



Assessing Health-Related Quality of Life Among Survivors of Testicular Cancer: A Systematic Review

American Journal of Men's Health
January-February 2021: 1–11
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1557988320982184
journals.sagepub.com/home/jmh


Michael J. Rovito, PhD, CHES^{1,2} , Agata Bruzzone, BS^{1,2},
Eunkyung Lee, PhD^{1,2}, Humberto López Castillo, MD, PhD^{1,2},
Walker Talton, MA^{1,2}, Lindsay Taliaferro, PhD²,
and Derek Falk, PhD³

Abstract

This study aimed to determine if the current health-related quality of life (HRQoL) tools created for survivors of testicular cancer are collecting the highest quality of data via a two-step methodological critique of both the seminal studies that produced a survivor of testicular cancer HRQoL tool (Phase 1) and the actual tool itself (Phase 2). It is the goal of this current article to present and discuss Phase 1.

A systematic review aimed to assess the methodological quality of studies conducted to create instruments used to measure survivors of testicular cancer HRQoL. Five reviewers independently assessed each study with the 20-item Appraisal Tool for Cross-Sectional Studies (AXIS). Inter-rater agreement and Fleiss' kappa was also assessed to ensure consistency in reported scores. Assessments for the EORTC QLQ-TC 26 and CAYA-T studies were low (AXIS 52.5%; IRA 95%; $\kappa = 0.779$) and fair (AXIS 65%; IRA 80%; $\kappa = 0.599$), respectively. Critical appraisal of the scales included issues within the three core AXIS domains. Primary concerns related to sampling methodology and the lack of a qualitative component of their core conceptual development phase.

Both reviewed seminal studies have significant methodological concerns that question the tools' quality. Next steps include extensive appraisal of the psychometric properties of the EORTC QLQ TC-26 and the CAYA-T to complete the comprehensive review. Accurate and reliable data are necessary to understand survivor of testicular cancer HRQoL and assist in building the bridge of communication between health care professionals and survivors to help to improve patient outcomes.

Keywords

Testicular cancer, quality of life, cancer survivorship, critical appraisal, AXIS

Received October 4, 2020; revised November 24, 2020; accepted November 30, 2020

Overview of Testicular Cancer Survivorship

Despite testicular cancer's (TC) relative rarity within the general male population, it is the most common form of cancer in young men, particularly among those aged 15–35 (Shanmugalingam et al., 2013; Smith et al., 2013). The 5-year survival rate for TC in developed nations is upwards of 95%–99% if the disease is detected early in the staging process (i.e., stages I to IIB) (Smith et al., 2013). Even though survival rates have increased significantly over the past few decades due to advancements in treatment, disparities exist with mortality outcomes between demographic groups. For example, although

White males experience the highest incidence of the disease, Black males have nearly twice the mortality rate

¹Department of Health Sciences, University of Central Florida, Orlando, FL, USA

²Department of Population Health Sciences, University of Central Florida, Orlando, FL, USA

³Department of Social Sciences and Health Policy, Wake Forest University, Winston Salem, NC, USA

Corresponding Author:

Michael J. Rovito, PhD, CHES, Department of Health Sciences, University of Central Florida, 4364 Scorpius St. HPA2-215, Orlando, FL 32816, USA.
Email: Michael.Rovito@ucf.edu



(Li et al., 2020). Further, within the next decade, Latinos are projected to have the highest incidence rates of TC among all races/ethnicities, surpassing White males (Ghazarian et al., 2017). Across all demographic subgroups, generally, TC cases are projected to increase by approximately 25% by 2025 (Matheson et al., 2016).

The fact that nearly one-third of TC cases are diagnosed at stage II or higher is a concern as it pertains to maintaining quality of life (QoL), or even more specific, health-related quality of life (HRQoL). The adverse health outcomes associated with TC survivorship, particularly cases diagnosed at later stages, are well documented (Kim et al., 2011). Survivors of TC experience significantly higher levels of stress related to peripheral neuropathy, the anxiety of relapse, sexual performance, as well as increased risks of depression, respiratory and cardiovascular disease, digestive diseases, infections, and diabetes relative to males not diagnosed with TC (Beard et al., 2013; Kim et al., 2011; Matheson et al., 2016). These individuals also have an increased risk of developing post-traumatic stress disorder and committing suicide (Bukavina et al., 2017; Gunnes et al., 2016). Survivors of TC treated with chemotherapy had a higher all-cause mortality than males who never had TC (Fosså et al., 2007; Klaassen et al., 2018; Oldenburg et al., 2007). It is apparent that survivors of TC are particularly vulnerable to long-term psychosocial outcomes due to the early-age onset of the disease and subsequent long-term (many times, decades) of stress and anxiety, among other outcomes stemming from their diagnosis and treatment (Rovito et al., 2018).

Measuring Health-Related Quality of Life

The World Health Organization (WHO) defines QoL (2014, p. 1) as

“an individual’s perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns. It is a broad ranging concept incorporating in a complex way the person’s physical health, psychological state, level of independence, social relationships, personal beliefs and their relationship to salient features of the environment.”

QoL serves as a primary outcome variable within a variety of disciplines, ranging from neoliberalism and globalization (“Neoliberalism, globalization, and inequalities: Consequences for health and quality of life,” 2020) to sustainability and community development (Vogt et al., 2020). HRQoL has garnered attention in the literature and in practice to serve as a more specific measure of QoL (Latas et al., 2014). This, too, has a far-reaching application in the literature, ranging from women’s health

(Zhu et al., 2016) to adolescent body image (Pollatos et al., 2020) and from stroke research (Martino Cinnera et al., 2020) to epilepsy (Reilly et al., 2020). HRQoL is also among the top measured variables in much of the cancer survivorship literature (Backemar et al., 2020; Bayley-Veloso et al., 2020; Grössmann et al., 2020; Rammant et al., 2020). However, these terms are many times used interchangeably. As a point of clarification for this present discussion, HRQoL will be used to reference both QoL and HRQoL.

