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 Pfälzer  
*University of Michigan - Flint*

Follow this and additional works at: [http://ecommons.udayton.edu/dpt_fac_pub](http://ecommons.udayton.edu/dpt_fac_pub)

Part of the Biomechanics Commons, Musculoskeletal System Commons, Orthopedics Commons, Other Analytical, Diagnostic and Therapeutic Techniques and Equipment Commons, Other Medical Sciences Commons, Other Medical Specialties Commons, Sports Sciences Commons, and the Therapeutics Commons

This Article is brought to you for free and open access by the Department of Physical Therapy at eCommons. It has been accepted for inclusion in Physical Therapy Faculty Publications by an authorized administrator of eCommons. For more information, please contact frice1@udayton.edu, msclangen1@udayton.edu.
In This Issue

Oncology Section EDGE Task Forces

Head & Neck Cancer: A Systematic Review of Outcome Measures for Temporomandibular-related Dysfunction

Head & Neck Cancer: A Systematic Review of Outcome Measures for Quantifying External Lymphedema

Breast Cancer: Systematic Review of Clinical Measures of Cardiorespiratory Fitness Tests

Prostate Cancer: A Systematic Review of Clinical Measures of Strength and Muscular Endurance

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
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/nuptms.

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

President’s Perspective ................................................................. 4
Lisa VanHoose, PT, PhD

Editor’s Message ............................................................................. 5
Lucinda Pfalzer, PT, PhD, FACSM, FAPTA

RESEARCH REPORTS
Oncology Section EDGE Task Forces
Head & Neck Cancers: A Systematic Review of Outcome
Measures for Temporomandibular-related Dysfunction .............. 6
Mary Lou Galantino, Melissa M. Eden, Bryan A. Spinelli, Ann Marie Flores

Head & Neck Cancers: A Systematic Review of Outcome
Measures for Quantifying External Lymphedema ....................... 15
Ann Marie Flores, Bryan A. Spinelli, Melissa M. Eden, Mary Lou Galantino

Breast Cancer: Systematic Review of Clinical
Measures of Cardiorespiratory Fitness Tests ............................. 24
Jacqueline Drouin, G. Stephen Morris

Prostate Cancer: A Systematic Review of Clinical
Measures of Strength and Muscular Endurance ..................... 37
Mary Insana Fisher, Claire C. Davies, Genevieve Colon, Hannah Geyer, Lucinda Pfalzer

PERSPECTIVES
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 .................... 45
Reyna Colombo, Christopher Wilson

Rehabilitation Oncology

Rehabilitation Oncology (ISSN 2168-3808) is the official publication of the Oncology Section, APTA. Copyright 2015 by 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:

<table>
<thead>
<tr>
<th>Size</th>
<th>Width</th>
<th>Height</th>
<th>Rate (Per Issue)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full page</td>
<td>7”</td>
<td>10”</td>
<td>$400 $350</td>
</tr>
<tr>
<td>1/2 page</td>
<td>7”</td>
<td>5”</td>
<td>$300 $250</td>
</tr>
<tr>
<td>1/4 page</td>
<td>3.5”</td>
<td>5”</td>
<td>$200 $150</td>
</tr>
<tr>
<td>1/8 page</td>
<td>3.5”</td>
<td>2.5”</td>
<td>$100 $75</td>
</tr>
<tr>
<td>Inside Front Cover</td>
<td>7”</td>
<td>10”</td>
<td>$650 $600</td>
</tr>
<tr>
<td>Inside Back Cover</td>
<td>7”</td>
<td>10”</td>
<td>$650 $600</td>
</tr>
<tr>
<td>Outside Back Cover</td>
<td>7”</td>
<td>5”</td>
<td>$650 $600</td>
</tr>
</tbody>
</table>

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
Lucinda (Cindy) Pfalzer, PT, PhD, FAPTA—Editor, University of Michigan, Flint, MI
Maggie Rinehart-Ayres, PT, PhD; Thomas Jefferson University, Philadelphia, PA
Charles McGarvey, PT, MS, DPT, FAPTA; CLM Consulting Services, Rockville, MD
Jacqueline Drouin, PT, PhD; Oakland University, Rochester, MI
Jeanette Lee, PT, PhD; UCSF/SFSU, San Francisco, CA
Antoinette Sander, PT, DPT, MS, CLT-LANA; Northwestern University, Chicago, IL
Kirsten Ness, PT, PhD; St. Jude Children’s Research Hospital, Memphis, TN
Steve Morris, PT, PhD; Wingate University, Wingate, NC
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. Multi-site 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. 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 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...”
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; Melissa M. Eden, PT, DPT, OCS; Bryan A. Spinelli, PT, MS, OCS, CLT-LANA; Ann Marie Flores, PT, PhD, CLT

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. 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%. 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. 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). Temporomandibular joint dysfunction is defined as “functional disturbances of the masticatory system.” 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 arthromechanics, 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. 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. 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. 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. 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.

