2015

Oncology Section EDGE Task Force on Prostate Cancer Outcomes: A Systematic Review of Clinical Measures of Strength and Muscular Endurance

Mary Insana Fisher
University of Dayton, mary.fisher@udayton.edu

Claire Davies
Baptist Health Lexington

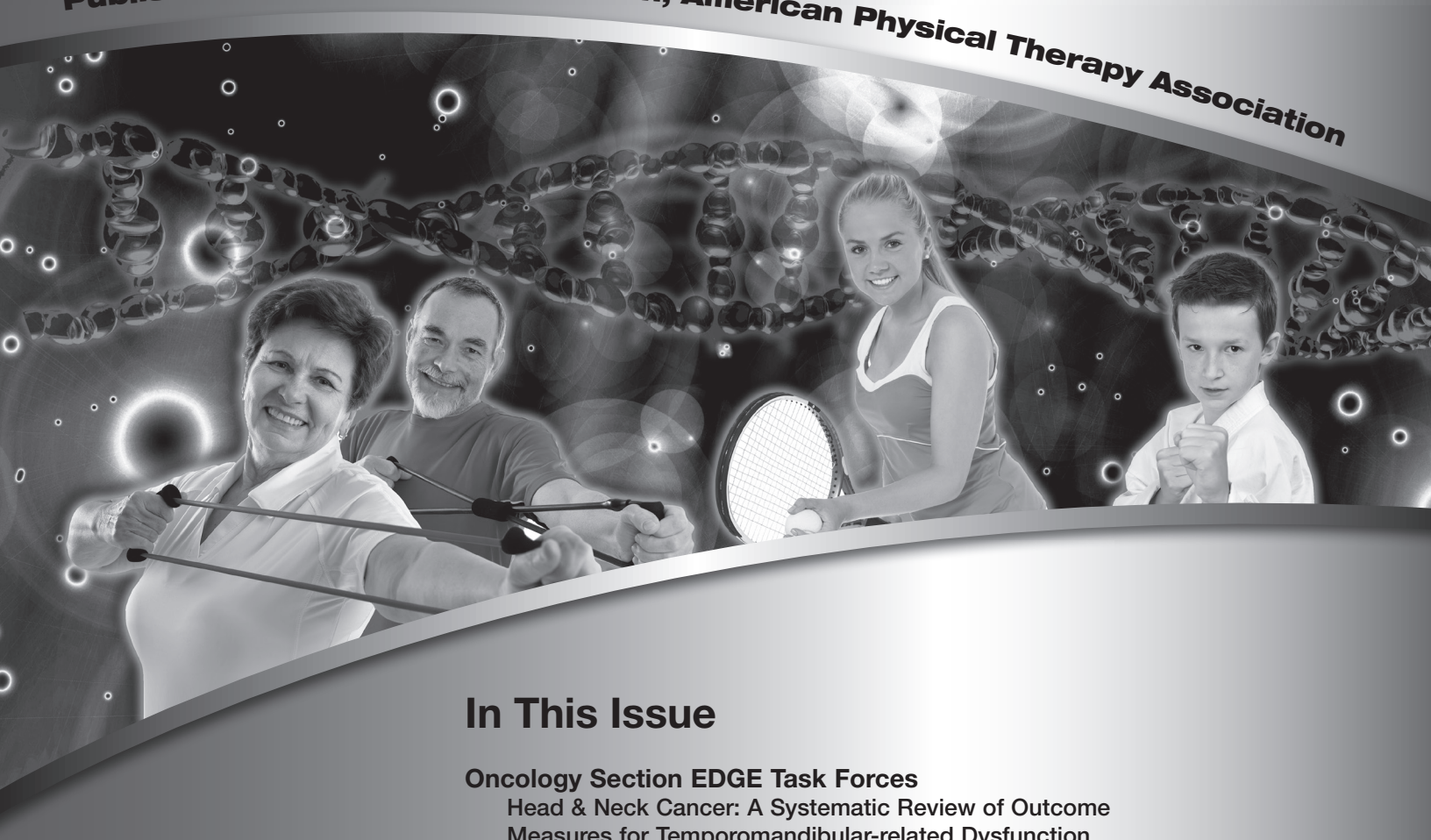
Genevieve Colon
University of Michigan - Flint

Hannah Geyer
University of Dayton

Lucinda Pfalzer
University of Michigan - Flint

Rehabilitation Oncology

Oncology Section
American Physical Therapy Association
Celebrate Living
Vol. 33, No. 2, 2015
Published by the Oncology Section, American Physical Therapy Association



In This Issue

APTA
American Physical Therapy Association

Assistant/Associate Professor in Oncology & Rehabilitation
Department of Physical Therapy and Human Movement Sciences &
Robert H. Lurie Comprehensive Cancer Center
Feinberg School of Medicine, Northwestern University

The Department of Physical Therapy and Human Movement Sciences (PTHMS) and Robert H. Lurie Comprehensive Cancer Center, Northwestern University Feinberg School of Medicine, are inviting applications for a full-time, tenure eligible faculty position in the area of oncology rehabilitation. Primary responsibilities will include developing a strong research program with strong multidisciplinary links to the Lurie Cancer Center. Furthermore the candidate will be involved in a Lymphatic and Integumentary Dysfunction related coursework in the Doctor of Physical Therapy (DPT) curriculum. The position also includes PhD and DPT student advising, service on departmental committees, and practice opportunities in clinical oncology rehabilitation. Depending on the experience and background of the applicant, this position is being offered at the assistant or associate professor level. Salary is commensurate with experience. Generous start-up package will be offered. Ideal applicants will hold the credentials of PT, PhD; have post-doctoral experience; have a record in obtaining federal training/research funding and expertise in the management of lymphedema.

PTHMS and the Lurie Cancer Center are a part of the vibrant educational and cultural community of Northwestern University. Both are located on the Chicago campus of Northwestern University, between Lake Michigan and North Michigan Avenue. The Robert H. Lurie Cancer Center is a National Cancer Institute-designated "Comprehensive" center dedicated to the highest standards of cancer research, patient care, education, and community outreach. Collaborations also exist with the Departments of Biomedical, Mechanical Engineering, Physical Medicine & Rehabilitation, Neurology, Physiology, Orthopedics, and Neurological Surgery. The successful applicant is expected to continue to build on existing and new collaborations. PTHMS and Lurie Cancer Center faculty members participate in PhD student training programs through the Northwestern University Interdepartmental Biological Sciences Graduate Program, Interdepartmental Neuroscience Program, and a joint DPT-PhD (Engineering) program. For more information please visit: http://www.cancer.northwestern.edu/home/index.cfm and http://www.feinberg.northwestern.edu/nupthms.

Northwestern University is an Equal Opportunity, Affirmative Action Employer of all protected classes, including veterans and individuals with disabilities. Women and minorities are encouraged to apply. Hiring is contingent upon eligibility to work in the United States.

Applications will be considered until position is filled with an anticipated start date of August 1, 2015. Interested applicants should send a letter of application, statement of research goals and teaching philosophy, curriculum vitae, and the names and addresses of three professional references to:

Ms. Erin Neal
Search Committee Coordinator
Department of Physical Therapy & Human Movement Sciences
645 N. Michigan Ave, Suite # 1100
Chicago, IL 60611
Email: e-neal@northwestern.edu
Phone: 312.503.3163

Table of Contents

Rehabilitation Oncology

Rehabilitation Oncology (ISSN 2168-3808) is the official publication of the Oncology Section, APTA. Nonmember and institution subscriptions are available for $70 per year (4 issues). Back issues are available for members ($5) and nonmembers ($20). The contents of articles appearing in this publication represent the thoughts and ideas of the authors and do not necessarily reflect the views of the Oncology Section, APTA. The editor reserves the right to edit submitted manuscripts or other material as necessary for publication. We encourage comments and opinions concerning the content of Rehabilitation Oncology through Letters to the Editor.

DEADLINES

Published 4 times per year in January, April, July, and October. Please submit all material for publication to the Editor by November 15, February 15, May 15 and August 15.

Please send all items for publication to the Editor:
Lucinda (Cindy) Pfalzer, PT, PhD, FAPTA
Editor, Rehabilitation Oncology
PT Department, Rm 2157 WSW Bldg
University of Michigan-Flint
303 E Kearsley St
Flint, MI 48502-1950
E-mail: cpfalzer@umich.edu

The contents of articles appearing in this publication represent the thoughts and ideas of the authors and are not necessarily those of the Editor, the Executive Committee of the Oncology Section, or general membership of the Section. The Editor reserves the right to edit submitted manuscripts or other material for publication. We encourage comments and opinions concerning the content of *Rehabilitation Oncology* through Letters to the Editor.

ADVERTISEMENTS

Individuals, health care organizations, businesses, and vendors are invited to advertise in this journal. Prices are as follows:

Size	Width	Height	Rate (Per Issue)	
			1x	4x
Full page	7"	10"	$400	$350
1/2 page	7"	5"	$300	$250
1/4 page	3.5"	5"	$200	$150
1/8 page	3.5"	2.5"	$100	$75
Inside Front Cover	7"	10"	$650	$600
Inside Back Cover	7"	10"	$650	$600
Outside Back Cover	7"	5"	$650	$600

For advertisements, contact 800.999.2782, ext. 8564.

Please send orders for subscriptions to:
Oncology Section, APTA
1055 N Fairfax, Suite 205, Alexandria, VA 22314

EDITORIAL BOARD

President's Perspective

Lisa VanHoose, PT, PhD

President of the Oncology Section;
Department Chairperson and Associate Professor, Hampton University, Hampton, VA

Mahatma Gandhi stated that "the best way to find yourself is to lose yourself in the service of others." As you treat your patients today and tomorrow, take a moment and ask if being a physical therapist (PT) or physical therapist assistant (PTA) is a service or a job for you. Your self-assessment will determine your role as the profession and the Section move forward with education, policy, research, and outreach initiatives. Service and employment are not synonymous. A job or employment will mandate that you continue with the status quo. You will treat your patients, document, and read just enough research to select a functional outcome measure and tool for billing. A job is just that—meeting the employment requirements so that one may receive money or compensation. Service is more dynamic and includes passion, courage, and innovation. The PT or PTA who views the profession as a service contemplates the needs of the field and the public in the future. He or she seeks out new knowledge to address those needs. He or she has the courage to implement strategies to address the needs. Service requires that we be an active part of the solution to address the movement needs of our communities, thereby transforming society.

I strongly believe that an intersection exists between service and research. Many times, data about the needs of a community or the generation of ideas or possibilities about an intervention kindles a flame for a new program, technique, or care delivery strategy. We also see the opposite, where passion about community needs drives the next research question and project. Research grants focus on societal impact and innovation. We are encouraged to develop projects across institutions and state lines. Multisite research studies have been utilized for years to improve recruitment and the impact of the studies. We have been able to freely enter into these courageous and creative ventures, because we were not restricted by licensing requirements and practice acts. We are excited to see this spirit of innovation and possibilities moving into the delivery of physical therapy services.

The Federation of State Boards of Physical Therapy (FSBPT) began conversations regarding the need for interstate licensure compacts in 2010. The FSBPT had the foresight to identify that our licensing processes would be a barrier to accessing physical therapy care. In 2014, the FSBPT Compact Advisory Team determined that interstate compacts are needed and feasible. The interstate licensure compact would allow for two or more states to agree on a "privilege to practice" model in remote states. The current model recommendation would allow a PT or PTA to hold an unrestricted license in a primary state and then he/she would notify remote states, participating in the licensure compacts, of practice. The notification, and possibly a fee, would allow for "privilege to practice" in remote states. This opportunity would have immediate impact for many of us providing oncology rehabilitation care. Cancer care is often a shared care model between academic medical centers, specialized medical care centers, and community health care centers. Often, our patients travel across state lines for care due to medical needs or personal factors. The proposed interstate licensure compacts would allow us to provide on-site care without the additional burdens of state licensing. Even more exciting is that interstate licensure compacts would remove barriers to providing telehealth and telerehabilitation. It will allow opportunities for many of us to provide essential services to those who need us most, in rural and underserved communities. I do believe that each and every one of us will answer the call to provide physical therapy care with a creative, passionate, and courageous spirit that can only be described as a service and not a job. As we serve others, I look forward to each of us finding ourselves.

Move forward,
Lisa

Editor's Message

Lucinda (Cindy) Pfalzer, PT, PhD, FACSM, FAPTA

Editor of *Rehabilitation Oncology* and Professor (Emerita), Physical Therapy Department,
University of Michigan-Flint, Flint, MI

The lazy days of summer—do we remember when we thought summer would last forever; but change is inevitable. There is major news from the Section regarding publication of the Journal—we have a new publisher beginning January 2016. Before I get to what this means for the Section and Journal, let me take a moment to acknowledge the integral role the Orthopaedic Section has played for our Journal for many years. Sharon Klinski has been our publisher since 1997 and a large part of the credit for the quality of the Journal is hers. Thank you Sharon for of all of your years of dedicated service to the Journal. We appreciate all of the work that has gone into making the Journal a success. The decision to change publishers was to bring the Journal into the multi-media age we now live in. The Journal's new publisher is Wolters Kluwer. This publisher will be able to provide improved access to the Journal in e-reader format, an iPhone app, a searchable portal for members to download individual pdfs of articles all the way back to the first published issue of the Journal, an improved webpage and web content for the Journal, and lastly, increase the reach of the Journal with indexing in OVID. The Journal is now available electronically through ProQuest, EBSCO Rehabilitation, Allen Press, and CINAHL. There are more benefits the members don't see; but are important for the quality and work of the Journal such as the electronic manuscript management system for the authors, reviewers, Associate Editors, and Editor. I look forward to hearing your feedback next year specifically regarding the new electronic features of the Journal as we transition the Journal to multi-media.

Back to this issue of the Journal, it was a busy first half of the year with two issues jam packed with research that continues the focus on evidence for practice. This issue has 4 systematic reviews from the Section's EDGE task forces for head and neck cancer, breast cancer, and prostate cancer. You have an issue full of information that translates to practice by addressing your needs to select better outcome measures to improve the examination and reporting of the care you provide. The Research Round-up (available via ePub at oncologypt.org/publications/rehabilitation-oncology-journal) column reports on the Cancer Rehabilitation Symposium that was hosted by the Rehabilitation Medicine Department (RMD) of the National Institutes of Health (NIH) Clinical Center and was co-sponsored by the National Cancer Institute and the National Center for Medical Rehabilitation Research this past June in Bethesda, Maryland. There was lively discussion, and we look forward to the publication of the working papers and research agenda. Please take a moment to read about this important initiative. As always, we are open to your feedback.

Erratum. The authors wish to correct an error in their article "Recommendations for patient-reported outcome measures for head and neck cancer-related shoulder dysfunction: A systematic review" volume 32, issue 3, 2014, p. 10. The statement on the scoring of the NDII should instead read "A lower score indicates greater impairment.[41,46]" The authors had this reversed and we (manuscript reviewers and I) did not catch it. It was published as "A higher score..." and it should read "A lower score..."

Oncology Section Celebrate Living
American Physical Therapy Association

2015 ONCOLOGY SECTION SLATE OF CANDIDATES

Vice President
Amy Litterini, PT, DPT

Secretary
Mary Fisher, PT, PhD, CLT Emily Hemingway, PT, DPT, ATC

Nominating Committee Member
Nancy Paddison, PTA, BA, CLT-LANA Laura Sheridan, PT, DPT, CLT Kathy Tasillo, PT, DPT

Go to *www.oncologypt.org* to View Each Candidate's Biography Form

EDGE Task Force on Head and Neck Cancer Outcomes

A Systematic Review of Outcome Measures for Temporomandibular-related Dysfunction

Mary Lou Galantino, PT, MS, PhD, MSCE[1]; Melissa M. Eden, PT, DPT, OCS[2]; Bryan A. Spinelli, PT, MS, OCS, CLT-LANA[3]; Ann Marie Flores, PT, PhD, CLT[4]

[1]Professor of Physical Therapy, School of Health Sciences, Stockton University, Galloway, NJ
[2]Assistant Professor of Physical Therapy, Mayo Clinic, Phoenix, AZ
[3]Rhode Island Hospital, Rehabilitation Services, Providence, RI
[4]Assistant Professor & Director, Center for Cancer Survivorship Studies, Department of Physical Therapy, Movement and Rehabilitation Sciences, Northeastern University, Boston, MA

ABSTRACT

Background: Patients with head and neck cancer (HNC) often experience significant postoperative limitations in temporomandibular joint (TMJ) function, facial pain, reduced nutritional intake, speech impairments, and compromised activities of daily living due to side effects of cancer treatment. Physical therapists treating these individuals must use valid and reliable patient-reported outcome measures to quantify change related to physical therapy intervention for the TMJ. **Purpose**: As part of the activities of the Oncology Section EDGE Task Force on Head and Neck Cancer Outcomes, we report evidence-based recommendations for patient-reported outcome measures for individuals with HNC-related temporomandibular dysfunction (TMD). **Methods**: A systematic literature review of TMD-related patient-reported outcome measures that are clinically feasible and relevant to the HNC patient population was conducted using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) model. Recommendations are based on the quality of psychometric properties, clinical utility, and previous use in HNC-related research. Twenty-two outcome measures were selected for review; 4 received a rating of 3, "recommended" for use in the HNC population. **Conclusions:** A variety of outcome measures have been reported in the literature for individuals with HNC-related TMD. Four measures, the Graded Chronic Pain Scale, 8 and 20-item Jaw Functional Limitation Scale and TMD Pain Screener, are recommended for clinical use by the researchers on this task force although it is important to note psychometric properties specific to the HNC population are lacking.

Systematic Review Registration Number: PROSPERO CRD42013004898

Address correspondence to: Mary Lou Galantino, PT, MS, PhD, MSCE, School of Health Sciences, Stockton University, 101 Vera King Farris Drive, Galloway, NJ 08025 (MaryLou.Galantino@stockton.edu).

Key Words: temporomandibular, impairment, outcome measure, head and neck cancer

INTRODUCTION

The incidence of head and neck cancer (HNC) has remained stable over the past decade due to improvements in the medical management of the disease, and a 5-year relative survival rate of 60% to 65%.[1,2] Treatment for HNC may include a combination of surgery, radiation, and/or chemotherapy. As of 2012, there were 364,000 HNC survivors in the United States, many of whom are dealing with the long-term sequelae of HNC treatment.[1] The location and nature of the medical management of HNC places survivors at risk for experiencing temporomandibular joint (TMJ)-related impairments and disability, which can lead to temporomandibular joint dysfunction (TMD).[2-4]

Temporomandibular joint dysfunction is defined as "functional disturbances of the masticatory system."[5] Changes in TMJ, occlusion, ligamentous, or soft tissue structures, or the muscles of mastication eventually can result in TMD. The TMJ and masticatory system is complex and requires a comprehensive understanding of the anatomy and physiology of the arthrokinematics, musculoskeletal, and neurological components in order to treat TMD. Trismus, restricted mouth or jaw opening, is a form of TMD that affects 10% to 50% of those with HNC.[4,6,7]

Radiation therapy for nasopharyngeal, base of tongue, salivary gland, maxillary, or mandibular often affect the TMJ and associated musculature including the pterygoid and the masseter muscles.[4,8] Loss of TMJ function and range of motion from radiation therapy appear to be related to fibrotic changes to TMJ soft tissue, including the muscles of mastication.[8,9] The HNC postoperative precautions resulting in changes to movement of the TMJ may also compound emotional aspects and guarding of the TMJ causing a vicious cycle of reduced movement, pain, and altered function.[8] Trismus in patients with HNC can have serious health implications, including reduced nutrition due to impaired mastication, difficulty speaking, impaired self-image, ineffective oral hygiene, and compromised assessments for cancer surveillance.[6,8]

In persons who receive radiation to the head and neck, trismus is often observed in conjunction with dysphagia.[10]

A systematic review published in 2014 highlights the following 12 core domains and symptoms most affected in HNC survivors: swallowing, oral pain, skin changes, dry mouth, dental health, opening mouth/trismus, taste, excess/thick mucus/saliva, shoulder disability/motion, voice/hoarseness, and social and functional domains.[11] It has been reported that individuals with HNC have a lower quality of life (QOL) than age-matched controls and in many cases a lower QOL than individuals with other common cancers.[12] However, little research has been conducted studying TMD in people with HNC that fall within the scope of physical therapy practice. The few studies related to TMD in HNC have suggested that pain and restricted motion are problems experienced by some HNC survivors.[3,13-15]

Physical therapy assessment must include outcome measures of physical function, which may include patient-reported outcome measures (PROs). Despite the risk of TMJ impairments and disability in the setting of HNC, physical therapists have very few HNC-specific PROs from which to choose. In fact, a panel of HNC experts were unable to recommend a PRO specific to TMD or trismus in the "outcomes toolbox for head and neck cancer research," a series of PROs addressing 18 main areas of concern in the HNC population. The authors only offer an objective measurement for trismus—the interincisal distance.[16]

When selecting an outcome measure in practice, a physical therapist must first consider the ability to interpret the test score in the population of interest. This requires that the psychometric properties of the outcome measure, such as reliability, validity, sensitivity, specificity, and the ability to detect change, are available in a population similar to the one of interest. Clinical utility, the ease of use, and accessibility of needed resources to administer the measure, are also important factors to consider when choosing an outcome measure. The Evaluation Database to Guide Effectiveness (EDGE) initiative by the American Physical Therapy Association (APTA) is currently identifying core sets of recommended tests and measures to be used within certain patient populations and diagnoses.[17] The Oncology Section of the APTA has established task forces for breast, prostate, lung, urogynecologic, and head and neck cancers. The purpose of this systematic review is to identify and provide recommendations of clinically feasible and relevant PROs that address pain and function for use in patients with HNC presenting with TMD.

DATA & METHODS

This systematic review is registered on PROSPERO (CRD42013004898). The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) model was utilized for this systematic review (Figure 1, Appendix 1).[18] The appendices for this article can be accessed via ePub at www.oncologypt.org

Two investigators (MLG and MME) independently completed the literature search in April 2014 using Ovid Medline. PubMed, PEDro, EBSCO Host, PsycInfo, and Cochrane Databases were subsequently explored using similar search strategies and terms (Figure 2). The literature search was limited to human subjects, availability of an abstract for screening, and publication in the English language. Because little research on this subject is available, we did not exclude studies based on publication date or level of evidence. Once duplicates were removed, 1068 total articles remained for review (Figure 1).

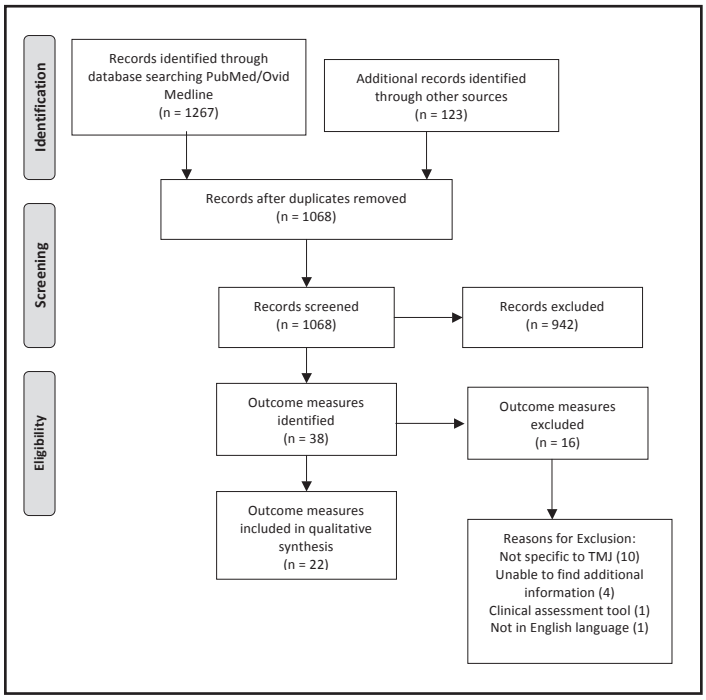


Figure 1. PRISMA flow diagram for systematic review.

Article titles and abstracts were reviewed, yielding 38 identified PROs. Of these, 22 were chosen for inclusion in the systematic review (Figure 1). To be included in the review, outcome measures had to be clinically feasible and patient-reported. Patient-reported outcome measures addressing health related QOL were excluded unless they were specifically developed for the HNC population. In addition, questionnaires assessing nonspecific impairment level constructs (ie, visual analog scale) and other joints/regions not specific to the TMJ region (ie, Dental Discomfort Questionnaire, 10-item cervical questionnaire, Sinonasal Assessment Questionnaire) were excluded. Outcomes specific to HNC that did not contain items related to TMJ function (ie, Cancer Problems in Living Scale, FACT H&N) were also excluded.

The 22 outcomes were divided between the two investigators for analysis. Relevant studies and full-text articles for each measure were retrieved to assess the clinical utility, psychometric properties, and relevance of the PROs to the HNC population. Clinical utility was based upon cost, ease of use and scoring, equipment requirements, and availability of normative data. Ease of use, based on responder burden, and ease of scoring were characterized as easy, moderate, or difficult. The primary reviewer completed the Head and Neck Cancer EDGE Task

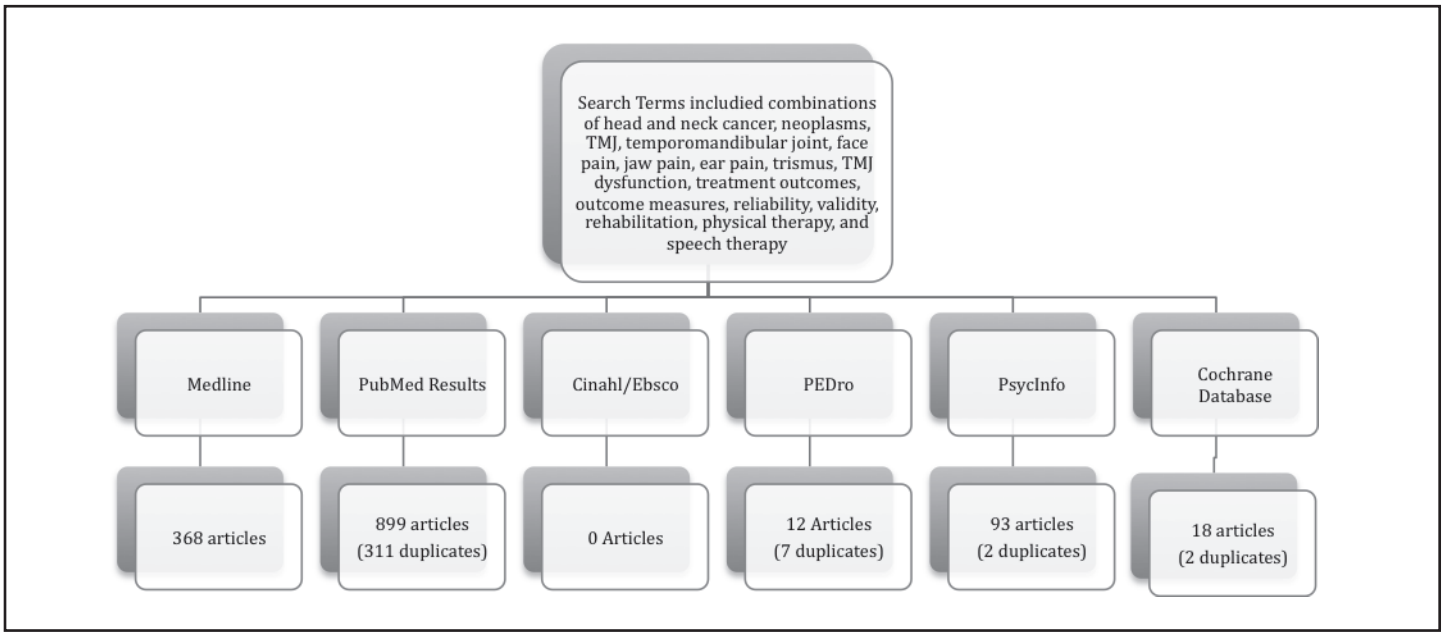


Figure 2. Terms for online database search. Literature search limits included the following: human subjects, English, full abstracts available.

Force Outcome Measure Rating Form (Appendix 2) for each of the assigned measures. Recommendations for each PRO were provided using a rating scale of 1 to 4, where a rating of 1 is not recommended and a rating of 4 is highly recommended (Table 1). The risk of bias at the study level was not considered when completing the Rating Form, however, we attempted to decrease the level of reporting errors through a review by the second investigator. In the case of disagreement between the two investigators, the rating in question would be brought to the entire HNC EDGE Task Force for consensus.

RESULTS

After a systematic review of the HNC literature, 22 TMD-related patient-reported outcome measures were identified for inclusion in this review. The HNC EDGE Task Force recommendations for TMD outcome measures can be found in Table 2.

"Highly Recommended" TMD Outcomes Measures

Of the 22 measures reviewed, no outcome measures merited the maximum score of 4 and the rating "Highly Recommended." The investigators were unable to find a PRO specific to TMD, which was previously utilized in the HNC research and demonstrates strong psychometric properties and clinical utility.

"Recommended" TMD Outcomes Measures

Four outcome measures received a score of 3, "Recommended." These outcome measures include the 8-item and 20-item Jaw Functional Limitation Scale (JFLS-8, JFLS-20), the TMD Pain Screener, and the Graded Chronic Pain Scale (GCPS). Although no published evidence of the psychometric properties in the HNC population were found, these measures have demonstrated good psychometric properties and clinical utility in other patient populations (Tables 3 and 4).