The Short Form 36-item Health Survey (SF-36) is arguably the gold standard tool to assess HRQoL with a generic approach to assessing impacts of disease and treatment. The SF-36 generally measures overall health function (e.g., Lins and Carvalho’s argument to use the SF-36 as a single unit measure of HRQoL (Lins & Carvalho, 2016)) after diagnosis of a disease and has demonstrated validity for certain contributors to HRQoL status (Guyatt, 1997). The measure does not include other HRQoL areas, such as emotional, spiritual, financial health, among others, thus failing to measure total HRQoL more comprehensively, and has issues with subscales being able to report accurately upon total physical and mental HRQoL (Guyatt, 1997; Taft et al., 2001). Despite criticism, the SF-36 provides a framework that many researchers use to guide their HRQoL studies (Arian et al., 2019; Lins & Carvalho, 2016).

In other attempts to measure HRQoL, the WHO developed two tools (WHOQOL, 2014): the 100-item version (WHOQOL-100) and the abbreviated, 26-item version (WHOQOL-BREF). The European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 (EORTC QLQ-C30) questionnaire was also developed to measure the overall health summary of QoL (Fayers et al., 2015). The EORTC QLQ-C30 consists of 30 items and must be completed with a complementary disease-specific module to gather information for specific individuals with specific diseases (Fayers et al., 2015). The use of these instruments, however, has not yet reached the same prominence as the SF-36.

There is contention on how best to capture HRQoL information from patients/participants, particularly as it pertains to survivors of TC. For example, Matheson et al. (2016) indicate that universal HRQoL measures tend to be too general to capture valid data on specific health issues (e.g., multiple drug interaction issues after cancer treatment and psychosocial concerns due to the effects of specific chemotherapies, among others). Hand (2016) further indicates that these more general HRQoL tools measure health and functional status, which may exclude influential variables upon HRQoL, such as patient/survivor values and the built environment. In summary, it appears that general HRQoL measures lack the capacity

to capture the specifics of an individual outcome or exposure and/or they lack the capacity to capture the multidimensional nature of HRQoL.

A Call for Better Tools to Assess HRQoL Among Survivors of TC

Matheson et al. (2020) suggest that TC-specific HRQoL assessments were created to understand and promote higher-quality extraction of HRQoL information (i.e., more specific to the disease/outcome in question, increased multidimensionality of core measures) within survivors of TC populations by building upon the framework of the SF-36. Such tools have yet to be fully and regularly incorporated into contemporary practice. It is not clear to these authors why there is a slower uptake in use of TC-specific HRQoL tools.

No formal systematic search has identified all existing survivors of TC-specific HRQoL tools to conduct a comprehensive review of the methodological soundness and practical fitness of each identified tool, although there have been one-off's in the past for one select scale (Sztankay et al., 2018). There is a lack of critique on the methodological quality of the seminal studies conducted to produce such tools. This is an important step that is often overlooked by such reviews. It is indeed important to assess the reliability and validity of the tools' data, but it is paramount to have a critical appraisal of the study that created and supports those tools. These authors feel that in order to truly capture the reality and complexity of survivorship, an appraisal of the seminal data necessary to build such tools should be subject to scrutiny. Such critique would provide more trustworthy findings, and thus offer more valid and applicable tools.

This study aimed to determine if the current survivors of TC HRQoL tools are collecting the highest quality of data via a two-step methodological critique of both the seminal studies that produced a survivors of TC HRQoL tool (Phase 1) and the actual tool itself (Phase 2). It is the goal of this current article to present and discuss the results of the first of two phases.

Methods

Design

A systematic review was conducted to assess the methodological quality of studies carried out with the sole purpose of producing a survivors of TC HRQoL tool (Phase 1). Each identified study was assessed with the Appraisal Tool for Cross-Sectional Studies (AXIS) (Downes et al., 2016). This information was used to inform the development and execution of the subsequent Phase 2 critique of the existing HRQoL tools' psychometric properties

discovered and verified in the systematic search process. This Phase 1 reporting was conducted based on the PRISMA guidelines (Moher et al., 2009). This study was performed with approval was granted by the Ethics Committee of (*information blinded for peer review purposes*).

Inclusion/Exclusion Criteria and Screening Procedure

Studies had to meet the following inclusion criteria to undergo AXIS assessment. First, there had to be a discussion of HRQoL assessment among a cohort of study participants within the study. Second, the study needed to have a discussion of the methods used to create a HRQoL tool (i.e., providing details on the development and implementation of such a tool). Lastly, the HRQoL tool had to be specifically designed for survivors of TC. The articles were screened in the following manner:

Primary Screening. Titles and abstracts were screened for relevance. Articles were moved forward if they discussed HRQoL assessment as its primary topic.

Secondary Screening. Full articles were screened for relevance. Articles discussing HRQoL assessment among a population of males who survived TC were moved forward.

Tertiary Screening. The third assessment further eliminated articles that did not discuss the method of development of an original TC survivorship HRQoL scale.

Data Sources and Search Strategy

The search strategy involved using the Ovid Medline (1950 to present), the Cumulative Index to Nursing and Allied Health Literature (1982 to present), PsycINFO (1806 to present), All EBM Reviews (2020), Ovid Healthstar (1966 to present), the Education Resources Information Center (1966 to present), and Google Scholar (2020) databases to locate relevant literature. Further, the review used ancestry (citation tracking) and grey literature (e.g., unpublished reports, PhD theses) searches to ensure full capture of relevant research. Keywords used for the search included several variants of TC, QoL, HRQoL, assessment, survey, scale, questionnaire, and survival.