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).
In persons who receive radiation to the head and neck, trismus is often observed in conjunction with dysphagia.\textsuperscript{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.\textsuperscript{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.\textsuperscript{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.\textsuperscript{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.\textsuperscript{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.\textsuperscript{17} The Oncology Section of the APTA has established task forces for breast, prostate, lung, urology, 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).\textsuperscript{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).

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 1. PRISMA flow diagram for systematic review.](image-url)
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

<table>
<thead>
<tr>
<th>Rating</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Highly Recommend</td>
</tr>
<tr>
<td>3</td>
<td>Recommend</td>
</tr>
<tr>
<td>2A</td>
<td>Unable to Recommend at This Time</td>
</tr>
<tr>
<td>2B</td>
<td>Unable to Recommend at This Time</td>
</tr>
<tr>
<td>1</td>
<td>Do Not Recommend</td>
</tr>
</tbody>
</table>
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. It includes 3 subscales (mastication, vertical dysfunction, and global functional limitation scores, and test score independence from constructs such as depression, somatization, anxiety, and other clinical findings. 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 present” 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.

Table 2. Outcome Measures Sorted by Task Force Rating

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

Table 3. Clinical Utility for Recommended Outcome Measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Equipment Needed</th>
<th>Cost</th>
<th>Ease of Use</th>
<th>Scoring</th>
<th>Normative Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>JFLS-8/ JFLS-20</td>
<td>Outcome measure, pen</td>
<td>Free</td>
<td>Easy</td>
<td>Easy</td>
<td>Not available</td>
</tr>
<tr>
<td>TMD Pain Screener</td>
<td>Outcome measure, pen</td>
<td>Free</td>
<td>Easy</td>
<td>Easy</td>
<td>Not available</td>
</tr>
<tr>
<td>Graded Chronic Pain Scale</td>
<td>Outcome measure, pen</td>
<td>Free</td>
<td>Easy</td>
<td>Moderate</td>
<td>Not available</td>
</tr>
</tbody>
</table>

Abbreviations: JFLS, Jaw Functional Limitation Scale; TMD, temporomandibular dysfunction joint; LDF-TMDQ, Limitations of Daily Function in Temporomandibular Dysfunction Questionnaire
<table>
<thead>
<tr>
<th>Outcome Measure</th>
<th>Psychometric Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Graded Chronic Pain Severity (GCPS)</strong></td>
<td>Internal consistency (Cronbach’s alpha):</td>
</tr>
<tr>
<td></td>
<td>Characteristic Pain Intensity Measure $\alpha = 0.84$</td>
</tr>
<tr>
<td></td>
<td>Interference Measure $\alpha = 0.95$</td>
</tr>
<tr>
<td></td>
<td>Temporal stability:</td>
</tr>
<tr>
<td></td>
<td>Characteristic Pain Intensity CCC = 0.91</td>
</tr>
<tr>
<td></td>
<td>Interference Measure CCC = 0.89</td>
</tr>
<tr>
<td></td>
<td>Convergent and discriminant validity of characteristic pain intensity was supported by</td>
</tr>
<tr>
<td></td>
<td>a substantial association with MPI pain severity (CCC = 0.65) and smaller associations</td>
</tr>
<tr>
<td></td>
<td>with measures of constructs other than pain (eg, depression [CES-D], somatic symptoms).</td>
</tr>
<tr>
<td></td>
<td>Such validity was supported for interference by substantial associations with the MPI</td>
</tr>
<tr>
<td></td>
<td>interference (CCC = 0.52) and dysfunctional measures (CCC = 0.51), and by smaller</td>
</tr>
<tr>
<td></td>
<td>associations with measures of other constructs (eg, the SF-12, CES-D).23</td>
</tr>
<tr>
<td></td>
<td>Sensitivity to change not tested.</td>
</tr>
<tr>
<td><strong>TMD Pain Screener</strong></td>
<td>Internal reliability $\alpha = 0.87$</td>
</tr>
<tr>
<td></td>
<td>Temporal stability of the individual items ($\kappa$) ranged from 0.52 to 0.78, indicating</td>
</tr>
<tr>
<td></td>
<td>fair to excellent agreement.22</td>
</tr>
<tr>
<td></td>
<td>Summary score ICC= 0.83</td>
</tr>
<tr>
<td></td>
<td>Validity for TMJ-related pain:</td>
</tr>
<tr>
<td></td>
<td>Sensitivity= 99%</td>
</tr>
<tr>
<td></td>
<td>Specificity= 97%</td>
</tr>
<tr>
<td></td>
<td>Researchers deemed the content validity to be excellent.</td>
</tr>
<tr>
<td></td>
<td>Sensitivity to change not tested.</td>
</tr>
<tr>
<td><strong>Jaw Functional Limitation Scale (JFLS versions 8 and 20)</strong></td>
<td>Construct validity was established via low correlations with depression, anxiety,</td>
</tr>
<tr>
<td></td>
<td>somatization, pain interference, pain-free opening, and palpation sensitivity, and via</td>
</tr>
<tr>
<td></td>
<td>moderate correlations with pain and jaw symptoms.20</td>
</tr>
<tr>
<td></td>
<td>Internal consistency (Cronbach’s alpha):</td>
</tr>
<tr>
<td></td>
<td>JFLS-8 - $\alpha = 0.87$</td>
</tr>
<tr>
<td></td>
<td>JFLS-20 - $\alpha = 0.95$</td>
</tr>
<tr>
<td></td>
<td>Temporal stability19:</td>
</tr>
<tr>
<td></td>
<td>JFLS-8 CCC= 0.81</td>
</tr>
<tr>
<td></td>
<td>JFLS-20 CCC= 0.87</td>
</tr>
<tr>
<td></td>
<td>Sensitivity to change not tested.</td>
</tr>
</tbody>
</table>

