed in a total of 5 studies, comparing a physical therapy intervention to a control group or comparing a lymphedema intervention to a control group, and the effect sizes ranged from 0.20 – 1.54. (Table 3) The confidence intervals were large for all of the 5 studies and 3 studies reviewed contained a zero in the calculated confidence

interval.^{31,32,37} Responsiveness values were widely variable for the DASH (ES = 0.20 – 1.13) when a physical therapy intervention was implemented and compared to a control group. The Wingate showed large responsiveness (ES = 1.12 and 1.54) during a physical therapy intervention compared to a control group. When an intervention specific to lymphedema was compared to a control group, moderate responsiveness ES = (0.45 and 0.71) of the DASH was found. Responsiveness was unable to be evaluated in the KAPS, Penn, SDQ-NL, SPADI, SRQ and ULDQ.

DISCUSSION

This systematic review identified eight patient-reported upper extremity outcome measures that have been administered to the breast cancer population. Of these, the KAPS, ULDQ and Wingate, were specifically developed for use in women who have been diagnosed with breast cancer. Based on the findings of this study, the DASH is the most favorable patient-reported outcome measure to use for the breast cancer population, however this recommendation is made with caution given the limitations discussed below.

Construct validity values indicate the DASH and Penn distinguished between a group of breast cancer survivors and healthy controls (ES = 1.47 & 1.46) in the two studies assessed.¹⁵ The DASH was also able to distinguish between a group of breast cancer survivors who were diagnosed with lymphedema compared to a group of breast cancer survivors who did not have lymphedema in one (ES = 1.63)³³ of the two studies examined.²⁹ Group differences were large (ES = 1.57) when comparing breast cancer survivors who had no pain and a normal ultrasound of the shoulder to breast cancer survivors who had pain and an abnormal ultrasound of the shoulder.³⁶ Additionally, large effects (ES = 0.92) were found when comparing breast cancer survivors who had pain and a normal ultrasound of the shoulder to breast cancer survivors who had pain and an abnormal ultrasound of the shoulder.³⁶

Construct validity of the DASH demonstrated moderate to large effect sizes in five of the six studies assessed. The two studies that had large sample sizes (>100) and incorporated the DASH^{29,33} showed greater precision with smaller confidence intervals than those studies with smaller sample sizes.^{15,36,38} Responsiveness of the DASH also showed moderate to large effect sizes in two of the four studies analyzed,^{32,35} however these two studies had relatively low sample sizes (29 and 21) and the corresponding effects sizes had larger confidence limits reflecting a lack of precision. We were unable to examine construct validity in the KAPS, ULDQ and Wingate due to a lack of reported total outcome score means and standard deviations or a lack of a comparison group. Of the three measures specific to breast cancer, responsiveness could only be calculated for the Wingate that did demonstrate the ability to identify changes over time.

While this is the first known systematic review to identify published patient-reported upper extremity outcome measures used in breast cancer research and to examine select psychometric properties of those patient-reported upper extremity outcome measures identified, it is important to point out that there are two prior individual reports discussing reliability and aspects of validity for the DASH³⁹ and KAPS in breast cancer survivors.⁴⁰

Mulero-Portela and colleagues³⁹ performed a cross-cultural adaptation of the DASH to Spanish for Puerto Rico and found a high level of internal consistency (Cronbach's $\alpha = 0.97$), but low to moderate correlations with the FACT-B subscales (Spearman rank correlation coefficient values ranging from 0.096 to 0.682). Nesvold et al assessed a variety of psychometric properties for KAPS in stage II breast cancer survivors.⁴⁰ Internal consistency was acceptable (Cronbach's $\alpha = 0.92$), and convergent validity found the total KAPS score explaining 32.5% of the variance in abduction on the affected shoulder, and concurrent validity explained 54.8% of the variance in arm symptom subscale of the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-Breast Cancer Module (EORTC QLQ-B23).⁴⁰ These two studies do have limitations that include an unclear definition of time since breast surgery³⁹ which can affect responsiveness estimates and assessing a very specific subset of women diagnosed with breast cancer (stage II) thereby limiting the generalizability.⁴⁰