Methodological Quality Assessment

Five raters independently appraised the methodological integrity of the qualified cross-sectional studies with the AXIS instrument (Downes et al., 2016). AXIS specifically

appraises the quality of study design, sample size and characteristics, measures, internal consistency, results, analysis, and limitations. The AXIS checklist is comprised of 20 questions in three primary categories: quality of reporting (7 questions), study design quality (7 questions), and possible introduction of biases (6 questions). AXIS does not provide an established rule for determining the quality of each study. The tool instead provides a means for a rater to critique each portion of a study to make a judgment on overall quality, which, according to Downes et al. (2016), may provide for an element of subjective assessment. For the present review, five raters (all professionals in the health sciences field) individually scored each identified cross-sectional study. The 20 AXIS questions were each originally scored by assigning numeric values to two categorical responses: Yes (scored as 1) and No (scored as 0).

After all raters completed their scoring, a series of peer conferences were held to help establish consensus on each of the 20 items if there were any discrepancies. Consensus was defined as 80% of the raters reaching an agreement (i.e., 4 out of 5 raters). If there was no consensus on an item, the item was scored as Partially Fulfilled (scored as 0.5). After each item had a determination of consensus on its score, the values were then summed for a total appraisal score value to determine the level of methodological quality of the study. Fleiss' kappa was calculated for assessing the consistency of observer agreement for each of the identified scales.

As observed in previous critical appraisals using AXIS, a percent value was predetermined at the beginning of this study to identify publications as high or low quality (Arab Alkabeya et al., 2019; Boxberger and Reimers, 2019; Henderson et al., 2019). For this study's critical appraisal, the authors predetermined the value of a high-quality publication as total appraisal scores equal to or exceeding 70% of the total (i.e., at least 14 out of 20 questions scored as 1, or a score ≥ 14 to compensate for some items where there was no consensus scored at 0.5). If the publication score was between 60% and 69.9%, it was considered fair quality, and below 60% was considered low quality.

Results

Tools Identified

Database searches identified 786 unique publications that passed primary screening (Figure 1). Articles were first excluded if the studies did not use specific HRQoL tools for survivors of TC ($n = 743$). For the remaining 43 articles, the full text was then reviewed, followed by exclusion of duplicates or different phases of the same HRQoL

tool specific to survivors of TC ($n = 41$). Overall, two studies met all inclusionary and exclusionary criteria out of the 786 total publications and were included in this critical appraisal review to be further examined for methodological rigor: (1) the European Organization for Research and Treatment of Cancer's Quality of Life Questionnaire for Testicular Cancer, 26 items (EORTC QLQ-TC 26) and (2) the Cancer Assessment for Young Adults for Men with Testicular Cancer (CAYA-T).

Holzner et al.'s (2013) study produced the EORTC QLQ-TC26 via in-depth literature reviews and qualitative interviews with 28 experts in the field (ranging from urologists to psychologists to nurses from Austria, the Netherlands, Italy, Canada, and England) and 62 survivors of TC (from Austria, Canada, and the Netherlands). This tool is a supplemental module questionnaire that is TC-specific and must be used alongside the core questionnaire, the EORTC QLQ-C30. This self-administered tool was published in 2012 and comprises 26 items grouped into seven multi-domain scales (treatment side effects, treatment satisfaction, future perspective, communication, sexual activity, functioning, and enjoyment) and six single items (job and education problems, physical limitations, family problems, infertility, body image problems, testicular transplant satisfaction).

Hoyt et al.'s (2013) study produced the CAYA-T via in-depth literature reviews and qualitative interviews with 21 patients aged 18–29 recruited from the California Cancer Care Registry. The prototype tool was then pilot tested on 15 demographically similar survivors of TC. This tool assesses seven biopsychosocial domains specific to survivors of TC. This questionnaire was published in 2013, it is self-administered, and comprises of 90 items grouped into 17 HRQoL domains: physical, sexual confidence, sexual functioning, body image strength, positive masculine self-image, positive adult self-image, cognitive-emotional regulation, disclosure ability, relationship maintenance, social connectedness, health-care confidence, goal navigation, goal facility, financial maintenance, recreational pursuit, spiritual stability, and finding meaning.

Critical Appraisal

The critical appraisal of the EORTC QLQ-TC 26 included the following scores: a total appraisal score of 10.5/20 (52.5%), which indicates a low-quality study, a 95% inter-rater agreement score, and a Fleiss kappa statistic of 0.779, which suggests substantial agreement among raters (Sim & Wright, 2005) (see Table 1). There was a discrepancy among the raters on whether the data were adequately described or not by the authors.