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 $\alpha = 0.84$ for the pain intensity subscale and Cronbach’s $\alpha = 0.95$ 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). 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. 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 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, University of Washington Quality of Life (UW-QOL), Liverpool Oral Rehabilitation Questionnaire version 3 (LORQ), University of California San Francisco (UCSF) Oral Cancer Pain Questionnaire, the Gothenburg Trismus Questionnaire (GTQ), and the Late Effects in Normal Tissues Subjective, Objective, Management and Analytic Scales (LENT/SOMA).

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. 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). 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. 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.

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

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. 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, the Oral Impacts on Daily Performance, TMD Disability Index, and the TMJ Score Questionnaire.

**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. 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 JFLS-8, JFLS-20, and GCPS.
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. 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. Ojo and colleagues 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.

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, 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. 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. 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.

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, shoulder, and lymphedema. 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.oncologypt.org/publications/rehabilitation-oncology-journal

REFERENCES

4. Bensadoun RJ, Riesenbeck D, Lockhart PB, Elting LS, Spijkervet FK, Brennan MT. A systematic review of tris-


EDGE Task Force on Head and Neck Cancer Outcomes
A Systematic Review of Outcome Measures for Quantifying External Lymphedema

Ann Marie Flores, PT, PhD, CLT1; Bryan A. Spinelli, PT, MS, OCS, CLT-LANA2; Melissa M. Eden, PT, DPT, OCS3; Mary Lou Galantino, PT, MS, PhD, MSCE4

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

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. Head and neck cancer represents 3% of all cancer survivors in the United States and is considered an understudied tumor type. 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) with approximately 291,108 people living with oral cavity and pharyngeal cancers in the United States in 2012—likely an underestimate since laryngeal, thyroid, and upper esophageal cancers are also head and neck cancers. 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. 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). 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.

Location of the tumor, tumor and lymph node resection, radiation, and multiple modality treatment are all associated with lymphedema of the head and neck. Other posttreatment factors such as muscle guarding, abnormal posture, and reduced movement in the affected areas may also contribute to lymphedema among HNC survivors. 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.  

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. Complete assessment of lymphedema involves patient subjective complaints, physical examination (ie, pitting, tissue texture), and objective tests and measures. 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).

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). 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. 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. 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.

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; forensics where it is used to estimate size, volume, and topography of missing body parts; manufacturing where it is used for machining; and dentistry where it is used in the creation of dental implants and dentures. 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. However, Harrison et al reported a measurement error range between 0.3–4.0 cm³ largely attributed to changes in positioning. The sensitivity and reliability of this method of edema measurement are high. Ramos and colleagues 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%. Kau et al 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. Together, these studies provide reference values for comparison for future studies. Limited research 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. 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. 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. 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 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. 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 and later modified by Schultze-Mosgau et al. Both were further adapted by Smith and Lewin to include neck measurements used in the MDACC-HNLRS. The facial measures rely on the use of bony landmarks of the face and head. Both MDACC-HNLRS and the ALOHA approach use linear distance measurements. The MDACC-HNLRS uses 12 measures adapted from Gabka et al. 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). 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 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). The only tape measurement found not to have high interrater reliability was the lip to lower neck circumference length (ICC = 0.420). Unlike the TDC, tape measures did not significantly discriminate between HNC survivors with lymphedema and healthy controls. 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. 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 measurements 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 measurements of a cone. Both methods are ideal for conical measures, and therefore, have been limited to use in the extremities. 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. 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. While these methods are useful for the extremities, only the frustum method has been used on a small sample of HNC survivors 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, specific types of HPV cause nearly 32% of all HNC cases and this incidence is expected to rise. 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. 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. Aside from surgical anatomy changes and associated side effects, the addition of side
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 firmness5,6 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

<table>
<thead>
<tr>
<th>Measure</th>
<th>Rating</th>
<th>Relevant Psychometric Properties</th>
<th>Clinical Utility</th>
</tr>
</thead>
</table>
| 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\(^{29,30,54}\)  
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\(^{45}\)  
ICC: 0.97  
Interrater reliability\(^{45}\)  
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\(\text{TM}\)) 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
cord damage, bleeding, oxygen desaturation, bronchospasm, fever, pneumonia, pneumothorax, and death.\textsuperscript{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