Table 1. Head and Neck Cancer Rating Scale

4	**Highly Recommend**	Highly recommended; the outcome has good psychometric properties and good clinical utility; the measure has been used in research on individuals with or post head and neck cancer.
3	**Recommend**	Recommended; the outcome measure has good psychometric properties and good clinical utility; no published evidence that the measure has been applied to research on individuals with or post head and neck cancer.
2A	**Unable to Recommend at This Time**	Unable to recommend at this time; there is insufficient information to support a recommendation of this outcome measure; the measure has been applied to research on individuals with or post head and neck cancer.
2B	**Unable to Recommend at This Time**	Unable to recommend at this time; there is insufficient information to support a recommendation of this outcome measure; no published evidence that the measure has been applied to research on individuals with or post head and neck cancer.
1	**Do Not Recommend**	Poor psychometric and/or poor clinical utility (time, equipment, cost, etc)

Table 2. Outcome Measures Sorted by Task Force Rating

Rating	Measure
3	Graded Chronic Pain Severity (GCPS)
3	Jaw Functional Limitation Scale (8-item version)
3	Jaw Functional Limitation Scale (20-item version)
3	TMD Pain Screener
2A	Late Effects Normal Tissue Task Force/Subjective, Objective, Management, Analytic (LENT/SOMA)
2A	European Organization of Research Treatment for Cancer Quality of Life Questionnaire-Head & Neck 35 (EORTC QLQ-HN35)
2A	Liverpool Oral Rehabilitation Questionnaire
2A	University of California San Francisco Oral Cancer Pain Questionnaire
2A	University of Washington QOL (UW-QOL)
2A	Gothenburg Trismus Questionnaire (GTQ)
2B	Manchester Orofacial Pain Disability Scale (MOPDS)
2B	Subjective Oral Health Status Indicators (SOHSI)
2B	TMJ Scale
2B	TMD questionnaire (LDF-TMDQ, 13 items), Jaw Function Scale
2B	10-item TMJ pain questionnaire
2B	Mandibular Function Impairment Questionnaire (MFIQ)
2B	Modified Symptom Severity Index (Mod-SSI)
2B	Tampa Scale for Kinesophobia
2B	TMD Self-efficacy Scale
2B	TMD Checklist
1	Conti Anamnestic Questionnaire
1	TMD Disability Index
1	TMJ Score Questionnaire
1	Oral Impacts on Daily Performances Inventory

Abbreviations: TMD, temporomandiublar dysfunction; TMJ temporomandibular joint; LDF-TMDQ, Limitations of Daily Function in Temporomandibular Dysfunction Questionnaire

The JFLS-20 is a 20-item measure derived through expert consensus and subsequently validated using Rasch methodologies in a sample of patients with TMD, primary Sjogren syndrome, burning mouth syndrome, skeletal malocclusion, and healthy controls.[19] It includes 3 subscales (mastication, vertical jaw mobility, and verbal and emotional expression) and a global scale of functional limitation of the jaw. The 8-item global scale of functional limitation of the jaw is known as the JFLS-8. The JFLS-8 and the JFLS-20 can both be appropriately utilized to measure TMD dysfunction in clinical and research settings.[19,20] Both scales require the patient to rank their level of difficulty on an 11-point Likert scale, where zero suggests "no limitation" and a response of 10 suggests "severe limitation." The recall period is one month. The total score is calculated by summing the individual item scores, with a higher score indicating greater limitation.[19] The JFLS-20 score can range from 0-200 and the JFLS-8 score can range from 0-80.

Internal consistency (Cronbach's alpha) for the JLFS-8 is 0.87 and for the JFLS is 0.95.[19] Test-retest reliability, assessed through temporal stability and the concordance correlation coefficient, (CCC rho), is 0.81 for the JFLS-8 and 0.87 for the JFLS-20.[19] An effect size for the JFLS-8 of 0.41 has been reported.[20] When analyzed separately, the 3 subscales of the JFLS-20 each exhibited excellent psychometric properties with respect to modeled variance, item fit, reliability, and internal consistency among the 5 diagnostic groups.[19] Construct and discriminant validity have been established through correlations with other measures, established between-group differences in the subscale and global functional limitation scores,[19] and test score independence from constructs such as depression, somatization, anxiety, and other clinical findings.[20] The JFLS-8 and the JFLS-20 are also highly correlated (0.9675).[19] Neither the JFLS-8 nor JFLS-20 has been used in published HNC research.

The TMD pain screener is a 3-item PRO with a recall period of the last 30 days. The first item addresses the frequency of pain ("no pain," "pain comes and goes," "pain is always present"). A response of "no pain" yields a score of zero, "pain comes and goes" yields one point, and "pain is always presents" yields two points. The second item addresses pain or stiffness upon awakening, and the third item is a 4-part question assessing a change in pain level for certain functional activities including chewing, opening/moving the jaw, grinding/clenching, and kissing/talking/yawning. Item two and the 4 components of item three have dichotomous response options of "no" and "yes." A "no" is awarded zero points and a "yes" is awarded one point. Item scores are then summed, with a test score of zero to 7 points possible. A higher score indicates greater limitation. In a population of subjects with pain-related TMD, internal consistency was deemed excellent (Cronbach's α= 0.87), and temporal stability of individual items (κ) ranged from 0.52 to 0.78, indicating fair to excellent agreement. The summary test score exhibited an intraclass correlation coefficient of 0.83. The tool was found to exhibit sensitivity of 99% and a specificity of 97%. The positive likelihood ratio ranged from 19.2 to 44.6, varying according to the comparison control group, whereas the negative likelihood ratio was 0.01 throughout. Both positive and negative likelihood ratio findings exceeded the accepted benchmarks of 10 or more and 0.1 or less for the positive and negative likelihood ratios, respectively. The TMD Pain Screener has not been used in HNC research.[21]

Table 3. Clinical Utility for Recommended Outcome Measures

Measure	Equipment Needed	Cost	Ease of Use	Scoring	Normative Data
JFLS-8/ JFLS-20	Outcome measure, pen	Free	Easy	Easy	Not available
TMD Pain Screener	Outcome measure, pen	Free	Easy	Easy	Not available
Graded Chronic Pain Scale	Outcome measure, pen	Free	Easy	Moderate	Not available

Abbreviations: JFLS, Jaw Functional Limitation Scale; TMD, temporomandibular dysfunction

Table 4. Psychometric Properties for Recommended Outcome Measures

Outcome Measure	Psychometric Properties
Graded Chronic Pain Severity (GCPS)	Internal consistency (Cronbach's alpha): Characteristic Pain Intensity Measure α = 0.84 Interference Measure α = 0.95 Temporal stability: Characteristic Pain Intensity CCC = 0.91 Interference Measure CCC = 0.89 Convergent and discriminant validity of characteristic pain intensity was supported by a substantial association with MPI pain severity (CCC = 0.65) and smaller associations with measures of constructs other than pain (eg, depression [CES-D], somatic symptoms). Such validity was supported for interference by substantial associations with the MPI interference (CCC = 0.52) and dysfunctional measures (CCC = 0.51), and by smaller associations with measures of other constructs (eg, the SF-12, CES-D).[23] Sensitivity to change not tested.
TMD Pain Screener	Internal reliability α= 0.87 Temporal stability of the individual items (κ) ranged from 0.52 to 0.78, indicating fair to excellent agreement.[22] Summary score ICC= 0.83 Validity for TMJ-related pain: Sensitivity= 99% Specificity= 97% Researchers deemed the content validity to be excellent. Sensitivity to change not tested.
Jaw Functional Limitation Scale (JFLS versions 8 and 20)	Construct validity was established via low correlations with depression, anxiety, somatization, pain interference, pain-free opening, and palpation sensitivity, and via moderate correlations with pain and jaw symptoms.[20] Internal consistency (Cronbach's alpha): JFLS-8 - α = 0.87 JFLS-20 - α = 0.95 Temporal stability[19]: JFLS-8 CCC= 0.81 JFLS-20 CCC= 0.87 Sensitivity to change not tested.

Abbreviations: CCC, concordance correlation coefficient; MPI, Multidimensional Pain Inventory; CES-D, Center for Epidemiological Studies-Depression Scale; SF-12, Short Form-12, TMD, temporomandibular dysfunction; TMJ, temporomandibular joint

The GCPS (version 2.0) is an 8-item PRO. In general, symptom recall is for the last 30 days; however, one item has a 24-hour recall and another item has a 6-month recall period. Six of the items are scored on an 11-point Likert scale, with a response of zero indicating no interference or pain and a response of 10 indicating "pain as bad as could be" or "unable to carry on any activities." Item one requires the patient to report the number of days the patient experienced facial pain in the past 6 months, and item 5 requires the patient to report the number of days facial pain limited performance of usual activities in the past 30 days. Scores are calculated for 3 subscales: the characteristic pain intensity score, which ranges from 0–100, is calculated as the mean intensity ratings for reported current, worst, and average pain; the disability score, which ranges from 0–100, is calculated as the mean rating for difficulty performing daily, social, and work activities; and the disability points score, which ranges from 0–3, is derived from a combination of ranked categories of number of disability days and disability score. Subscale scores for pain intensity and disability are combined to calculate a chronic pain grade that enables classification of chronic pain patients into 5 hierarchical categories: grades 0 (no pain) to IV (high disability-severely limiting).[22]

The GCPS is a reliable and valid instrument that assesses the constructs of pain intensity and pain-related disability. The measure's psychometric properties were studied in a population of patients with back pain, headache, and TMJ disorders. The internal consistency was measured with Cronbach's α = 0.84 for the pain intensity subscale and Cronbach's α = 0.95 for the

pain-related disability subscale. The temporal stability (2–7 days) was high for pain intensity (CCC = 0.91), pain-related disability (CCC = 0.89), and chronic pain grade (weighted kappa = 0.87).[22] Among patients with moderate to severe chronic musculoskeletal pain, the GCPS has been shown to be modestly responsive to changes after 12 months of treatment with an efficacious pain intervention, with standardized moderate effect sizes for the intensity and disability sub-scales of 0.41 and 0.43, respectively. Among participants with chronic knee or hip pain, the standardized effect size for the GCPS intensity was 0.32.[23] The GCPS has not been used in HNC research.

"Unable to Recommend at This Time" TMD Outcomes Measures

Sixteen outcome measures received a score of 2 and are not recommended at this time for use by physical therapists treating individuals with TMD in the setting of HNC. A score of 2 is further categorized as 2A if the measure has been used in the HNC literature or 2B if there is no published evidence that the tool has been used in HNC-related research. Six PROs were designated a rating of 2A, including the EORTC QLQ H&N43,[24] University of Washington Quality of Life (UW-QOL),[25,26] Liverpool Oral Rehabilitation Questionnaire version 3 (LORQ),[27] University of California San Francisco (UCSF) Oral Cancer Pain Questionnaire,[28] the Gothenburg Trismus Questionnaire (GTQ),[29] and the Late Effects in Normal Tissues Subjective, Objective, Management and Analytic Scales (LENT/SOMA).[30]

Two HNC-specific QOL measures were included in the review, the EORTC QLQ H&N 43 and the UW-QOL. The EORTC QLQ H&N43 is the most recent revision of the EORTC QLQ-H&N 35. It currently lacks psychometric studies to support its use. More importantly, similar to the UW-QOL, the EORTC QLQ-H&N43 is a measure designed to quantify health-related QOL and therefore provides a test score interpreted as an overall QOL score for many constructs, not a score specific to TMD. As a result, these two HNC-specific QOL measures cannot be recommended solely for quantification of TMD dysfunction.

The LORQ is a 40-item measure developed through expert opinion. It addresses issues related to oral function, oro-facial appearance, social interaction, and dentures/dental prostheses. While the tool is relevant to the HNC population, it lacks strong psychometric properties and easy accessibility to scoring and interpretation. In addition, over half of the questions pertain to oral prosthetics, therefore decreasing its utility for a large portion of the HNC population.[27] The UCSF Oral Cancer Pain Questionnaire is an 8-item PRO used to quantify patients' pain levels before and after surgical resection for oral cancer. Items one, 3, and 5 evaluate the intensity, sharpness and throbbing nature of pain when the patient is not engaged in oral function. Items 2, 4, and 6 measure the intensity, sharpness, and throbbing nature of pain during oral function (talking, eating, and drinking).[28] Although some of these items are related to TMJ function, this tool is not specific to TMD and has not been validated for patients with TMD. In addition, the measure is not easily accessible to clinicians. The GTQ is the only measure reviewed specifically designed to measure trismus.[29] Unfortunately, the tool is not easily accessible and lacks supporting research. The LENT/SOMA scale addresses radiotherapy toxicities in patients with HNC and is therefore not relevant for use by physical therapists.[30]

Ten PROs reviewed received a 2B rating, including: Subjective Oral Health Status Indicators (SOHSI),[31,32] 10-item TMJ Pain Questionnaire,[33] Mandibular Function Impairment Questionnaire (MFIQ),[34] Modified Symptom Severity Index (SSI),[35] TMD Checklist,[36] TMJ Scale,[19] TMD Questionnaire (LDF-TMDQ), which is also known as the Jaw Function Scale,[37] the Manchester Orofacial Pain Disability Scale (MOPDS),[38] Tampa Scale for Kinesiophobia,[39] and the TMD Self-efficacy Scale.[40]

Although the SOHSI, 10-item TMJ Pain Questionnaire, MFIQ, Modified SSI, and TMD Checklist may appropriately measure the construct of TMD dysfunction, the questionnaires and information on scoring and interpretation are not easily accessible by clinicians and therefore cannot be recommended. The TMJ Scale is, however, accessible to clinicians, but at a financial cost. It is also 97-items in length and therefore not feasible for use in a busy clinical setting.[19] Our review also found that the LDF-TMDQ and the MOPDS have been developed in other countries and may lack validity in the United States. Two PROs, Tampa Scale for Kinesiophobia, and TMD Self-Efficacy Scale, may be of benefit for use in a holistic assessment of TMD, however, do not specifically address the construct of TMD and therefore cannot be recommended at this time.

"Not Recommended" TMD Outcome Measures

Four outcome measures were given the lowest score of 1 and are not recommended secondary to limited to no available research supporting the psychometric properties of the tools across all patient populations. These outcome measures include the Conti Anamnestic Questionnaire,[41] the Oral Impacts on Daily Performance,[42,43] TMD Disability Index,[44] and the TMJ Score Questionnaire.[45]

DISCUSSION

The APTA's Oncology Section Head and Neck EDGE Task Force recommends that physical therapists use the GCPS, JFLS-8 or JFLS-20, and/or the TMD Pain Screener to quantify TMD in patients with HNC. These measures merited a score of 3, "recommended," because they demonstrate good psychometric properties, but have not been tested or utilized in HNC literature. Each of the recommended measures can be found in the Diagnostic Criteria for Temporomandibular Disorders (DC/TMD), which is a compilation of clinical and patient-reported measures designed to assess TMD in clinical and research settings.[46] The DC/TMD has two components, Axis I and Axis II. Axis I of the DC/TMD includes the TMD Pain Screener, a symptom questionnaire, demographics, and a clinical examination. Axis II includes various PROs addressing constructs related to QOL, anxiety, depression, and TMD, including the JLFS-8, JLFS-20, and GCPS.[47]

Only 6 of the 22 measures reviewed in this study have been used in the HNC literature, all of which were not recommended at this time due to concerns regarding tool accessibility, clinical utility, test score interpretation specifically for TMD dysfunction, and availability of psychometric data to merit their use. Given the lack of PROs related to TMJ dysfunction designed for the HNC population, physical therapists are left to use TMJ-related outcome measures that have been designed for other patient populations whose needs may be quite different than the particular concerns of a patient with HNC. Additional research is therefore required to determine the psychometric properties of the measures recommended by the HNC EDGE Task Force in the HNC population, and to further study relevant PROs that could not be recommended by the Task Force at this time. For example, the TMD Disability Index is similar in format to the Oswestry Disability Index and Neck Disability Index, both frequently used PROs in the physical therapy profession for back and neck pain, and has been used in physical therapy research.[48-51] Despite its use, the TMD Disability Index received a rating of "not recommended" because psychometric properties for the measure have not been reported.

A systematic review of the literature to identify publications utilizing PROs suitable for oral and maxillofacial surgery found there are numerous questionnaires available, however selecting the most appropriate one can be difficult.[52] Ojo and colleagues[53] explored HNC QOL instruments and determined that cross-study comparisons have been hampered by the heterogeneity of measures used in research and the fact that reviews of HNC instruments have not been comprehensively studied. Given the volume and heterogeneity of PROs there is no gold standard HNC-specific TMD questionnaire.[53]

Similar to the Oncology Section's HNC EDGE Task Force's reviews on neck and shoulder PROs and measurement of external lymphedema for HNC survivors, [54,55,58] the obvious limitation in this review, was that very few of the measures reviewed have actually been utilized to evaluate TMJ-related impairment and disability in the HNC patient population. Referral to physical therapy for HNC related-TMD is underutilized.[56] In addition, there are a limited number of investigators with expertise in TMJ-related impairments and function conducting research in the HNC population. Prolonged survivorship due to improved medical management of HNC increases the importance of restoring optimal function in order to improve QOL and function.[56] Given the rise in HPV-associated HNCs, improvements in early detection, prevention, and medical management strategies, patients with HNC may have unmet physical therapy needs to reduce or prevent TMJ disability.

To our knowledge, this is the first attempt at providing recommendations for TMD-specific PRO utilization by physical therapists treating patients with HNC. Future research should consider the psychometric properties of PROs, including establishment of normative values and responsiveness to change, addressing TMD in the HNC population. Although great strides have been made in the assessment of PROs and QOL in HNC, more work is needed to improve the clinical utility of these measures in order to link research to clinical practice.[9,53,57]

CONCLUSION

The HNC EDGE Task Force recommendations provide physical therapists with evidence-based outcomes measure recommendations for use with patients with HNC presenting with TMD. The HNC EDGE Task Force recommends 4 measures that are easy to use in the clinic with good psychometric properties. These measures fulfill Medicare functional outcome reporting requirements and can provide reliable and valid data from which to quantify patient response to physical therapy interventions.

Our review reveals a gap in psychometric research evaluating properties of available TMJ-related PROs applied in the HNC population. The HNC EDGE Task Force reviews include recommendations for other areas that affect HNC patients that fall under the scope of practice for physical therapy such as neck,[54] shoulder,[55] and lymphedema.[58] Further research is needed to evaluate psychometric properties of PROs for use in the HNC patient population. Studies evaluating HNC-specific PROs' responsiveness to change, generation of normative values, and reliability and construct validity in the HNC patient population are needed. The HNC EDGE Task Force recommendations are a first-step to fill the gap in knowledge of useful, relevant, and patient friendly TMJ-related outcome measures for the HNC patient population.

ACKNOWLEDGMENTS

No funding was available for this project. All Task Force member efforts were on a volunteer basis. We would like to thank Kelsey Watkins, Student Physical Therapist (Class of 2015) from Stockton University, for her research assistance in this manuscript, and Eliane Purchase, Medical Librarian at the Mayo Clinic in Arizona, for her assistance in developing the research search strategy.

Please note: The appendices for this article can be viewed online at: http://www.ocnologypt.org/publications/rehabilitation-oncology-journal

REFERENCES

1. National Cancer Institute. SEER Cancer Statistics Review, 1975-2010. http://seer.cancer.gov/csr/1975_2010/. Accessed based on November 2012 SEER data submission, posted to the SEER web site, April 2013.
2. Shone GR, Yardley MP. An audit into the incidence of handicap after unilateral radical neck dissection. *J Laryngol Otol.* 1991;105(9):760-762.
3. Oz B, Memis A. Development of musculoskeletal complaints and functional disabilities in patients with laryngeal carcinoma after neck dissection sparing spinal accessory nerve. *Eur J Cancer Care (Engl).* 2009;18(2):179-183.
4. Bensadoun RJ, Riesenbeck D, Lockhart PB, Elting LS, Spijkervet FK, Brennan MT. A systematic review of tris-

mus induced by cancer therapies in head and neck cancer patients. *Support Care Cancer.* 2010;18(8):1033-1038.
5. Okeson J. Etiology of functional disturbances in the masticatory system. *Management of Temporomandibular Disorders and Occlusion*. 6th ed. St Louis, MO: Mosby Year Book Publication, Inc; 2008:130-163.
6. Weber C, Dommerich S, Pau HW, Kramp B. Limited mouth opening after primary therapy of head and neck cancer. J *Oral Maxillofac Surg.* 2010;14(3):169-173.
7. Louise Kent M, Brennan MT, Noll JL, et al. Radiation-induced trismus in head and neck cancer patients. *Support Care Cancer.* 2008;16(3):305-309.
8. National Cancer Institute. Late Complications of Head and Neck Radiation. 2014; http://www.cancer.gov/cancertopics/pdq/supportivecare/oralcomplications/HealthProfessional/page16 - Section_675. Accessed December 10, 2014.
9. Cousins N, MacAulay F, Lang H, MacGillivray S, Wells M. A systematic review of interventions for eating and drinking problems following treatment for head and neck cancer suggests a need to look beyond swallowing and trismus. *Oral Oncol.* 2013;49(5):387-400.
10. Pauli N, Johnson J, Finizia C, Andrell P. The incidence of trismus and long-term impact on health-related quality of life in patients with head and neck cancer. *Acta Oncol.* 2013;52(6):1137-1145.
11. Chera BS, Eisbruch A, Murphy BA, et al. Recommended patient-reported core set of symptoms to measure in head and neck cancer treatment trials. *J Natl Cancer Inst.* 2014;106(7).
12. Tschiesner U, Rogers S, Dietz A, Yueh B, Cieza A. Development of ICF core sets for head and neck cancer. *Head Neck.* 2010;32(2):210-220.
13. van Wilgen CP, Dijkstra PU, van der Laan BF, Plukker JT, Roodenburg JL. Morbidity of the neck after head and neck cancer therapy. *Head Neck.* 2004;26(9):785-791.
14. Speksnijder CM, van der Bilt A, Slappendel M, de Wijer A, Merkx MA, Koole R. Neck and shoulder function in patients treated for oral malignancies: a 1-year prospective cohort study. *Head Neck.* 2013;35(9):1303-1313.
15. Ahlberg A, Nikolaidis P, Engstrom T, et al. Morbidity of supraomohyoidal and modified radical neck dissection combined with radiotherapy for head and neck cancer: a prospective longitudinal study. *Head Neck.* 2012;34(1):66-72.
16. Ringash J, Bernstein L, Cella D, et al. Outcomes toolbox for head and neck cancer research. *Head Neck.* 2015;37:425-439.
17. Field-Fote E. Towards Optimal Practice: What can we gain from assessment of patient progress with standardized outcome measures? http://www.ptresearch.org/article/14/edge-taskforce. March 24, 2014.
18. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med.* 2009;151(4):264-269, w264.
19. Ohrbach R, Larsson P, List T. The jaw functional limitation scale: development, reliability, and validity of 8-item and 20-item versions. *J Orofacial Pain.* 2008;22(3):219-230.
20. Ohrbach R, Granger C, List T, Dworkin S. Preliminary development and validation of the Jaw Functional Limitation Scale. *Community Dent Oral Epidemiol.* 2008;36(3):228-236.
21. Gonzalez YM, Schiffman E, Gordon SM, et al. Development of a brief and effective temporomandibular disorder pain screening questionnaire: reliability and validity. *J Am Dent Assoc.* 2011;142(10):1183-1191.
22. Ohrbach R, Turner JA, Sherman JJ, et al. The Research Diagnostic Criteria for Temporomandibular Disorders. IV: evaluation of psychometric properties of the Axis II measures. *J Orofacial Pain.* 2010;24(1):48-62.
23. Elliott AM, Smith BH, Hannaford PC, Smith WC, Chambers WA. Assessing change in chronic pain severity: the chronic pain grade compared with retrospective perceptions *Br J Gen Pract.* 2002;52:269-274.
24. EORTC Quality of Life. EORTC QOL Module for Head and Neck Cancer (QLQ-H&N43)(Revision of EORTC QLQ-H&N35). http://groups.eortc.be/qol/eortc-qol-module-head-and-neck-cancer-qlq-hn43-revision-eortc-qlq-hn35. Accessed August 5, 2014.
25. Merseyside Regional Head and Neck Cancer Centre. UW-QOL v4 Questionnaire - translated versions. http://www.headandneckcancer.co.uk/For+professionals/Quality+of+Life+(QOL)/UW-QOLv4+Translations.aspx. Accessed August 2, 2014.
26. Lowe D, Rogers SN. University of Washington Quality of Life Questionniare (UW-QOL v4): Guidance for scoring and presentation. 5/25/2012; http://www.headandneckcancer.co.uk/File.ashx?id=10285. Accessed August 2, 2014.
27. Pace-Balzan A, Cawood JI, Howell R, Butterworth CJ, Lowe D, Rogers SN. The further development and validation of the Liverpool Oral Rehabilitation Questionnaire: a cross-sectional survey of patients attending for oral rehabilitation and general dental practice. *Int J Oral Maxillofac Surg.* 2006;35(1):72-78.
28. Kolokythas A, Connelly ST, Schmidt BL. Validation of the University of California San Francisco Oral Cancer Pain Questionnaire. *J Pain.* 2007;8(12):950-953.
29. Johnson J, Carlsson S, Johansson M, et al. Development and validation of the Gothenburg Trismus Questionnaire (GTQ). *Oral Oncol.* 2012;48(8):730-736.
30. The Christie NHS Foundation Trust. CTCAE/LENT SOMA questionnaires. 2014; http://www.christie.nhs.uk/the-foundation-trust/treatments-and-clinical-services/clinical-oncology-scoring-treatment-effects/ctcaelent-soma-questionnaires.aspx. Accessed August 3, 2014.
31. Locker D. Subjective Oral Health Status Indicators. In: Slade G, ed. *Measuring Oral Health and Quality of Life*. Chapel Hill: University of North Carolina, Dental Ecology; 1997:105-112.

32. Locker D, Miller Y. Evaluation of Subjective Oral Health Status Indicators. *J Public Health Dent.* 1994;54(3):167-176.
33. Braun B, Schiffman EL. The validity and predictive value of four assessment instruments for evaluation of the cervical and stomatognathic systems. *J Craniomandib Disord.* 1991;5(4):239-244.
34. Stegenga B, de Bont LG, de Leeuw R, Boering G. Assessment of mandibular function impairment associated with temporomandibular joint osteoarthrosis and internal derangement. *J Orofac Pain.* 1993;7(2):183-195.
35. Nixdorf DR, John MT, Wall MM, Fricton JR, Schiffman EL. Psychometric properties of the modified Symptom Severity Index (SSI). *J Oral Rehabil.* 2010;37(1):11-20.
36. Zhao NN, Evans RW, Byth K, Murray GM, Peck CC. Development and validation of a screening checklist for temporomandibular disorders. *J Orofac Pain.* 2011;25(3):210-222.
37. Sugisaki M, Kino K, Yoshida N, Ishikawa T, Amagasa T, Haketa T. Development of a new questionnaire to assess pain-related limitations of daily functions in Japanese patients with temporomandibular disorders. *Community Dent Oral Epidemiol.* 2005;33(5):384-395.
38. Aggarwal V, Zakrewska J, Macfarlane G, Macfarlane TV. Development and validation of the Manchester orofacial pain disability scale. *Community Dent Oral Epidemiol.* 2005;33:141-149.
39. Visscher CM, Ohrbach R, van Wijk AJ, Wilkosz M, Naeije M. The Tampa Scale for Kinesiophobia for Temporomandibular Disorders (TSK-TMD). *Pain.* 2010;150(3):492-500.
40. Brister H, Turner JA, Aaron LA, Mancl L. Self-efficacy is associated with pain, functioning, and coping in patients with chronic temporomandibular disorder pain. *J Orofac Pain.* 2006;20(2):115-124.
41. Valle-Corotti K, Pinzan A, do Valle CV, Nahas AC, Corotti MV. Assessment of temporomandibular disorder and occlusion in treated class III malocclusion patients. *J Appl Oral Sci.* 2007;15(2):110-114.
42. Slade GD. Derivation and validation of a short-form oral health impact profile. *Community Dent Oral Epidemiol.* 1997;25(4):284-290.
43. Slade GD, Spencer AJ. Development and evaluation of the Oral Health Impact Profile. *Community Dent Health.* 1994;11(1):3-11.
44. Steigerwald D. The Steigerwald/Maher TMD Disability Questionnaire. *Today's Chiropractic.* 1997;26(4):86-91.
45. Ness GM, Laskin DM. Global doctor opinion versus a patient questionnaire for the outcome assessment of treated temporomandibular disorder patients. *J Oral Maxillofac Surg.* 2012;70(7):1531-1533.
46. International RDC-TMD Consortium. Diagnostic Criteria for Temporomandibular Disorders (2014). http://www.rdc-tmdinternational.org/TMDAssessmentDiagnosis/DCTMD.aspx. Accessed December 28, 2014.
47. International RDC/TMD Consortium Network. Diagnostic Criteria for Temporomandibular Disorders: Clinical Protocol and Assessment Instruments. 3Dec2014; http://www.rdc-tmdinternational.org/LinkClick.aspx?fileticket=h8mJRFXb70k%3d&tabid=1210&portalid=18&mid=3296. Accessed December 28, 2014.
48. Gonzalez-Iglesias J, Cleland JA, Neto F, Hall T, Fernandez-de-las-Penas C. Mobilization with movement, thoracic spine manipulation, and dry needling for the management of temporomandibular disorder: a prospective case series. *Physiother Theory Pract.* 2013;29(8):586-595.
49. Furto ES, Cleland JA, Whitman JM, Olson KA. Manual physical therapy interventions and exercise for patients with temporomandibular disorders. *Cranio.* 2006;24(4):283-291.
50. Cleland J, Palmer J. Effectiveness of manual physical therapy, therapeutic exercise, and patient education on bilateral disc displacement without reduction- of the temporomandibular joint: a single-case design. *J Orthop Sports Phys Ther.* 2004;34(9):535-548.
51. Khuman R, Chavda D, Surbala L, Chanundhary E, Bhatt U, Nambi G. Physical therapy in temporomandibular dysfunction following maxillo-mandibular fixation in sub-condylar mandibular fracture-a single case study. *Int J Health Sci Res.* 2013;3(9):45-55.
52. Kanatas AN, Rogers SN. A systematic review of patient self-completed questionnaires suitable for oral and maxillofacial surgery. *Br J Oral Maxillofac Surg.* 2010;48(8):579-590.
53. Ojo B, Genden EM, Teng MS, Milbury K, Misiukiewicz KJ, Badr H. A systematic review of head and neck cancer quality of life assessment instruments. *Oral Oncol.* 2012;48(10):923-937.
54. Spinelli BA, Galantino ML, Eden MM, Flores AM. Recommendations for patient-reported outcome measures for head and neck cancer-related neck dysfunction: a systematic review. *Rehabil Oncol.* 2013;32(3):20-31.
55. Eden MM, Flores AM, Galantino ML, Spinelli BA. Recommendations for patient-reported outcome measures for head and neck cancer-related shoulder dysfunction: a systematic review. *Rehabil Oncol.* 2014;32(3):6-19.
56. Guru K, Manoor UK, Supe SS. A comprehensive review of head and neck cancer rehabilitation: physical therapy perspectives. *Indian J Palliat Care.* 2012;18(2):87-97.
57. Djan R, Penington A. A systematic review of questionnaires to measure the impact of appearance on quality of life for head and neck cancer patients. *J Plast Reconstr Aesthet Surg.* 2013;66(5):647-659.
58. Flores AM, Spinelli BA, Eden MM, Galantino ML. EDGE Task Force on Head and Neck Cancer Outcomes: a systematic review of outcome measures for quantifying external lymphedema. *Rehabil Oncol.* 2015;33(2):15-23.