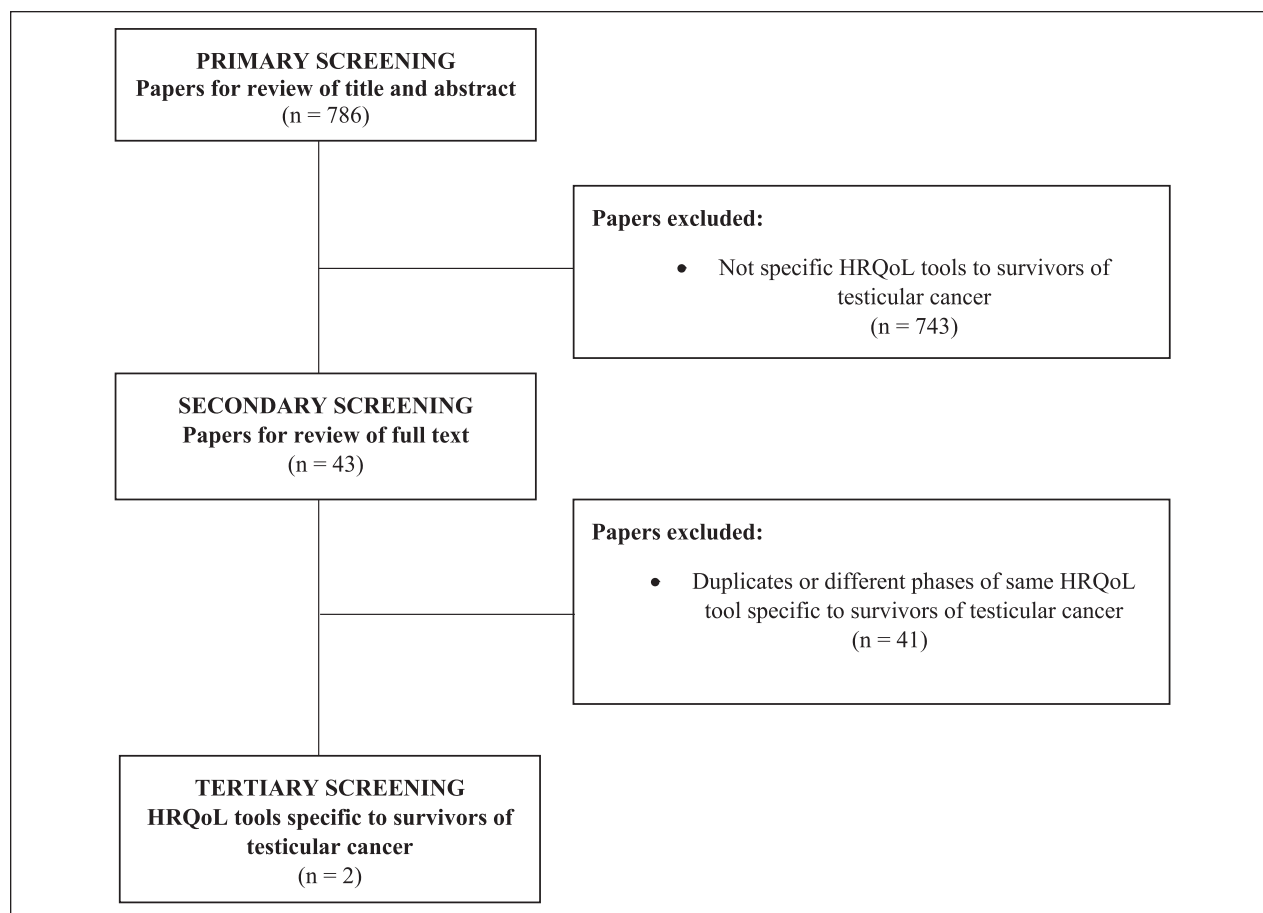


Figure 1. PRISMA flow chart.

The critical appraisal of the CAYA-T resulted in the following comprehensive scores: a total appraisal score of 13/20 (65%), which indicates a fair-quality study, an 80% inter-rater agreement score, and a Fleiss kappa statistic of 0.599, which suggests moderate agreement among raters (Sim & Wright, 2005). There was a discrepancy on the discussion primarily concerning the sampling methodologies employed by the authors, including the representativeness of the study population and possible non-response biases.

Discussion

Key Findings

Improved treatment and prognosis of TC is increasing not only the number of survivors of the disease, but also the length of their lives. The literature outlines unique challenges survivors of TC have in improving and maintaining HRQoL outcomes post-treatment. There are some concerns that existing tools may not to capture the most comprehensive HRQoL information possible among the

population. It was, therefore, imperative to initiate a formal two-phased process to determine the methodological rigor of the original studies that produced survivors of TC HRQoL tools. Phase 1's critical appraisal of the seminal study from which these TC-specific HRQoL tools stemmed provided an opportunity for us to begin the process of determining if the tools are built upon the actuality and involvedness of the survivor experience and the overall methodological quality of existing tools. This information provides a necessary foundation to conduct a critique of each tool's applicability in the field.

This research identified two survivors of TC HRQoL tools (the EORTC QLQ-TC26 and the CAYA-T) and provided a critical appraisal to determine the soundness of their seminal cross-sectional study design. Phase 1 findings suggest that there are significant methodological concerns for each reviewed seminal study. The Hoyt et al. (2013) (CAYA-T) and the Holzner et al. (2013) (EORTC QLQ TC-26) appraisals indicate a failure to: (1) recruit a diverse sample that represents a spectrum of perceptions and attitudes to inform the development of the scale, (2) provide an explanation for the possible presence of

Table 1. Study Quality Scores Obtained for the Appraisal Tool for Cross-Sectional Studies (AXIS).