<table>
<thead>
<tr>
<th>Section/topic</th>
<th>#</th>
<th>Checklist item</th>
<th>Reported on page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>1</td>
<td>Identify the report as a systematic review, meta-analysis, or both.</td>
<td>15</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structured summary</td>
<td>2</td>
<td>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.</td>
<td>15</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rationale</td>
<td>3</td>
<td>Describe the rationale for the review in the context of what is already known.</td>
<td>15-16</td>
</tr>
<tr>
<td>Objectives</td>
<td>4</td>
<td>Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).</td>
<td>16</td>
</tr>
<tr>
<td>METHODS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protocol and registration</td>
<td>5</td>
<td>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.</td>
<td>16</td>
</tr>
<tr>
<td>Eligibility criteria</td>
<td>6</td>
<td>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.</td>
<td>16</td>
</tr>
<tr>
<td>Information sources</td>
<td>7</td>
<td>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.</td>
<td>16</td>
</tr>
<tr>
<td>Search</td>
<td>8</td>
<td>Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.</td>
<td>Figure 1 Please also see the Data &amp; Methods section on page 16</td>
</tr>
<tr>
<td>Study selection</td>
<td>9</td>
<td>State the process for selecting studies (ie, screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).</td>
<td>16</td>
</tr>
<tr>
<td>Data collection process</td>
<td>10</td>
<td>Describe method of data extraction from reports (eg, piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.</td>
<td>16</td>
</tr>
<tr>
<td>Data items</td>
<td>11</td>
<td>List and define all variables for which data were sought (eg, PICOS, funding sources) and any assumptions and simplifications made.</td>
<td>16</td>
</tr>
<tr>
<td>Risk of bias in individual studies</td>
<td>12</td>
<td>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.</td>
<td>16</td>
</tr>
<tr>
<td>Summary measures</td>
<td>13</td>
<td>State the principal summary measures (eg, risk ratio, difference in means).</td>
<td>NA</td>
</tr>
<tr>
<td>Synthesis of results</td>
<td>14</td>
<td>Describe the methods of handling data and combining results of studies, if done, including measures of consistency (eg, I^2) for each meta-analysis.</td>
<td>NA</td>
</tr>
</tbody>
</table>
REFERENCES


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 usable 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.

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 recurrence 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, arrhythm-
mias, bradycardias, progressive declines in left ventricular function, and heart failure.\textsuperscript{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.\textsuperscript{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.\textsuperscript{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.\textsuperscript{3} 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.\textsuperscript{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.\textsuperscript{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.

\section*{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\textsubscript{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 \textit{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.

\section*{RESULTS}

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

\begin{table}[h]
\centering
\begin{tabular}{|c|c|}
\hline
\textbf{4} & Highly Recommend \tabularnewline
\hline
3 & Recommend \tabularnewline
\hline
2A & Unable to Recommend At this time \tabularnewline
\hline
2B & Unable to Recommend At this time \tabularnewline
\hline
1 & Do not Recommend \tabularnewline
\hline
\end{tabular}
\caption{Breast Cancer EDGE* Evaluation Categories}
\end{table}

*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. Concurrent validity is the degree to which SET measures agree with those obtained using gold standard measures. The statistical analyses used to determine concurrent validity are Intraclass Correlation Coefficients [ICC (model 2, 1 or 2, k)] and Coefficients of Variation (COV). The ICC provides a measure of the agreement between two or more tests expressed as values between 0 and 1.0. Higher values represent better agreement; however, the values are difficult to interpret and there is controversy regarding the level viewed as acceptable. 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. 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. 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.

The COV presents the typical error as the percent change in the mean. 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. 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. 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. The recommended statistics again are the ICC (model 3, 1) and the COV as described previously.

Standard errors of a measure are clinically useful for determining both concurrent validity and test-retest reliability. Like the COV, the SEM is a calculation of the typical error and it is robust to small samples. The SEM is clinically useful as it is reported in the same units as the test data. 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. 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. Both methods require the subject to perform a maximal physical effort and analyze expired gases for O₂ and CO₂ content which is then used to calculate maximal oxygen consumption or VO₂max. 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 (CO₂ released/O₂ consumed) of 1.15 or greater. However, not all individuals are able to reach maximal cardiorespiratory levels due to musculoskeletal, neuromuscular or other physical limitations. It is also not safe for many individuals to perform maximal exercise, particularly if they have known or suspected cardiovascular conditions.

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 assess changes in CRF. The MCID derived from studies in

**Figure 1.** Flow of literature search process for psychometric measures.
### Table 2. Standardized Exercise Test Outcomes: EDGE Task Force Rating & Clinical Utility