EDGE Task Force on Head and Neck Cancer Outcomes

A Systematic Review of Outcome Measures for Quantifying External Lymphedema

Ann Marie Flores, PT, PhD, CLT[1] ; Bryan A. Spinelli, PT, MS, OCS, CLT-LANA[2] ; Melissa M. Eden, PT, DPT, OCS[3]; Mary Lou Galantino, PT, MS, PhD, MSCE[4]

[1]Assistant Professor & Director, Center for Cancer Survivorship Studies, Department of Physical Therapy, Movement and Rehabilitation Sciences, Northeastern University, Boston
[2]Rhode Island Hospital, Rehabilitation Services, Providence, RI
[3]Assistant Professor of Physical Therapy, Mayo Clinic, Phoenix, AZ
[4]Professor of Physical Therapy, School of Health Sciences, Stockton University, Galloway, NJ

ABSTRACT

Background: Survivors of head and neck cancer (HNC) and its treatment experience high rates of lymphedema. Unlike the extremities, the head and neck is difficult to measure and does not easily lend itself to having a contralateral side for comparison. Being an irregularly shaped part of the body, measures of edema for the extremities cannot be adapted for the head and neck. The need exists for outcome measures to objectively quantify head and neck lymphedema using evidence-based practice guidelines. **Purpose:** The purpose of this study is to identify and recommend external edema outcome measures for lymphedema in the HNC population. **Methods:** A systematic review of the literature on edema measures for use in the HNC patient population was conducted. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) model was used to guide which articles were chosen for inclusion, determination of eligibility, screening, and identification for the final review. Recommendations are based on the quality of psychometric properties, clinical feasibility, and previous use in HNC-related research. Six edema outcome measures were reviewed and none are recommended at this time; however, several hold great promise for future use in the clinic. **Conclusions:** This systematic review provides an overview for physical therapists on measures of external edema for the HNC patient population. The edema measures included in this review have been tested on HNC patients but have not been rigorously tested due to their novelty. At this time, no outcome measures for objectively quantifying external edema for the HNC population can be recommended. There is need for more research on this topic prior to providing definitive recommendations.

Address correspondence to: Ann Marie Flores, PT, PhD, CLT, Dept of Physical Therapy, Movement and Rehabilitation Sciences, 301 Robinson Hall, Northeastern University, Boston, MA 02115 (am.flores@neu.edu).

Systematic Review Registration Number: PROSPERO CRD42013004898

Key Words: lymphedema, head and neck cancers, measurement

INTRODUCTION

Cancers of the head and neck (HNC) involve the oral cavity, pharynx, larynx, paranasal sinuses, nasal cavity, salivary glands, upper esophagus, face, associated musculature, and/or bone and some thyroid cancers.[1] Head and neck cancer represents 3% of all cancer survivors in the United States and is considered an understudied tumor type.[2] The 5-year relative survival rate of these cancers widely varies—94% (lip cancer), 74% (salivary gland), and 30% (hypopharynx, and "other" cancers of the oral cavity and pharynx)[2] with approximately 291,108 people living with oral cavity and pharyngeal cancers in the United States in 2012[3]– likely an underestimate since laryngeal, thyroid, and upper esophageal cancers are also head and neck cancers.[3,4] Despite the rarity of HNC, disability and other side effects are highly prevalent. Lymphedema is known to affect the majority of HNC survivors (60-80%) affecting external structures (eg, outward and obvious swelling of the neck, head, and face), internal structures (eg, tongue, oral cavity, pharynx, hypopharynx, larynx, trachea, and esophagus), or a combination of both.[5,6] Head and neck cancer survivors usually have one or multiple treatment modalities. For example, tumor and lymph node resection and/or radiation, chemotherapy, or combined chemotherapy and radiation (CCR).[7,8] These treatments can be a grueling experience with serious side effects (eg, mucositis, cachexia, fatigue, oral and non-oral pain) that may require feeding tube placement.[9]

Location of the tumor, tumor and lymph node resection, radiation, and multiple modality treatment are all associated with lymphedema of the head and neck.[5] Other posttreatment factors such as muscle guarding, abnormal posture, and reduced movement in the affected areas may also contribute to lymphedema among HNC survivors.[10] Lymphedema, left untreated, can have serious consequences resulting in risk for recurrent infection such

as cellulitis, and interference with the ability to breathe, swallow, and/or speak—key functions of life.[6]

The 5-year relative survival rate from HNC, particularly for those with human papillomavirus (HPV) positive tumors, is improving. Survival rates (relative and absolute) are high. For example, survival rates for oral cavity and pharyngeal tumors are 62.7% and for laryngeal cancers, 60%. With increased survival, it stands to reason that more HNC survivors will be living with long-term cancer-related lymphedema.[6,12] Head and neck cancer survivors diagnosed with lymphedema are commonly referred to physical therapy for lymphedema management.[13,14] Complete assessment of lymphedema involves patient subjective complaints, physical examination (ie, pitting, tissue texture), and objective tests and measures.[15-17] Reliable and valid measures are necessary for clinicians to accurately assess and document patient progress and guide clinical decision-making.

This systematic review is part of the Evaluation Database to Guide Effectiveness (EDGE) activities of the American Physical Therapy Association Oncology Section's Head and Neck Cancer Task Force. The purpose of this systematic review is to identify and provide recommendations of clinically feasible and relevant objective tests and measures of edema for HNC survivors with cancer-related lymphedema.

DATA & METHODS

The authors registered this systematic review on PROSPERO (CRD42013004898). The search was limited only to studies that report on measurement of edema due to cancer treatment-related lymphedema of the head and neck. Because little research exists on concerns of HNC survivors that falls within the scope of physical therapy practice, we placed no limits on publication dates or level of evidence. We conducted a search of literature databases (PubMed, PEDro, EBSCO Host, Medline, PsycInfo, and Cochrane Database) using the following key words in all combinations: edema, human, lymphedema, cancer, head, neck, radiation, radiotherapy, surgery, neck dissection, face, measurement, outcome, outcome measure, radiation fibrosis, lymphostasis, inflammation, reliability, and validity. Articles were limited to studies on humans and published in English. Duplicates were removed and abstracts for all citations were screened to determine if they were to be included in the final review. Figure 1 illustrates the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) process used for the literature search.

Two investigators (AMF and BAS) independently completed a search of the literature using the databases and personal files of published articles not included in the literature databases as mentioned earlier in June 2014. This provided verification of the search results. The investigators then independently reviewed the abstracts of the search results and eliminated non-relevant articles based on the predetermined inclusion criteria (English language, any date of publication, any level of evidence) and exclusion criteria (no requirement of a physician prescription or performance of a medical procedure for measuring edema, non-human sample, quality of life study). The measures of edema identified by the authors were divided between reviewers, who independently retrieved relevant studies and full-text articles for each measure to assess clinical feasibility, psychometric properties, and relevance to the HNC patient population. The primary reviewer completed the HNC Task Force Outcome Measure Rating Scale Form (Table 1) for each assigned measure. The risk of bias of measure assessment was addressed by ensuring a secondary review for all measures. In the case of a discrepancy in rating between reviewers, two other investigators were available to evaluate the reviews and determine the decision outcome, along with the other two investigators, until total agreement was reached. Outcome measure recommendations were agreed upon using the EDGE criteria resulting in an ordinal score with anchors at 1 (measure is not recommended) and 4 (measure is highly recommended) (see Table 1).[18]

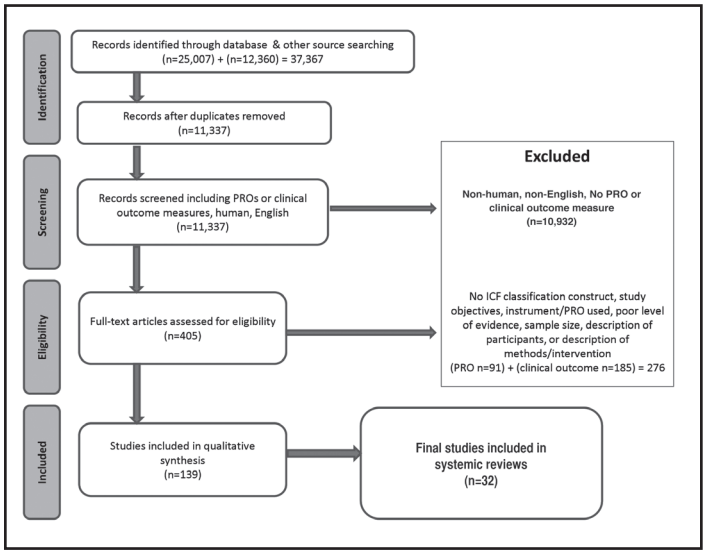


Figure 1. PRISMA model.

RESULTS

Using our keywords, we identified 25,007 records. All duplicates were removed which resulted in 11,337 articles. Eighty-six outcome measures of lymphedema were identified; 40 were patient reported outcomes (PRO) (represented in 141 articles) and the remainder were clinical measures (represented in 254 articles). We excluded 46 measures which had only one article addressing the measure. Using the remaining articles, the investigators generated a list of 12 measures of edema for HNC treatment-related lymphedema. Of these clinical measures, only 7 represented measurements of external edema for lymphedema of the head and neck that can be performed by physical therapists: real-time ultrasound, 3-dimensional (3D) laser scanning, tissue dielectric constant (TDC) method, M.D. Anderson Cancer Center – head and neck lymphedema rating scale (MDACC-HNLRS), and the ALOHA method (circumferential volume revised for head and neck), frustum volume estimation, and disk model method. One hundred thirty-nine references were included in the review (see Figure 1 and Table 1).

"Highly Recommended" and
"Recommended" Edema Outcome Measures

In order to reach this level of recommendation, the measure must have evidence supporting excellent psychometric properties and clinical utility. The measure should also have been tested in the HNC population. Of the 6 measures, none reached the level of "highly recommended" or "recommended," per the requirements outlined in Figure 2.

"Unable to recommend at this time"
Edema Outcome Measures

This recommendation has two levels depending on whether the measure has been used in the HNC patient population (2A) or not (2B). Four types of measures reached this level of recommendation. Three of these use technological devices to measure external edema—ultrasound, 3D scanning, and TDC. The final type is standardized tape measurement of the head, neck, and face. Included in this final type are two different approaches—the MDACC-HNLRS and the ALOHA tape measurement system. All have been used in the HNC population.

Real-time ultrasound measures skin-to-bone distance, skin thickness, subcutaneous thickness, and/or resistance to compression. Measurement of skin to bone distance at the mandibular ramus, angle and hyoid using a 7.5 MHz linear transducer was found to be reliable in healthy individuals (intraclass correlation coefficients ICC, ranged from 0.88 to 0.97).[14] Skin thickness (r = 0.95) and subcutaneous thickness (r = 0.84) have been shown to be strongly correlated with duration of lymphedema in women with breast cancer-related lymphedema.[19] At the time of this writing and to the best of our knowledge, no studies have formally investigated the responsiveness of ultrasound as a measure of edema among those with HNC-related lymphedema. However, one study showed a significant reduction in skin-to-bone distance after patients with HNC-related lymphedema received a course of manual lymphatic drainage and compression therapy.[14] Significant differences in skin thickness were found in patients who received systemic enzyme therapy after undergoing bimaxillary orthognathic surgery compared to those who did not receive systemic enzyme therapy.[20]

Three-dimensional scanning is a more recent application in its use for edema measurement. It has been used in a variety of industries—engineering where it is used for design and building projects;[21,22] forensics where it is used to estimate size, volume, and topography of missing body parts;[23,24] manufacturing where it is used for machining;[25,26] and dentistry where it is used in the creation of dental implants and dentures.[27,28] Three-dimensional scanning uses laser technology to scan the body part to be measured. It is a noninvasive approach with no radiation exposure and excellent reproducibility, with an error of less than one tenth of a millimeter.[29-31] However, Harrison et al reported a measurement error range between 0.3-4.0 cm^3 largely attributed to changes in positioning.[32] The sensitivity[33] and reliability[29-40] of this method of edema measurement are high. Ramos and colleagues[33] report that 3D imaging is sensitive, reproducible, and significant in detecting change in those with breast cancer-related lymphedema with initial volumes of 500 mL or less. The mean error of measurement for volume measured using 3D imaging has been found to be within 3.5%.[39] Kau et al[40] measured facial morphology in 473 male and female subjects using this technology. The average linear distance among male subjects ranged from –6.30 to 4.44 mm, with similar measurements among females ranging from –6.32 to 4.25 mm.[40] Together, these studies provide reference values for comparison for future studies. Limited research[33,39,40] using this method in the HNC survivor population exists at this time. However, this method allows one to measure any linear distance from any point to another in the affected and unaffected areas.

Like the 3D scanning technique, the TDC has also been used in health research and other industries. The Moisture Meter D (Delfin Technologies, Kuopio, Finland), is the only TDC device that has been tested with HNC survivors to measure cancer-related lymphedema. Inter-observer reliability, measured by the ICC has been found to be 0.973 in previous research.[41-43] Internal consistency has not been reported nor has the measurement error for the head and neck. The TDC device measures physiologic properties of biological fluid to a depth of between 0.5 mm – 5 mm for the ventral forearm and 2.5 mm on the lateral thorax. These depths not only vary by body location but also depend on

Table 1. EDGE Head and Neck Cancer Task Force Outcome Measure Rating Form

EDGE Head and Neck Cancer Task Force Outcome Measure Rating Form		
4	Highly Recommend	The outcome has excellent psychometric properties and clinical utility; the measure has been used in research on individuals with or post head and neck cancer.
3	Recommend	The outcome measure has good psychometric properties and good clinical utility; no published evidence that the measure has been applied to research on individuals with or post head and neck cancer.
2A	Unable to Recommend at this time	There is insufficient information to support a recommendation of this outcome measure; the measure has been used in research in individuals with or post head and neck cancer.
2B	Unable to Recommend at this time	There is sufficient information to support the recommendation of this outcome measure; no published evidence that the measure has been applied to research on individuals with or post head and neck cancer.
1	Do Not Recommend	Poor psychometrics &/or poor clinical utility (time, equipment, cost, etc).

the diameter of the probe used for TDC.[41] Recently, the TDC has been tested on the forehead and just below the maxilla (cheek) to measure the TDC up to a depth of 0.5 – 2.5 mm, but the thickness of the skin was a mediator in the measurement. This means that as tissue thickness increases in the forehead, the TDC value is reduced, but the opposite is true for the maxillary TDC measurement. However, in general, as tissue thickness increases, so does the TDC value indicating increased water content.[44] Only a few studies have used this method with HNC survivors, but it has not yet been tested for safety, efficacy, diagnostic reliability, and validity with large samples. Nixon et al[45] report interrater reliability as 0.97. They also report that the TDC, as measured by the Moisture Meter D, was able to discriminate between HNC survivors and healthy controls ($t = 8.97$, $p < 0.001$). The TDC was also correlated with linear tape measurements of the head and neck area (rho = 0.37-0.38), but not with the rating scale used by the MDACC-HNLRS. This study used a small sample size of 40—20 HNC survivors and 20 healthy controls.[43] At the time of the writing of this paper, this device is experimental in the United States in terms of its application in assessing edema of the head and neck. Unfortunately, none of the studies examined for this review considered obesity or tissue fibrosis as possibly leading to inaccurate estimates of fluid in the head and neck.

The tape measurements are based on dental and maxillofacial surgery measurement standards first described by Gabka et al[46] and later modified by Schultze-Mosgau et al.[47,48] Both were further adapted by Smith and Lewin to include neck measurements used in the MDACC-HNLRS.[17] The facial measures rely on the use of bony landmarks of the face and head.[17,43,44] Both MDACC-HNLRS and the ALOHA approach use linear distance measures. The MDACC-HNLRS uses 12 measures adapted from Gabka et al.[46] Two are point-to-point measures (mandibular angle to angle and tragus to tragus), 7 are facial measures (tragus to mental protuberance; tragus to mouth angle; mandibular angle to nasal wing; mandibular angle to medial canthus; mandibular angle to lateral canthus; mental protuberance to medial canthus; mandibular angle to mental protuberance), and 3 are circumferential neck measures (superior neck immediately beneath the mandible; medial neck midway between the superior and inferior neck; inferior neck at the base of the neck).[17,46] The ALOHA tape measurements utilize one point-to-point measure (ear to ear beginning at the inferior ear lobe/face junction of the left and right meeting at a point 8 cm inferior to the lower lip); 3 circumferential measures (lower neck circumference at the base of the neck; upper neck circumference inferior to the mandible; length from lip to lower neck circumference along the midline inferior lower lip to lower neck circumference). Purcell et al[43] compared 3 measurement methods—MDACC-HNLRS, ALOHA, and TDC alone. The ALOHA trial showed high interrater reliability for 3 tape measures with ICCs ranging from 0.948 (ear to ear length), 0.969 (upper neck circumference) and 0.979 (lower neck circumference).[43] The only tape measurement found not to have high interrater reliability was the lip to lower neck circumference length (ICC = 0.420).[43] Unlike the TDC, tape measures did not significantly discriminate between HNC survivors with lymphedema and healthy controls.[43] However, the authors recommend the tape measures as sensitive enough to measure within-person change over time, but did not provide evidence to support this claim.[43] Moreover, from a practical perspective, tape measurement takes a great deal of time to perform, making it difficult to measure irregularly shaped parts of the face and head such as the edema in the mandibular, auricular, nasal, orbital and submandibular areas. Table 2 provides a summary of the psychometric properties of these measures.

"Not Recommended" Edema Outcomes Measures

The remaining two measurements—frustum and disk model methods—merited a score of "not recommended." This means that poor psychometrics and lack of clinical utility exists for the frustum and disk model methods. Both of these edema measures were tested together in the studies we examined; however, it is important to note that each is a distinct method. The frustum method is an ancient method used to estimate the volume of multi-sided pyramidal shapes. A frustum is a geometric solid volume measurement of a cone or pyramidal structure and is created by at least two planes that bisect one another to form a side of a cone. In contrast, the disk method relies on the sum of multiple circumferential measures of a cone. Both methods are ideal for conical measures, and therefore, have been limited to use in the extremities.[48,49] An exception to this is one study that used a sample of 4 patients with head and neck lymphedema, in which the frustum method was calculated based on circumferential tape measurements of the head.[50] The frustum and disk model methods have also been used to measure volume of the prostate—another irregularly shaped body part. The frustum method was found to underestimate prostate volume by 50% when compared to planimetry and overestimated volume compared to ultrasound guided MRI.[45] While these methods are useful for the extremities, only the frustum method has been used on a small sample of HNC survivors[50] at the time of this writing. Unfortunately, no psychometric data are reported for the HNC patient population using these measures (see Table 2).

DISCUSSION

While smoking and/or alcohol abuse are common causes of HNC,[1] specific types of HPV cause nearly 32% of all HNC cases and this incidence is expected to rise.[12] With this growth, we expect that physical therapists will experience increased referrals for patients with HNC. In addition to lymphedema, HNC survivors undergoing treatment will have reduced dietary intake, with 37% to 68% experiencing moderate to severe pain, impaired swallowing, xerostomia, taste, and hoarseness.[52] Conical measures using circumferential measures or formulae dependent on circumferential measures (eg, frustum, disk model methods) are easy to use to measure edema of the extremities, but they do not adequately measure edema in the head and neck. These methods may underestimate volume of the head and neck.[51] Aside from surgical anatomy changes and associated side effects, the addition of side

Table 2. Psychometric Properties and Clinical Utility of External Edema Measures

Measure	Rating	Relevant Psychometric Properties	Clinical Utility
Ultrasound	2A	Test-retest reliability[14] ICC = 0.88-0.97 SEM: 0.50-0.64 mm $MDC_{95\%}$: 1.37 – 1.78 mm MCID, reference values: NA	Intensive training is required. Equipment (ultrasound unit) is costly, but is becoming more readily available in physical therapy clinics due to usage with musculoskeletal conditions. However, a higher frequency probe is typically used to measure skin thickness.
Three-dimensional Scanning	2A	Mean coefficient of variation for volume[29] 2.02 – 5.09% Mean error[29,30,34] 0.31 – 0.86 mm 0.01 – 0.23 cm^3 SEM, MDC, MCID: NA Reference values[40] • Males: –6.30 to 4.44 mm • Females: –6.32 to 4.25 mm	Intensive training is required. Equipment (3-dimensional imaging system) and computer software is costly and not readily available in clinic. Requires standardized patient positioning.
Tape Measurement	2A	Interrater reliability[43] ICC: 0.42 – 0.95 Concurrent validity with Moisture Meter[43] r = 0.17 –0.38, p > 0.05 for all measurements except lip to lower neck circumference SEM, MDC, MCID, reference values: NA	Easy to use. Minimal training required. Equipment (tape measure) is inexpensive and readily available. Requires standardized patient positioning.
Tissue Dielectric Constant	2A	Intrarater reliability[43] ICC: 0.97 Interrater reliability[43] ICC: 0.97 Concurrent validity with MDACC level [43] r = 0.58 SEM, MDC, MCID, reference values: NA	Easy to use. Minimal training required. Equipment (Moisture Meter™) is costly and not readily available in clinic.

Abbreviations: ICC: intraclass correlation coefficient; SEM: standard error of measure; MDC: minimal detectable change; MCID: minimal clinical important difference; NA: not available; MDACC: MD Anderson Cancer Center Head and Neck Lymphedema Rating Scale

effects associated with chemotherapy, radiation, or CCR increases the risk of lymphedema of any of the structures targeted by cancer treatment. For example, radiation-induced tissue fibrosis will include damage to the remaining lymphatic structures in the treated area.[54]

Subjective information through self-reports such as heaviness, tightness, and firmness[5,6] along with objective information provided by clinical measurement of lymphedema can provide information to determine successful outcomes when treating HNC survivors. Lymphedema is a multi-faceted diagnosis based on the presence or absence of skin integrity, color, edema, pitting, positive Stemmer's sign, and malformations associated with lymphostasis and infection.[15] Therefore, using edema measurement as a singular indicator of lymphedema can be misleading in determining the severity of the lymphedema.[16,55] However, in the head, face, and neck, edema drastically changes one's appearance and may impair the ability to breathe, swallow, and vocalize, making external measures of edema a crucial part of any physical therapy examination, evaluation, and assessment of the HNC survivor. It is important that physical therapists treating the HNC survivor with lymphedema have reliable and valid measures of edema that are clinically feasible and have realistic and reasonable associated costs. While there are also internal measures of lymphedema, these are invasive and require technical procedures that fall outside of the scope of physical therapist practice. For example, a bronchoscopy is performed by an otolaryngologist to assess internal swelling as well as check for cancer recurrence in the oropharyngeal cavity, as Deng and colleagues have done. Bronchoscopy is not risk-free. These risks include aspiration, vocal

cord damage, bleeding, oxygen desaturation, bronchospasm, fever, pneumonia, pneumothorax, and death.[56,57] Physical therapists specializing in HNC would benefit from having well-formed relationships with physician or speech therapy practitioners to take advantage of information resulting from imaging procedures outside the PT scope of practice. Additional measurements could be taken and internal and external volume could then, hypothetically, be calculated. Three-dimensional scanning, ultrasound, and TDC can be performed by physical therapists; however, each require additional training on the use and application of these methods. Three-dimensional scanning has a steep learning curve because it requires that the user know how to use the scanner and requires multiple images to render a 3D image. The software to create these images is complex and difficult to learn. Further, the cost of a 3D scanner, software and software training is cost-prohibitive for the average clinical setting. Real-time ultrasound, on the other hand, is increasingly utilized in physical therapy clinics and those who already use it for evaluation of treatment of patients (such as urinary incontinence or adhesive capsulitis) may find it easy to use in the HNC patient population. Finally, TDC appears to be the easiest to use, set-up, and interpret.

Quantifying edema reduction is crucial to gauge the effect of treatment for lymphedema in the head and neck. It presents a unique measurement challenge in effectively treating HNC survivors. Edema measurement will fulfill new Medicare outcomes documentation requirements to demonstrate reduction in volume due to external lymphedema. At this point, there is no gold standard to measure edema in the head and neck. Common practice involves linear distances of the face and circumferential measures of the neck—such as those used in the MDACC-HNLRS and ALOHA methods. However, neither of these methods actually measure volume but provide a way to document reduction in linear distance that is reasonably expected to reflect a reduction in volume in response to treatment. The advantage of the linear distances used in these measures is that clinically speaking the baseline of these measures could be measured using diagnostic MRI or CT scan. However, this approach involves the cooperation of physicians—particularly radiologists—and is not typically measured as part of clinical standards of care. Another advantage of the MDACC-HNLRS and ALOHA methods is that both provide a systematic way to measure edema. The ALOHA method also incorporates measurement of the TDC and, similar to the MDACC-HNLRS, also uses tape measurements and classification of HNC lymphedema characteristics. These approaches are clinically feasible, cost-effective, and not time-intensive. The linear measurements are estimated by the authors to take approximately 5 minutes to measure by clinicians experienced with HNC survivors, making them not time intensive. However, the inclusion of the TDC increases the time to complete the measurement and, in total (linear and TDC measurement), we estimate these could take up to 10 minutes. Ten minutes could take up one third of the time to evaluate or treat a non-complex patient, even more time may be required if the physical therapist needs to re-measure. At this point though, TDC is not commonly used clinically and is not likely to be found in the average clinic.

The techniques of ultrasound, 3D scanning, and TDC methods hold great promise for clinical translation but all need more research investigating their psychometric properties of reliability and validity, sensitivity to change, and references for comparison. All 3 are used to measure external edema and, while none addresses the measurement of internal edema, they are noninvasive and not dependent on a skill set possessed only by physicians. Like MRI and CT-scanning methods, 3D laser scanning can precisely measure external lymphedema and does not require a medical referral. The disadvantage is that the entire volume of the head and neck region is assumed to be included in the overall volume measure. To illustrate, it assumes that the scanned area is a solid with no "empty" spaces such as the oral cavity, sinuses, pharynx, larynx, and upper esophagus. Of these 3 methods dependent on the use of technology, 3D scanning is prohibitively expensive and presents a steep learning curve for the physical therapist. In addition to the 3D laser scanner (of which there are several types to choose), the clinic will also need to purchase 3D computer software to measure volume. The computer programs are complicated and require intensive training. Other costs associated with the use of the 3D scanner include software training, service agreements, software updates, repair costs, and costs for personnel to take the images and render them usable in 3D software.

In contrast, the TDC is a small hand-held device that provides fast and readily available measurements. The TDC measures the electrical properties of lymphatic fluid—not necessarily edema like 3D scanning—but it is comparable to the MDACC-HNLRS and ALOHA systems. Likewise, ultrasound holds promise and further research is warranted. We caution that the 3D scanners, real-time ultrasound, and TDC devices may be cost-prohibitive for a clinic to purchase.

Measurement of head and neck lymphedema poses many challenges to physical therapists. Methods that are clinically realistic are limited and few have had in-depth testing with patients with HNC. Head and neck cancer is a rare tumor type and patients tend to be very sick during active treatment, making it difficult to recruit and retain large samples for research. The MDACC-HNLRS and the ALOHA measurement systems are the best available and feasible measures for use in the physical therapy clinic. Collected prospectively and across clinics, these measurement systems may provide important data for future research and comparability across clinics.

While we focused on measurement of external edema for the HNC survivor with cancer-related lymphedema, it is clear that much more research needs to be conducted to address the complex cluster of issues typical of lymphedema. Future research will address this through the review of PROs related to QOL and lymphedema.

CONCLUSIONS

Measurement of lymphedema of the head and neck is in its

infancy. Few methods exist to calculate volume and/or physiological properties that characterize lymphedema of the head and neck. Volume assessment using tape measures and specific landmarks are the most commonly used method currently, but other methods such as 3D scanning, ultrasound, and TDC hold great promise. However, because of the lack of good psychometric properties or clinical utility, no outcome measures for objectively quantifying external edema for the HNC population can be recommended at this time. The future development and application of these devices to measure edema in HNC survivors with lymphedema merits systematic and comparative research to evaluate measurement properties such as reliability, validity, sensitivity to change, and normative values. Much research needs to be conducted to develop and compare a variety of methods for edema of the head and neck. We recommend that physical therapists fully weigh the advantages and disadvantages of any measure of edema while treating HNC survivors.

ACKNOWLEDGMENTS

No funding was available for this project. All member efforts were on a volunteer basis. We wish to thank Elyse Bloomfield and Mariesa Gonzalez (students in the Northeastern University DPT program) for assisting with downloading and screening of abstracts.