Items	Consensus scores*		
	Holzner et al. (EORTC QLQTC-26)	Hoyt et al. (CAYA-T)	
1	Were the aims/objectives of the study clear?	1	1
2	Was the study design appropriate for the stated aim(s)?	1	1
3	Was the sample size justified?	0	0.5
4	Was the target/reference population clearly defined? (Is it clear who the research was about?)	1	1
5	Was the sample frame taken from an appropriate population base so that it closely represented the target/reference population under investigation?	0	0.5
6	Was the selection process likely to select subjects/participants that were representative of the target/reference population under investigation?	0	0
7	Were measures undertaken to address and categorize non-responders?	0	0.5
8	Were the risk factor and outcome variables measured appropriate to the aims of the study?	1	1
9	Were the risk factor and outcome variables measured correctly using instruments/measurements that had been trialed, piloted or published previously?	1	1
10	Is it clear what was used to determine statistical significance and/or precision estimates? (e.g., <i>p</i> -values, confidence intervals)	0	1
11	Were the methods (including statistical methods) sufficiently described to enable them to be repeated?	0	1
12	Were the basic data adequately described?	0.5	0
13	Does the response rate raise concerns about non-response bias?	0	0
14	If appropriate, was information about non-responders described?	0	0.5
15	Were the results internally consistent?	1	1
16	Were the results presented for all the analyses described in the methods?	1	1
17	Were the authors' discussions and conclusions justified by the results?	1	1
18	Were the limitations of the study discussed?	1	0
19	Were there any funding sources or conflicts of interest that may affect the authors' interpretation of the results?	0	0
20	Was ethical approval or consent of participants attained?	1	1
	Total appraisal score	10.5/20 (52.5%)	13/20 (65%)
	Study quality	Low	Fair
	Inter-rater agreement	19/20 (95%)	16/20 (80%)
	Fleiss' kappa	0.779	0.599

Note. *Score meaning: 1 = Yes; 0 = No; 0.5 = Partial.

non-response and ascertainment biases, nor provide ways these limitations were mitigated, and/or (3) report on the qualitative nature of how interviews/focus groups were conducted (i.e., there was a lack of any comprehensive discussion on the qualitative methods used to establish a thematic structure). Each study had unique limitations in each of the AXIS tool's categories: *quality of reporting*, *study design quality*, and *possible introduction of biases*. Hoyt et al.'s (CAYA-T) study was the more rigorous of both tools as it scored higher in each of the three categories, but it did have its fair share of limitations.

For Category 1: Quality of reporting, both the Hoyt et al. (2013) and Holzner et al. (2013) studies discussed the aims of their research, clearly stated the target population, and reported on the results that were outlined in their methods. There was some variation related to the reporting of significance testing, data, and limitations. Unlike Hoyt et al.'s (2013) study, there was a lack of in-depth discussion in Holzner et al.'s (2013) study on the qualitative methods used to establish a thematic structure from the semistructured interviews conducted with patients. This is problematic as it serves as a very important part of

the scale-developmental process. Proper, meaningful qualitative data collection, reporting, and analysis serve as fundamental building blocks for any assessment tool. We support Nassar-McMillan and Borders (Nassar-McMillan & Borders, 2002) assertion that qualitative methods in survey design “provide an efficient means for the purposes of both item generation and refinement.” (p. 2) Not providing that information in any significant detail is detrimental to fully understanding the conceptual framework of the tool in question, which casts doubt on its applicability in the field.

For Category 2: Study design quality, the primary concerns stemmed from each study’s sample demographics. Holzner et al. (2013) and Hoyt et al.’s (2013) studies both recruited non-diverse demographic samples, which affect the generalizability of their produced tools. Holzner et al.’s (2013) sample ($n = 62$ in Phase 1 and $n = 156$ in Phase 3) consisted of what appeared to be upper-middle class, nearing-middle aged males from Australia, Austria, Canada, Italy, the Netherlands, and Spain. There was no mention of racial/ethnic diversity, which lets us assume there was a lack of it within the sample. The authors, interestingly, made an emphasis to mention which language each country primarily spoke. Language diversity does not guarantee a diversified sample based on race/ethnicity. In fairness, we acknowledge this was designed via the European Organization for Research and Treatment of Cancer, which would help explain the use of their limited selection pool.

Hoyt et al. (2013) was more thorough in ensuring a representative sample ($n = 171$) than Holzner et al. (2013). Although the sample was recruited from the California Cancer Registry, which may lead to generalizability concerns, the authors made mention of race/ethnicity, sexual orientation, and more in-depth socioeconomic indicators as part of their recruitment methodology. Why is this important if TC manifests itself most of the time among White males? Minority males tend to be disproportionately affected by late-stage discovery and suffer from worse HRQoL outcomes as compared to their White counterparts (Ghazarian et al., 2017; Li et al., 2020). For example, Black males have nearly twice the mortality rate and Latinos are projected to have the highest incidence rates of TC among all race/ethnicities by the year 2026 (Ghazarian et al., 2017; Li et al., 2020). These data indicate a significant shift in the burden associated with TC now resides with minority males. The data on, and perspectives of, these men, however, are largely absent from the current peer-reviewed survivors of TC literature (Oldenburg et al., 2007), particularly pertaining to HRQoL. Hoyt et al. (2013), for example, had two Black/African Americans enrolled in their study, representing 1.2% of the total study population. So, although selecting from a demographically varied cancer registry

has the best chances of soliciting a diverse sample, this sometimes fails to result in a true demographic representation.

The literature calls for research to expand into minority populations to understand this shift in TC incidence, prevalence, and mortality, yet such requests remain mostly unanswered (Klaassen et al., 2018). Bukavina et al.’s (2017) work, for example, is a clear call to action for HRQoL research to expand further into minority populations. As it pertains to survivors of TC HRQoL, we are failing to understand the importance of promoting more diversity in our samples, let alone achieving even the most modest of inclusion standards. The field must put forth a better effort at recruiting more inclusive samples.

Sztankay et al. (2018) attempted to expand the sampling pool in their attempt to validate the EORTC QLQ TC-26 but fell short. The authors followed what appeared to be Holzner et al.’s (2013) sampling methodology but swapped representation from Australia and Canada for Serbia and Poland. Hoyt et al., on the other hand, had a more representative sample but it came only from California, thus also questioning its ability to generalize to other populations, also at the global scale. For survivors of TC HRQoL tools to be truly generalizable, we need to broaden our sampling pools, especially into traditionally underserved communities, both globally and in the United States. These would include populations in which TC, historically, was not a concern but is now a rising point of interest due to emerging disparities in mortality and late-stage morbidity outcomes.