<table>
<thead>
<tr>
<th>Standardized Exercise Test</th>
<th>Task Force Rating</th>
<th>Psychometrics</th>
<th>Clinical Utility and Safety</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Maximal Tests with Oxygen Analysis</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Treadmill</strong></td>
<td>1 Do not Recommend</td>
<td>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</td>
<td>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</td>
</tr>
<tr>
<td><strong>Cycle Ergometer</strong></td>
<td>1 Do not Recommend</td>
<td>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</td>
<td>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</td>
</tr>
<tr>
<td><strong>Submaximal Tests with Prediction Equations</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Treadmill</strong></td>
<td>2A Unable to Recommend At this time</td>
<td>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.</td>
<td>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</td>
</tr>
<tr>
<td><strong>Cycle Ergometer</strong></td>
<td>2A Unable to Recommend At this time</td>
<td>Used in 5 breast cancer research studies, Psychometrics poor for sensitivity and concurrent validity with VO$_2$ max 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</td>
<td>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</td>
</tr>
<tr>
<td><strong>Field Tests</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>12-Min Run</strong></td>
<td>1 Do Not Recommend</td>
<td>Not used in breast cancer research, MDD exceeds MCID, Validity of CRF prediction equations not assessed in breast cancer research</td>
<td>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</td>
</tr>
<tr>
<td><strong>12-Min Walk</strong></td>
<td>2A Unable to Recommend At this time</td>
<td>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</td>
<td>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</td>
</tr>
<tr>
<td><strong>1-Mile Walk</strong></td>
<td>2A Unable to Recommend At this time</td>
<td>Used in breast cancer research, MDD and SEM not defined, Validity of CRF prediction equations not assessed in breast cancer research</td>
<td>Easy to administer, Can assess multiple subjects at same time, Requires vital signs monitoring for safety as potential to perform a maximal test</td>
</tr>
<tr>
<td><strong>Step Test (Canadian Aerobic Fitness Test protocol)</strong></td>
<td>2A Unable to Recommend At this time</td>
<td>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.</td>
<td>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</td>
</tr>
</tbody>
</table>

Abbreviations: CRF, cardiorespiratory fitness; MDD, minimal detectable difference; MCID, minimal clinically important difference; SEM, standard of measurement
apparently healthy people is one metabolic equivalent (MET), which is equal to 3.5 ml/kg/min of oxygen. 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.

VO\textsubscript{2max} Treadmill Tests

Extensive study and scientific evidence from other populations supports the VO\textsubscript{2max} treadmill SETs as the gold standard for accurate measurement of CRF. These tests are performed on a motor-driven treadmill with expired gases analyzed for O\textsubscript{2} and CO\textsubscript{2} content. Administration time is about 45 to 60 minutes including time for set-up, preparation, warm-up, testing, and cool-down. 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 has been achieved.

Twenty-three articles were initially found that used VO\textsubscript{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.

### Table 3. Concurrent Validity

<table>
<thead>
<tr>
<th>Exercise Test</th>
<th>Population Assessed</th>
<th>Concurrent Validity\textsuperscript{a}</th>
<th>COV or Percent Difference between Means\textsuperscript{b}</th>
<th>Standard Error of the Measure\textsuperscript{c}</th>
<th>Minimal Detectable Difference\textsuperscript{d} ml/kg/min</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Maximal Tests with Oxygen Analysis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treadmill Test</td>
<td>Healthy Males &amp; Females, Athletes\textsuperscript{17-19}</td>
<td>ICC = .97</td>
<td>&lt;1.5–2.0% (with Douglas Bag)</td>
<td>± 0.05-0.06</td>
<td>± 0.15 - 0.17</td>
</tr>
<tr>
<td></td>
<td>Women with Breast Cancer</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
</tr>
<tr>
<td>Cycle Ergometer</td>
<td>Healthy Males &amp; Females, Athletes\textsuperscript{16-19}</td>
<td>( R^2 = .60-96 )</td>
<td>COV = 1.5 (1.1-2.0)</td>
<td>- 0.7 to 1.3</td>
<td>1.9-3.6</td>
</tr>
<tr>
<td></td>
<td>Women with Breast Cancer (Reported as raw difference and percent)</td>
<td>Not found</td>
<td>- 16.7%\textsuperscript{b} (4.8 ml/kg/min)</td>
<td>Not found</td>
<td>Not found</td>
</tr>
<tr>
<td><strong>Submaximal Tests used with Prediction Equations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treadmill Test</td>
<td>Healthy Males &amp; Females, Athletes\textsuperscript{17-19}. ( R^2 = .76-.88 )</td>
<td>Not found</td>
<td>± 3.4 – 5.3</td>
<td>9.4-14.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Women with Breast Cancer</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Cycle Ergometer</td>
<td>Healthy Males &amp; Females, Athletes, Individuals with Cancer\textsuperscript{16-19, 27, 34, 71} ( R^2 = .61-.96 ) ( R^2 = .50-62 ) \textsuperscript{10}</td>
<td>9.4 - 16.6% -9.0%\textsuperscript{11} (Compared to VO\textsubscript{2max} cycle ergometer)</td>
<td>SEE = 3.12-4.23</td>
<td>11.6-15.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Women with Breast Cancer</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
</tr>
<tr>
<td>12 Minute Run</td>
<td>Healthy Males &amp; Females, Athletes\textsuperscript{17-19, 27} ( R^2 = .02-.85 ) \textsuperscript{not a typo}</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Women with Breast Cancer</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
</tr>
<tr>
<td>12 Minute Walk</td>
<td>Individuals with Heart and Lung Disease \textsuperscript{27, 37} ( R^2 = .24-.27 )</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Women with Breast Cancer</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
</tr>
<tr>
<td>1 Mile Walk</td>
<td>Healthy Males &amp; Females\textsuperscript{16, 34, 38-40} ICC=96-97 ( R^2 = .62-86 )</td>
<td>Not found</td>
<td>SEE = 5.68 SEE 0.325 L/min</td>
<td>Not found</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Women with Breast Cancer</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
</tr>
<tr>
<td>Step Test</td>
<td>Healthy Males &amp; Females, Athletes\textsuperscript{17-19, 33-35} ( R^2 = .62-.85 )</td>
<td>-3.8%</td>
<td>SEM = 4.08 SEE 2.9–4.1</td>
<td>11.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Women with Breast Cancer</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
</tr>
</tbody>
</table>