Appendix 1. PRISMA Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	15
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	15
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	15-16
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	16
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (eg, Web address), and, if available, provide registration information including registration number.	16
Eligibility criteria	6	Specify study characteristics (eg, PICOS, length of follow-up) and report characteristics (eg, years considered, language, publication status) used as criteria for eligibility, giving rationale.	16
Information sources	7	Describe all information sources (eg, databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	16
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Figure 1 Please also see the Data & Methods section on page 16
Study selection	9	State the process for selecting studies (ie, screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	16
Data collection process	10	Describe method of data extraction from reports (eg, piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	16
Data items	11	List and define all variables for which data were sought (eg, PICOS, funding sources) and any assumptions and simplifications made.	16
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	16
Summary measures	13	State the principal summary measures (eg, risk ratio, difference in means).	NA
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (eg, I^2) for each meta-analysis.	NA

REFERENCES

1. National Cancer Institute. Head and neck cancers. National Cancer Institute at the National Institute of Health Web site. http://www.cancer.gov/cancertopics/factsheet/Sites-Types/head-and-neck. Updated February 1, 2013. Accessed November 11, 2013.
2. Ries LAG, Young JL, Keel GE, Eisner MP, Lin YD, Horner MJ, eds. SEER survival monograph: Cancer survival among adults: U.S. SEER program, 1988-2001, patient and tumor characteristics. Bethesda, MD: National Cancer Institute, SEER Program; 2007.
3. Howlader N, Noone AM, Krapcho M, et al. SEER Cancer Statistics Review, 1975-2012, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2012/, based on November 2014 SEER data submission, posted to the SEER web site, April 2015.
4. De Moor JS, Mariotto AB, Parry C, et al. Cancer survivors in the United States: Prevalence across the survivorship trajectory and implications for care. *Cancer Epidemiol Biomarkers Prev*. 2013;22(4):561-700.
5. Deng J, Ridner SH, Dietrich MS. Factors associated with external and internal lymphedema in patients with head-and-neck cancer. *Int J Radiat Oncol Biol Phys*. 2012;84(3):e319-328.
6. Deng J, Ridner SH, Dietrich MS, Wells N, Murphy BA. Assessment of external lymphedema in patients with head and neck cancer: A comparison of four scales. *Oncol Nurs Forum*. 2013;40(5):501-506.
7. Liu M, Shi X, Guo X, et al. Long-term outcome of irradiation with or without chemotherapy for esophageal squamous cell carcinoma: A final report on a prospective trial. *Radiat Oncol*. 2012;7:142.
8. Thariat J, Hamoir M, Garrel R, et al. Management of the neck in the setting of definitive chemoradiation: Is there a consensus? A GETTEC study. *Ann Surg Oncol*. 2012;19(7):2311-2319.
9. Romesser PB, Romanyshyn JC, Schupak KD, et al. Percutaneous endoscopic gastrostomy in oropharyngeal cancer patients treated with intensity-modulated radiotherapy with concurrent chemotherapy. *Cancer*. 2012;118(24):6072-6078.
10. Fialka-Moser V, Crevenna R, Korpan M, Quittan M. Cancer rehabilitation: Particularly with aspects on physical impairments. *J Rehabil Med*. 2003;35(4):153-162.
11. Howlader N, Noone AM, Krapcho M, et al. SEER cancer statistics review, 1975-2008. National Cancer Institute Web site. http://seer.cancer.gov/csr/1975_2008. Updated 2014.
12. Näsman A, Nordfors C, Holzhauser S, et al. Incidence of human papillomavirus positive tonsillar and base of tongue carcinoma: A stabilisation of an epidemic of viral induced carcinoma? *Eur J Cancer*. 2015;51(1):55-61.
13. Tacani PM, Franceschini JP, Tacani RE, et al. Retrospective study of the physical therapy modalities applied in head and neck lymphedema treatment. *Head Neck*. 2014 Oct 21. Epub ahead of print.
14. Piso DU, Eckardt A, Liebermann A, Gehrke A. Reproducibility of sonographic soft-tissue measurement of the head and neck. *Am J Phys Med Rehabil.* 2002;81(1):8-12.
15. Foldi M, Foldi E, Kublik S. *Textbook of Lymphology*. 6th ed. Munich, Germany: Elsevier; 2005.
16. Zuther JE, Norton S. *Lymphedema Management: The Comprehensive Guide for Practitioners*. 3rd ed. New York, NY: Thieme; 2013.
17. Smith BG, Lewin JS. Lymphedema management in head and neck cancer. *Curr Opin Otolaryngol Head Neck Surg*. 2010;18(3):153-158.
18. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *J Clin Epidemiol*. 2009;62(10):e1-34.
19. Mellor RH, Bush NL, Stanton AW, Bamber JC, Levick JR, Mortimer PS. Dual-frequency ultrasound examination of skin and subcutis thickness in breast cancer-related lymphedema. *Breast J*. 2004;10:496-503.
20. Shetty V, Mohan A. A prospective, randomized, double-blind, placebo-controlled clinical trial comparing the efficacy of systemic enzyme therapy for edema control in orthognathic surgery using ultrasound scan to measure facial swelling. *J Oral Maxillofac Surg*. 2013;71(7):1261-1267.
21. Baaklini G. Optical instrumentation and nondestructive evaluation branch research and technology advances at NASA Glenn Research Center since 1941. *J Aerosp Eng*. 2013;26:395-408.
22. Huang Y, Leu MC, Mazumder J, Donmex A. Additive manufacturing: Current state, future potential, gaps and needs, and recommendations. *Manuf Sci Eng.* 2015(Online 2014 Nov);137(1):014001.
23. Buck U, Naether S, Räss B, Jackowski C, Thali MJ. Accident or homicide--virtual crime scene reconstruction using 3D methods. *Forensic Sci Int*. 2013;225(1-3):75-84.
24. Davy-Jow SL, Lees DM, Russell S. Virtual forensic anthropology: Novel applications of anthropometry and technology in a child death case. *Forensic Sci Int*. 2013;224(1-3):e7-10.
25. Son H, Kim C, Kim C. Automatic 3D reconstruction of as-built pipeline based on curvature computations from laser-scanned data. *Construction Research Congress*. 2014:925-934.
26. Song I, Yang J, Shimada K. Development of sketch-based 3D modeling system for rapid generation and evaluation of automotive seat shape using reference models. *J Mech Des*. 2014;136(5):051001.
27. Kim KB, Kim JH, Kim WC, Kim JH. Three-dimensional evaluation of gaps associated with fixed dental prostheses fabricated with new technologies. *J Prosthet Dent.* 2014;112(6):1432-1436.
28. Zhao XZ, Xu WH, Tang ZH, Wu MJ, Zhu J, Chen S. Accuracy of computer-guided implant surgery by a CAD/CAM and laser scanning technique. *Chin J Dent Res*. 2014;17(1):31-36.

29. Yip E, Smith A, Yoshino M. Volumetric evaluation of facial swelling utilizing a 3D range camera. *Int J Oral Maxillofac Surg*. 2004;33(2):179-182.
30. Kau CH, Richmond S, Zhurov AI, et al. Reliability of measuring facial morphology with a 3-dimensional laser scanning system. *Am J Orthod Dentofacial Orthop*. 2005;128(4):424-430.
31. Kau CH, Richmond S, Incrapera A, English J, Xia JJ. Three-dimensional surface acquisition systems for the study of facial morphology and their application to maxillofacial surgery. *Int J Med Robot*. 2007;3(2):97-110.
32. Harrison JA, Nixon MA, Fright WR, Snape L. Use of hand-held laser scanning in the assessment of facial swelling: A preliminary study. *Br J Oral Maxillofac Surg*. 2004;42(1):8-17.
33. Ramos SM, O'Donnell LS, Knight G. Edema volume, not timing, is the key to success in lymphedema treatment. *Am J Surg*. 1999;178(4):311-315.
34. Ort R, Metzler P, Kruse AL, et al. The Reliability of a Three-Dimensional Photo System- (3DMDface-) Based Evaluation of the Face in Cleft Lip Infants. *Plast Surg Int*. 2012;2012:138090.
35. Rana M, Gellrich NC, Ghassemi A, Gerressen M, Riediger D, Modabber A. Three-dimensional evaluation of postoperative swelling after third molar surgery using 2 different cooling therapy methods: a randomized observer-blind prospective study. *J Oral Maxillofac Surg*. 2011;69(8):2092-2098.
36. Rana M, Gellrich NC, Joos U, Piffko J, Kater W. 3D evaluation of postoperative swelling using two different cooling methods following orthognathic surgery: a randomised observer blind prospective pilot study. *Int J Oral Maxillofac Surg*. 2011;40(7):690-696.
37. Rana M, Gellrich NC, von See C, et al. 3D evaluation of postoperative swelling in treatment of bilateral mandibular fractures using 2 different cooling therapy methods: a randomized observer blind prospective study. *J Craniomaxillofac Surg*. 2013;41(1):e17-23.
38. Maus EA, Tan IC, Rasmussen JC, et al. Near-infrared fluorescence imaging of lymphatics in head and neck lymphedema. *Head Neck*. 2012;34(3):448-453.
39. Bowskill J, Baldock C, Booth PW. Measuring facial swelling using 3D imaging. *Med Inform (Lond)*. 1997;22(2):155-164.
40. Kau CH, Richmond S, Zhurov A, et al. Use of 3-dimensional surface acquisition to study facial morphology in 5 populations. *Am J Orthod Dentofacial Orthop*. 2010;137(4 Suppl):S56 e1-9.
41. Mayrovitz HN, Davey S, Shapiro E. Suitability of single tissue dielectric constant measurements to assess local tissue water in normal and lymphedematous skin. *Clin Physiol Funct Imaging*. 2009;29:123–127.
42. Jensen MR, Birkballe S, Noerregaard S, Karlsmark T. Validity and interobserver agreement of lower extremity local tissue water measurements in healthy women using tissue dielectric constant. *Clin Physiol Funct Imaging*. 2012;32:317–322.
43. Purcell A, Nixon J, Fleming J, McCann A, Porceddu S. Measuring head and neck lymphoedema: The "ALOHA" trial. *Head Neck*. 2014 Sep 12.
44. Mayrovitz HN, Bernal M, Brlit F, Desfor F. Biophysical measures of skin tissue water: Variations within and among anatomical sites and correlations between measures. *Skin Res Technol*. 2013;19:47-54.
45. Nixon J, Purcell A, Fleming J, McCann A, Porceddu S. Pilot study of an assessment tool for measuring head and neck lymphoedema. *Br J Community Nurs*. 2014 Apr;Suppl S6, S8-11.
46. Gabka J MT. Measuring techniques and clinical testing of an anti-inflammatory agent (tantum). *Munch Med Wochenschr*. 1971;13:198-203.
47. Schultze-Mosgau S, Schmelzeisen R, Fröhlich JC, et al. Use of ibuprofen and methylprednisolone for prevention of pain and swelling after removal of impacted third molars. *J Oral Maxillofac Surg*. 1995;53:2-7.
48. Stizia J. Volume measurement in lymphoedema treatment: Examination of formulae. *Eur J Cancer*. 1995;4:11-16.
49. Deltombe T, Jamart J, Recloux S, et al. Reliability and limits of agreement of circumferential, water displacement, and optoelectronic volumetry in the measurement of upper limb lymphedema. *Lymphology*. 2007;40(1):26-34.
50. Ayestaray B, Bekara F, Andreoletti JB. π-Shaped lymphaticovenular anastomosis for head and neck lymphoedema: A preliminary study. *J Plast Reconstr Aesthet Surg*. 2013;66(2):201-206.
51. Liu D, Usmani N, Sloboda RS. Transrectal ultrasound based prostate volume determination: Is the frustum algorithm more accurate than planimetry? *Med Phys*. 2013;40(3):031705.
52. Vokes EE, Kies MS, Haraf DJ, et al. Concomitant chemoradiotherapy as primary therapy for locoregionally advanced head and neck cancer. *J Clin Oncol*. 2000;18(8):1652-1661.
53. Guru K, Manoor UK, Supe SS. A comprehensive review of head and neck cancer rehabilitation: Physical therapy perspectives. *Indian J Palliat Care*. 2012;18(2):87-97.
54. Habbous S, Chu KP, Qiu X, et al. The changing incidence of human papillomavirus-associated oropharyngeal cancer using multiple imputation from 2000 to 2010 at a comprehensive cancer centre. *Cancer Epidemiol*. 2013;37(6):820-829.
55. U.S. Department of Health and Human Services (National Institutes of Health National Cancer Institute). Common Terminology Criteria for Adverse Events (CTCAE). Version 4.0. Published: May 28, 2009 (v4.03: June 14, 2010) NIH Publication No. 09-5410
56. Medford AR. Endobronchial ultrasound and extended roles: Know thy limitations. *Chest.* 2013;143(5):1516.
57. Haga T, Fukuoka M, Morita M, Cho K, Tatsumi K. Indications and complications associated with fiberoptic bronchoscopy in very elderly adults. *J Ameri Ger Soc*. 2014;62(9):1803-1805.

Oncology Section EDGE Task Force Breast Cancer Outcomes:
A Systematic Review of Clinical Measures of Cardiorespiratory Fitness Tests

Jacqueline S. Drouin, PT, PhD[1]; G. Stephen Morris, PT, PhD, FACSM[2]

[1]Oakland University, School of Health Sciences, Physical Therapy Program, Rochester, MI
[2]Wingate University, Department of Physical Therapy, Wingate, NC

*Presented as an Oncology Section Educational Session at the American Physical Therapy Association's Combined Sections Meeting in Las Vegas, Nevada; February 2014.

ABSTRACT

Background: Outcomes from standardized exercise tests (SET) are used by physical therapists (PTs) to define cardiorespiratory fitness (CRF) and develop safe and effective prescriptions for aerobic exercise training programs; however, the psychometric properties and clinical utility of standard exercise testing has not been evaluated in the breast cancer survivor population. **Objective:** To evaluate the psychometric properties, safety, and clinical utility of SET and provide informed recommendations for their use in clinical practice involving female breast cancer survivors. **Design:** The study was a systematic review with a measurement focused design. **Methods:** A comprehensive search was performed with a health sciences librarian to identify articles that reported psychometric data on SET among women with breast cancer. Criterion articles were evaluated using the Cancer *EDGE Task Force Outcome Measure Rating Form*. **Results:** Sixty-eight articles met initial inclusion criteria, but only 5 were found that reported useable psychometric data. Maximal treadmill and cycle ergometer tests involving expired gas analysis were valid and accurate in this population, but are not safe or efficient for use in a clinical setting. Submaximal tests (treadmill, cycle ergometer, step, and walk/run tests) that use prediction equations to estimate CRF had large errors for minimal detectable differences (3.32-15.80 ml/kg/min) which approached or exceeded the minimal clinically important difference (3.5 ml/kg/min). Furthermore, these prediction equations have not been validated in this population. **Discussion:** Limited evidence was found describing the psychometric properties of SETs used to determine CRF in breast cancer survivors. Available studies suggest clinically efficient SETs have moderate to poor concurrent validity with a moderate to high standard error; however, standardized exercise testing can serve to screen for exercise safety and provide estimates of CRF. Exercise prescriptions developed from the outcomes of SET will require follow up for potential modifications. Appreciating SET limitations enables proper client education and training adjustments needed for safety and exercise efficacy. **Conclusions:** Further understanding of the psychometric properties of SET used in the breast cancer survivor population is needed in order to make these tests safe, accurate, and clinically useful.

Address correspondence to: Jacqueline S. Drouin, PT, PhD, Oakland University, School of Health Sciences, Physical Therapy Program, 3155 Human Health Building, Rochester, MI 48309-4482, Ph: (248) 364-8684, Fax: (248) 364-3079 (drouin@oakland.edu).

Key Words: aerobic exercise testing, psychometrics, clinical utility, cardiorespiratory fitness, submaximal exercise, maximal exercise

INTRODUCTION

Aerobic exercise (AE) is recommended as a safe and effective method for reducing treatment related side effects and promoting cardiorespiratory fitness (CRF) in women with breast cancer.[1,2] Available evidence suggests that AE may enhance cancer treatment efficacy and induce changes in gene expression and cancer biology, which might prevent an initial occurrence or a reoccurrence of cancer.[1-4] Based on this positive evidence for health promotion and cancer prevention, it becomes the physical therapist's role to develop safe and effective exercise prescriptions predicated on using accurate and clinically appropriate clinical tests of CRF.[1-3,5-8] Current consensus statements recommend using standardized exercise tests (SET) according to established professional guidelines to determine CRF.[2,5-10] However, available consensus statements acknowledge that the psychometric properties of established SETs were developed using non-cancer populations; therefore, the accuracy and applicability of CRF measures in breast cancer survivors is not known.[1,2,9-11]

Accurate cardiovascular screening and exercise testing is particularly important for breast cancer survivors since cancer treatments and adjuvant therapies cause varying amounts of acute and/or long-term cardiovascular disorders.[1,2,9,12-14] These disorders include hypertension, endocardial fibrosis, arrhyth-

mias, bradycardias, progressive declines in left ventricular function, and heart failure.[13,14] Furthermore, a review of exercise testing outcomes from cancer patients found markedly reduced CRF levels compared to apparently healthy subjects that were attributed to cancer treatments, aging, and sedentary lifestyles.[3] This elevated cardiovascular risk signals the need for appropriate cardiovascular screening, and, when indicated, CRF testing to insure that exercise prescriptions are safe and effective.[2,9] Admittedly, there are screening tools such as surveys and walking tests (6 minute walk test) that are associated with physical function and have proven to be clinically useful.[5] However, these assessments do not measure or readily predict CRF, an outcome required for the development of effective exercise prescriptions or for screening for disease. Therefore, these assessments will not be evaluated in this paper.[2,5]

While a considerable body of evidence exists describing the psychometric properties of SET measures of CRF in apparently healthy individuals and individuals with cardiovascular disease, the psychometric properties of SET used in assessing the CRF of breast cancer survivors has not been previously evaluated.[1,9,15-42] Therefore, the purpose of this study was to evaluate the psychometric properties, safety, and clinical utility of SETs in order to make informed recommendations regarding their use in clinical practice for the purpose of measuring CRF in female survivors of breast cancer.

METHODS

A systematic review using a measurement focused design was used to find and evaluate articles that contained data on the psychometric properties of SETs used in female survivors of breast cancer. The clinical utility and safety of these tests was also examined. The literature search was initially performed by graduate students in a Doctor of Physical Therapy program with guidance from a University Health Sciences Librarian. Criterion articles were written in English or had a published abstract written in English. The search began online using CINAHL, PEDro, Pubmed (Medline), Cochrane, Science Direct, Hooked on Evidence, Web of Science, Scopus, and Sport Discus databases. References listed in the selected criterion articles were also evaluated for potential inclusion. Key search words included: breast neoplasm, breast cancer, cancer, SET, aerobic capacity, aerobic exercise, aerobic endurance, cardiorespiratory fitness, cardiovascular endurance, maximal oxygen consumption, VO_2max, psychometrics, reliability, validity, standard error of the measure (SEM), minimal detectable difference (MDD), minimal clinically important difference (MCID), SET, standardized SET, treadmill test, cycle ergometer test, walk tests, run tests, and step tests.

The psychometric properties, safety, and clinical utility of the SETs were reviewed using the Cancer EDGE (*Evidence Database to Guide Effectiveness*) Task Force Outcome Measure Rating Form. This evaluation tool uses a format adapted from the *Section on Research EDGE Form* and recommended by the American Physical Therapy Association EDGE task force. Although this tool recommends limiting the search to articles published in the previous 10 years, the SET psychometric studies began in the 1930s, so no date limitation was applied to articles in the search. The evaluators were two physical therapists with advanced academic degrees in exercise physiology, research design, and statistics. The evaluations ratings used were the categories from the Breast Cancer EDGE tool seen in Table 1. When evaluation ratings did not agree, the authors reviewed the evidence to achieve consensus.

Criterion articles were from original studies that presented the psychometrics, safety and clinical utility of SETs used in the breast cancer population. Articles were also included if the data provided could be used to calculate concurrent validity, test-retest reliability, SEM, MDD, and MCID. Articles were removed that did not include data from a SET, did not clearly explain the type of SET performed, or that used functional assessments or reported fitness derived from surveys or self-reports. Systematic reviews, commentaries or recommendations, or practice guidelines were not reviewed; however, reference lists from these sources were examined for potential references.

RESULTS

Initially, 3837 articles were identified through key word searches of online data bases. Following removal of duplicates (n = 1803) and non-criterion articles through title and abstract

Table 1. Breast Cancer EDGE* Evaluation Categories

4	Highly Recommend	Highly Recommend; the outcome has excellent psychometric properties and clinical utility; the measure has been used in research on individuals with or post breast cancer.
3	Recommend	Recommend; the outcome measure has good psychometric properties and good clinical utility; no published evidence that the measure has been applied to research on individuals with or post breast cancer.
2A	Unable to Recommend At this time	Unable to recommend at this time; there is insufficient information to support a recommendation of this outcome measure; the measure has been used in research on individuals with or post breast cancer.
2B	Unable to Recommend At this time	Unable to recommend at this time; there is insufficient information to support a recommendation of this outcome measure; no published evidence that the measure has been applied to research on individuals with or post breast cancer.
1	Do not Recommend	Poor psychometrics &/or poor clinical utility (time, equipment, cost, etc.).

* Cancer EDGE (Evidence Database to Guide Effectiveness) Task Force Outcome Measure Rating Form; adapted from the Section on Research EDGE Form recommended by the APTA EDGE Task Force.

reviews (n = 1984), and then adding 18 articles from reference checks, 68 articles were selected for review (Figure 1). Of these articles, 5 were identified that included psychometric data or data that could be used to calculate psychometric values in the breast cancer population. Two studies were also found that examined the safety of exercise testing conducted in female breast cancer survivors.

Evidence from sports medicine and science journals recommend that concurrent validity and test-retest reliability are the psychometrics most appropriate for evaluating SETs.[15,43,44] Concurrent validity is the degree to which SET measures agree with those obtained using gold standard measures.[17,43,44] The statistical analyses used to determine concurrent validity are Intraclass Correlation Coefficients [ICC (model 2, 1 or 2, k)] and Coefficients of Variation (COV).[15,44,45] The ICC provides a measure of the agreement between two or more tests expressed as values between 0 and 1.0.[20,44] Higher values represent better agreement; however, the values are difficult to interpret and there is controversy regarding the level viewed as acceptable.[15,43-45] Some older articles suggest that correlation coefficients of .75 or higher are excellent, while more current reviews indicate that the ICC must be .90 or higher for a test to be useful in clinical practice and research.[15,43,45,46] Prior studies have also used the Pearson Product Moment (r) to determine associations between tests; however, this statistic is not the preferred method as it is not able to account for systematic error and it overestimates associations when samples are small.[15,43-45] Based on the evidence, a rating of excellent in this review required a value of .90 or above. Measures of concurrent validity for various SET appear in Table 3 and Table 4; when the ICC was not available, the Pearson Product Moments was instead reported.

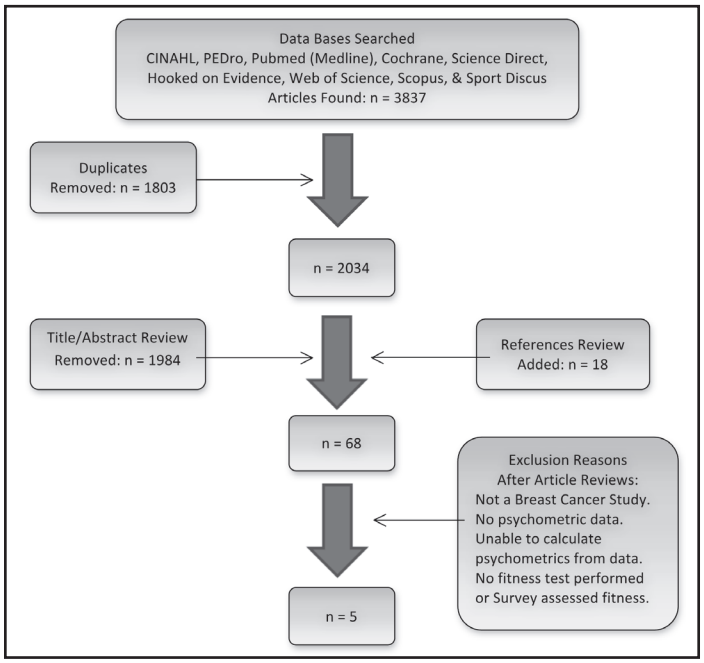


Figure 1. Flow of literature search process for psychometric measures.

The COV presents the typical error as the percent change in the mean.[15,45] The COV is a straightforward, dimensionless measure that enables comparisons of scores from different types of subjects (males vs females) as well as different tests (treadmill vs cycle ergometer). As such, the COV is easier to interpret and may be more useful in clinical and research settings.[15] The lower the COV value, the higher the agreement with the gold standard measure; however, some authors suggest that 5% or higher represents poor reliability and this standard was used in the current evaluation.[20] The COV values derived from sport and performance tests appear in Table 3 and Table 4. When the COV was not available, the percent difference between the tests was reported.

Test-retest reliability provides a measure of the precision or reproducibility of SET measures taken over time and it can be used to assess the exercise device, the test protocol, the subject, as well as the raters.[15] The recommended statistics again are the ICC (model 3, 1) and the COV as described previously.[46]

Standard errors of a measure are clinically useful for determining both concurrent validity and test-retest reliability.[44] Like the COV, the SEM is a calculation of the typical error and it is robust to small samples.[15,44] The SEM is clinically useful as it is reported in the same units as the test data.[15,44] For example, an SEM of 0.05 ml/kg/min for an SET for an MCID of 3.5 ml/kg/min would be considered excellent agreement. However, an SEM for an SET that approaches or exceeds the MCID would be invalid and unreliable for measuring CRF and subsequently developing safe and effective exercise prescription intensities. The SEM is also used to calculate the MDD which allows clinicians to determine whether changes in a patient's score were real or simply within the error of the measure; therefore, lower MDD values were associated with higher levels of accuracy.[43] The values for test-retest reliability appear in Table 3 and Table 4. When the SEM was not reported, the standard error of the estimate (SEE) was provided.

Maximal Exercise Tests

Maximal SETs complete with gas analysis can be performed using a treadmill or a cycle ergometer.[2] Both methods require the subject to perform a maximal physical effort and analyze expired gases for O_2 and CO_2 content which is then used to calculate maximal oxygen consumption or VO_2max.[2] The criteria defining a maximal CRF SET includes plateauing of the heart rate and/or oxygen consumption despite increases in the workload and a respiratory exchange ratio (CO_2 released/O_2 consumed) of 1.15 or greater.[17-19] However, not all individuals are able to reach maximal cardiorespiratory levels due to musculoskeletal, neuromuscular or other physical limitations.[2,6-8] It is also not safe for many individuals to perform maximal exercise, particularly if they have known or suspected cardiovascular conditions.[2,6-8]

Maximal oxygen consumption is measured in milliliters per kilogram body weight per minute (ml/kg/min) which is a standardized value that allows for comparisons between individuals or across time for a single individual which can be used to asses changes in CRF.[17-19] The MCID derived from studies in

Table 2. Standardized Exercise Test Outcomes: EDGE Task Force Rating & Clinical Utility

Standardized Exercise Test	Task Force Rating	Psychometrics	Clinical Utility and Safety
Maximal Tests with Oxygen Analysis			
Treadmill	1 Do not Recommend	• Used in 14 breast cancer research intervention studies • Psychometrics for valid CRF measures found in one breast cancer study • Excellent validity and reliability from non-breast cancer studies	• Very expensive • Time intensive (45-60 min) • Specialized training • Requires 2 clinicians to perform • Often requires physician onsite • Contraindicated in national practice standards due to safety risks of performing maximal exercise
Cycle Ergometer	1 Do not Recommend	• Used in 15 breast cancer intervention studies • Psychometrics found adequate for test-retest reliability in one breast cancer study • Psychometrics from non-breast cancer studies were poor to excellent for concurrent validity and moderate to poor for clinical relevance as MDD may exceed MDIC	• Very expensive • Time intensive (45-60 min) • Specialized training • Requires 2 clinicians to perform • Often requires physician onsite • Contraindicated in national practice standards due to safety risks of performing maximal exercise
Submaximal Tests with Prediction Equations			
Treadmill	2A Unable to Recommend At this time	• Used in 7 breast cancer research studies • Prediction equations for heart rate maximum to determine submaximal test endpoints and CRF found invalid in 2 breast cancer studies • Psychometrics not assessed in breast cancer research studies • Psychometrics from non-breast cancer studies are fair to good for concurrent validity but standard errors were high and the MDD exceeded the MDIC contributing to poor reliability.	• Low cost and easy to administer • Less time to administer (20-30 minutes) • Requires 2 clinicians: one performs test and other onsite for safety concerns • Requires following standardized protocol and accurate heart rate measures • Safe to administer with proper pre-screening. Monitoring and patient instructions
Cycle Ergometer	2A Unable to Recommend At this time	• Used in 5 breast cancer research studies • Psychometrics poor for sensitivity and concurrent validity with VO_2max cycle ergometer tests in 2 breast cancer studies • Psychometrics adequate for test-retest reliability in one breast cancer study • Prediction equations for heart rate maximum to determine submaximal test endpoints and CRF were invalid in 2 breast cancer studies	• Low cost, portable, and easy to administer • Less time to administer (20-30 minutes) • Requires 2 clinicians: one performs test and other onsite for safety concerns • Requires following standardized protocol and accurate heart rate measures • Safe to administer with proper pre-screening. Monitoring and patient instructions
Field Tests			
12-Min Run	1 Do Not Recommend	• Not used in breast cancer research • MDD exceeds MCID • Validity of CRF prediction equations not assessed in breast cancer research	• Easy to administer • Low cost • Can assess multiple subjects at same time • Contraindicated due to safety as vital signs not monitored during test and there is potential to perform a maximal test
12-Min Walk	2A Unable to Recommend At this time	• Used in 5 breast cancer research studies • MDD and SEM not defined • Reliability and Validity not determined for breast cancer survivors • Validity of CRF prediction equations not assessed in breast cancer research	• Easy to administer • Low cost • Can assess multiple subjects at same time • Contraindicated due to safety as subjects vital signs are not monitored during testing and there is potential to perform a maximal test
1-Mile Walk	2A Unable to Recommend At this time	• Used in breast cancer research • MDD and SEM not defined • Validity of CRF prediction equations not assessed in breast cancer research	• Easy to administer • Can assess multiple subjects at same time • Requires vital signs monitoring for safety as potential to perform a maximal test
Step Test (Canadian Aerobic Fitness Test protocol)	2A Unable to Recommend At this time	• Used in one breast cancer research study • Psychometrics for not assessed in breast cancer research studies • Validity of CRF prediction equations not assessed in breast cancer research • Standard errors were high and the MDD exceeded the MDIC contributing to poor reliability.	• Easy to administer • Low cost • Can assess multiple subjects at same time • Requires fixed step height and cadence • Safety issues related to balance and falls, and vital signs not monitored during test leading to potential to perform a maximal test • Requires high MET levels not appropriate for sedentary or low fitness individuals

Abbreviations: CRF, cardiorespiratory fitness; MDD, minimal detectable difference; MCID, minimal clinically important difference; SEM, standard of measurement