For Category 3: Introduction of biases, much of the concern for both the Hoyt et al. (2013) and the Holzner et al. (2013) studies stemmed from the presence of non-response and ascertainment biases, which relates to Category 2’s sampling methodology weaknesses. However, what is unclear in both studies is not necessarily the presence of non-response bias, but the influence. There is a lack of discussion about non-responders and how that may have affected the results. Both studies do, however, provide information on the inclusionary/exclusionary processes that narrowed the sample size down to its analyzed level. Hoyt et al. provided more information on this phenomenon than Holzner et al. (2013), thus helping legitimize their sample. Finally, there are reports of internal consistency on the part of Hoyt et al. (2013). Holzner et al. (2013) offers up a bit less but both studies allude to their scales being reliable and valid. Phase 2 of this two-phased critique of both scales will address this issue more in-depth.

Comment on HRQoL Assessment

The quality of general HRQoL tools has been debated for the past few decades. Gill and Feinstein (1994) conducted

a critical appraisal of HRQoL and suggested that many, if not most, tools assess patient health status rather than actually measuring HRQoL. Guyatt (1997) indicated that most general measures of HRQoL (in this case, the SF-36) “focuses to a large extent on how patients are functioning” (p. 720) instead of HRQoL. These works indicate that HRQoL is more than physical factors of health and wellness, comprised of mental, functional, and emotional well-being. We cannot agree more.

This idea is supported in the TC-specific HRQoL literature by the works of Smith et al. (2016) and Alexis et al. (2020). Those specific studies offer a perspective of multidimensionality related to TC-related HRQoL instead of the traditional unidimensional approach to measuring QoL among survivors of TC, such as Lockett et al.'s (2008) study that called for continued use of the SF-36. Primarily assessing physical health status with minimal effort to measure any other dimension (i.e., mental, emotional, spiritual, environmental, financial) would be shortsighted.

One other concern is the incorporation, or lack thereof, of TC survivor voices in the development of TC-specific HRQoL tools. Gill and Feinstein (1994) asserted that HRQoL can only be appropriately assessed via incorporating the opinions of actual patients and not rely so much upon the word of, as they put it, “experts.” It appears that Holzner et al.'s (2013) EORTC QLQ TC-26 counters this notion almost completely. According to Holzner et al. (2013), “. . . comments by patients were very rare and therefore did not have a substantial impact on item selection and item wording.” (p. 373) Further, the authors state that “Overall, patients made only a low number of comments indicating good acceptance and understandability of the items.” (p. 373) This is a stretch. One cannot assume just because a participant does not comment on a tool or an individual item that it implies comprehension. This appears counterintuitive to a patient-centered approach necessary for measuring HRQoL.

Hoyt et al. (2013), on the other hand, reported a fairly inclusive method to incorporate the patient voice into the development of the CAYA-T. There was no indication of the role health and wellness professionals played in the creation and/or validation of the tool's thematic structure. A more responsible approach to creating such a tool would be to strike a balance between the patients' and providers' voices in the tool's developmental phases. We believe that the incorporation of more representative voices would increase the scope of the tool to include topic otherwise unknown if not for the expansion. For example, there are some important topics not covered by either tool, including fear of recurrence and body image concerns beyond the lack of a testicle, such as overweight/obesity, which many survivors of TC experience post-treatment.

Strengths and Limitations

This critical appraisal is not without its own limitations. For example, the use of AXIS is a limited measuring approach, especially when it refers to an unstandardized interpretation of the final score. AXIS does not provide an established rule of determining the quality of each study. The tool instead provides a means for a rater to critique each portion of a study to make a judgment on overall quality, which, according to Downes et al. (2016) may provide for an element of subjective assessment. However, our numeric rating system of percentage of agreement helped provide a more objective approach to assessing overall quality.

It was, in our collective opinion, essential to employ a critical appraisal analysis via the AXIS tool for this review, as Downes et al. (2016) suggests, due to its focus on an integration of the “best external evidence with clinical care” (p. 1). They further note that “when reading any type of evidence, being critical of all aspects of the study design, execution, and reporting is vital for assessing its quality before being applied to practice.” (p. 1) It bears repeating that tools/instruments/methods employed in a clinical setting, used to measure clinical outcomes, and/or inform clinical care decisions, demand rigorous testing to ensure quality data. This lends rationale for why we opted to employ our two-phased process in this review. Critiquing the design of a study that produced an assessment tool is just as vital to determining the overall quality of said tool by testing for its psychometric properties. Considering the information collected from these scales have clinical implications, we found the use of AXIS to be appropriate for this review.

The scope of this review, admittedly, is limited. However, this is due to the dearth of survivors of TC research in the field. There are some notable exceptions (e.g., Saab et al., 2014). There is an ever scarcer amount of literature on survivors of TC HRQoL scale development and use. This, unfortunately, is one of the inherent challenges of scale development, particularly with a relatively rarer disease in a select population.

Our consensus scoring technique method could have possibly introduced bias into the study. For example, we decided to conference rather than use raw ratings in the calculation of agreement, which could possibly inflate the % agreement. However, we had a 3-point scoring system to account for non-consensus instead of a compete forcing of compliance or noncompliance.

Conclusion

A critical appraisal of an assessment tool is a necessity to ensure the collection of accurate and consistent data. One step that is overlooked, is an assessment of the study that

originally produced the tool. This critical appraisal can offer the field a more comprehensive analysis so as to properly comment on a tool's quality and determine its applicability within the field.