A recommendation for future research includes conducting studies to assess the psychometric properties of disease/condition specific or joint specific patient-reported outcome measures in the breast cancer survivor population, as well as those outcome measures that were specifically designed for use in women who have received treatment for breast cancer. Additionally, we recommend future studies use a larger sample size of at least 100 participants in order for more precise estimates to be examined and reported. Clinicians and researchers utilizing patient-reported upper extremity outcome measures in the breast cancer population need to recognize that these measures lack important psychometric information. Having this information will help to provide clinicians and researchers with evidence for best and appropriate use of these patient-reported outcome measures in breast cancer survivors.

Study Limitations

Several limitations exist with this systematic review and the results should be interpreted with caution. As with any systematic review, it is possible that some related papers could not be identified, such as unpublished work. Our search was limited to English-language journals, and studies using other languages were not considered. Although attempts were made to contact the authors, for aim 2 we were only able to examine the psychometric properties in 4 of the 8 patient-reported measures that have been administered to a breast cancer survivor population due to a lack of reported means and standard deviations or comparison group. Therefore, we cannot exclude the possibility that a superior questionnaire exists. Additionally, there were several limitations of the studies included in this review. First, studies showing large (>0.80) effect sizes for construct validity and responsiveness had relative small sample sizes.^{15,34,35} The impact of this was that several estimates were imprecise, including one example for construct validity (Table 2) and several examples where the responsiveness of measures could not distinguish between control and treatment groups (Table 3). Only two studies had greater than 100 participants, Smoot and colleagues included 144 participants and Dawes and colleagues included 204 participants.^{29,33} If the outcome measures were investigated in larger cohorts, the results could be different and estimates would be more precise (e.g. smaller 95% CI's). The intervention studies that reported large responsiveness had varied assessment times, 1, 3 and 6 months, from initial

assessment to follow up assessment. Other limitations include that the data could not be pooled due to the heterogeneity between studies with regards to the interventions used, variability of breast cancer related treatment (ALND, SLND, reconstruction, etc.), as well as a wide range of assessment periods (4 weeks to 5 years). Although the results may be relevant to the decision-making process for administering a patient-reported outcome measure in the breast cancer survivor population, our findings provide an updated summary of the state of the literature but should not be viewed as definitive. Our data provides complementary information on which patient-reported outcome measures could be administered to the breast cancer survivor population, but the ultimate choice must be made in the context of a clinician's experience and factors specific to the breast cancer survivor. These factors include, but are not limited to, whether the breast cancer survivors seen in a clinical practice are similar to those enrolled in a respective study and the clinicians' familiarity with the patient-reported outcome measures.

CONCLUSIONS

There are over 30 upper extremity patient-reported upper extremity outcome measures but only 8 have been administered to the breast cancer survivor population. This review reinforces the importance of understanding the constructs and psychometric properties of patient-reported outcome measures for accurate use and interpretation in a specific patient population. Based on the available evidence from this review we recommend administering the DASH to assess patient-reported upper extremity function in breast cancer survivors, although we acknowledge further work needs to be done incorporating sample sizes larger than 100 to assess the psychometric properties of these outcome measures. The results of this review do provide a clear direction for those researchers wishing to conduct future research investigating the psychometric properties and usefulness of patient-reported outcome measures administered to a breast cancer survivor cohort.