Table 3. Concurrent Validity

Exercise Test	Population Assessed	Concurrent Validity[a]	COV or Percent Difference between Means[b]	Standard Error of the Measure[c] ml/kg/min	Minimal Detectable Differenced ml/kg/min
Maximal Tests with Oxygen Analysis					
Treadmill Test	Healthy Males & Females, Athletes[17-19]	ICC = .97	<1.5-2.0% (with Douglas Bag)	± 0.05-0.06	± 0.15 - 0.17
	Women with Breast Cancer	*Not found*	*Not found*	*Not found*	*Not found*
Cycle Ergometer	Healthy Males & Females, Athletes[16-19]	*R2* = .60-96	COV = 1.5 (1.1-2.0)	- 0.7 to 1.3 SEE = - 4.5	1.9-3.6
	Women with Breast Cancer[55] (Reported as raw difference and percent)	*Not found*	- 16.7%[B] (-4.8 ml/kg/min)	*Not found*	*Not found*
Submaximal Tests used with Prediction Equations					
Treadmill Test	Healthy Males & Females, Athletes[17-19,]	*R2 =.76-.88*	*Not found*	± 3.4 – 5.3	9.4-14.7
	Women with Breast Cancer	*n/a*	*n/a*	*n/a*	*n/a*
Cycle Ergometer	Healthy Males & Females, Athletes, Individuals with Cancer[16-19, 27, 34, 71]	*R2 =.61-.96* *R2 = .50-.62* [10]	9.4 - 16.6% -9.0%[71] (Compared to V̊O2max cycle ergometer)	SEE = 3.12-4.23	11.6-15.8
	Women with Breast Cancer	*Not found*	*Not found*	*Not found*	*Not found*
12 Minute Run	Healthy Males & Females, Athletes[17-19, 27]	*R2 = .02-.85 (not a typo)*	*n/a*	*n/a*	*n/a*
	Women with Breast Cancer	*Not found*	*Not found*	*Not found*	*Not found*
12 Minute Walk	Individuals with Heart and Lung Disease [27, 37]	*R2 = .24-.27*	*n/a*	*n/a*	*n/a*
	Women with Breast Cancer	*Not found*	*Not found*	*Not found*	*Not found*
1 Mile Walk	Healthy Males & Females[16, 34, 38-40]	ICC=.96-.97 *R2 = .62-86*	*Not found*	SEE = 5.68 SEE 0.325 L/min	*Not found*
	Women with Breast Cancer	*Not found*	*Not found*	*Not found*	*Not found*
Step Test	Healthy Males & Females, Athletes[17-19, 33-35]	*R2 = .62-*.85	-3.8%	SEM = 4.08 SEE = 2.9–4.1	11.3
	Women with Breast Cancer	*Not found*	*Not found*	*Not found*	*Not found*

[a] The Pearson Product Moment (r) was reported when Intraclass Correlation Coefficients (ICC) were not found.
[b] Percent difference was reported when the Coefficient of Variation was not found.
[c] Standard Error of the Estimate (SEE) was reported when Standard Error of the Measure (SEM) was not found.
[d] Minimal Detectable Difference (MDD) at 95% Confidence Level.

apparently healthy people is one metabolic equivalent (MET), which is equal to 3.5 ml/kg/min of oxygen.[2,42,43] The 3.5 ml/kg/min definition for the MET is lower in women than in men and it also gradually decreases with aging; however, the validity of the MCID measure has not been determined among breast cancer survivors.[2,42,43]

VO_2max Treadmill Tests

Extensive study and scientific evidence from other populations supports the VO_2max treadmill SETs as the gold standard for accurate measurement of CRF.[2,9,15,17-21] These tests are performed on a motor-driven treadmill with expired gases analyzed for O_2 and CO_2 content. Administration time is about 45 to 60 minutes including time for set-up, preparation, warm-up, testing, and cool-down.[2,9,15] There are a number of standardized protocols available for use in different populations, all involving a step-wise, progressive increase in exercise intensity until the criteria demonstrating maximal exertion the maximal criteria has been achieved.[2,17-19]

Twenty-three articles were initially found that used VO_2max treadmill testing to determine the CRF of breast cancer survivors; however, several of these papers were secondary analysis of original data sets resulting in only 14 original sets of exercise test data.[47-69] Unfortunately, the majority of the studies did not

Table 4. Test-retest Reliability: Standardized Exercise Tests

Exercise Test	Population	Test-Retest Correlation[a]	Percent Difference[b]	Standard Error of the Measure[c] ml/kg/min	Minimal Detectable Difference
Maximal Tests with Oxygen Analysis					
Treadmill Test	Healthy Males & Females[16-20,25,26, 107]	ICC = .95	2.0-5.0% COV = 0.9[D]	*2.58*[107]	*Not found*
	Women with Breast Cancer	*Not found*	*Not found*	*Not found*	*Not found*
Cycle Ergometer	Healthy Males & Females, Athletes[17-20]	*R2 = .74-94*	COV = 0.9[d]	- 8.1%	*n/a*
	Women with Breast Cancer	*Not found*	*Not found*	*Not found*	*Not found*
Submaximal Tests with Prediction Equations					
Treadmill Test	Healthy Males & Females, Athletes[17-19, 34]	*R2 = .85*	*Not found*	*Not found*	*Not found*
	Women with Breast Cancer	*Not found*	*Not found*	*Not found*	*Not found*
Cycle Ergometer	Healthy Males & Females, Athletes, individuals with Cancer[17-19, 27, 71.96]	ICC = .95-.99[17-19, 97] ICC = .873[71]	*Not found*	1.5[17-19, 27] -.02 ± 3.29[96, e]	3.32
	Women with Breast Cancer	*Not found*	*Not found*	*Not found*	*Not found*
Step Test	Healthy Males & Females[36]	*R2* = .62	*Not found*	*Not found*	*Not found*
	Women with Breast Cancer	*Not found*	*Not found*	*Not found*	*Not found*
12 Minute Run	Healthy Males & Females[27]	Φ = .96	*Not found*	1.6	4.54
	Women with Breast Cancer	*Not found*	*Not found*	*Not found*	*Not found*
12 Minute Walk	Individuals with Heart and Lung Disease [37]	ICC=.96-.99	COV= 4.2-8.6	*Not found*	*Not Found*
	Women with Breast Cancer	*Not found*	*Not found*	*Not found*	*Not found*
1 Mile Walk	Healthy Males & Females[40]	ICC = .97	*Not found*	*Not found*	*Not found*
	Women with Breast Cancer	*Not found*	*Not found*	*Not found*	*Not found*

[a] When Intraclass Correlation Coefficients (ICC) were not found; *R2* was reported.
[b] Percent difference was reported when the Coefficient of Variation was not found.
[c] Standard Error of the Estimate (SEE) was reported when Standard Error of the Measure (SEM) was not found.
[d] Calculated for peak VO2 measures, not always with oxygen analysis.
[e] Limits of Agreement.

explain whether the subjects' performance met the criteria for maximal exertion and many did not report oxygen consumption in ml/kg/min. These reporting problems have been cited in prior reviews as they prevent scientific interpretation of the results as well as calculation of psychometric measures.[9,10] Only one article provided data sufficient to demonstrate that the subjects met the VO_2max criteria, which suggests that this mode of SET is valid in this population.[57] Although the psychometric properties for VO_2max treadmill tests undertaken by healthy populations and individuals with heart disease are available, no studies were found that presented correlation coefficients for test-retest reliability, COV, SEM, MDD, or MCID specifically for VO_2max treadmill tests performed by breast cancer survivors.

VO_2max treadmill tests are the recommended method for evaluating CRF because of the accuracy of its measurements.[2,9,15,17-21] Despite its status as the gold standard measure, VO_2max treadmill testing is not recommended in national standards written for exercise testing in the clinic.[2,5-8,19] These tests are time-consuming, expensive to perform, and require specialized and costly equipment.[2,5-8] These tests also require advanced training for the test administrators and the presence of at least two clinic staff members for safety during testing.[2,5-8] Depending on the subject's physical status, a physician may be required to be on-site during testing.[2,5-8] Most importantly, national guidelines for clinical practice do not deem maximal exercise testing as being safe for many individuals when performed in the clinical setting.[2,5-8] Therefore, based on clinical efficiency and safety concerns, the *EDGE* rating for maximal treadmill tests with oxygen analysis was: *"1 - Do not Recommend."*

VO_2max Cycle Ergometer Tests

Maximal oxygen consumption testing can also be performed using a bicycle ergometer. As with maximal treadmill exercise testing, VO_2max cycle ergometer tests require subjects to perform a maximal physical effort to criterion endpoints and measures CRF (VO_2/ml/kg).[2] A mechanically-braked cycle ergometer provides standardized pedal resistance that can be used in various single and multi-stage protocols.[2] Since the subject is seated, the test can be used for individuals who are unable or are unsafe to walk on a treadmill.[2] The test takes approximately 45 to 60 minutes to administer.[2,34] In studies on apparently healthy subjects, test-retest reliability can be good if standardized testing guidelines are followed.[2,15,16] However, cycle ergometer measures are approximately 10% lower than gold standard measures, which reduces the test's concurrent validity.[2,15,20,28,46]

Fifteen studies were found that used VO_2max cycle ergometer tests to determine the CRF levels of female breast cancer survivors.[55,70-83] Only one of these studies presented the psychometric properties of maximal cycle ergometer SETs among breast cancer survivors. Dolan et al found poor concurrent validity between VO_2max cycle ergometer measures and VO_2max treadmill measures among 12 breast cancer survivors.[55] The cycle ergometer measures were 4.8 ml/kg/min lower than measures obtained from the treadmill test and this difference exceeded the MCID for CRF measures.[55]

Studies from other populations found acceptable test-retest reliability measures; however, studies from women with breast cancer and other populations found the concurrent validity with gold standard measures was poor.[16.34,55] In addition, similar to VO_2max treadmill tests, VO_2max cycle ergometer tests are not recommended in national standards for clinical exercise testing due to the poor clinical efficiency and safety concerns for patients.[2,5-8,19] Therefore, the *EDGE* rating for VO_2max cycle ergometer tests was: *"1 - Do not Recommend."*

Submaximal Exercise Tests

Because of the safety concerns associated with maximal exercise testing, exercise testing protocols requiring submaximal exertional efforts are available and widely used clinically. These tests do not involve analysis of expired gases and are stopped at a predetermined exertional level. Resulting physiologic data (heart rate, time) as well as anthropometric data are entered into derived equations or prediction equations which generate estimates of VO_2max.[2,34] Therefore, subject preparation, precise heart rate measurement, and valid prediction equations are required for estimates of CRF to be accurate.[2,28,34,39,41] Since subjects do not perform a maximal physical effort, submaximal tests offer a safer alternative to maximal testing as long as proper screening, monitoring, and patient instructions are provided.[2,5-8,34] Also, the costs associated with a metabolic cart and specialty training are not required since oxygen analysis is not performed.

Submaximal Treadmill Tests

Submaximal treadmill testing protocols often use the same protocols as VO_2max treadmill tests.[2,28,34] Testing takes approximately 20 to 30 minutes and typically the test ends when individuals reach 70% of their heart rate reserve, 85% of their predicted maximal heart rate, or they develop adverse signs and symptoms.[2,28,34] Eight published articles were found that appeared to consist of 7 original data sets derived from female breast cancer survivors who had undergone submaximal treadmill testing.[84-92] The tests primarily tracked changes in CRF following subject participation in exercise training interventions. Studies on non-breast cancer subjects found the SEM could exceed the MCID for concurrent validity and test-retest reliability if standardized testing protocols were not followed.[2,34,41] Two studies examining survivors of breast cancer found that current equations used to estimate maximum heart rate overestimated maximal heart rate by approximately 10 beats per minute.[57,58] One of these studies also suggested that reported ratings of perceived exertion were significantly lower than predicted values for specified exercise intensities.[57] A study by Evans et al examined differences in heart rate, ratings of perceived exertion, and blood lactate concentrations between women with breast cancer and apparently healthy women at various exertional intensities.[92] Results demonstrated that these physiologic measures were similar between these groups at low to moderate exertional levels, but at 70% of VO_2max blood lactate levels were significantly different, suggesting differences existed in the physiological responses of survivors of breast cancer to moderate to high intensity exercise. In turn, these findings suggest that a better understanding of these differences is needed in order to write safe, appropriate, and effective exercise prescriptions.

In summary, submaximal treadmill tests have been used in breast cancer studies to provide an estimate of CRF and are relatively easy to administer and safe with proper screening and monitoring.[1,2,11,83] These tests offer improved clinical utility relative to time spent in testing, equipment costs, and specialty training. However, psychometric measures may be unacceptable and the equations used to predict CRF require further validation. Therefore, the *EDGE* rating for submaximal treadmill tests was: *"2A - Unable to Recommend at this time."*

Submaximal Cycle Ergometer Tests

Submaximal cycle ergometer tests are also available and utilize prediction equations to estimate CRF and to determine test endpoints.[2,16,28,34] Therefore, the costs of a gas analysis equipment and specialty training are eliminated. Because these are submaximal tests, subjects are not required to perform a maximal effort; therefore, with proper screening, the tests are safer to perform.[2,34] The test is clinically efficient and takes approximately 20 to 30 minutes to administer.[2,16,28,34]

Five studies were found that used submaximal cycle ergometer tests in breast cancer research and two studies measured psychometric properties.[71,78,93-95] Debacker et al compared measures from a submaximal cycle ergometer test to a maximal short exercise capacity test (steep ramp test) and a VO_2max cycle ergometer test to detect changes in CRF after an 18-week training program among male and female cancer survivors.[71] The submaximal cycle ergometer tests proved invalid for detecting changes in CRF following exercise training compared with the two maximal tests (r = .71 to .79*)*. The study did calculate test-

retest reliability for the submaximal cycle ergometer test in this population as good as ICC = .873 (CI_{95} =.72 to .95). A subsequent study by May et al also examined the sensitivity of submaximal cycle ergometer tests to detect changes in CRF following training in subjects with cancer.[78] The study found a modest, but poor correlation [r = -.51 (p = .006)] for the sensitivity of submaximal cycle ergometer measures compared with VO_2max cycle measures. The subjects in both the Debacker et al and the May et al studies had a variety of cancer diagnoses and data was reported in aggregate, so results specific for female breast cancer survivors were not available.

One additional study was found that examined test-retest reliability for the submaximal cycle ergometer test with oxygen analysis in subjects with lymphohematopoietic cancers.[96] Although this study did not include women with breast cancer, it did provide psychometrics for test-retest reliability for a submaximal cycle ergometer test of ICC = .96 among individuals with cancer which was excellent. Limits of agreement (LOA) and random error measures were also provided as LOA = -.02 ± 3.29 which can exceed the MCID for CRF measures.

Submaximal cycle ergometer tests have been used in breast cancer research and they appear to offer improved safety with proper screening. The tests also have good clinical utility relative to equipment cost and administration time.[74,83] There is also some psychometric evidence collected from women with breast cancer.[71, 78] Currently available evidence suggests test-retest reliability for submaximal cycle ergometers is moderate to good; however, standard errors continue to be large which may contribute to inappropriate categorization of CRF and subsequent errors in the exercise prescription intensities.[71,78] Both studies also suggest that submaximal cycle ergometer tests lack sensitivity to identify changes in CRF following training.[71,78] Finally, similar to submaximal treadmill tests, ratings of perceived exertion and prediction equations for test endpoints and CRF have not been validated in the breast cancer population. The *EDGE* rating for the submaximal cycle ergometer test was: *"2A - Unable to Recommend at this Time."*

Field Tests

Field tests also use derived prediction equations to estimate CRF using distance, time, or heart rate responses during walking or running tests.[2,34] Common field tests include the 12 minute walk (12-MWT) or run test, the 1.5 mile walk or run test, and the 1 mile walk test (1-MWT).[2,34,38-40] Field tests are easy to administer, require little equipment, and often can be used to assess multiple individuals simultaneously.[2,34] The prediction equations that estimate CRF from field test outcomes may require factors such as age, gender, weight, and training status to improve the estimate's accuracy; however, the prediction equations must be validated within each specific population.[2,34,38-40] Field tests suffer from the fact that vital signs are not monitored during testing and individuals with low fitness levels or cardiovascular risk may perform a maximal physical effort and compromise their safety.[2,34]

Twelve Minute Walk/Run Tests

The 12 minute walk and run tests use prediction equations based on distance covered in 12 minutes to estimate CRF.[2,34] The objective of 12 min walk and run tests is to cover as much distance as possible within the allotted time. Cooper's 12 minute run test was the original assessment tool of this type and it was developed using young athletic males.[2,34] The 12-MWT is a modification of Cooper's run test; however, the walk test may be more appropriate for individuals in the rehabilitation setting.[2, 34] Both tests require a stop-watch or clock and a track with marked distances. The test-retest reliability for the 12 minute run test was reported as r =.90; however, concurrent validity is not clear as values range from r = .13 to .90.[2,27,34] Since vital signs are not monitored during either test, there are safety concerns for individuals with low fitness or cardiovascular risk factors.[2,34] There were no articles found that validated the CRF prediction equations in breast cancer survivors for either of these tests.

No articles were found that used or validated the 12 minute run test in breast cancer survivors. The test is also not recommended for individuals with low fitness, cardiovascular risks, or in rehabilitation due to safety concerns. Therefore the *EDGE* rating for the 12 minute run test was: *"1 – Do not recommend."*

Five articles were found that used the 12-MWT among breast cancer survivors.[97-101] The studies all referenced the concurrent validity and reliability given for the Cooper 12 minute run test, but no articles were found that provided psychometrics for the 12-MWT administered specifically to breast cancer survivors. Since there were no psychometric measures or validated prediction equations for the 12-MWT for use with breast cancer survivors and there are safety concerns related to lack of vital signs measurement during testing, the *EDGE* rating for the 12-MWT was: *"2A - Unable to Recommend at this Time."*

Distance Walk or Run Tests

The 1-MWT, also known as the Rockport Test, predicts CRF fitness categories based on the heart rate obtained at or near the end of the test. This test has been used and reported with walking and running and has variations in the distances covered including: 1.0, 1.5, and 2.0 miles. This test appears to have good concurrent validity (ICC = .96-.97) and test-retest agreement values (ICC = .97) from other populations, although the SEM is not known. The prediction equations derived from this test have not been validated among breast cancer survivors. The test may also have issues with safety since since again vital signs are not monitored during testing.

Two studies were found that used a 1-MWT in breast cancer research, although it appeared that the same data set was used for the two different analyses.[80,104] The test appears to be clinically efficient and may be safe for the breast cancer survivors as long as subjects are appropriately screened and vital signs are monitored during testing for safety.[1,9,34] The test appears to have good concurrent validity and test-retest reliability in other populations; however, the prediction equations have not been validated for use with breast cancer survivors. The 1-MWT may become a clinically useful assessment; however, at the present time, the

evaluators did not feel there was adequate evidence available to recommend this test. Therefore, the *EDGE* rating for the 1-MWT was *"2A - Unable to Recommend at this Time."*

Step Tests

Step tests estimate CRF through equations that use the heart rate response to stepping up and down stairs of a fixed height at a fixed stepping rate.[2,34] The test requires a criterion height step, a metronome or audio cadence signal, and a stopwatch. The text takes only a few minutes to perform and several subjects can be tested at the same time. The test has good clinical utility since it is low cost, requires minimal equipment, and is time-efficient. Based on data from other populations, the test has moderate to weak agreement with gold standard measures and the test-retest reliability exceeds the MCID for low accuracy. The test also requires energy expenditures of 7-9 METs, rendering it unsafe or not possible for individuals with low fitness or cardiovascular disease to perform.[2,34,103] The requirement for stepping up and down from a specific height poses performance and safety concerns for individuals with musculoskeletal, balance, or sensory deficits (peripheral neuropathies) in the lower extremities.

One study was found that used a step test [Canadian Aerobic Fitness Test (CAFT)] to assess CRF of women with breast cancer.[102] However, no studies were found which presented the psychometric properties of step tests or the CAFT in the breast cancer survivor population. The test is safe and clinically efficient as long as individuals have a high fitness level, but issues with accuracy and safety are a concern in the clinical setting. The *EDGE* rating for the step test was: *"2A - Unable to Recommend at this Time."*

Exercise Testing and Training Safety

Two articles were found that examined the safety of maximal cycle ergometer tests for women with breast cancer. Hornsby et al found 12 non-significant ECG changes and 3 non-life threatening events during 30 VO_2max cycle ergometer tests performed by 10 women in treatment for breast cancer.[74] Jones et al examined the safety of VO_2max cycle ergometer test among 85 individuals with advanced cancers, 39 of whom had metastatic breast cancer.[83] One female subject experienced a significant ECG change (ST segment depression) and another experienced an exercise-induced, asymptomatic right bundle branch block that normalized at test termination. Both studies concluded that exercise testing is relatively safe and effective for use in research among female breast cancer survivors as long as proper screening, and vital signs and ECG monitoring are performed. A few additional articles and reviews provided summary information based on the use of exercise testing performed in research on breast cancer survivors that supported the safety of exercise testing and training as long as standardized screening and monitoring is performed.[1,70,78]

DISCUSSION

The purpose of this study was to evaluate the psychometric properties, safety, and clinical utility of tests used to assess or estimate CRF for their use in clinical practice among women with breast cancer. A systematic literature search yielded only 5 studies that assessed psychometric values for CRF measures among women with breast cancer and two studies that assessed safety.[55-58,71,74,78,83]

Three studies were found that reported the psychometric measures derived from maximal tests and involving analysis of expired gases. One study provided evidence that women undergoing treatment for breast cancer could meet the criteria for performing valid VO_2max treadmill tests.[57] This study and a follow up study also found that currently used heart rate prediction equations significantly overestimate maximal heart rates for women with breast cancer.[57,5] In the event that maximal heart rates are being overestimated with current equations, exercise test endpoints and exercise prescriptions would be calculated at higher than desirable intensities and pose safety and motivation concerns. A third study estimated the concurrent validity of VO_2max cycle ergometer tests and found CRF measures were 4.8 ml/kg/min lower, which exceeded the MCID of 3.5 ml/kg/min. This suggests that individuals would be incorrectly placed in lower CRF fitness categories and subsequently receive less effective exercise prescriptions, particularly if they were not exercising on a cycle ergometer.

Two studies were found that examined submaximal cycle ergometer psychometric measures.[71,78] Both studies found that submaximal cycle ergometer measures were not sensitive to detect changes in CRF following an exercise training intervention.[71,78] Additionally, one of the studies did find that submaximal cycle ergometer test-retest reliability (ICC = .873) was acceptable, although lower than values reported for apparently healthy subjects (ICC = .95-.99).

This evaluation found very few studies on the psychometric properties of CRF measures among women with breast cancer. Furthermore, the psychometrics determined from other populations for tests with good clinical utility were not sufficiently valid or reliable.[2,16,34,38,40] Therefore, the highest ratings given in this evaluation were: *"2A - Unable to recommend at this time."* This rating was given to submaximal treadmill and cycle ergometer tests, step tests, the 12-MWT, and the 1-MWT. While the step test may not be appropriate for all individuals, the decision to use this test should be made by the clinician based on the individual's evaluation results. Three of the CRF measures received a rating of: *"1 – Do not recommend."* The CRF measures that received this rating were the VO_2max cycle ergometer and treadmill tests, due to poor clinical utility and safety concerns based on national guidelines.[2,5-8] The 12 minute run test also received a rating of *"1 – Do not recommend,"* since it had not been used in research studies involving breast cancer survivors, the psychometrics were not adequately determined, and due to safety concerns related to the potential to perform a maximal test.[2,27,34]

Despite the current ratings presented in this evaluation, physical therapist practice guidelines stipulate that some form of screening or testing be performed prior to developing activity or exercise prescriptions.[5] Recommendations from other studies and reviews suggest that current guidelines developed for apparently

healthy subjects and individuals with cardiovascular and pulmonary conditions be used to guide exercise screening and testing among individuals with cancer; however, the articles also recommend further research to support the validity, efficacy, and safety of CRF testing for women with breast cancer.[1,9,11,74,83]

In terms of clinical practice, following professional guidelines for screening and testing will promote as much as currently possible, valid, reliable and safe results.[9] However, in the event that exercise screening or testing protocols require modifications to accommodate the patient's needs or preferences, these adaptations should be noted in the physical therapy report to enable accurate test replication and interpretation at a subsequent time.[5] To promote accuracy and safety, clinicians need to appreciate screening and testing risks and be aware of the individual's health status and possible cancer treatment co-morbidities. Some of these conditions include: pain, nausea, dehydration, lymphedema or shoulder pathology, breathing or movement difficulties related to the surgery, reconstruction methods that may compromise core-stability and place the individual at risk for back pain, cognitive deficits (chemo-brain) or depression that may affect adherence to exercise or judgment, a prior sedentary lifestyle, cardiovascular toxicity from cancer medications, chest wall or pulmonary fibrosis from radiation treatments, metastasis to the bone which may result in fractures or nerve compression, and secondary cancers such as leukemia following treatments.[13,14,105,106] Furthermore, since breast cancer treatments combined with aging increase a woman's risk for cardiovascular disease, it is recommended that CRF measures include vital signs monitoring to improve safety or uncover abnormal responses to exercise.[2,5,14]

Finally, since submaximal exercise tests can have large SEM and MDD that exceed the MCID, clinicians may wish to recognize that measures of CRF in female breast cancer survivors may provide outcomes that over- or underestimate their CRF. Under estimating CRF would lead to an ineffective activity prescription while overestimating CRF could lead to excessive fatigue, discouragement, and disincentive to exercise. Furthermore, overestimating the exercise intensity could lead to excessive cardiovascular demands and promote arrhythmia, dyspnea, or even death.[2] Recognizing the limitations of CRF measures, clinicians can promote safe and effective exercise training through patient education on normal and abnormal exercise responses, information on when to contact the physical therapist regarding modifications to the exercise prescription, or when to contact a physician or seek emergent care.

Limitations

The search process was limited to articles written or translated into the English language; therefore, the search process has not been exhaustive. A second limitation was the small number of articles found on the validity and reliability of CRF tests in female breast cancer survivors. This lack of evidence prevented substantive and comprehensive evaluations and judgments. Confidence in the findings was also limited by the quality of the reported data which has been addressed in detail in prior reviews.[9,10] Improved reporting will allow for successful future evaluation studies and clinical determinations.

Future Study Recommendations

Recommendations for research on the psychometric properties of SETs used to assess the CRF in women with breast cancer would include the following. First, additional studies are required that determine concurrent validity, test-retest reliability, and sensitivity to changes in CRF measures following an intervention. Second, cardiorespiratory fitness must be reported as oxygen consumption in milliliters per kilogram per minute since this measure can be compared over time for a single individual, as well as across subjects. Third, SET endpoints need to be clearly reported so that readers will know that the study subjects met the required test criteria for accurate outcomes. Third, the proper statistical analyses need to be used in order to have confidence in the findings. The ICC or the COV are the recommended for concurrent validity and test-retest reliability; however, the SEM and MDD are also necessary to assess clinical relevance.[15,44,46] Fourth, studies on concurrent validity require that comparisons be made against gold standard measures as comparisons between two non-criterion tests compounds errors. Fifth, to establish reasonable precision for test-retest reliability, the recommendation is to have 50 subjects perform 3 trials each approximately 2.5 days apart.[15,20] Finally, prediction equations that estimate CRF, maximal heart rates, and perceived exertion ratings require validation within this population.[57,58] Ideally, evidence for the psychometric properties of SETs would be useful across all stages of breast cancer, all types of treatments, and through all survivorship periods.

CONCLUSION

This study evaluated the psychometric properties, safety, and clinical utility of SETs used to determine or estimate CRF in breast cancer survivors and provided informed recommendations for their use in clinical practice among women with breast cancer. However, the evaluation found limited evidence on the psychometric properties of CRF tests for this population. Suggestions were made for adapting the current evidence for testing into clinical practice as well as for designing future research on this topic among women with breast cancer.

ACKNOWLEDGEMENTS

From the Oakland University School of Health Sciences Physical Therapy Program: Allegra Adams, DPT; Kendra Sereno, DPT; Luiza Iervolino, DPT; Jeffrey Duncan, SPT; Kristina Hylland SPT, Breanne Walkowiak, SPT; and Katie Peterson SPT for assistance with the literature search. Kenneth Hightower, PhD, Dean of the School of Health Sciences (retired); Richard Rozek, PhD, Interim Dean of the School of Health Sciences; Kristine Thompson, PT, PhD, Director of the Physical Therapy Program; and Julia Rodriguez, MS, Nursing, Health Sciences & Faculty Research Support Librarian for consulting and support.

REFERENCES

1. Schmitz KH, Courneya KS, Matthews C, et al. American College of Sports Medicine's roundtable on exercise guidelines for cancer survivors. *Med Sci Sports Exerc*. 2010;42(7):1409-1426.