In our collective opinion, both reviewed studies have significant enough concerns pertaining to the seminal studies' methodology to question the data's reliability and validity. The next step would be an extensive joint review of the psychometric properties of the EORTC QLQ TC-26 and the CAYA-T to compete the comprehensive review. Considering that these are the only two TC-specific HRQoL scales available, there is also a need to determine the generalizability and applicability of the scales, which the findings of this appraisal call out as a major concern.

Implications for Cancer Survivors

The development of a multidimensional tool is needed to focus on these areas of concern to provide researchers, health-care professionals, and patients the highest and most relevant information regarding survivors of TC HRQoL. Having access to specific HRQoL information for this population will help us understand how to properly approach and address HRQoL setbacks for these individual's post-treatment. Thus, allowing medical personnel and others the ability to ultimately increase the HRQoL for survivors of TC.

The limitations surrounding the ability to confidentially measure HRQoL specific to TC presents researchers and practitioners the opportunity to develop a true gold standard tool to accurately measure the HRQoL among survivors of TC. A HRQoL assessment created specifically for survivors of TC that has been thoroughly tested for methodological rigor (both the tool and its seminal study) can assist researchers and practitioners in understanding the needs and wants of this population from diagnosis through, potentially, the rest of their lives. The creation of such a tool should (1) lead to the understanding of both health function and health status and (2) capture the multidimensional factors of HRQoL among survivors of TC. This, in turn, will allow researchers and practitioners to collect the highest quality of data and provide the opportunity to improve the overall HRQoL of survivors of TC.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this

article: Funding for this work was granted by the University of Central Florida's School of Social Work's Center for Behavioral Health Research and Training (CBHRT). This publication was also supported in part by Grant Number (T32CA122061-12) from the NCI (Dr. Falk).

ORCID iD

Michael J. Rovito  <https://orcid.org/0000-0001-8086-3460>

References

- Alexis, O., Adeleye, A. O., & Worsley, A. J. (2020). Men's experiences of surviving testicular cancer: An integrated literature review. *Journal of Cancer Survivorship, 14*(3), 284–293.
- Arab Alkabeya, H., Hughes, A. M., & Adams, J. (2019). Factors associated with hand and upper arm functional disability in people with rheumatoid arthritis: A systematic review. *Arthritis Care & Research, 71*(11), 1473–1481.
- Arian, M., Mirmohammadhani, M., Ghorbani, R., & Soleimani, M. (2019). Health-related quality of life (HRQoL) in beta-thalassemia major (β -TM) patients assessed by 36-item short form health survey (SF-36): A meta-analysis. *Quality of Life Research, 28*(2), 321–334.
- Backemar, L., Johar, A., Wikman, A., Zylstra, J., Gossage, J., Davies, A., Lagergren, J., & Lagergren, P. (2020). The influence of comorbidity on health-related quality of life after esophageal cancer surgery. *Annals of Surgical Oncology, 27*(8), 2637–2645.
- Bayley-Veloso, R., Szabo, Y. Z., Cash, E., Zimmaro, L., Siwik, C., Kloecker, G., Salmon, P., van der Gryp, K., & Sephton, S. E. (2020). The association between history of traumatic events and health-related quality of life among lung cancer patients. *Journal of Psychosocial Oncology, 38*(5), 627–634.
- Beard, C. J., Travis, L. B., Chen, M. H., Arvold, N. D., Nguyen, P. L., Martin, N. E., Kuban, D. A., Ng, A. K., & Hoffman, K. E. (2013). Outcomes in stage I testicular seminoma: A population-based study of 9193 patients. *Cancer, 119*(15), 2771–2777.
- Boxberger, K., & Reimers, A. K. (2019). Parental correlates of outdoor play in boys and girls aged 0 to 12—A systematic review. *International Journal of Environmental Research and Public Health, 16*(2), 190.
- Bukavina, L., Zaramo, C. E., Tarabonata, L., & Modlin, Jr, C. S. (2017). Perceptions of health-related quality of life: A minority men's perspective. *American Journal of Men's Health, 11*(4), 990–998.
- Downes, M. J., Brennan, M. L., Williams, H. C., & Dean, R. S. (2016). Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS) *BMJ Open, 6*(12), e011458.
- Fayers, P., Aaronson, N. K., Bjordal, K., Groenvold, M., Curran, D., & Bottomley, A. (2015). *The EORTC QLQ-C30 scoring manual 2001*, 3rd edn. European Organisation for Research and Treatment of Cancer.
- Fosså, S. D., Gilbert, E., Dores, G. M., Chen, J., McGlynn, K. A., Schonfeld, S., Storm, H., Hall, P., Holowaty, E.,