Acknowledgments

We certify that no party having a direct interest in the results of the research supporting this article has or will confer a benefit on us or on any organization with which we are associated AND, if applicable, we certify that all financial and material support for this research (eg, NIH or NHS grants) and work are clearly identified in the title page of the manuscript. (Shana Harrington, Lori A Michener, Tiffany Kendig, Susan Miale, Steven Z George)

Abbreviations

ALND	Axillary lymph node dissection
CI	Confidence Interval
DASH	Disabilities of Arm, Shoulder and Hand
ES	Effect Size
Penn	Penn Shoulder Score
SPADI	Shoulder Pain and Disability Index
SDQ-NL	Shoulder Disability Questionnaire-Dutch

SLND	Sentinel lymph node dissection
Wingate	10 Questions by Wingate

References

1. How many women get breast cancer?. 2013. <http://www.cancer.org/Cancer/BreastCancer/OverviewGuide/breast-cancer-overview-key-statistics>
2. Society, AC. What Are the Key Statistics About Breast Cancer?. 2012. <http://www.cancer.org/Cancer/BreastCancer/DetailedGuide/breast-cancer-key-statistics>. Accessed 8/23/12, 2012
3. Gordon LG, Scuffham P, Battistutta D, Graves N, Tweeddale M, Newman B. A cost-effectiveness analysis of two rehabilitation support services for women with breast cancer. *Breast Cancer Res Treat*. Nov; 2005 94(2):123–133. [PubMed: 16261411]
4. Yap KP, McCready DR, Narod S, Manchul LA, Trudeau M, Fyles A. Factors influencing arm and axillary symptoms after treatment for node negative breast carcinoma. *Cancer*. Mar 15; 2003 97(6): 1369–1375. [PubMed: 12627499]
5. Vinokur AD, Threatt BA, Vinokur-Kaplan D, Satariano WA. The process of recovery from breast cancer for younger and older patients. Changes during the first year. *Cancer*. Mar 1; 1990 65(5): 1242–1254. [PubMed: 2302673]
6. Tengrup I, Tennvall-Nittby L, Christiansson I, Laurin M. Arm morbidity after breast-conserving therapy for breast cancer. *Acta Oncol*. 2000; 39(3):393–397. [PubMed: 10987237]
7. Tengrup I, Nittby LT, Christiansson I, Laurin M. [Problems with arms are common after breast surgery. Lymphedema is a frequent complication in elderly women treated for breast cancer]. *Lakartidningen*. Nov 17; 1999 96(46):5089–5091. [PubMed: 10608134]
8. Tasmuth T, von Smitten K, Kalso E. Pain and other symptoms during the first year after radical and conservative surgery for breast cancer. *Br J Cancer*. Dec; 1996 74(12):2024–2031. [PubMed: 8980408]
9. Sugden EM, Rezvani M, Harrison JM, Hughes LK. Shoulder movement after the treatment of early stage breast cancer. *Clin Oncol (R Coll Radiol)*. 1998; 10(3):173–181. [PubMed: 9704180]
10. Springer BA, Levy E, McGarvey C, et al. Pre-operative assessment enables early diagnosis and recovery of shoulder function in patients with breast cancer. *Breast Cancer Res Treat*. Feb; 2010 120(1):135–147. [PubMed: 20054643]
11. Shamley DRSR, Weatherall R, Oskrochi R, Watson M, Ostlere S, Sugden E. Changes in Shoulder Muscle Size and Activity Following Treatment for Breast Cancer. *Breast Cancer Res Treat*. 2006
12. Ridner SH. Quality of life and a symptom cluster associated with breast cancer treatment-related lymphedema. *Support Care Cancer*. Nov; 2005 13(11):904–911. [PubMed: 15812652]
13. Kuroi K, Shimozuma K, Ohsumi S, Imai H, Ono M. Current status of health outcome assessment of medical treatment in breast cancer. *Breast Cancer*. 2007; 14(1):74–80. [PubMed: 17244999]
14. Herd-Smith A, Russo A, Muraca MG, Del Turco MR, Cardona G. Prognostic factors for lymphedema after primary treatment of breast carcinoma. *Cancer*. Oct 1; 2001 92(7):1783–1787. [PubMed: 11745250]
15. Harrington S, Padua D, Battaglini C, et al. Comparison of shoulder flexibility, strength, and function between breast cancer survivors and healthy participants. *J Cancer Surviv*. Jun; 2011 5(2):167–174. [PubMed: 21225372]
16. Dawson J, Hill G, Fitzpatrick R, Carr A. Comparison of clinical and patient-based measures to assess medium-term outcomes following shoulder surgery for disorders of the rotator cuff. *Arthritis Rheum*. Oct 15; 2002 47(5):513–519. [PubMed: 12382300]
17. Brennan GP, Parent EC, Cleland JA. Description of clinical outcomes and postoperative utilization of physical therapy services within 4 categories of shoulder surgery. *J Orthop Sports Phys Ther*. Jan; 2010 40(1):20–29. [PubMed: 20044703]