\textsuperscript{a} The Pearson Product Moment (r) was reported when Intraclass Correlation Coefficients (ICC) were not found.

\textsuperscript{b} Percent difference was reported when the Coefficient of Variation was not found.

\textsuperscript{c} Standard Error of the Estimate (SEE) was reported when Standard Error of the Measure (SEM) was not found.

\textsuperscript{d} Minimal Detectable Difference (MDD) at 95% Confidence Level.
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. Only one article provided data sufficient to demonstrate that the subjects met the VO2 max criteria, which suggests that this mode of SET is valid in this population. Although the psychometric properties for VO2 max 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 VO2 max treadmill tests performed by breast cancer survivors.

VO2 max treadmill tests are the recommended method for evaluating CRF because of the accuracy of its measurements. Despite its status as the gold standard measure, VO2 max treadmill testing is not recommended in national standards written for exercise testing in the clinic. These tests are time-consuming, expensive to perform, and require specialized and costly equipment. These tests also require advanced training for the test administrators and the presence of at least two clinic staff members for safety during testing. Depending on the subject’s physical status, a physician may be required to be on-site during testing. 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. Therefore, based on clinical efficiency and safety concerns, the EDGE rating for maximal treadmill tests with oxygen analysis was: “I - Do not Recommend.”

Table 4. Test-retest Reliability: Standardized Exercise Tests

<table>
<thead>
<tr>
<th>Exercise Test</th>
<th>Population</th>
<th>Test-Retest Correlation</th>
<th>Percent Difference</th>
<th>Standard Error of the Measure ml/kg/min</th>
<th>Minimal Detectable Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Maximal Tests with Oxygen Analysis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treadmill Test</td>
<td>Healthy Males &amp; Females</td>
<td>ICC = .95</td>
<td>2.0-5.0%</td>
<td>COV = 0.99</td>
<td>2.58**</td>
</tr>
<tr>
<td></td>
<td>Women with Breast Cancer</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
</tr>
<tr>
<td>Cycle Ergometer</td>
<td>Healthy Males &amp; Females, Athletes</td>
<td>R2 = .74-94</td>
<td>COV = 0.90</td>
<td>- 8.1%</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>Women with Breast Cancer</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
</tr>
<tr>
<td><strong>Submaximal Tests with Prediction Equations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treadmill Test</td>
<td>Healthy Males &amp; Females, Athletes</td>
<td>R2 = .85</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
</tr>
<tr>
<td></td>
<td>Women with Breast Cancer</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
</tr>
<tr>
<td>Cycle Ergometer</td>
<td>Healthy Males &amp; Females, Athletes, individuals with Cancer</td>
<td>ICC = .95-99**</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
</tr>
<tr>
<td></td>
<td>ICC = .873</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
</tr>
<tr>
<td></td>
<td>Women with Breast Cancer</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
</tr>
<tr>
<td>Step Test</td>
<td>Healthy Males &amp; Females</td>
<td>R2 = .62</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
</tr>
<tr>
<td></td>
<td>Women with Breast Cancer</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
</tr>
<tr>
<td>12 Minute Run</td>
<td>Healthy Males &amp; Females</td>
<td>Φ = .96</td>
<td>Not found</td>
<td>1.6</td>
<td>4.54</td>
</tr>
<tr>
<td></td>
<td>Women with Breast Cancer</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
</tr>
<tr>
<td>12 Minute Walk</td>
<td>Individuals with Heart and Lung Disease</td>
<td>ICC= .96-99</td>
<td>COV= 4.2-8.6</td>
<td>Not found</td>
<td>Not Found</td>
</tr>
<tr>
<td></td>
<td>Women with Breast Cancer</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
</tr>
<tr>
<td>1 Mile Walk</td>
<td>Healthy Males &amp; Females</td>
<td>ICC = .97</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
</tr>
<tr>
<td></td>
<td>Women with Breast Cancer</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
<td>Not found</td>
</tr>
</tbody>
</table>

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.
**VO\textsubscript{max} Cycle Ergometer Tests**