2. Pescatello LS, ed. *American College of Sports Medicine: ACSM's Guidelines for Exercise Testing and Prescription.* 9th ed. Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins Health; 2014.
3. Jones LW, Eves ND, Haykowsky M, Freedland SJ, Mackey JR. Exercise intolerance in cancer and the role of exercise therapy to reverse dysfunction. *Lancet Oncol*. 2009;10(6):598-605.
4. Scott JM, Lakoski S, Mackey JR, Douglas PS, Haykowsky MJ, Jones LW. The potential role of aerobic exercise to modulate cardiotoxicity of molecularly targeted cancer therapeutics. *Oncologist*. 2013;18(2):221-231.
5. American Physical Therapy Association. *Guide to Physical Therapist Practice.* 3rd ed. Alexandria, VA: American Physical Therapy Association; 2014. http://guidetoptpractice.apta.org/. Accessed September 1, 2014.
6. American Thoracic Society/American College of Chest Physicians. ATC/ACCP statement on cardiopulmonary exercise testing. *Am J Respir Crit Care Med*. 2003;167(2):211-277.
7. Balady GJ, Arena R, Sietsema K, et al. Clinician's guide to cardiopulmonary exercise testing in adults: A scientific statement from the American Heart Association. *Circulation*. 2010;122(2):191-225.
8. Myers J, Arena R, Franklin B, et al. Recommendations for clinical exercise laboratories: A scientific statement from the American Heart Association. *Circulation*. 2009;119(24):3144-3161.
9. Jones LW, Eves ND, Haykowsky M, Joy AA, Douglas PS. Cardiorespiratory exercise testing in clinical oncology research: Systematic review and practice recommendations. *Lancet Oncol*. 2008;9(8):757-765.
10. Steins Bisschop CN, Velthuis MJ, Wittink H, et al. Cardiopulmonary exercise testing in cancer rehabilitation: A systematic review. *Sports Med*. 2012;42(5):367-379.
11. Jones LW, Eves ND, Peppercorn J. Comment: Pre-exercise screening and prescription guidelines for cancer patients. *Lancet Oncol*. 2010:Published online.
12. Lakoski SG, Eves ND, Douglas PS, Jones LW. Exercise rehabilitation in patients with cancer. *Nat Rev Clin Oncol*. 2012;9|:288-296.
13. Smith LA, Cornelius VR, Plummer CJ, et al. Cardiotoxicity of anthracycline agents for the treatment of cancer: Systematic review and meta-analysis of randomised controlled trials. *BMC Cancer*. 2010;10:337-351.
14. Jones LW, Haykowsky MJ, Swartz JJ, Douglas PS, Mackey JR. Early breast cancer therapy and cardiovascular injury. *J Am Coll Cardiol*. 2007;50(15):1435-1441.
15. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med*. 2000;30:375-381.
16. Akalan C, Robergs RA, Kravitz L. Prediction of VO_2max from an individualized submaximal cycle ergometer. *J Exerc Physiol*. 2008;11(2):1-17.
17. Howley ET, Bassett DR, Welch HG. Criteria for maximal oxygen uptake: Review and commentary. *Med Sci Sports Exerc*. 1995;27(9):1292-1301.
18. Taylor HL, Buskirk E, Henschel A. Maximal oxygen intake as an objective measure of cardio-respiratory performance. *J Appl Physiol*. 1955;8:73-80.
19. Shephard RC, Allen C, Benade AVS, et al. The maximal oxygen intake. an international reference standard of cardiorespiratory fitness. *Bull WHO*. 1968;38:757-764.
20. Hopkins WG, Schabort EJ, Hawley JA. Reliability of power in physical performance tests. *Sports Med*. 2001;31(3):211-234.
21. Bruce RA, Fisher LD, Cooper MN, Gey GO. Separation of effects of cardiovascular disease and age on ventricular function with maximal exercise. *Am J Cardiol*. 1974;34:757-763.
22. Bruce RA, McDonough JR. Stress testing in screening for cardiovascular disease. *Bull N Y Acad Med*. 1969;45(12):1288-1305.
23. Naughton J, Balke B, Nagle F. Refinements in method of evaluation and physical conditioning before and after myocardial infarction. *Am J Cardiol*. 1964;14:837-843.
24. Bruce RA, Gey GO, Cooper MN, Fisher LD, Peterson DR. Seattle heart watch: Initial clinical circulatory and electrocardiographic responses to maximal exercise. *Am J Cardiol*. 1974;33(4):459-470.
25. Froelicher Jr VF, Brammell H, Davis G, Noguera I, Stewart A, Lancaster MC. A comparison of the reproducibility and physiologic response to three maximal treadmill exercise protocols. *Chest*. 1974;65(5):512-517.
26. Balke B, Ware R. An experimental study of air force personnel. *US Armed Forces Med J*. 1959;10:675-688.
27. Penry JT, Wilcox AR, Yun J. Validity and reliability analysis of Cooper's 12-minute run and the multistage shuttle run in healthy adults. *J Strength Cond Res*. 2011;25(3):597-605.
28. George JD, Vehrs PR, Babcock GJ, Etchie MP, Chinevere TD, Fellingham GW. A modified submaximal cycle ergometer test designed to predict treadmill VO_2max. *Meas Phys Educ Exerc Sci*. 2000;4(4):229-243.
29. Hawkins SA, Wiswell RA. Rate and mechanism of maximal oxygen consumption decline with aging: Implications for exercise training. *Sports Med*. 2003;33(12):877-888.
30. Ogawa T, Spina R, Martin III WH, et al. Effects of aging, sex, and physical training on cardiovascular responses to exercise. *Circulation*. 1992;86:494-503.
31. Eskurza I, Donato AJ, Moreau KL, et al. Changes in maximal aerobic capacity with age in endurance-trained women: 7-year follow-up. *J Appl Physiol*. 2002;92:2303-2308.
32. Pollock ML, Foster C, Knapp D, et al. Effect of age and training on aerobic capacity and body composition of master athletes. *J Appl Physiol*. 1987;62:725-731.
33. Singh VN. The role of gas analysis with exercise testing. *Prim Care*. 2001;28(1):159-179.
34. Noonan V, Dean E. Submaximal exercise testing: Clinical application and interpretation. *Phys Ther*. 2000;80:782-807.
35. McArdle WD, Katch FI, Pechar GS. Reliability and interrelationships between maximal oxygen intake, physical work

capacity and step-test scores in college women. *Med Sci Sports Exerc*. 1972;4(182-186).
36. Jette M, Campbell J, Mongeon J, Routhier R. The Canadian Home Fitness Test as a predictor for aerobic capacity. *Can Med Assoc J*. 1976;114:680–682.
37. Mungall PF, Hainsworth R. Assessments of respiratory function in patients with chronic obstructive airway disease. *Thorax*. 1979;34:254–258.
38. Kline GM, Porcari JP, Hintermeister R, et al. Estimation of VO_2max from a one-mile track walk, gender, age, and body weight. *Med Sci Sports Exerc*. 1987;19:253–259.
39. Coleman AE. Validation of a submaximal test of maximal oxygen intake. *J Sports Med Phys Fitness*. 1976;16:106-111.
40. Hageman PA, Walker SN, Pullen CH, Pellerito, P. Test-retest reliability of the Rockport fitness walking test and other fitness measures in women ages 50-69 years. *J Geriatr Phys Ther*. 2001;24(2):7-11.
41. Marsh C. Evaluation of the ACSM submaximal treadmill running test for predicting VO_2max. *J Strength Cond Res*. 2012;26(2):548-554.
42. Ainsworth BE, Haskell WL, Herrmann SD, et al. 2011 compendium of physical activities: A second update of codes and MET values. *Med Sci Sports Exerc*. 2011;43(8):1575-1581.
43. Heinemann AW. Rehabilitation measures database: The rehabilitation clinician's place to find the best instruments to screen patients and monitor their progress. http://www.rehabmeasures.org/rehabweb/rhstats.aspx. Updated 2010. Accessed July 21, 2014, 2014.
44. Weir JP. Quantifying test-retest reliability using the Intraclass Correlation Coefficient and the *SEM*. *J Strength Cond Res*. 2005;19(1):231-240.
45. Kottner J, Audige L, Brorson S, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *J Clin Epidemiol*. 2011;64:96-106.
46. Hopkins WG, Hawley JA, Burke LM. Design and analysis of research on sport performance enhancement. *Med Sci Sports Exerc*. 1999;31:472-85.
47. Bertz H, Dimeo FC, Kanz L, Keul J, Mertelsmann R, Tilmann MHM. Aerobic exercise in the rehabilitation of cancer patients after high dose chemotherapy and autologous peripheral stem cell transplantation. *Cancer*. 1997;79(9):1717.
48. Burnham TR, Wilcox A. Effects of exercise on physiological and psychological variables in cancer survivors. *Med Sci Sports Exerc*. 2002;34:1863-7.
49. Courneya KS, McKenzie DC, Mackey JR, et al. Moderators of the effects of exercise training in breast cancer patients receiving chemotherapy: A randomized controlled trial. *Cancer*. 2008;112:1845-53.
50. Courneya KS, Segal RJ, Mackey JR, et al. Effects of aerobic and resistance exercise in breast cancer patients receiving adjuvant chemotherapy: A multicenter randomized controlled trial. *J Clin Oncol*. 2007;25(28):4396-4404.
51. Dimeo FC. Aerobic exercise following bone marrow transplantation. *Rehabil Oncol*. 2000;18(3):6-7.
52. Dimeo FC, Stieglitz RD, Novelli-Fischer U, Fetscher S, Mertelsmann R, Keul J. Correlation between physical performance and fatigue in cancer patients. *Ann Oncol*. 1997;8(12):1251-1255.
53. Dimeo FC, Fetscher S, Lange W, Mertelsmann R, Keul J. Effects of aerobic exercise on the physical performance and incidence of treatment-related complications after high-dose chemotherapy. *Blood*. 1997;90(9):3390-3394.
54. Dimeo FC, Bertz H, Finke J, Fetscher S, Mertelsmann R, Keul J. An aerobic exercise program for patients with haematological malignancies after bone marrow transplantation. *Bone Marrow Transplant*. 1996;18(6):1157-1160.
55. Dolan LB, Lane K, McKenzie DC. Optimal mode for maximal aerobic exercise testing in breast cancer survivors. *Integr Cancer Ther*. 2012;11(4):321-326.
56. Dolan LB, Gelmon K, Courneya KS, et al. Hemoglobin and aerobic fitness changes with supervised exercise training in breast cancer patients receiving chemotherapy. *Cancer Epidemiol Biomarkers Prev*. 2010;19:2826-32.
57. Drouin JS, Birkmeier L, Jones K, Miller S, Gellish R, Goslin BR. Evaluation of maximal heart rate prediction equations for women during breast cancer treatment: A measurement focused study. *Rehabil Onc*. 2013;31(2):11-16.
58. Harbac RM, Drouin JS, Kerrigan D, Gellish RL, Goslin BR. Evaluation of maximal heart rate prediction equations for individuals during and following cancer treatment: A measurement focused study. *Midwest ACSM Conference Proceedings*. 2012.
59. Drouin JS, Birk TJ, Wirth JC. Random control clinical trial on effect of aerobic exercise training on weight management during radiation treatment for breast cancer. *Rehabil Oncol*. 2006;24(3):6-10.
60. Drouin JS, Young TJ, Beeler J, et al. Random control clinical trial on the effects of aerobic exercise training on erythrocyte levels during radiation treatment for breast cancer. *Cancer*. 2006;107(10):2490-2495.
61. Drouin JS, Armstrong H, Krause S, et al. Effects of aerobic exercise training on peak aerobic capacity, fatigue, and psychological factors during radiation for breast cancer. *Rehabil Oncol*. 2005;23(1):11-17.
62. Fairey AS, Courneya KS, Field CJ, Bell GJ, Jones LW, Mackey JR. Effects of exercise training on fasting insulin, insulin resistance, insulin-like growth factors, and binding proteins in postmenopausal breast cancer survivors: A randomized controlled trial. *Cancer Epidemiol Biomarkers Prev*. 2003;12:721-727.
63. Fairey AS, Courneya KS, Field CJ, et al. Effect of exercise training on C-reactive protein in postmenopausal breast cancer survivors: A randomized controlled trial. *Brain Behav Immun*. 2005;19:381-388.
64. Fairey AS, Courneya KS, Field CJ, Bell GJ, Jones LW, Mackey JR. Randomized controlled trial of exercise and blood immune function in postmenopausal breast cancer survivors. *J Appl Physiol*. 2005;98:1534-1540.

65. Hutnick NA, Williams NI, Kraemer WJ, et al. Exercise and lymphocyte activation following chemotherapy for breast cancer. *Med Sci Sports Exerc*. 2005;37(11):1827-1835.
66. Kim C, Kang D, Smith BA, Landers KA. Cardiopulmonary responses and adherence to exercise in women newly diagnosed with breast cancer undergoing adjuvant therapy. *Cancer Nurs*. 2006;29(2):156-165.
67. Rahnama N, Nouri R, Rahmaninia F, Damirchi A, Emami H. The effects of exercise training on maximum aerobic capacity, rest heart rate, blood pressure and anthropometric variables of postmenopausal women with breast cancer. *J Res Med Sci.* 2010;15(2):78-83.
68. Garner D, Erck EG. Effects of aerobic exercise and resistance training on stage I and II breast cancer survivors: A pilot study. *Am J Health Educ*. 2008;39(4):200-205.
69. Cheema BSB, Gaul C. Full-body exercise training improves fitness and quality of life in survivors of breast cancer. *J Strength Cond Res*. 2006;20(1):14-21.
70. Courneya KS, Mackey JR, Bell GJ, et al. Randomized controlled trial of exercise training in postmenopausal breast cancer survivors: Cardiopulmonary and quality of life outcomes. *J Clin Oncol*. 2003;21:1600-1608.
71. De Backer IC, Schep G, Hoogeveen A, Vreugdenhil G, Kester AD, van Breda E. Exercise testing and training in a cancer rehabilitation program: The advantage of the steep ramp test. *Arch Phys Med Rehabil*. 2007;88(5):610-616.
72. Herrero F, San Juan AF, Fleck SJ, et al. Effects of detraining on the functional capacity of previously trained breast cancer survivors. *Int J Sports Med*. 2007;28:257-264.
73. Herrero F, San Juan AF, Fleck SJ, et al. Combined aerobic and resistance training in breast cancer survivors: A randomized, controlled pilot trial. *Int J Sports Med*. 2006;27:573-580.
74. Hornsby WE, Douglas PS, West MJ, Kenjale AA, et al. Safety and efficacy of aerobic training in operable breast cancer patients receiving neoadjuvant chemotherapy: A phase II randomized trial. *Acta Oncol*. 2014;53(1):65-74.
75. MacVicar MG, Winningham ML, Nickel JL. Effects of aerobic interval training on cancer patients' functional capacity. *Nurs Res*. 1989;38:348-351.
76. Winningham ML, MacVicar MG, Bondoc M, Anderson JI, Minton J. Effect of aerobic exercise on body weight and composition in patients with breast cancer on adjuvant chemotherapy. *Oncol Nurs Forum*. 1989;16:683-689.
77. Winningham ML, MacVicar MG. The effect of aerobic exercise on patient reports of nausea. *Oncol Nurs Forum*. 1988;15:447-450.
78. May AM, van Weert E, Korstjens I, Hoekstra-Weebers JE, et al. Monitoring training progress during exercise training in cancer survivors: A submaximal exercise test as an alternative for a maximal exercise test? *Arch Phys Med Rehabil.* 2010;91(3):351-357.
79. Neil SE, Klika RJ, Garland SJ, McKenzie DC, Campbell KL. Cardiorespiratory and neuromuscular deconditioning in fatigued and non-fatigued breast cancer survivors. *Support Care Cancer*. 2013;21:873-881.
80. Pinto BM, Clark MM, Maruyama NC, Feder SI. Psychological and fitness changes associated with exercise participation among women with breast cancer. *Psychol Oncol.* 2003;12(2):118-126.
81. Schmidt K, Vogt L, Thiel C, Jäger E, Banzer W. Validity of the six-minute walk test in cancer patients. *Int J Sports Med*. 2013;34:631–636.
82. Wood WW, Dea AM, Reeve BB, et al. Cardiopulmonary fitness in patients undergoing hematopoietic SCT: A pilot study. *Bone Marrow Transplant.* 2013:1-8.
83. Jones LW, Eves ND, Mackey JR, et al. Safety and feasibility of cardiopulmonary exercise testing in patients with advanced cancer. *Lung Cancer*. 2007;55:225-232.
84. Battaglini CL, Bottaro M, Dennehy C, et al. A randomized study on the effects of a prescribed exercise interventionon lean mass and fatigue changes in breast cancer patients during treatment. *Sao Paulo Med. J*. 2007;125:22-28.
85. Courneya KS, Friedenreich CM, Sela RA, Quinney HA, Rhodes RE. Correlates of adherence and contamination in a randomized controlled trial of exercise in cancer survivors: An application of the theory of planned behavior and the five factor model of personality. *Ann Behav Med*. 2002;24:257–268.
86. Courneya KS, Friedenreich CM, Sela RA, Quinney HA, Rhodes RE, Handman M. The group psychotherapy and home-based physical exercise (group-hope) trial in cancer survivors: Physical fitness and quality of life outcomes. *Psychooncol*. 2003;12:357-374.
87. Hsieh CC, Sprod LK, Hydock DD, Carter SD, Hayward R, Schneider CM. Effects of a supervised exercise intervention on recovery from treatment regimens in breast cancer survivors. *Oncol Nurs Forum*. 2008;35(6):909-915.
88. Kolden GG, Strauman TJ, Ward A, et al. A pilot study of group exercise training (GET) for women with primary breast cancer: Feasibility and health benefits. *Psychol Oncol.* 2002;11(5):447-456.
89. Nieman DC, Cook VD, Henson DA, et al. Moderate intensity exercise training and natural killer cell cytotoxic activity in breast cancer patients. *Int J Sports Med*. 1995;16(5):334-337.
90. Schneider CM, Hsieh CC, Sprod LK, Carter SD, Hayward R. Effects of supervised exercise training on cardiopulmonary function and fatigue in breast cancer survivors during and after treatment. *Cancer*. 2007;110(4):918-925.
91. Sprod LK, Hsieh CC, Hayward R, Schneider CM. Three versus six months of exercise training in breast cancer survivors. *Breast Cancer Res Treat*. 2010;121:413-419.
92. Evans ES, Battaglini CL, Groff DG, Hackney AC. Aerobic exercise intensity in breast cancer patients: A preliminary investigation. *Integr Cancer Ther*. 2009;8:139-147.
93. Adamsen L. Effect of a multimodal high intensity exercise intervention in cancer patients undergoing chemotherapy: Randomised controlled trial. *BMJ*. 2009;339:b3410.

(Continued on page 51)

Oncology Section EDGE Task Force on Prostate Cancer Outcomes:
A Systematic Review of Clinical Measures of Strength and Muscular Endurance

Mary Insana Fisher, PT, PhD, OCS, CLT[1]; Claire C. Davies, PT, PhD, CLT[2]; Genevieve Colon, SPT[3]; Hannah Geyer, SPT[4]; Lucinda Pfalzer, PT, PhD, FACSM, FAPTA[5]

[1]Assistant Professor, Department of Physical Therapy, University of Dayton, Dayton, OH
[2]Physical Therapist, Rehabilitation Services, Baptist Health Lexington, Lexington, KY
[3]Student Physical Therapist, University of Michigan – Flint, Flint, MI
[4]Student Physical Therapist, University of Dayton, Dayton, OH
[5]Professor Emerita, Physical Therapy Department, University of Michigan – Flint, Flint, MI

ABSTRACT

Background: Strength deficits are a common morbidity following treatment for prostate cancer. Accurate assessment of strength and muscular endurance following prostate cancer treatments is essential to identify deficits and plan rehabilitation. **Purpose:** To identify strength and muscular endurance outcome measures that possess strong psychometric properties and are clinically useful for examination of men treated for prostate cancer. **Methods:** Multiple electronic databases were searched for articles published after 1995. Studies of tools used to assess strength and muscular endurance were included if they reported psychometric properties, were clinically feasible methods, performed on adults, and published in the English language. Each outcome measure was independently reviewed and rated by two reviewers. A single Cancer EDGE Task Force Outcome Measure Rating Form was completed for each category of strength or endurance assessment, and a recommendation was made using the 4-point Cancer EDGE Task Force Rating Scale. **Results:** Of the original 683 articles found, 30 were included in this review. Hand-grip strength and hand-held dynamometry were rated 3, recommended for clinical use. One repetition maximum was rated 2A, unable to recommend at this time but the measure has been used in research on individuals with prostate cancer. Manual muscle testing was rated 2B, unable to recommend at this time due to lack of psychometric support, and muscular endurance testing was not recommended (1). **Conclusions:** Utilizing objective dynamometry for hand grip and muscle strength testing provides precise measurement to assess baseline status and monitor change among men treated for prostate cancer.

Address correspondence to: Mary I. Fisher, PT, PhD, OCS, CLT, Department of Physical Therapy, University of Dayton, 300 College Park, Dayton, OH 45469, Ph: (937) 229-5617, Fax: (937) 229-5601 (mary.fisher@udayton.edu).

Key Words: prostate neoplasms, muscle performance, dynamometry, outcome measures, psychometrics

INTRODUCTION

Prostate cancer is estimated to be the most common form of cancer in American men. After lung cancer, it is the leading cause of cancer death among males.[1] The American Cancer Society estimates that approximately 221,000 new cases of prostate cancer will be diagnosed in the year 2015 alone, and approximately 1 in 7 men will be diagnosed with prostate cancer during their lifetime.[1] Incidence rates of prostate cancer have changed dramatically over the past 20 years; rapidly increasing from 1988 to 1992, declining sharply from 1992 to 1995, remaining relatively stable from 1995 to 2000, and again decreasing from 2000 to 2010.[1] This unpredictable trend primarily reflects the change in the utilization of prostate-specific antigen (PSA) blood testing by health care providers for the detection of prostate cancer.[1] Prostate cancer may be a fatal disease, but most men diagnosed with prostate cancer do not die from it. The relative United States 5-year survival rate for all stages of prostate cancer is nearly 100%, while the 10-year and 15-year survival rates are 99% and 94%, respectively, with more than 2.9 million men still living.[1]

As the number of men living beyond a prostate cancer diagnosis rises, focus of care has broadened to include quality of life (QOL) issues. Recent research provides evidence that the majority of cancer survivors have significant impairments that often go undetected and/or untreated, and therefore may result in disability.[2] Androgen deprivation therapy (ADT) is a common treatment method for the early stages of prostate cancer. During the first year of ADT, survivors of prostate cancer (PCS) often experience a deficiency in sex hormones, insulin resistance, an increased central/visceral adiposity, a decrease in bone density, lean muscle mass, and whole body muscle strength.[3] What is significant for PCS is that the impairments can often be seen in the whole body, rather than just the area treated for cancer. Prostate cancer survivors receiving ADT had 40% less upper body strength

than a group of non-ADT PCS and 22% less upper and lower body strength than a healthy control group, and 27% reduction in strength compared to a PCS non-ADT group.[4] Strength is reduced on isotonic testing such as chest press and leg extension, isokinetic knee extension testing and isometric testing with grip dynamometer.[4-7] The long term effects of ADT persist over time with a decrease in lean muscle mass.[6] Adverse changes in muscle composition may exacerbate normal sarcopenia, thereby further impacting muscular strength and endurance as well as physical function and independent living. The decrease in muscle mass and subsequent strength is associated with impaired functional mobility as indicated by increased times to complete a 5 repetition sit-to-stand test and 6 meter walk test.[6,8]

Diminished muscular endurance and fatigue are also increasingly recognized as a troublesome complaint among patients with cancer.[7,9] Cancer-related fatigue has been hypothesized to be both a central phenomenon as well as a peripheral occurrence. Centrally mediated fatigue is thought to arise from the loss of voluntary activation of muscles due to processes proximal to the neuromuscular junction, while peripheral fatigue has been attributed to failure of muscular contraction or metabolic changes within the muscle.[10] Muscular changes associated with ADT use can influence muscular endurance in PCS and have a significant negative impact on QOL and patients' self-care abilities. Researchers have reported impairments among PCS to be as high as 24% for activities of daily living (ADLs) and 42% for instrumental activities of daily living (IADLS).[8] Among such patients, the prevalence of fatigue is generally reported to be greater than 65%.[11] Furthermore, complaints of diminished endurance and of fatigue persist beyond the treatment timeframe.[7,12,13]

Impairments in strength and muscular endurance have been linked to declines in independence, functional mobility, and subsequent QOL. Activities of daily living deficits, the use of an assistive device, and abnormal functional screen findings are associated with an increased risk of falling.[7] Falls can lead to more serious injuries such as an increased risk of fractures and hospitalizations, thereby decreasing the QOL and level of independence for survivors.[8] It is important, therefore, to accurately identify impairments in muscular performance in order to initiate early intervention to mitigate the effects of ADT and subsequent functional decline among PCS.

In 2010, the American Physical Therapy Association's (APTA) Oncology Section created the EDGE (Evaluation Database to Guide Effectiveness) Task Force to develop recommendations for outcome measures to be used when assessing the status of survivors of cancer.[14] This systematic review evaluates the ways in which strength and muscular endurance are measured clinically in individuals with prostate cancer. The reliability, validity, minimal detectable change (MDC), and/or minimally clinically important difference (MCID) are important psychometric properties that need to be evaluated to justify clinical use of outcome measures.[15] Tools used to track and measure patient outcomes should be validated in the population in which they are used to be most beneficial. Additionally, these tools need to be assessed in light of clinical utility, including the availability of resources, cost, ease of use, and availability of normative data. The purpose of this systematic review is to identify commonly used methods of evaluating strength and muscular endurance in PCS and to make recommendations of the best methods based on psychometric properties and clinical utility.

METHODS

Search strategy

The authors systematically searched the literature for outcome measures that directly measured strength and muscular endurance to evaluate the psychometric properties and clinical utility of such measures. The primary search was conducted in February 2014 in PubMed/Medline and CINAHL, with secondary searches occurring through July 2014 using Web of Science, Ovid, Google Scholar, Sports Discus, Cochrane Review, PEDro, and Academic Search Premier. Search terms used alone and in combination included: Prostate cancer or neoplasm and; strength measure/measurement/test, muscular endurance measure/measurement/test, manual muscle test, psychometric properties, clinometrics, dynamometer/dynamometry, power, and energy, along with the following MESH terms: "Muscle strength dynamometer" OR "Muscle Strength" OR "Hand Strength." Relevant articles and journals focusing on orthopedics or fitness measures were reviewed recursively for other potential studies. The prostate cancer population took first priority within the search, however, if no studies included this population, patients with other cancers, geriatric patients, and the general population were considered for review.

Included studies of tests of muscle strength and muscular endurance had to report psychometric properties, present clinically feasible methods, have adults (preferably male) as participants, and be published in the English language. The publication dates were limited to 1/1/1995 and after, as long as the inclusion criteria were met. Studies were excluded if they focused on non-clinical measures of strength and muscular endurance, or were functional mobility measures (eg, Timed Up and Go, sit-to-stand, gait speed, etc.).

After completion of the literature search, relevant articles were classified into 4 strength categories and one additional category for muscular endurance. The 4 strength categories were: manual muscle test (MMT), 1 repetition maximum (1-RM) testing, hand-grip strength (HGS) using dynamometry, and hand-held dynamometry (HHD). These categories for strength measurement tools were selected based on characteristics of each measurement tool described in the available literature. Each outcome measure was appraised by two reviewers independently using the Cancer EDGE Outcome Measure Rating Form.[14] Outcome measures were then rated on the 1-4 Cancer EDGE Task Force Rating Scale taking into consideration both psychometric properties and clinical utility (Figure 1).[14] If an outcome measure rating was found to be in disagreement between the two independent reviewers, the disagreement was resolved by discussion with all 5 reviewers until consensus was obtained. Finally, all articles reviewed for an outcome measure were included in a reference section of the EDGE form for each appropriate measure.

4	**Highly Recommend**	Highly recommended; the outcome has good psychometric properties and good clinical utility; the measure has been used in research on individuals with or post prostate cancer.
3	**Recommend**	Recommended; the outcome measure has good psychometric properties and good clinical utility; no published evidence that the measure has been applied to research on individuals with or post prostate cancer.
2A	**Unable to Recommend at this time**	Unable to recommend at this time; there is insufficient information to support a recommendation of this outcome measure; the measure has been used in research on individuals with or post prostate cancer.
2B	**Unable to Recommend at this time**	Unable to recommend at this time; there is insufficient information to support a recommendation of this outcome measure; no published evidence that the measure has been applied to research on individuals with or post prostate cancer.
1	**Do Not Recommend**	Poor psychometrics &/or poor clinical utility (time, equipment, cost, etc.).

Figure 1. Cancer EDGE Rating Scale.

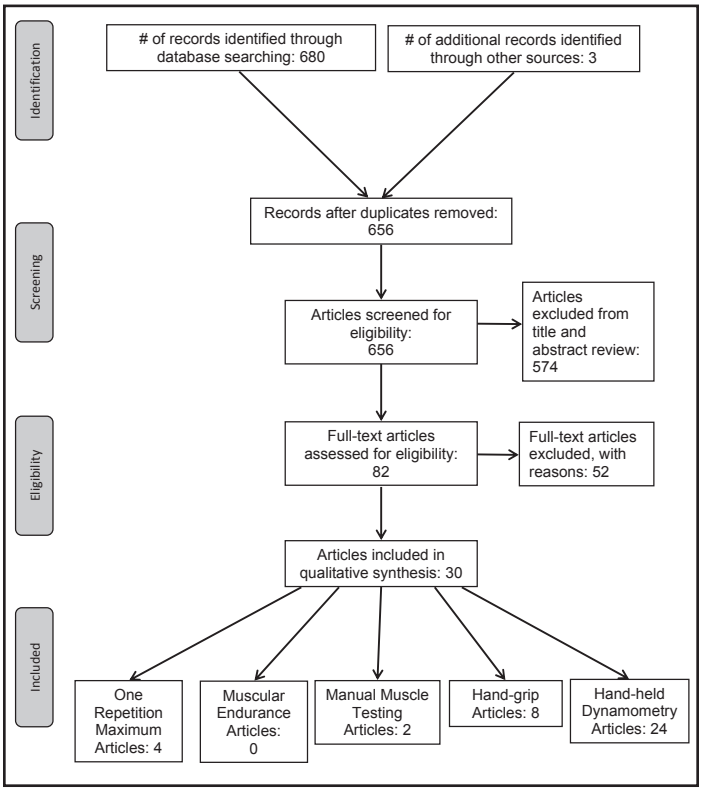


Figure 2. PRISMA flow of literature search.

Data Extraction and Synthesis

Relevant psychometric data, when available, were extracted and recorded on the Cancer EDGE Task Force Outcome Measure Rating Form for each study. This data included: intra-, inter-, and test-retest reliability values, with confidence intervals as available, validity, MDC, standard error of measurement (SEM), and MCID. Reliability and validity were determined by either the Pearson (r) or Intraclass Correlation Coefficient (ICC), or Kappa values (K). Correlation coefficients of greater than 0.75 are considered good to excellent, 0.5-0.74 moderate, and below 0.5 considered poor.[16] Kappa values greater than 80% demonstrated excellent agreement, 61% to 80% substantial agreement, 41% to 60% adequate agreement, and less than 40% showed poor agreement.[16] Clinical utility was assessed using the criteria of: availability of resources, cost, ease of use including time necessary to complete testing and clinician training, scoring and interpretation, and availability of normative data for comparison.

RESULTS

The initial literature search for muscle strength and endurance resulted in 683 articles. The titles were screened and any duplicates removed by the assessors. Article titles and abstracts were then reviewed to identify studies that specifically addressed the purpose of this review. Eighty-two articles were retrieved and assessed for eligibility. Thirty articles were included in the study after exclusions were applied. Figure 2, the PRISMA Flow diagram, details the literature search process.

By category, the number of articles reviewed were: MMT = 2, 1-RM = 4, HGS = 8, and HHD = 24. No articles were found which met inclusion criteria to assess muscular endurance measures, although such tests have been used in prostate cancer research. Note that some research studies evaluated multiple tools, such that the number of articles for each category is not mutually exclusive. Table 1 demonstrates the clinical usefulness of strength and muscle endurance testing methods.