18. Gabel CP, Yelland M, Melloh M, Burkett B. A modified QuickDASH-9 provides a valid outcome instrument for upper limb function. *BMC musculoskeletal disorders*. 2009; 10:161. [PubMed: 20021677]
19. Cleland JA, Childs JD, Whitman JM. Psychometric properties of the Neck Disability Index and Numeric Pain Rating Scale in patients with mechanical neck pain. *Arch Phys Med Rehabil*. Jan; 2008 89(1):69–74. [PubMed: 18164333]
20. Beaton DE, Boers M, Wells GA. Many faces of the minimal clinically important difference (MCID): a literature review and directions for future research. *Curr Opin Rheumatol*. Mar; 2002 14(2):109–114. [PubMed: 11845014]
21. Wright RW, Baumgarten KM. Shoulder outcomes measures. *J Am Acad Orthop Surg*. Jul; 2010 18(7):436–444. [PubMed: 20595136]
22. Slobogean GP, Noonan VK, O'Brien PJ. The reliability and validity of the Disabilities of Arm, Shoulder, and Hand, EuroQol-5D, Health Utilities Index, and Short Form-6D outcome instruments in patients with proximal humeral fractures. *J Shoulder Elbow Surg*. Apr; 2010 19(3):342–348. [PubMed: 20189839]
23. Michener LA, McClure PW, Karduna AR. Anatomical and biomechanical mechanisms of subacromial impingement syndrome. *Clinical biomechanics*. Jun; 2003 18(5):369–379. [PubMed: 12763431]
24. Beaton DE, Katz JN, Fossel AH, Wright JG, Tarasuk V, Bombardier C. 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. *J Hand Ther*. Apr-Jun; 2001 14(2):128–146. [PubMed: 11382253]
25. Marinus J, Ramaker C, van Hilten JJ, Stiggelbout AM. Health related quality of life in Parkinson's disease: a systematic review of disease specific instruments. *Journal of neurology, neurosurgery, and psychiatry*. Feb; 2002 72(2):241–248.
26. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. 2. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc; 1988.
27. Wiebe S, Rose K, Derry P, McLachlan R. Outcome assessment in epilepsy: comparative responsiveness of quality of life and psychosocial instruments. *Epilepsia*. Apr; 1997 38(4):430–438. [PubMed: 9118848]
28. Portney, LG.; Watkins, MP. *Foundations of Clinical Research: Applications to Practice*. 3. Prentice hall; 2008.
29. Smoot B, Wong J, Cooper B, et al. Upper extremity impairments in women with or without lymphedema following breast cancer treatment. *J Cancer Surviv*. Jun; 2010 4(2):167–178. [PubMed: 20373044]
30. Rietman JS, Dijkstra PU, Geertzen JH, et al. Treatment-related upper limb morbidity 1 year after sentinel lymph node biopsy or axillary lymph node dissection for stage I or II breast cancer. *Ann Surg Oncol*. Nov; 2004 11(11):1018–1024. [PubMed: 15525832]
31. Lee SA, Kang JY, Kim YD, et al. Effects of a scapula-oriented shoulder exercise programme on upper limb dysfunction in breast cancer survivors: a randomized controlled pilot trial. *Clinical rehabilitation*. Jul; 2010 24(7):600–613. [PubMed: 20530648]
32. Lau RW, Cheing GL. Managing postmastectomy lymphedema with low-level laser therapy. *Photomedicine and laser surgery*. Oct; 2009 27(5):763–769. [PubMed: 19878027]
33. Dawes DJ, Meterissian S, Goldberg M, Mayo NE. Impact of lymphoedema on arm function and health-related quality of life in women following breast cancer surgery. *Journal of rehabilitation medicine : official journal of the UEMS European Board of Physical and Rehabilitation Medicine*. Aug; 2008 40(8):651–658.
34. Cinar N, Seckin U, Keskin D, Bodur H, Bozkurt B, Cengiz O. The effectiveness of early rehabilitation in patients with modified radical mastectomy. *Cancer Nurs*. Mar-Apr; 2008 31(2): 160–165. [PubMed: 18490892]
35. Beurskens CH, van Uden CJ, Strobbe LJ, Oostendorp RA, Wobbes T. The efficacy of physiotherapy upon shoulder function following axillary dissection in breast cancer, a randomized controlled study. *BMC Cancer*. 2007; 7:166. [PubMed: 17760981]