Maximal oxygen consumption testing can also be performed using a bicycle ergometer. As with maximal treadmill exercise testing, VO\textsubscript{max} cycle ergometer tests require subjects to perform a maximal physical effort to criterion endpoints and measures CRF (VO\textsubscript{max}/ml/kg).\textsuperscript{2} A mechanically-braked cycle ergometer provides standardized pedal resistance that can be used in various single and multi-stage protocols.\textsuperscript{2} Since the subject is seated, the test can be used for individuals who are unable or are unsafe to walk on a treadmill.\textsuperscript{2} The test takes approximately 45 to 60 minutes to administer.\textsuperscript{2,3,4} In studies on apparently healthy subjects, test-retest reliability can be good if standardized testing guidelines are followed.\textsuperscript{2,15,16} However, cycle ergometer measures are approximately 10% lower than gold standard measures, which reduces the test’s concurrent validity.\textsuperscript{2,15,20,28,46}

Fifteen studies were found that used VO\textsubscript{max} cycle ergometer tests to determine the CRF levels of female breast cancer survivors.\textsuperscript{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\textsubscript{max} cycle ergometer measures and VO\textsubscript{max} treadmill measures among 12 breast cancer survivors.\textsuperscript{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.\textsuperscript{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.\textsuperscript{16,34,55} In addition, similar to VO\textsubscript{2}max treadmill tests, VO\textsubscript{max} cycle ergometer tests are not recommended in national standards for clinical exercise testing due to the poor clinical efficiency and safety concerns for patients.\textsuperscript{2,5-8,10} Therefore, the EDGE rating for VO\textsubscript{max} 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\textsubscript{max}.\textsuperscript{2,3,4} Therefore, subject preparation, precise heart rate measurement, and valid prediction equations are required for estimates of CRF to be accurate.\textsuperscript{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.\textsuperscript{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\textsubscript{max} treadmill tests.\textsuperscript{2,28,34} Testing takes approximately 20 to 30 minutes and typically the test ends when individu-
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. 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. 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. 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 $.90$. Since vital signs are not monitored during either test, there are safety concerns for individuals with low fitness or cardiovascular risk factors. 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: “I – Do not recommend.”

Five articles were found that used the 12-MWT among breast cancer survivors. 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. 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. 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

Field Tests

Field tests also use derived prediction equations to estimate CRF using distance, time, or heart rate responses during walking or running tests. 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). Field tests are easy to administer, require little equipment, and often can be used to assess multiple individuals simultaneously. 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. 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.
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 test 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 poor accuracy and safety are a concern in the clinical setting. The test also requires energy expenditures of 7-9 METs, rendering it unsafe 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 VO2 max cycle ergometer tests performed by 10 women in treatment for breast cancer.74 Jones et al. examined the safety of VO2 max 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 VO2 max 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.57,58 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 VO2 max 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 = .87) 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 VO2 max 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.8 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.3 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

Rehabilitation Oncology
Vol. 33, No. 2, 2015


39. McArdle WD, Katch FI, Pecgor GS. Reliability and interrelationships between maximal oxygen intake, physical work


(Continued on page 51)
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.

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. 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. 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. 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. 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.

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. 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. 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...
Impairments in strength and muscular endurance have been linked to declines in independence, functional mobility, and subsequent QOL. Activities of daily living (ADLs) and 42% for instrumental activities of daily living (IADLS). Among such patients, the prevalence of fatigue is generally reported to be greater than 65%. Furthermore, complaints of diminished endurance and fatigue persist beyond the treatment timeframe. Diminished muscular endurance and fatigue are also increasingly recognized as a troublesome complaint among patients with cancer. 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. 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). Among such patients, the prevalence of fatigue is generally reported to be greater than 65%. Furthermore, complaints of diminished endurance and fatigue persist beyond the treatment timeframe.

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. 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. 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 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. 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.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. 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). 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.
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. 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.

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. 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, and the loss of strength and muscular endurance impairs functional mobility and subsequent QOL. 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

<table>
<thead>
<tr>
<th>Measure</th>
<th>Equipment Needed</th>
<th>Cost</th>
<th>Ease of Use</th>
<th>Scoring/Interpretation</th>
<th>Normative Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMT</td>
<td>No</td>
<td>Free</td>
<td>High</td>
<td>Easy</td>
<td>Yes*</td>
</tr>
<tr>
<td>1-RM</td>
<td>Yes – weights or machines</td>
<td>Minimal</td>
<td>Moderate</td>
<td>Easy</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>HGS</td>
<td>Yes – dynamometer</td>
<td>Moderate</td>
<td>High</td>
<td>Easy</td>
<td>Yes</td>
</tr>
<tr>
<td>HHD</td>
<td>Yes – dynamometer</td>
<td>Moderate</td>
<td>High</td>
<td>Easy</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Muscle Endurance