Two measures were recommended (rated 3) by the Prostate Cancer EDGE Task Force members: HGS and HHD.[5] These measures are recommended for clinical use to objectify strength measures. One repetition maximum testing was scored a 2A, unable to recommend at this time, because of a lack of high clinical feasibility, although there is evidence of use in prostate cancer research for chest and leg press strength assessment. Manual muscle testing and muscle endurance were scored a 2B, unable to recommend at this time, due to lack of psychometric support. Muscular endurance testing lacks psychometric support and is difficult to perform in a clinical setting, and was rated by the Task Force as 1, do not recommend. See Table 2 for Task Force ratings and clinical utility comments. Table 3 details the psychometric properties of the clinical measures of strength.

DISCUSSION

The measurement of strength and muscular endurance in men who have been treated with ADT for prostate cancer is essential to the rehabilitation continuum. The effect of ADT on muscular tissue is well documented,[4-6] and the loss of strength and muscular endurance impairs functional mobility[6] and subsequent QOL.[4,7] Therefore, valid and reliable measures of strength and muscular endurance are critical for this population in order to identify deficits, to establish a comprehensive picture of the patient's functional goals and needs, and to monitor progress throughout the course of treatment and beyond.

Table 1. Clinical Usefulness for Strength and Muscular Endurance Measures

Measure	Equipment Needed	Cost	Ease of Use	Scoring/Interpretation	Normative Data
MMT	No	Free	High	Easy	Yes*
1-RM	Yes – weights or machines	Minimal	Moderate	Easy	Inconsistent
HGS	Yes – dynamometer	Moderate	High	Easy	Yes
HHD	Yes – dynamometer	Moderate	High	Easy	Yes
Muscle Endurance	Yes – weights or machines	Minimal	Not established	Moderate	No

Abbreviations: MMT, manual muscle test; 1- RM, one repetition maximum; HGS, hand-grip strength; HHD, hand-held dynamometry
*Based on the rating of a 5 being "normal" strength.

Table 2. Prostate Cancer EDGE Task Force Ratings and Clinical Utility

Measure	Prostate Cancer EDGE Task Force Rating	Clinical Utility
Hand-grip Strength	3	Equipment is easy to use clinically and staff training is simple. Good clinical utility.
Hand-held Dynamometry	3	Easy to use clinically; methodology similar to manual muscle testing. Normative data available.
1 Repetition Maximum	2A	Not often used clinically. Psychometric support is limited.
Manual Muscle Test	2B	Highly useful in the clinic, but poor psychometric properties do not support use.
Muscle Endurance	1	Not often tested clinically. Used in research, but lacks psychometric assessment.

Findings from this systematic review indicate that the measurement of strength is best performed using objective dynamometry for both hand grip and extremity measures. No recommendations for the clinical measurement of muscular endurance can be made at this time.

Strength

Accurate measurement of strength using dynamometry is achieved through a method that is valid, reliable, and sensitive to change. Importantly, by quantifying force output as a measure of strength, clinicians can measure strength objectively to determine deficits, plan treatment, and measure progress. Although used widely, MMT, which ranks strength on a 0-5 scale (0 representing no muscular contraction and a 5 indicating full strength),[17] has limitations which need to be considered in light of emerging affordable and clinically feasible alternatives which provide greater validity and reliability. Manual muscle testing is a subjective measure of strength. This is particularly true for the antigravity grades of 3 or greater, which lend themselves to personal interpretation of the evaluator. Although reliability measures indicate that there is adequate intrarater consistency within a single evaluator, the amount of force exerted by multiple testers of a 3+ for the same participant are quite variable.[18] Another important limitation to be considered is that the MMT scale is ordinal rather than interval; the difference between a muscle graded a 3 and one graded a 4 is not necessarily the same as the difference between a 4 and a 5. This limitation in grading and lack of precision of measurement does not allow the clinician to accurately describe strength gains made through rehabilitative measures, and generally lacks the sensitivity needed to appreciate small gains in strength.

Tools which are considered accurate possess a small level of error. The SEM of the two HHD examined in this review varies from 4.9-12.5N.[19,20] One kilogram is equivalent to 2.2 pounds or 9.8N. With a SEM of no greater than 12.5N, the error of measurement in the HHD is less than 1.3 kg (2.9 pounds). The hand grip dynamometers evaluated have a SEM of 0.76 – 1.25 kg.[21] Any amount of change in strength measures greater than the SEM, 1.3 kg, is real change. Research and analysis establishing the MDC or MCID for dynamometry is slight, but studies reported MDC values of 1.75-5.58 kg in cancer populations,[20,22] and up to 7.3 kg in a healthy population.[19] What the actual amount of change in force output that is clinically meaningful will vary depending on the muscle group tested, the age and gender of the individual, and the functional needs of that person. This clinically meaningful change will require the judgment of the clinician.

The validity and reliability of HHD is well established in the literature. These psychometric properties have been described for multiple populations: healthy individuals, chronically ill, and those with cancer.[20,23-27] Overall, validity with strength measured using isokinetic dynamometry is good to excellent.[25,26] Although reliability is reported as good to excellent in most studies,[23,24,26,27] it can be improved through the use of a fixation method. Because research shows that the tester gender, body weight, or grip strength can influence the force values obtained using HHD,[28] it is important to create a mechanism of consistent resistance. Research supports using some external fixation for the dynamometer to improve the interrater reliability of dynamometry in a clinical setting. Studies have investigated different devices including brackets,[29-31] or straps.[32,33] The studies whose psychometric properties are reported in this review did not use external stabilization, and it is reasonable to conclude that reliability measures would improve with this use. In a clinical situation, a mobilization belt can be strapped around the dynamometer and fixed in opposition to the force vector to provide a consistent resistance for maximal voluntary contractions.

Table 3. Psychometric Properties of Strength and Muscle Endurance Methods or Tools

Measure	Intrarater Reliability (ICC)	Interrater Reliability (ICC)	Test/Retest Reliability (ICC)	Responsiveness to Change	Validity	Clinical Utility
Manual Muscle Testing		ICU Patients: [42] Knee ext, hip flex, ankle DF = 0.99 (0.97-1.00) Agreement for detecting significant weakness = 1.00 (0.55-1.00) [43]Knee ext, hip flex, ankle DF = 0.31-0.75 (0.17-0.87) Sum Score = 0.83 (0.67-0.91) Using Cohen's kappa = .38 (0.44-1.0)				Easy to do with no equipment required; interrater reliability highly variable. Validity not established outside of neurological population. Not used in prostate cancer population.
1-RM Testing	Healthy Population: Squat/Knee ext: ICC=0.64-0.90[44]	Healthy Population: Squat/Knee ext: ICC=0.94-0.96[44]	Healthy Untrained Population: Leg press/knee ext: ICC=0.97-0.99[45] Type 2 DM Leg press/knee ext: 0.98-0.99 (0.95-0.99)[46]	**Standard Error of Measurement** Healthy Population: Squat/Knee ext: 3.2-13.1[44] Type 2 DM (using est 1-RM): 1.24-3.56[46]		Good psychometric values, however, clinical utility makes this measure difficult to use. Time consuming to perform, and if done on equipment, expensive.
Hand-grip Strength	Healthy Population: ICC = 0.94-0.98[47]	Healthy Population: ICC = 0.99[48] Mean of 2 trials: ICC=.93-.95[21] ICC=0.78-0.82 (0.65-0.89)[49] Highest Score of 3 Trials ICC= 0.82-0.92[48]	Community-dwelling Older Adults: ICC = 0.94-0.98[26] Advanced Cancer Patients: Spearman's rho =.097[50]	**Standard Error of Measurement** Older Adults in Care 0.76 – 1.25 kg.[21,48]	Advanced Cancer Patients: %CV as a measure of precision = 6.3[50] %CV=10.59-11.61[51]	Hand-grip strength has good psychometric properties and clinical utility; it has been tested in healthy patients and advanced cancer patients, but not tested in individuals with prostate cancer.
Hand-held Dynamometry	Healthy Population: [30] Male examiners (knee flex/ext) ICC=0.79-0.93 (0.64-0.97) Female examiners (knee flex/ext) ICC=0.88-0.91 (0.81-0.95)	Healthy Population: Knee ext = 0.91-0.97 (0.68-0.99) Critically Ill: Knee ext = 0.78-0.79 (0.32-0.95)[23] Hip flex, knee ext, ankle DF = 0.76-0.94 (0.33-0.97)[24] Cancer Population: Pearson r=0.96 ICC=0.96 (0.92-0.98)[20] Advanced Cancer: ICC=0.83[51]	Healthy Population: Knee ext = 0.96-0.97 (0.85-0.99) Critically Ill: Knee ext = 0.89-0.92 (0.69-0.97)[23] Community-dwelling Older Adults: Hip abd/flex; knee ext ICC=0.88-0.94[52] Knee ext ICC=0.84-0.95[26] Advanced Cancer: ICC=0.90[51]	**Standard Error of Measure:** Healthy Pop: 5.5-6.9 kg[23] Critically Ill: 2.1-2.6 kg[23] Cancer Pop (with a pull-gauge): 2.02 kg (+/- 3.96)[20] **Minimal Detectable Change:** Cancer Pop: 5.59 kg[20] 1.75-2.77kg[22] Healthy Pop: .57 – 1.27[19] 6.5-7.3 kg[23] Critically Ill: 5.7-7.1 kg[23] **Responsiveness (Effect Size):** Lower extremity strength using multiple devices/ populations: ~0.40-1.0[53]	HHD with Biodex: r=0.91[25] With TUG: r=-0.71-0.86 With Gait Speed: r=0.79-0.83[26]	Hand-held dynamometry has good psychometric properties and clinical utility; it has been tested in healthy patients and advanced cancer patients, but not tested in individuals with prostate cancer.

Abbreviations: ICU, intensive care unit; flex, flexion; ext, extension; abd, abduction; pop, population; DF, dorsiflexion; ICC, intraclass correlation coefficient; r, Pearson's correlation coefficient; SEM, standard error of measurement; MDC, minimal detectable change; MCID, minimally clinical important difference; DM, diabetes mellitus; HHD, hand-held dynamometer; 1-RM, 1-repetition maximum; CV, coefficient of variation; TUG, Timed Up and Go; kg, kilograms

Sensitivity to change is impacted by the tool used, as well as the unit of measure for that tool. Manual muscle testing lacks properties of measurement which are sensitive. Force is often measured by Newtons (N), pounds (lb), or kilograms (kg). The unit of force output for MMT remains an ordinal number whereas the output on dynamometers is in pounds or kilograms. Muscles graded a 4 may have as little as 10% of the maximum strength of a muscle.[34] Hand-held dynamometry uses a unit of measure that is an interval scale; the amount of difference between a 3 and a 4 is the same as between a 4 and a 5. Sensitivity to change over time can then be accurately described. Furthermore, clinicians consistently evaluate patient performance against an expected normal level of performance. Use of HHD allows this comparison to be made as normative values have been established for human strength measures. Although outside the scope of this paper, the reader is encouraged to reference the numerous studies reporting these values.[35-37]

Measurement of strength is most accurate using dynamometry. The use of dynamometry in the prostate cancer population is limited to two smaller studies[20,22] which used a strain-gauge rather than a force gauge typically seen clinically. This limited the authors' ratings of HHD and HGS to a 3 (recommended) in this review, however, both HGS and HHD offer the clinician a clinically feasible method to measure strength that has the necessary psychometric properties to support good validity, reliability, MDC, MCID, and sensitivity to change, and have been used in other cancer populations. The use of 1-RM cannot be recommended secondary to low clinical utility and weaker psychometric properties.

Muscular Endurance

Clinically feasible methods of measuring muscular endurance with accompanying sound psychometric properties and normative values remain elusive. Because of this, muscular endurance, the ability to sustain force output over time, is seldom assessed in a clinical setting. Yet understanding overall muscular fitness after treatment with ADT is an important consideration given the effects of ADT on muscle tissue, including sarcopenia.[38] A component of muscular fitness is muscular endurance. Research is emerging suggesting that muscular endurance is lower in men treated for prostate cancer with ADT.[7,9] Therefore, finding appropriate means to assess this clinically is important for monitoring patient status.

The most available method to measure muscular endurance is some variation of a repetition to failure loading test. An early study examining muscular fitness among men treated with ADT compared a group engaged in a resistance exercise program to a group without exercise using a fixed load repeatedly lifted at a standard rate, and counted the number of repetitions correctly completed.[7] Findings from this study showed an increase in the number of repetitions after 3 months of resistance exercise training, with an accompanying decrease in self-reported fatigue using the Functional Assessment of Cancer Therapy – Fatigue.[7] Another study recorded the number of repetitions of 70% of 1RM lifted until failure comparing a group of men on ADT to a group of healthy controls. This study found no differences between groups for muscular endurance repetitions to failure using 70% of 1 RM.[6] To better understand the implications of these results, it is important to examine how muscular endurance should be measured.

The American College of Sports Medicine recommends that lifting 40-60% of a maximum resistance repeatedly in training will increase muscular endurance.[39] Intuitively, then, measuring endurance should be completed using repetitions to failure of 40-60% of 1RM. Neither study purporting to measure muscular endurance utilized this method, although findings that an increase in repetitions suggest an increase in muscular endurance among men using ADT. The limitation to this study was the lack of a healthy control comparison to determine whether deficits in muscular endurance were present at baseline.

Measuring muscular endurance with a repetition to failure using 40-60% of 1 RM is not without merit. Establishing a baseline measure for an individual is possible, and repeating the measure postintervention can inform change. The limitation of this methodology is the lack of normative data for age, gender, and muscle group. Such data is difficult to gather, as the number of repetitions to failure is largely dependent upon the muscle mass of the individual.[40]

What is needed is a test for muscular endurance which possesses good clinical feasibility, along with strong psychometric properties. Isokinetic dynamometry offers a more reliable and valid method to measure muscular endurance, but lacks clinical feasibility. However, a study of muscular endurance using a Biodex stationary dynamometer, measuring maximal voluntary isometric contraction (MVIC) levels pre- and post-endurance activity, shows promise.[41] Findings from this study suggest that rather than measuring repetitions to failure as a unit to quantify endurance, perhaps measuring MVIC pre- and post-activity may provide a more reliable and valid method to measure endurance. It may be possible that using HHD to measure MVIC before and after some fatiguing activity holds promise for a more clinically realistic measure of muscular endurance. At this time, measuring muscle endurance clinically is not feasible and this systematic review does not support it.

Further investigation is needed in designing a clinically feasible, reliable, valid, and standardized method to measure muscular endurance. A clinical method of measuring muscle endurance should utilize the guiding principles of 40% to 60% of maximum resistance lifted over time. The clinical method should also be responsive enough to detect differences between healthy and injured tissue, as well as have a reliable and quantifiable normative unit of measure.

Other research in cancer outcome measures should focus on the specific needs of the population. More studies with men who have been treated for prostate cancer examining reliability and validity as well as responsiveness to change are needed to determine intervention effectiveness. Cutoff scores should be established to assess the severity of impairment and functional limitations. Tools for specific practice settings across the continuum of care need to be explored; it is reasonable to believe that responses of

individuals will vary based on whether they are in the acute stage of recovery or a more long-term stage, as the impact of treatment changes with time.

CONCLUSION

Psychometrically strong and clinically feasible outcome measures need to be utilized in evidence-based practice of physical therapy. Measuring strength and muscular endurance precisely in men with prostate cancer allows clinical decision-making to accurately identify impairments in body structures which may impact activity and participation. Both HGS and HHD are recommended as valid and reliable methods to assess strength in PCS. No clinical measures for muscle endurance could be recommended at this time. Further research is necessary to devise a clinically feasible muscular endurance test with sound psychometric properties for clinical use in this population.

ACKNOWLEDGEMENTS

Support was provided through Graduate Research Assistants at the University of Michigan-Flint and the University of Dayton. Library Search Assistance was provided at University of Michigan-Flint, Baptist Health Lexington, and University of Dayton.

REFERENCES

1. American Cancer Society. Cancer Survivorship Facts & Figures. 2015; http://www.cancer.org/research/cancer-factsstatistics/index. Accessed April14, 2015.
2. Silver JK, Baima J, Mayer RS. Impairment-driven cancer rehabilitation: An essential component of quality care and survivorship. *CA Cancer J Clin.* 2013;63(5):295-317.
3. Algotar AM, Thompson PA, Ranger-Moore J, et al. Effect of aspirin, other NSAIDS, and statins on PSA and PSA velocity. *Prostate*. 2010;70:883-888.
4. Basaria S, Lieb J 2nd, Tang AM, et al. Long-term effects of androgen deprivation therapy in prostate cancer patients. *Clin Endocrinol (Oxf)*. 2002;56(6):779-786.
5. Girard D, Marino FE, Cannon J. Evidence for reduced neuromuscular function in men with a history of androgen deprivation therapy for prostate cancer. *Clin Physiol Funct Imaging*. 2013;34(3):209-217.
6. Galvão DA, Taaffe DR, Spry N, Joseph D, Turner D, Newton RU. Reduced muscle strength and functional performance in men with prostate cancer undergoing androgen suppression: A comprehensive cross-sectional investigation. *Prostate Cancer Prostatic Dis*. 2009;12(2):198-203.
7. Segal RJ, Reid RD, Courneya KS, et al. Resistance exercise in men receiving androgen deprivation therapy for prostate cancer. *J Clin Oncol*. 2003;21(9):1653-1659.
8. Bylow K, Dale W, Mustian K, et al. Falls and physical performance deficits in older patients with prostate cancer undergoing androgen deprivation therapy. *Urology*. 2008;72(2):422-427.
9. Alt CA, Gore EM, Montagnini ML, Ng AV. Muscle endurance, cancer-related fatigue, and radiotherapy in prostate cancer survivors. *Muscle Nerve*. 2011;43(3):415-424.
10. Yavuzsen T, Davis MP, Ranganathan VK, et al. Cancer-related fatigue: Central or peripheral? *J Pain Symptom Manage*. 2009;38(4):587-596.
11. Hansen PA, Dechet CB, Porucznik CA, LaStayo PC. Comparing eccentric resistance exercise in prostate cancer survivors on and off hormone therapy: A pilot study. *PM R*. 2009;1(11):1019-1024.
12. Stone P, Hardy J, Huddart R, A'Hern R, Richards M. Fatigue in patients with prostate cancer receiving hormone therapy. *Eur J Cancer*. 2000;36(9):1134-1141.
13. Galvão DA, Taaffe DR, Spry N, Joseph D, Newton RU. Combined resistance and aerobic exercise program reverses muscle loss in men undergoing androgen suppression therapy for prostate cancer without bone metastases: A randomized controlled trial. *J Clin Oncol*. 2010;28(2):340-347.
14. Levangie PK, Fisher MI. Oncology Section Task Force on Breast Cancer Outcomes: An introduction to the EDGE TASK Force and clinical measures of upper extremity function. *Rehabil Oncol*. 2013;31(1):6-10.
15. Field-Fote E, Levangie P, Craik R. Towards Optimal Practice-How Can Students Contribute? *Student Assembly Pulse*. March 2007:4.
16. Portney LG WM. *Foundations of Clinical Research: Applications to Practice*. 3rd ed. Old Tappen, NJ: Pearson Prentice Hall; 2009.
17. Hislop HJ, Avers D, Brown MB. *Daniels and Worthingham's Muscle Testing: Techniques of Manual Examination*. 9th ed. St. Louis: Elsevier Saunders; 2013.
18. Knepler C, Bohannon RW. Subjectivity of forces associated with manual-muscle test grades of 3+, 4-, and 4. *Percept Mot Skills*. 1998;87(3):1123-1128.
19. Dollings H, Sandford F, O'Conaire E, Lewis JS. Shoulder strength testing: the intra- and inter-tester reliability of routine clinical tests, using the PowerTrack™ II Commander. *Shoulder Elbow*. 2012;4(2):131-140.
20. Knols RH, Stappaerts KH, Fransen J, Uebelhart D, Aufdemkampe G. Isometric strength measurement for muscle weakness in cancer patients: Reproducibility of isometric muscle strength measurements with a hand-held pull-gauge dynamometer in cancer patients. *Support Care Cancer*. 2002;10(5):430-438.
21. Mawdsley RH, Ferrara JM, Ferrar RP, Urban JJ Halbach HL. Reliability of an alternative position of measuring grip strength in elderly females. *Issues on Aging*. 2001;24(1):7-10.
22. Knols RH, Aufdemkampe G, de Bruin ED, Uebelhart D, Aaronson NK. Hand-held dynamometry in patients with haematological malignancies: Measurement error in the clinical assessment of knee extension strength. *BMC Musculoskelet Disord*. 2009;10:31-31.
23. Baldwin CE, Paratz JD, Bersten AD. Muscle strength assessment in critically ill patients with handheld dynamometry: An investigation of reliability, minimal detectable change, and time to peak force generation. *J Crit Care*. 2013;28(1):77-86.
24. Vanpee G, Segers J, Van Mechelen H, et al. The interob-

server agreement of handheld dynamometry for muscle strength assessment in critically ill patients. *Crit Care Med.* 2011;39(8):1929-1934.
25. Martin HJ, Yule V, Syddall HE, Dennison EM, Cooper C, Aihie Sayer A. Is hand-held dynamometry useful for the measurement of quadriceps strength in older people? A comparison with the gold standard Bodex dynamometry. *Gerontology*. 2006;52(3):154-159.
26. Schaubert KL, Bohannon RW. Reliability and validity of three strength measures obtained from community-dwelling elderly persons. *J Strength Cond Res*. 2005;19(3):717-720.
27. Bohannon RW, Schaubert KL. Test-retest reliability of grip-strength measures obtained over a 12-week interval from community-dwelling elders. *J Hand Ther*. 2005;18(4):426-427.
28. Wadsworth CT, Krishnan R, Sear M, Harrold J, Nielsen DH. Intrarater reliability of manual muscle testing and hand-held dynametric muscle testing. *Phys Ther.* 1987;67(9):1342-1347.
29. Kobler MJ, Cleland JA. Strength Testing Using Hand-Held Dynamometry. *Phys Ther Rev*. 2005;10(2):99-112.
30. Lu T, Hsu H, Chang L, Chen H. Enhancing the examiner's resisting force improves the reliability of manual muscle strength measurements: Comparison of a new device with hand-held dynamometry. *J Rehabil Med*. 2007;39(9):679-684.
31. Tate AR, McClure PW, Kareha S, Irwin D. Effect of the Scapula Reposition Test on shoulder impingement symptoms and elevation strength in overhead athletes. *J Ortho Sports Phys Ther.* 2008;38(1):4-11.
32. McGirr KA, Kennedy T, Molgaard MA, Rathleff MS. Intra-Tester Reliability of Hand-Held Dynamometry and Strap-Mounted Dynamometry for Assessment of Ankle Strength. *Int J Athl Ther Train.* 2014;19(2):14-19.
33. Bohannon RW, Pritchard RO, Glenney SS. Portable belt-stabilized hand-held dynamometry set-up for measuring knee extension force. *Isokinet Exerc Sci* 2013;21(4):325-329.
34. Dvir Z. Grade 4 in manual muscle testing: The problem with submaximal strength assessment. *Clin Rehabil.* 1997;11(1):36-41.
35. Andrews AW, Thomas MW, Bohannon RW. Normative values for isometric muscle force measurements obtained with hand-held dynamometers. *Phys Ther*. 1996;76(3):248-259.
36. Bohannon RW. Reference values for extremity muscle strength obtained by hand-held dynamometry from adults aged 20 to 79 years. *Arch Phys Med Rehabil.* 1997;78(1):26-32.
37. Hughes RE, Johnson ME, O'Driscoll SW, An KN. Normative values of agonist-antagonist shoulder strength ratios of adults aged 20 to 78 years. *Arch Phys Med Rehabil.* 1999;80(10):1324-1326.
38. Storer TW, Miciek R, Travison TG. Muscle function, physical performance and body composition changes in men with prostate cancer undergoing androgen deprivation therapy. *Asian J Androl.* 2012;14(2):204-221.
39. *ACSM's Guidelines for Exercise Testing and Prescription*. 7th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2006.
40. Shimano T, Kraemer WJ, Spiering BA, et al. Relationship between the number of repetitions and selected percentages of one repetition maximum in free weight exercises in trained and untrained men. *J Strength Cond Res*. 2006;20(4):819-823.
41. Roy J, Ma B, Macdermid JC, Woodhouse LJ. Shoulder muscle endurance: The development of a standardized and reliable protocol. *Sports Med Arthrosc Rehabil Ther Technol*. 2011;3(1):1-1.
42. Ciesla N, Dinglas V, Fan E, Kho M, Kuramoto J, Needham D. Manual muscle testing: a method of measuring extremity muscle strength applied to critically ill patients. *J Vis Exp*. 2011(50).
43. Hough CL, Lieu BK, Caldwell ES. Manual muscle strength testing of critically ill patients: Feasibility and interobserver agreement. *Crit Care.* 2011;15(1):R43-R43.
44. Tagesson SKB, Kvist J. Intra- and interrater reliability of the establishment of one repetition maximum on squat and seated knee extension. *J Strength Cond Res*. 2007;21(3):801-807.
45. Levinger I, Goodman C, Hare DL, Jerums G, Toia D, Selig S. The reliability of the 1RM strength test for untrained middle-aged individuals. *J Sci Med Sport.* 2009;12(2):310-316.
46. Abdul-Hameed U, Rangra P, Shareef MY, Hussain ME. Reliability of 1-repetition maximum estimation for upper and lower body muscular strength measurement in untrained middle aged type 2 diabetic patients. *Asian J Sports Med*. 2012;3(4):267-273.
47. Peolsson A, Hedlund R, Oberg B. Intra- and inter-tester reliability and reference values for hand strength. *J Rehabil Med*. 2001;33(1):36-41.
48. MacDermid JC, Alyafi T, Richards RS, Roth JH. Test-retest reliability of isometric strength and endurance grip tests performed on the Jamar and NK devices. *Physiother Can*. 2001;53(1):48.
49. Trutschnigg B, Kilgour RD, Reinglas J, et al. Precision and reliability of strength (Jamar vs. Biodex handgrip) and body composition (dual-energy X-ray absorptiometry vs. bioimpedance analysis) measurements in advanced cancer patients. *Appl Physiol Nutr Metab.* 2008;33(6):1232-1239.
50. Stone CA, Nolan B, Lawlor PG, Kenny RA. Hand-held dynamometry: Tester strength is paramount, even in frail populations. *J Rehabil Med.* 2011;43(9):808-811.
51. Ottenbacher KJ, Branch LG, Ray L, Gonzales VA, Peek MK, Hinman MR. The reliability of upper- and lower-extremity strength testing in a community survey of older adults. *Arch Phys Med Rehabil.* 2002;83(10):1423-1427.
52. Bohannon RW. Responsiveness of hand-held dynamometry to changes in limb muscle strength: A retrospective investigation of published research. *Isokinet Exerc Sci*. 2009;17(4):221-225.

PRevention, Intervention, and Sustained wellness Model (PRISM) Care Philosophy in Cancer Survivorship, Palliative Care, and Chronic Disease Management in the Era of Healthcare Reform: A Perspective Paper

Reyna Colombo PT, MA[1]; Christopher Wilson PT, DPT, GCS[2]

[1]Director of Rehabilitation Services – Troy, Beaumont Health, Troy, MI
[2]Coordinator of Clinical Education, Rehabilitation Services – Troy, Beaumont Health, Troy, MI; and Clinical Assistant Professor, Oakland University, Rochester, MI

ABSTRACT

Background: Multiple agencies have recognized the increasing care demands and the associated costs of a growing aging population with incurable or life-threatening conditions, including cancer. Although the core concepts of physical therapy (PT), cancer survivorship and palliative care (PC) appear to be congruent and complimentary, there is little evidence in the literature demonstrating a consistent role of PT in these settings. This article will outline a care philosophy to use for educating PTs, patients/clients, interdisciplinary team members, and as a guide for PT care for the longitudinal management of the patient with a cancer diagnosis, incurable illness, or a life threatening illness. **Methods:** Literature review and perspective regarding PT within cancer survivorship, PC and chronic disease management in the context of healthcare reform to provide a perspective of the PRevention, Intervention, and Sustained wellness Model (PRISM) care philosophy. **Discussion:** The PRISM promotes the longitudinal management by the physical therapist for a patient with a number of long-term conditions, including cancer survivorship, chronic disease or life threatening illness and its side effects. *Prevention* includes primary, secondary, and tertiary prevention in the presence of a disease. *Intervention* is generally considered conventional or traditional physical therapy, while *Sustained wellness* incorporates concepts related to maintaining health or slowing the decline of a progressive illness, injury prevention, and anticipation and management of potential medical crises. **Conclusion:** The PRISM may have utility to guide and educate care providers on PT's role in managing individuals with these conditions.

Address correspondence to: Reyna Colombo (rcolombo@beaumont.edu) or Christopher Wilson (Christopher.wilson@beaumont.edu); Beaumont Health System, Rehabilitation Services, Suite 203, 44201 Dequindre, Troy, MI 48085, Ph: (248) 964-4014, Fax: (248) 964-8099.

Key Words: side effects, prevention, navigator, wellness, prospective

INTRODUCTION

Palliative care (PC) and physical therapy (PT) are congruent with similar treatment philosophies, including anticipating and mitigating predicable and unforeseen medical events, optimizing and maintaining quality of life (QOL) in the existence of a disease process or impairment, and supporting the patient holistically.[1] Despite these similarities, many physical therapists (PTs) who work with patients with a chronic disease or a life-threatening illness are not aware that they are in essence, and in reality, providing a component of PC. Palliative care, while a relatively new term to many practitioners, is the holistic, longitudinal management of an individual with a chronic, life threatening, or incurable illness.[2] Cancer survivorship and QOL concepts are closely tied to the care philosophy of PC and cancer survivors note that QOL is a high priority throughout the survivorship journey, of which PT has the potential to play an important role across the spectrum of the disease process and treatment of side effects.[3] For example, chemotherapy treatment is associated with pain, fatigue, weakness, and the magnitude of symptoms directly affected QOL.[4] Rather than focusing on the clinical aspects of patient management for the patient with a life threatening illness, this critical review of the literature and perspective paper examines various aspects as it relates to how cancer survivorship and PC treatment philosophy are congruent with the overarching concepts of health care reform. This literature review examines how PTs and PT practice is best positioned to assist in meeting the goals of cancer survivorship and PC through the PRISM care model.