36. Jeong HJ, Sim YJ, Hwang KH, Kim GC. Causes of shoulder pain in women with breast cancer-related lymphedema: a pilot study. *Yonsei Med J.* Jul; 2011 52(4):661–667. [PubMed: 21623610]
37. Portela AL, Santaella CL, Gomez CC, Burch A. Feasibility of an Exercise Program for Puerto Rican Women who are Breast Cancer Survivors. *Rehabilitation oncology.* Apr 1; 2008 26(2):20–31. [PubMed: 20664723]
38. Crosbie J, Kilbreath SL, Hollmann L, York S. Scapulohumeral rhythm and associated spinal motion. *Clinical biomechanics.* Feb; 2008 23(2):184–192. [PubMed: 17981379]
39. Mulero-Portela AL, Colon-Santaella CL, Cruz-Gomez C. Cross-cultural adaptation of the Disability of Arm, Shoulder, and Hand questionnaire: Spanish for Puerto Rico version. *International journal of rehabilitation research. Internationale Zeitschrift fur Rehabilitationsforschung. Revue internationale de recherches de readaptation.* Dec; 2009 32(4):287–293. [PubMed: 19901616]
40. Nesvold IL, Fossa SD, Naume B, Dahl AA. Kwan's arm problem scale: psychometric examination in a sample of stage II breast cancer survivors. *Breast Cancer Res Treat.* Sep; 2009 117(2):281–288. [PubMed: 19112616]
41. Hudak PL, Amadio PC, Bombardier C. Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand) [corrected]. The Upper Extremity Collaborative Group (UECG). *Am J Ind Med.* Jun; 1996 29(6):602–608. [PubMed: 8773720]
42. McClure PW, Michener LA. Measures of Adult Shoulder Function. *Arthritis Rheum.* 2003; 49(5S):S50–58.
43. McClure PW, Michener LA. Measures of Adult Shoulder Function: The American Shoulder and Elbow Surgeons Standardized Shoulder Form Patient Self-Report Section (ASES), Disabilities of Arm, Shoulder and Hand (DASH), Shoulder Disability Questionnaire, Shoulder Pain and Disability Index (SPADI), and Simple Shoulder Test. *Arthritis Care & Research.* 2003; 49(5S):S50–58.
44. Dowrick ASBJ, Gabbe BJ, et al. Outcome Instruments For The Assessment Of The Upper Extremity Following Trauma: A Review. *International Journal of the Care of the Injured.* 2004; 36:468–476. [PubMed: 15755426]
45. Kwan W, Jackson J, Weir LM, Dingee C, McGregor G, Olivotto IA. Chronic arm morbidity after curative breast cancer treatment: prevalence and impact on quality of life. *J Clin Oncol.* Oct 15; 2002 20(20):4242–4248. [PubMed: 12377968]
46. Leggin BG, Michener LA, Shaffer MA, Brennehan SK, Iannotti JP, Williams GR Jr. The Penn shoulder score: reliability and validity. *J Orthop Sports Phys Ther.* Mar; 2006 36(3):138–151. [PubMed: 16596890]
47. de Winter AF, van der Heijden GJ, Scholten RJ, van der Windt DA, Bouter LM. The Shoulder Disability Questionnaire differentiated well between high and low disability levels in patients in primary care, in a cross-sectional study. *Journal of clinical epidemiology.* Nov; 2007 60(11):1156–1163. [PubMed: 17938058]
48. van der Windt DA, van der Heijden GJ, de Winter AF, Koes BW, Deville W, Bouter LM. The responsiveness of the Shoulder Disability Questionnaire. *Annals of the rheumatic diseases.* Feb; 1998 57(2):82–87. [PubMed: 9613336]
49. Rietman JS, Dijkstra PU, Hoekstra HJ, et al. Late morbidity after treatment of breast cancer in relation to daily activities and quality of life: a systematic review. *Eur J Surg Oncol.* Apr; 2003 29(3):229–238. [PubMed: 12657232]
50. Roach KE, Budiman-Mak E, Songsiridej N, Lertratanakul Y. Development of a shoulder pain and disability index. *Arthritis care and research : the official journal of the Arthritis Health Professions Association.* Dec; 1991 4(4):143–149. [PubMed: 11188601]
51. Roddey TS, Olson SL, Cook KF, Gartsman GM, Hanten W. Comparison of the University of California-Los Angeles Shoulder Scale and the Simple Shoulder Test with the shoulder pain and disability index: single-administration reliability and validity. *Phys Ther.* Aug; 2000 80(8):759–768. [PubMed: 10911414]
52. MacDermid JC, Solomon P, Prkachin K. The Shoulder Pain and Disability Index demonstrates factor, construct and longitudinal validity. *BMC musculoskeletal disorders.* 2006; 7:12. [PubMed: 16472394]