<table>
<thead>
<tr>
<th>Measure</th>
<th>Equipment Needed</th>
<th>Cost</th>
<th>Ease of Use</th>
<th>Scoring/Interpretation</th>
<th>Normative Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes – weights or machines</td>
<td>Minimal</td>
<td>Not established</td>
<td>Moderate</td>
<td>No</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Measure</th>
<th>Prostate Cancer EDGE Task Force Rating</th>
<th>Clinical Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hand-grip Strength</td>
<td>3</td>
<td>Equipment is easy to use clinically and staff training is simple. Good clinical utility.</td>
</tr>
<tr>
<td>Hand-held Dynamometry</td>
<td>3</td>
<td>Easy to use clinically; methodology similar to manual muscle testing. Normative data available.</td>
</tr>
<tr>
<td>1 Repetition Maximum</td>
<td>2A</td>
<td>Highly useful in the clinic, but poor psychometric properties do not support use.</td>
</tr>
<tr>
<td>Muscle Endurance</td>
<td>1</td>
<td>Not often used clinically. Psychometric support is limited.</td>
</tr>
</tbody>
</table>

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), 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. 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. 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. 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, and up to 7.3 kg in a healthy population. 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. Overall, validity with strength measured using isokinetic dynamometry is good to excellent. Although reliability is reported as good to excellent in most studies, 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, 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, or straps. 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

<table>
<thead>
<tr>
<th>Measure</th>
<th>Intrarater Reliability (ICC)</th>
<th>Interrater Reliability (ICC)</th>
<th>Test/Retest Reliability (ICC)</th>
<th>Responsiveness to Change</th>
<th>Validity</th>
<th>Clinical Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Manual Muscle Testing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICU Patients:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Knee ext, hip flex, ankle DF = 0.98 (0.97-1.00)</td>
<td>Agreement for detecting significant weakness = 1.00 (0.55-1.00)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-RM Testing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Healthy Population:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Squat/Knee ext:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICC=0.64-0.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Healthy Population:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Squat/Knee ext:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICC=0.94-0.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Healthy Untrained Population:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leg press/knee ext:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICC=0.97-0.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Type 2 DM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leg press/knee ext:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.98-0.99</td>
<td>(0.95-0.99)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Standard Error of Measurement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Healthy Population:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Squat/Knee ext:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICC=3.2-13.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Type 2 DM (using est 1-RM):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.24-3.56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hand-grip Strength</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Healthy Population:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICC = 0.94-0.98</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean of 2 trials:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICC=.93-.95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICC = 0.78-0.82</td>
<td>(0.65-0.89)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Highest Score of 3 Trials</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICC=0.82-0.92</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Community-dwelling Older Adults:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICC = 0.94-0.98</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Advanced Cancer Patients:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spearman’s rho =0.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Standard Error of Measurement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Healthy Population:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Knee ext</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICC=0.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Critically Ill:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Knee ext = 0.78-0.79</td>
<td>(0.32-0.95)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hip flex, knee ext, ankle DF = 0.76-0.94</td>
<td>(0.33-0.97)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Community-dwelling Older Adults:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hip abd/flex; knee ext ICC=0.88-0.94</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Knee ext ICC=0.84-0.95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Advanced Cancer:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICC=0.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Healthy Population:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Knee ext = 0.96-0.97</td>
<td>(0.68-0.99)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Critically Ill:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Knee ext = 0.85-0.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cancer Population:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pearson r=0.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICC=0.96 (0.92-0.98)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Advanced Cancer:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICC=0.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Standard Error of Measure:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Healthy Pop:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Knee ext</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICC=5.5-6.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Critically Ill:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Knee ext = 2.1-2.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cancer Pop (with a pull-gauge):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.02 kg (+/- 3.96)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Minimal Detectable Change:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cancer Pop:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.59 kg</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Healthy Pop:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.7 - 7.1 kg</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Responsiveness (Effect Size):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lower extremity strength using multiple devices/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>populations: ~0.40-1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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.\textsuperscript{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.\textsuperscript{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\textsuperscript{20,22} which used a strain-gauge using the Functional Assessment of Cancer Therapy – Fatigue.\textsuperscript{7} The number of repetitions after 3 months of resistance exercise is lower in men treated for prostate cancer with ADT.\textsuperscript{7,9} Therefore, findings from this study showed an increase in muscular fitness among men treated with ADT using a fixed load repeatedly lifted at a standard rate, and counted the number of repetitions correctly completed.\textsuperscript{7} 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 clinically is not feasible and this systematic review does not support it.

The American College of Sports Medicine recommends that lifting 40-60% of a maximum resistance repeatedly in training will increase muscular endurance.\textsuperscript{39} Intuitively, then, measuring endurance should be completed using repetitions to failure of 40-60% of 1RM. Neither study purporting to measure muscular endurance utilizing this method, although findings that an increase in repetitions suggest an increase in muscular endurance among men undergoing 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 1RM 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.\textsuperscript{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.\textsuperscript{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


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, MA1; Christopher Wilson PT, DPT, GCS2

1Director of Rehabilitation Services – Troy, Beaumont Health, Troy, MI
2Coordinator 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 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.

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 predictable 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. 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. 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. For example, chemotherapy treatment is associated with pain, fatigue, weakness, and the magnitude of symptoms directly affected QOL. 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.

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. 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
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. 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.
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.

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. 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.
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. 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. 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. 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.” 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.” 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.

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. 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%. 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. 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 appropriate. 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.

**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. 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 “aftercare” 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 from 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. 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. 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 PT’s 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


(Continued from page 36)