Health care reform and the Affordable Care Act, despite its political controversy, brought to light the need for better management of life-threatening diseases and chronic illness in the interest of cost-effectiveness and quality outcomes. Some of the core concepts of this management include:

- ☐ improved access to care,
- ☐ earlier management,
- ☐ preparation for medical emergencies and events,
- ☐ holistic support of the patient and family, and
- ☐ an integrated care team.

Reported benefits of this improved management include:

- ☐ the possibility for improved functional capacity,
- ☐ improved longevity,
- ☐ reduced medical costs,
- ☐ reduced stress to patients and caregivers, and
- ☐ the patient's active participation in their disease management and medical care.[5]

Despite the congruency and compatibility of the PT and PC philosophies, no articles were located that demonstrate strong evidence that PT had achieved full integration into care teams or were consistently consulted. This lack of evidence outlines a research opportunity to provide evidence for the role of PT within PC to best determine where PT principles may be applied with the most benefit to the patient, the profession, and to society at large. The purpose of this review is to outline the PRISM care philosophy and provide a perspective as to the future role and opportunities that the physical therapist is optimally positioned to provide. PRISM is a care philosophy and educational tool to assist PTs and other stakeholders in understanding PTs role in the management of the cancer survivor or the patient with chronic disease or life-threatening illness. This perspective is provided in the context of a literature review highlighting issues related to the aging population, the disproportionate health care costs of this patient population, especially nearing the end of life, and the focus on reform of the healthcare system. These factors provide a significant opportunity for PTs to provide resources to manage these patients.

METHODS

The key databases examined in the literature review included the APTA's Open Door portal, PEDro database, CINAHL, and Google Scholar from 2003 to 2013. Key terms searched included "physical therapy," "palliative care," "chronic disease," and "health care reform"/"affordable care act." Six articles were selected that best outlined the core concepts of the objectives of this review. Although research articles were preferred as well as articles published within the past 10 years, the sample of research articles providing information related to the topic of PT involvement in PC in the context of healthcare reform were relatively limited. This prompted utilization of some meta-analyses, position papers by think-tanks, and expert opinion medical manuscripts on the topic, as this level of articles is more common and applicable to this review.

RESULTS

In a systematic review by Meier entitled, "*Increased Access to Palliative Care and Hospice Services: Opportunities to Improve Value in Healthcare,"* PC services were examined in the context of public health policy.[6] Meier clarified the definition of PC to provide best-possible QOL for patients and family caregivers, work in an interdisciplinary and community-based fashion to promote seamless models of care across a range of settings. A key aspect of PC is that "ideally PC should be initiated concurrently with a diagnosis of serious illness and at the same time as curative or disease modifying treatments." A MedPAC report from 2010 that found that "10 percent—of the sickest Medicare beneficiaries accounted for about 57 percent of total program spending, which was more than $44,220 per capita per year," with the costliest individuals being those with a chronic disease.[6] In a 2010 article by the federal Office of Assistant Secretary for Planning and Evaluation, individuals with a life-threatening illness or chronic disease and their associated conditions and functional impairment constitute about 10% of all patients in the United States (US), but account for well over half of the nation's health care costs.[6] Meier found 15 articles that provided evidence that PC programs positively affected "physical and psychosocial symptoms, family caregiver well-being, bereavement outcomes, and patient, family, and physician satisfaction." A few studies noted that PC treatment philosophies may even be associated with a prolongation of life. Meier highlighted an article by Morrison et al in 2008 describing a net per-patient savings of $2,659. Common barriers to PC include variable access, inconsistent services (especially in rural areas), and an inadequate workforce with expertise in PC.[6]

In a white paper outlined by the RAND Corporation entitled, "*Living Well at the End of Life,"* Lynn and Adamson discussed the issue of aging baby boomers, which will significantly add to the demand upon healthcare resources.[7] By the year 2030, 9 million Americans who were born in the 1950s will be reaching 85 years of age and are likely to face significant disability.[7] When living with a life-threatening illness, a disproportionate amount of health care costs are concentrated in the last several weeks and months of life. Nine out of 10 people who die have a chronic life-threatening illness including cancer, cardiac, respiratory illness, dementia, and stroke. About 25% of those with chronic illness may experience disability from their condition at any one time. Three different common scenarios of chronic, life-threatening illness were outlined:

1. "Short period of evident decline" which is typical of cancer. In this case, there is a longer term preservation of comfort and function until the disease process becomes overwhelming to the systems, then a steady, rapid decline in function may occur.
2. "Longer term limitations with intermittent exacerbations and sudden dying." This is more common with organ system failure pathologies such as COPD and CHF.
3. "Prolonged dwindling." This is more typical of central nervous system failure that is generally of a slow decline where institutional long-term care facilities are beneficial.

Based on these descriptions, methods to innovate and reform care for these chronically ill adults were recommended, including integrating care across settings, quality improvement programs for

pain management, advanced directive planning, and PC consultations. Some longer term suggestions by the RAND Corporation include addressing the shortage of caregivers, reforming federal finance policy, evaluating the cost-effectiveness of treatments, re-evaluating life possibilities for dementia (having the person make life choices before decisional capacity is absent), and an emphasis on strategic planning. These recommendations help lay the framework for what gaps PTs may fill in the realm of PC and chronic disease management.

The Agency on Healthcare Research and Quality (AHRQ) performed a meta-analysis entitled, "*Palliative Care for Adults",* to analyze the evidence related to practice recommendations and clinical guidelines for PC practitioners.[8] The quality of evidence and strength of recommendations were categorized to provide the reader with improved information to provide evidence-based best practices in PC. Some of the highlights of the findings included that planning for PC should begin relatively soon after diagnosis of a life-threatening illness. Emphasis was on the physical symptoms of a disease causing suffering and that controlling these symptoms should be a priority of PC practitioners. Communication with patients and families was a critical component to set realistic goals, but provide realistic hope. The authors provided additional evidence that PC was compatible, and not mutually exclusive, with all other medical treatments, including curative measures. Even after the patient's death, the healthcare team plays a key role in the grief and bereavement process. When discussing which patients would be appropriate candidates for PC, considerations include disease progression with functional decline, pain or symptoms not responding to treatment, and a need for advance care planning. Functional decline and pain control are two categories of interventions that PTs may be able to assist in the management of. The AHRQ document listed several diagnoses that may prompt PC consideration including "debility/failure to thrive, cancer, heart disease, pulmonary disease, dementia, liver disease, renal disease, and neurologic diseases such as stroke, Parkinson's, amyotrophic lateral sclerosis (ALS), and multiple sclerosis (MS)." Several neuromuscular, musculoskeletal, cardiopulmonary, and neoplastic conditions are within the scope of PTs practice.

An important component to determine the needs of the individual with PC or with a life-threatening illness is the ability to quantify and predict the level of disability, extent of disease, and possibility of death. An article by Lau et al entitled, "*Use of Palliative Performance Scale (PPS) for End-of-Life Prognostication in a Palliative Medicine Consultation Service"* was a prospective analysis of 513 patients evaluated by a British Columbia PC team.[9] The PPS as described by Wilner and Arnold evaluates 5 observer-related domains to rate a patient in various categories to attempt to describe the level of disease and to identify and track potential care needs of PC patients.[10] The 5 categories included:

- ☐ Ambulation;
- ☐ Activity level;
- ☐ Evidence of disease;
- ☐ Self-care, intake; and
- ☐ Level of consciousness.

Besides the *evidence of disease* component, all components are consistently evaluated by PTs as a regular part of their scope of practice. Lau et al report that initial PPS scoring is a significant predictor of survival, challenges are present due to the "ambiguity and difficulty when assessing patients at higher PPS because of the subjective nature of the tool."[9] This tool and its potential usefulness in predicting survival rates may be useful to PTs who are working with individuals facing life-threatening illness in various stages of disease progression. This may provide PTs with improved data on anticipated lifespan and may assist in improved activity prescription and anticipatory equipment prescription for future disease progression.

When shifting focus more specifically toward PT involvement within PC, a key finding in a manuscript based on expert opinions was that "Rehabilitative services are underutilized in the PC setting, and more research is needed to address how patients may benefit as they approach the end of their lives."[11] The role of rehabilitation in PC is described and the evidence for the benefit of rehabilitation for several major diagnoses reviewed. With regard to PT, 3 facets of PT in PC were outlined: (1) Direct patient care, (2) Educating the patient, family, and caregivers, and (3) Functioning as a team member within the interdisciplinary group. In addition to the physical benefits provided, the authors discussed the psychological benefit of PT, including several citations providing evidence for a reduction of psychological suffering when participating in rehabilitation in patients with terminal cancer. In later stages of disease progression, this reduction in psychological distress and suffering may outweigh physical gains and pain reduction that the PTs may be intending to treat. Specific disease conditions and the clinical indications for treatment were described including ALS, advanced dementia, chronic heart failure, COPD, and outlined specific treatment regimens. Some limitations to rehabilitation and PC described included infrastructure of the medical system not supporting the utilization or cost-effectiveness of PT within PC. In addition, functional outcome measures were advocated, such as the Palliative Performance Scale.[10]

Palliative Care and Management Principles in Older Patients with Advanced Chronic Obstructive Pulmonary Disease, a systematic review by Yohannes, a physiotherapist in the United Kingdom examined the care principles for end stage COPD.[12] Psychiatric disorders, in addition to dyspnea from the disease process, are a significant issue. With regard to dyspnea management for this patient population, as expected, this was a key symptom that required management and reported as 90% prevalence of dyspnea at rest or minimal exertion with end stage COPD as compared to end stage CHF with 60%.[12] Application of a fan pointed in the direction of the face displayed early evidence of reducing the symptoms of dyspnea, in addition to commonly utilized medical treatments including supplemental oxygen and medications. Fatigue was a common symptom noted by nearly all patients. Yohannes recommended applying a fatigue rating

scale such as the Manchester COPD Fatigue Scale. This scale is valid and reliable in quantifying fatigue. Pulmonary rehabilitation displayed evidence of reducing and controlling fatigue in this patient population.[12] Specifically, home-based physical rehabilitation was useful in maintaining physical functioning in advanced COPD. Finally, unlike late stage cancer, patients with COPD are less likely to have access to or be aware of PC services, creating a health disparity in this underserved population.[12]

PRevention, Intervention, Sustained wellness Model (PRISM) Care Philosophy

The PRISM care philosophy encapsulates and summarizes multiple concepts of chronic disease management. The intention of the PRISM care philosophy is to provide an educational tool and guiding care philosophy toward care of the patient with a chronic disease or life-threatening illness (Figure 1). The PRISM may be employed to increase the awareness, understanding, and communication of the varied roles that PTs offer in the management of the cancer survivor or patient with a chronic or life-threatening disease. The visual representation of a spectrum provides illustration of the relative continuum of roles that PT is able to offer, as opposed to a dichotomous perspective of "PT or no PT." This spectrum provides for an illustration of a variety of different involvement levels for the physical therapist from an annual check-up with a PT in the early stages of a newly-diagnosed disease through early intervention for an impairment or functional imitation, all the way through providing psychosocial support and comfort measures at the transition to hospice care. To illustrate the clinical applicability and utility of the PRISM care philosophy, the 3 core concepts of *Prevention, Intervention, and Sustained wellness* are summarized with patient examples. These examples are not meant to be all-inclusive, only to highlight common opportunities that PTs may apply each phase of the PRISM care philosophy.

Prevention. Preventative care, a growing area of physical therapist practice, is not commonly considered in the presence of an already diagnosed incurable illness. Physical therapist care in the realm of prevention may take multiple forms in chronic disease. A common consideration of primary prevention is the initial education in avoiding at-risk behaviors, such as poor diet or smoking, or encouraging disease-preventing behaviors, such as exercise and advocating for appropriate cancer screenings. In conditions such as COPD and CHF, a periodic or annual re-evaluation and exercise or activity prescription may be appropriate to maintain strength and functional capacity. These periodic re-evaluations (prospective surveillance) may be able to establish baseline functional levels and identify an early exacerbation of a disease process and prompt further management as appropriation. Prevention in an acute or in-patient setting may include prospective screenings by PTs of all patients within a specific nursing unit to provide early identification and treatment of potential rehabilitation needs before several days of bedrest causes unwanted medical sequelae.

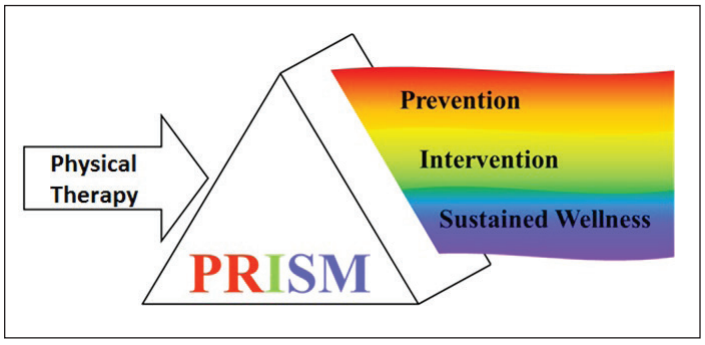


Figure 1. PRISM visual depiction.

Intervention. The intervention phase of the PRISM philosophy is what is commonly considered by interdisciplinary team members, patients, and some PTs as traditional or conventional PT. Although this is what is commonly known as PT, there is perceived to be an underutilization of traditional PT services in the presence of a cancer diagnosis, chronic disease, or other life-threatening illness. A portion of this underutilization may be the perception that in order for traditional PT to meet the tenets of medical necessity, reasonable progress must be made. As many chronic or life-threatening illnesses require the skill of a physical therapist to maintain or slow the decline of function or impairments, conventional intervention-based PT may be indicated. This point was clarified in the US Supreme Court case *Jimmo vs Sebelius* that ruled rehabilitation services may be applicable and medically necessary to maintain or slow the decline of the functional status of a patient with a degenerative illness or an incurable condition.[13] This may include concepts such as cancer rehabilitation or impairment-based interventions for side effects of treatments or surgical procedures. This may also include intervention-based treatments for conditions such as pelvic floor rehabilitation for genitourinary cancers.

Sustained wellness. The last concept entails the relatively stable period after an exacerbation or a change in an incurable or chronic condition where a certain level of activity, exercise, or health maintenance behaviors are required or beneficial to maintain an optimal level of activity and participation and allow the patient/client to enjoy the highest level of QOL for as long of a duration as possible. In the absence of these *Sustained wellness* activities, a patient's condition may worsen; however, within the *Sustained wellness* phase, skilled PT may not be medically necessary, feasible, or appropriate. An example of this is a facilitated exercise regimen provided by an exercise and wellness program for cancer survivors in active treatment and who prefer supervision, guidance, or encouragement with a prescribed exercise regimen. This may include a cancer survivor who has achieved remission or cure but is still working through after effects of chemotherapy, radiation, or surgical intervention, or an "after-care" program for patients with residual deficits after a stroke to maintain a certain level of functioning when their recovery has stabilized and their care no longer meets medical necessity for physical therapist services. These after-care programs may begin with an exercise prescription for a certain level of activity from a physical therapist and facilitated by a healthcare or exer-

cise professional with training in the medical condition and the complexities of the condition. In the *Sustained wellness* phase, the physical therapist may provide a consultative or supportive role with periodic re-evaluations to monitor, correct, or progress an exercise prescription.

DISCUSSION

Based on the literature review, the PRISM care philosophy and the philosophical compatibility with cancer survivorship and PC, several key outcomes should be pursued to assist in PTs managing the patient with a cancer diagnosis, chronic disease, or life-threatening illness. These are discussed in the context of the APTA's Position Statement passed by the House of Delegates in 2011 entitled THE ROLE OF PHYSICAL THERAPY IN HOSPICE AND PALLIATIVE CARE HOD P06-11-14-11.[14]

"Resolved, The American Physical Therapy Association (APTA) endorses the inclusion of the following concepts in hospice and PC:

- ☐ *Continuity of care and the active, compassionate role of PTs and physical therapist assistants in hospice and PC.*
- ☐ *Respect for the rights of all individuals to have appropriate and adequate access to physical therapy services, regardless of medical prognosis or setting.*
- ☐ *An interdisciplinary approach, including timely and appropriate physical therapist and physical therapist assistant involvement, especially during transitions of care or during a physical or medical change in status.*
- ☐ *Education of PTs, physical therapist assistants, and respective students in the concepts related to treating an individual while in hospice and PC.*
- ☐ *Appropriate and comparable coverage and payment for physical therapy services for individuals who have transitioned to hospice or PC in all clinical settings."*

Each component of this position is examined as to the PT profession's capacity, opportunities, and limitations in achieving each of these tenets.

Continuity of care and the active, compassionate role of PTs and physical therapist assistants in hospice and PC.

In this tenet, an emphasis is placed on the continuity of care with the PTs. Currently PTs and PTAs are well-educated in compassionate care in a more episodic manner of care with a clear "evaluation" and "discharge" with the hope of a successful discharge to where the physical therapist will deem a treatment session "successful" when the patient does not have to return to the care of the physical therapist. Although this thought process is changing, it is changing in a slow manner and is incompatible with the care concepts of PRISM and chronic disease management. The management of the individual with a life-threatening illness and chronic disease requires a more longitudinal management of patient care that may require periodic re-assessment, early identification of disease progression, anticipation of crisis events or medical exacerbation, and some periods of conventional intervention-based PT management. The APTA endorses positions that highlight this model of care, including endorsement of an *Annual Visit with a Physical Therapist* and *Physical Therapist of Record and "Hand Off" Communication*.[15,16] These positions provide examples of the PRISM concepts of prevention, early identification, and continuity of services. It is recommended that future professional education incorporate these concepts into curricula. A suggested emerging role of a physical therapist is that of a "rehabilitation navigator," with analogous structures to an attending physician or nurse navigator. This rehabilitation navigator would coordinate care of a variety of rehabilitation professionals throughout the disease process and maintain a long-term relationship with the patient. In addition to providing components of disease-specific treatment, the navigator would assist in symptom monitoring as well as facilitation and consultation of specialist PTs who focus in one area of specialty, while the therapist navigator provides for longitudinal management and coordination of care throughout the disease continuum.

Respect for the rights of all individuals to have appropriate and adequate access to physical therapy services, regardless of medical prognosis or setting.

This objective of the APTA's position on the role of PT in hospice and PC focuses on both ends of the disease spectrum, touching on the awareness and referral process requirements among all stakeholders as to the PTs role in management in *Prevention,* early *Intervention,* and *Sustained wellness* across the continuum of care.[17,18] Some patients may not understand the role of exercise, prevention, wellness behaviors, and early intervention that a physical therapist may offer, especially in the presence of chronic or life-threatening illness. The health care team members who are commonly considered referral sources for PTs, including physicians, mid-level providers, nurse navigators, and social workers, may not have increased awareness of the role of PT in the continuity of care beyond conventional PT care. Historically, the APTA has endorsed improving access to PT services through legislative efforts including direct consumer access, especially as it relates to wellness services. As an example, Michigan was the most recent, and 50th state in the US, to obtain direct consumer access effective January 2015. The legislative rules did place a time limitation on conventional PT visits, but no visit limits on wellness or preventative visits to the PT services.[19]

An interdisciplinary approach, including timely and appropriate physical therapist and physical therapist assistant involvement, especially during transitions of care or during a physical or medical change in status.

In an unpublished doctoral dissertation by Wilson entitled, "*Perceptions of PTs Regarding the Role of Physical Therapy Within Hospice and Palliative Care in the USA and Canada: A Qualitative Study",* a common theme noted by the participants in the study was the importance of integration into the care team to provide an interdisciplinary approach to patient management. There was wide variability and frequent inconsistencies to physical therapist involvement into the Interdisciplinary Team (IDT)/

Interdisciplinary Group (IDG). This emphasizes the importance of continuity of services and hand-off of care described in the APTA position when a patient transfers form one setting to another and that the physical therapist of record should provide a comprehensive description of the overall care philosophy of the patient, recommendations for provision of treatment, precautions, contraindications, and individual patient preferences.[16] Several examples of this hand-off of care includes transition to and from home into hospitalization, discharge from sub-acute rehabilitation or in-patient rehabilitation into the home setting, and transition from outpatient therapy to a cancer survivorship exercise program to wellness or monitoring settings, to name a few.

Education of PTs, physical therapist assistants, and respective students in the concepts related to treating an individual while in hospice and palliative care.

In promoting and proliferating the PRISM role of PT within cancer survivorship, PC and chronic disease management, emphasis is placed on the potential circumstances that PTs and PTAs may be self-limiting their own involvement in cancer survivorship, chronic disease management, and PC services. The authors' clinical observations note that PTs and PTAs lacked awareness of their role or potential role in the longitudinal management of a cancer survivor or the patient with a chronic or life-threatening illness. This lack of understanding of the role of PT within cancer survivorship or PC among PTs and PTAs may cause hesitation among referring physicians or practitioners, even those who are strong advocates of the role of rehabilitation professionals in management of their patients. The evidence of PT within cancer survivorship and PC is continuing to grow; however, a slow transition is expected toward educating PT/PTAs in this role as much of the institutional practice changes have historically occurred through professional education as opposed to post-professional education. Although forward-thinking programs may be adopting cancer survivorship, chronic disease management, and PC principles into curricula, it may take an external credentialing body like CAPTE (Commission on Accreditation of Physical Therapy Education) to successfully drive more widespread change.

Appropriate and comparable coverage and payment for physical therapy services for individuals who have transitioned to hospice or PC in all clinical settings.

Another administrative barrier to this care setting is that current payment structures that do not consistently pay for PT services in direct access, self-referral environments, or near the end of life. The Medicare Hospice per diem rate is noted as a potential limiting factor in PT provision in the end of life care. One added benefit of PC services that creates significant practice opportunities for PTs, is most insurances cover PC services in the same payment structure that they cover traditional PT care. The APTA is creating an Alternative Payment System working with the American Medical Association and the American Occupational Therapy Association to update the PT CPT codes.[20] This new billing, coding, and payment model better positions the PT profession to provide an improved spectrum of care, especially in the interest of wellness and prevention models to better capture the skills of PT in the context of PRISM. This important step to improve support and access to PT includes properly reflecting the decision making capacity of PT as a consultative service as compared to a time-based, intervention-oriented service.

CONCLUSION

The PRISM philosophy incorporates several professional PT positions and encapsulates current evidence on cancer survivorship, management of chronic conditions, and longitudinal care of terminal or life-threatening illness and may have utility to guide and educate care providers on PTs role in managing individuals with these conditions. These core concepts are visualized in the PRISM care philosophy and are outlined in the APTA's position on hospice and PC are not just localized to patients within PC or even those with a chronic disease or a life-threatening illness. These concepts are a philosophical evolution of the profession of PT from an episodic, intervention-based practice, toward a primary care service that incorporates all aspects of participation in life events across the entire spectrum with a focus on QOL. To achieve these lofty goals, the physical therapist and profession must advocate for appropriate policy, education, and payment structures for PTs of the future to manage these individuals in a longitudinal manner.

REFERENCES

1. World Health Organization. WHO Definition of Palliative Care. http://www.who.int/cancer/palliative/definition/en/. Accessed October 7, 2014.
2. Center to Advance Palliative Care. Defining Palliative Care. http://www.capc.org/building-a-hospital-based-palliative-care-program/case/definingpc. Accessed October 7, 2014.
3. Dow KH, Ferrell BR, Haberman MR, Eaton L. The meaning of quality of life in cancer survivorship. *Oncol Nurs Forum*. 1999;26(3): 519-528.
4. Heydarnejad MS, , Hassanpour Dehkordi A, Solati Dehkordi K. Factors affecting quality of life in cancer patients undergoing chemotherapy. *Afr Health Sci.* 2011;11(2):266-270.
5. Koh HK, Sebelius KG. Promoting prevention through the Affordable Care Act. *N Engl J Med*. 2010;363:1296-1299.
6. Meier DE. Increased access to palliative care and hospice services: opportunities to improve value in healthcare. *Milbank Quarterly*. 2011;89:343-380.
7. Lynn J, Adamson DM. *Living Well at the End of Life*. White Paper. Santa Monica, CA: RAND Corporation; 2003.
8. McCusker M, Ceronsky L, Crone C, Epstein H, Green B, et al. *Palliative Care for Adults*. Agency for Healthcare Research and Quality National Guideline Clearinghouse. 2013; 81:92.
9. Lau F, Maida V, Downing M, Lesperance M, Karlson N, Kuziemsky C. Use of the Palliative Performance Scale (PPS) for End-of-Life Prognostication in a Palliative Medicine Consultation Service. *J Pain Symptom Manage.* 2009;37:964-982.

10. Wilner LS, Arnold RM. The Palliative Performance Scale #125. *J Palliat Med.* 2006;9(4):994-994. doi:10.1089/jpm.2006.9.994.
11. Barawid E, Covarrubias N, Tribusio B, Liao S. The Benefits of Rehabilitation for Palliative Care Patients. *Am J Hosp Palliat Care.* December 2, 2013: 104990911351447.
12. Yohannes AM. Palliative care and management principles in older patients with advanced chronic obstructive pulmonary disease. *Aging Health*. 2011;7:401-421.
13. Center for Medicare and Medicaid Services. *Jimmo v. Sebelius Settlement Agreement Fact Sheet.* http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/SNFPPS/Downloads/Jimmo-FactSheet.pdf. Accessed October 7, 2014.
14. American Physical Therapy Association. THE ROLE OF PHYSICAL THERAPY IN HOSPICE AND PALLIATIVE CARE HOD P06-11-14-11. [Position]. http://www.apta.org/uploadedFiles/APTAorg/About_Us/Policies/Health_Social_Environment/RoleHospicePalliativeCare.pdf. Accessed November 14, 2014.
15. American Physical Therapy Association. ANNUAL VISIT WITH A PHYSICAL THERAPIST. HOD P05-07-19-20 [Position] http://www.apta.org/uploadedFiles/APTAorg/About_Us/Policies/Practice/AnnualVisit.pdf. Accessed November 14, 2014.
16. American Physical Therapy Association. PHYSICAL THERAPIST OF RECORD AND "HAND OFF" COMMUNICATION HOD P06-08-16-16 [Position] http://www.apta.org/uploadedFiles/APTAorg/About_Us/Policies/Practice/RecordHandOffCommunication.pdf. Accessed November 14, 2014.
17. American Physical Therapy Association. PHYSICAL THERAPISTS' ROLE IN PREVENTION, WELLNESS, FITNESS, AND DISEASE MANAGEMENT BOD P02-14-02-01 [Position] http://www.apta.org/uploadedFiles/APTAorg/About_Us/Policies/Practice/-calTherapistsRoleInPreventionWellnessFitnessDiseaseMgmtBODP02-14-02-01.pdf. Accessed November 14, 2014.
18. American Physical Therapy Association. PHYSICAL THERAPY FOR OLDER ADULTS HOD P06-06-08-04. http://www.apta.org/uploadedFiles/APTAorg/About_Us/Policies/Practice/OlderAdults.pdf
19. of Michigan. *Public Act 260 of 2014: Occupations; physical therapists; prescription requirement; eliminate under certain circumstances. (Effective: 1/1/2015).* http://www.legislature.mi.gov/documents/2013-2014/publicact/pdf/2014-PA-0260.pdf. Accessed October 7, 2014
20. American Physical Therapy Association. Alternative Payment System. http://policy.apta.org/uploadedFiles/policyaptaorg/National_Issues/APS/PositionPaper_AlternativePaymentSystem.pdf. Accessed November 14, 2014.

(Continued from page 36)

94. Tosti KP, Hackney AC, Battaglini CL, Evans ES, Groff D. Exercise in patients with breast cancer and healthy controls: Energy substrate oxidation and blood lactate responses. *Integr Cancer Ther*. 2011;10(1):6-15.
95. Turner J, Hayes S, Reul-Hirche H. Improving the physical status and quality of life of women treated for breast cancer: A pilot study of a structured exercise intervention. *J Surg Oncol*. 2004;86(3):141-146.
96. Furzer BJ, Wallman KE, Ackland TR, Joske DJL. Test-retest reliability of the aerobic power index submaximal exercise tests in cancer patients. *J Sport Sci Med*. 2012;11(4):690-694.
97. Campbell A, Mutrie N, White F, McGuire F, Kearney N. A pilot study of a supervised group exercise programme as a rehabilitation treatment for women with breast cancer receiving adjuvant treatment. *Eur J Oncol Nurs*. 2005;9(1):56-63.
98. Mock V, Dow KH, Meares CJ, et al. An exercise intervention for management of fatigue and emotional distress during radiotherapy treatment for breast cancer. *Oncol Nurs Forum*. 1997;24:991–1000.
99. Mock V, Burke MB, Sheehan P, et al. A nursing rehabilitation program for women with breast cancer receiving adjuvant chemotherapy. *Oncol Nurs Forum*. 1994;21:899-907.
100. Schwartz AL, Mori M, Gao R, Nail LM, King ME. Exercise reduces daily fatigue in women with breast cancer receiving chemotherapy. *Med Sci Sports Exerc*. 2001;33(5):718-723.
101. Schwartz AL, Nail LM, Chen SL, et al. Fatigue patterns observed in patients receiving chemotherapy and radiotherapy. *Cancer Invest*. 2000;18:11-19.
102. Segal R, Evans W, Johnson D, et al. Structured exercise improves physical functioning in women with stages I and II breast cancer: Results of a randomized controlled trial. *J Clin Oncol*. 2001;19:657-665.
103. Hansen D, Jacobs N, Bex S, D'Haene G, Dendale P, Claes N. Are fixed-rate step tests medically safe for assessing physical fitness? *Eur J Appl Physiol*. 2011;111(10):2593-2599.
104. Ingram C. A home based, physical activity intervention increased physical activity, fitness, and vigour and reduced fatigue in sedentary women with early stage breast cancer. *Evid Based Nurs*. 2006;9(1):19.
105. Stubblefield M, O'Dell MW. *Cancer Rehabilitation Principles and Practice.* Demos Medical; 2009:No. 1.
106. Jones LW, Haykowsky MJ, Swartz JJ, Douglas PS, Mackey J. Early breast cancer therapy and cardiovascular injury. *J Am Coll Cardiol*. 2007;50:1435–1441.
107. Vickers RR. Measurement Error in Maximal Oxygen Uptake Tests. *Naval Health Research Center* and *Bureau of Medicine and Surgery* (M2). Report No. 04-03 San Diego, CA. November 3, 2014.

Rehabilitation Oncology

American Physical Therapy Association
2920 East Avenue South, Suite 200
La Crosse, WI 54601-7202