53. Bot SD, Terwee CB, van der Windt DA, Bouter LM, Dekker J, de Vet HC. Clinimetric evaluation of shoulder disability questionnaires: a systematic review of the literature. *Annals of the rheumatic diseases*. Apr; 2004 63(4):335–341. [PubMed: 15020324]
54. Beaton D, Richards RR. Assessing the reliability and responsiveness of 5 shoulder questionnaires. *J Shoulder Elbow Surg*. Nov-Dec;1998 7(6):565–572. [PubMed: 9883415]
55. Williams JW Jr, Holleman DR Jr, Simel DL. Measuring shoulder function with the Shoulder Pain and Disability Index. *The Journal of rheumatology*. Apr; 1995 22(4):727–732. [PubMed: 7791172]
56. L'Insalata JC, Warren RF, Cohen SB, Altchek DW, Peterson MG. A self-administered questionnaire for assessment of symptoms and function of the shoulder. *J Bone Joint Surg Am*. May; 1997 79(5):738–748. [PubMed: 9160947]
57. Wingate L. Efficacy of Physical Therapy for Patients Who Have Undergone Mastectomies: A Prospective Study. *Phys Ther*. 1985; 65(6):896–900. [PubMed: 4001167]

Figure 1.
Flow chart depicting search and selection process

Table 1
Patient-rated Outcome Measures for the upper extremity used or developed for patients with breast cancer.

Outcome Measure	Population Developed	Validated Populations	Instrument Properties	Scale	Use in BCS
DASH	Individuals with a upper extremity musculoskeletal problems including surgical cases. ^{41,42}	Diverse populations, for single or multiple disorders, throughout the upper extremity, and at a single or multiple points in time. ²⁴	30 questions relating to symptoms (5 items) and functional tasks (25 items). Work and sports-specific optional modules (4 items each). ⁴³	0–100, a higher score reflects greater upper extremity disability. ⁴⁴	25 studies
KAPS	Women with breast cancer. ^{40,45}	Stage II breast cancer. ⁴⁰	13 items consisting of 2 scales. Problem subscale (8 items) rating problems with upper extremity/shoulder function, pain, stiffness and swelling; ADL subscale (5 items) rating impairment of ADLs. ⁴⁰	8–40, Problem subscale; 5–25, ADL subscale. Higher scores reflect increased difficulty. ⁴⁰	4 studies
Penn	Musculoskeletal surgical and non-surgical disorders. ⁴⁶	Shoulder impingement/tendonitis, rotator cuff tear, instability, adhesive capsulitis, proximal humeral fracture, acromioclavicular joint osteoarthritis (OA) and glenohumeral joint OA. ⁴⁶	Contains 3 subscales: function (20 items), pain (3 items) and satisfaction with use of shoulder (1 item). ⁴⁶	0–100, with 3 subscales: Pain = 0–30, Satisfaction = 0–10, and Function = 0–60. Lower scores indicate lower function, lower satisfaction, and higher pain. ⁴⁶	1 study
SDQ-NL	Patients with non-surgical, soft-tissue disorders of the shoulder. ^{47,48}	Patients presenting to general practitioners, orthopedic practice or secondary care rheumatologists with a variety of shoulder complaints, excluding those related to neurological, vascular or internal disorders; systemic rheumatic diseases; fractures; and dislocations. ⁴⁷	Contains 16 statements that describe the situations in which patients experience pain and what some of the effects may be. ⁴⁹	0–100, a higher score reflects greater upper extremity disability. ⁴⁸	5 studies
SPADI	Males with shoulder pain. Designed to be used in an outpatient clinic as a means to measure the “impact of shoulder pathology in terms of pain and disability”. ⁵⁰	Shoulder pain of many origins (subacromial shoulder pain, rotator cuff disease) and many different clinical settings (ambulatory care clinic, medical centre, upper extremity clinic, physical therapy clinic, and private practice surgeon). ^{51–54}	Numeric rating scale with 13 items (5 for pain and 8 for disability). ⁵⁵	0–100, a higher score reflects greater upper extremity pain & disability. ⁵⁵	3 studies
SRQ	Developed in patients with a variety of shoulder disorders, including glenohumeral instability, impingement syndrome, complete rotator cuff tear, osteoarthritis, and adhesive capsulitis. ⁵⁶	Glenohumeral instability, rotator cuff tear, osteoarthritis, adhesive capsulitis. ⁵⁶	21 items comprised of six separately scored domains: global assessment, pain, daily activities, recreational and athletic activities, work, and satisfaction. ⁵⁶	The total possible score ranges from 17–100, with lower scores indicating poorer function. ⁵⁶	1 study
ULDQ	Women with breast cancer. ¹⁰	Women with unilateral, stage I–III breast cancer 12+ months after surgery. ¹⁰	80 items organized into 6 categories: symptom severity in the past 2 weeks; problem/symptom frequency in the past 2 weeks; current activity limitation; time relating to interference with normal work in the past 2 weeks; time with specified problems with work or daily activity performance in the past 2 weeks; and interference	Not reported	3 studies

Outcome Measure	Population Developed	Validated Populations	Instrument Properties	Scale	Use in BCS
Wingate	Questions were developed to assess functional status of women status post mastectomy over time. ⁵⁷	Not reported	with weekly social, sports, or recreational activities in the last 2 weeks. with weekly social, sports, or recreational activities in the last 2 weeks. 10 questions, each relating to degree of difficulty performing specified upper extremity tasks with the affected extremity. ⁵⁷	0–40 where lower scores are indicative of greater upper limb function. ⁵⁴	4 studies

DASH= Disabilities of the Upper extremity, Shoulder, and Hand; KAPS = Kwan's Upper extremity/Shoulder Problem Scale; Penn – University of Pennsylvania Shoulder Score; SPADI = Shoulder Pain and Disability Index; SDQ-NL = Shoulder Disability Questionnaire-Dutch; SRQ = Shoulder Rating Questionnaire; ULDDQ = Upper Limb Disability Questionnaire; Wingate = 10 Questions by Wingate.

Table 2

Construct Validity of the DASH, Penn, SDQ-NL, and Wingate patient-rated outcome measures for Breast Cancer Survivors by cohort, assessed with effect sizes where $(M_1 - M_2)/\text{pooled SD}$, where M is the mean for each cohort group indicated and the pooled standard deviation is the root mean square of the two standard deviations for the groups.

Measure	Author	Study Design	Sample Size	Comparison Groups	Effect Size (Cohen's d)	95% Confidence Interval
<i>Lymphedema vs No lymphedema</i>						
DASH	Smoot et al 2010 ²⁹	Cross-Sectional	BCS=144	Lymphedema (n=71) No lymphedema (n=73)	0.65	0.31 – 0.98
	Dawes et al 2008 ³³	Cross-Sectional	BCS=204	Lymphedema (n=72) No lymphedema (n=132)	1.63	1.30 – 1.96
<i>Dominant vs Non Dominant Extremity</i>						
DASH	Crosbie et al 2010 ³⁸	Observational	BCS=53	Dominant (n=29) Non-dominant (n=24)	0.27	-0.30 – 0.83
<i>BCS vs Healthy Controls</i>						
DASH	Harrington et al 2011 ¹⁵	Case-Control	BCS=23	BCS (n=23) Healthy Control (n=23)	1.47	0.80 – 2.10
PENN	Harrington et al 2011 ¹⁵	Case-Control	BCS=22	BCS (n=22) Healthy Control (n=22)	1.46	0.77 – 2.10
<i>Miscellaneous</i>						
DASH	Jeong et al 2011 ³⁶	Cross-Sectional	BCS=38	I = No pain & Normal US (n=11) II = Pain & Normal US (n=12) III = Pain & Abnormal US (n=15)	I & II = 0.73 I & III = 1.57 II & III = 0.92	0.15 – 1.54 0.64 – 2.40 0.08 – 1.67
SDQ-NL	Rietman et al 2004 ³⁰	Prospective	BCS=204	ALND (n=131) SLND (n=58)	0.39	0.05 – 0.68

BCS = Breast Cancer Survivor; US = ultrasound; ALND = axillary lymph node dissection; SLND = sentinel lymph node dissection

Table 3

Responsiveness: Effect sizes for the DASH, and Wingate; coefficient of responsiveness was calculated as mean $(test_1 - test_2)_{treatment} - \text{mean } (test_1 - test_2)_{control} / SD_{test_1 \text{ pooled}}$

Measure	Author	Study Design	Sample Size	Cohort	Effect Sizes	Confidence Interval
<i>Physical Therapy Intervention</i>						
DASH	Beurskens et al 2007 ³⁵	Randomized Control	BCS = 29	Intervention = 15 BCS Control = 14	3 months = 1.13 6 months = 0.94	0.33 – 1.86 0.15 – 1.68
	Lee et al 2010 ³¹	Randomized Control	BCS = 31	Intervention = 13 BCS Control = 18	2 months = 0.20	-0.57 – 0.97
	Portela et al 2008 ³⁷	Randomized Control	BCS = 22	Intervention = 13 BCS Control = 9	6 months = 0.37	-0.51 – 1.21
	Cinar et al 2008 ³⁴	Randomized Control	BCS = 57	Intervention = 27 BCS Control = 30	1 month = 1.12 6 months = 1.54	0.52 – 1.63 0.93 – 2.11
Wingate						
<i>Lymphedema Intervention</i>						
DASH	Lau et al 2009 ³²	Randomized Control	BCS = 21	Intervention = 11 BCS Control = 10	1 month = 0.45	-0.44 – 1.30
					2 months = 0.71	-0.20 – 1.56