- Andersen, A., & Joensuu, H. (2007). Noncancer causes of death in survivors of testicular cancer. *Journal of the National Cancer Institute*, *99*(7), 533–544.
- Ghazarian, A. A., Kelly, S. P., Altekruze, S. F., Rosenberg, P. S., & McGlynn, K. A. (2017). Future of testicular germ cell tumor incidence in the United States: Forecast through 2026. *Cancer*, *123*(12), 2320–2328.
- Gill, T. M., & Feinstein, A. R. (1994). A critical appraisal of the quality of quality-of-life measurements. *Journal of the American Medical Association*, *272*(8), 619–626.
- Grössmann, N., Robausch, M., Rothschedl, E., Wild, C., & Simon, J. (2020). Publicly accessible evidence of health-related quality of life benefits associated with cancer drugs approved by the European Medicines Agency between 2009 and 2015. *European Journal of Cancer*, *129*, 23–31.
- Gunnes, M. W., Lie, R. T., Bjørge, T., Ghaderi, S., Ruud, E., Syse, A., & Moster, D. (2016). Reproduction and marriage among male survivors of cancer in childhood, adolescence, and young adulthood: A national cohort study. *British Journal of Cancer*, *114*(3), 348–356.
- Guyatt, G. (1997). Insights and limitations from health-related quality-of-life research. *Journal of General Internal Medicine*, *12*(11), 720.
- Hand, C. (2016). Measuring health-related quality of life in adults with chronic conditions in primary care settings: Critical review of concepts and 3 tools. *Canadian Family Physician*, *62*(7), e375–e383.
- Henderson, S. E., Brady, E. M., & Robertson, N. (2019). Associations between social jetlag and mental health in young people: A systematic review. *Chronobiology International*, *36*(10), 1316–1333.
- Holzner, B., Efficace, F., Basso, U., Johnson, C. D., Aaronson, N. K., Arraras, J. I., Smith, A. B., Chow, E., Oberguggenberger, A. S., Bottomley, A., & Steiner, H. (2013). Cross-cultural development of an EORTC questionnaire to assess health-related quality of life in patients with testicular cancer: The EORTC QLQ-TC26. *Quality of Life Research*, *22*(2), 369–378.
- Hoyt, M. A., Cano, S. J., Saigal, C. S., & Stanton, A. L. (2013). Health-related quality of life in young men with testicular cancer: Validation of the Cancer Assessment for Young Adults (CAYA). *Journal of Cancer Survivorship*, *7*(4), 630–640.
- Kim, C., McGlynn, K. A., McCorkle, R., Erickson, R. L., Niebuhr, D. W., Ma, S., Graubard, B., Aschebrook-Kilfoy, B., Barry, K. H., & Zhang, Y. (2011). Quality of life among testicular cancer survivors: A case-control study in the United States. *Quality of Life Research*, *20*(10), 1629–1637.
- Klaassen, Z., Chandrasekar, T., Goldberg, H., Arora, K., Sayyid, R. K., Victor, J. C., Fleshner, N. E., Hamilton, R. J., & Kulkarni, G. S. (2018). Mp84-16 factors associated with early mortality and survival outcomes among patients with testicular neoplasms: A population-level analysis. *The Journal of Urology*, *199*(4S), e1131–e1131.
- Latas, M., Stojković, T., Ralić, T., Jovanović, S., Špirić, Ž., & Milovanović, S. (2014). Medical students' health-related quality of life: A comparative study. *Vojnosanitetski Pregled*, *71*(8), 751–756.
- Li, Y., Lu, Q., Wang, Y., & Ma, S. (2020). Racial differences in testicular cancer in the United States: Descriptive epidemiology. *BMC Cancer*, *20*, 1–10.
- Lins, L., & Carvalho, F. M. (2016). SF-36 total score as a single measure of health-related quality of life: Scoping review. *SAGE Open Medicine*, *4*, 2050312116671725.
- Luckett, T., Butow, P. N., King, M. T., & Olver, I. N. (2008). Psycho-social issues in long-term survivors of testicular cancer: Directions for future research. *Asia-Pacific Journal of Clinical Oncology*, *4*(3), 125–131.
- Martino Cinnera, A., Bonni, S., Pellicciari, M. C., Giorgi, F., Caltagirone, C., & Koch, G. (2020). Health-related quality of life (HRQoL) after stroke: Positive relationship between lower extremity and balance recovery. *Topics in Stroke Rehabilitation*, 1–7.
- Matheson, L., Boulton, M., Lavender, V., Protheroe, A., Brand, S., Wanat, M., & Watson, E. (2016). Dismantling the present and future threats of testicular cancer: A grounded theory of positive and negative adjustment trajectories. *Journal of Cancer Survivorship*, *10*(1), 194–205.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, *151*(4), 264–W64.
- Nassar-McMillan, S. C., & Borders, L. D. (2002). Use of focus groups in survey item development. *The Qualitative Report*, *7*(1), 1–12.
- Oldenburg, J., Kraggerud, S. M., Brydøy, M., Cvancarova, M., Lothe, R. A., & Fossa, S. D. (2007). Association between long-term neuro-toxicities in testicular cancer survivors and polymorphisms in glutathione-S-transferase-P1 and-M1, a retrospective cross sectional study. *Journal of Translational Medicine*, *5*(1), 70.
- Pollatos, O., Georgiou, E., Kobel, S., Schreiber, A., Dreyhaupt, J., & Steinacker, J. M. (2020). Trait-based emotional intelligence, body image dissatisfaction, and HRQoL in children. *Frontiers in Psychiatry*, *10*, 973.
- Rammant, E., Van Wilder, L., Van Hemelrijck, M., Pauwels, N. S., Decaestecker, K., Van Praet, C., Bultijnc, R., Ost, P., Van Vaerenbergh, T., Verhaeghe, S., & Van Heck, A. (2020). Health-related quality of life overview after different curative treatment options in muscle-invasive bladder cancer: An umbrella review. *Quality of Life Research*, *29*, 2887–2910.
- Reilly, C., Hallböök, T., Viggedal, G., Rydenhag, B., Uvebrant, P., & Olsson, I. (2020). Parent reported health related quality of life (HRQoL) and behaviour in young people with epilepsy before and two years after epilepsy surgery. *Seizure*, *74*, 1–7.
- Rovito, M. J., Manjelievskaia, J., Leone, J. E., Lutz, M., Cavayero, C. T., & Perlman, D. (2018). Recommendations for treating males: An ethical rationale for the inclusion of testicular self-examination (TSE) in a standard of care. *American Journal of Men's Health*, *12*(3), 539–545.
- Saab, M., Noureddine, S., Huijter, H. A. S., & DeJong, J. (2014). Surviving testicular cancer: The Lebanese lived experience. *Nursing Research*, *63*(3), 203–210.

- Shanmugalingam, T., Soultati, A., Chowdhury, S., Rudman, S., & Van Hemelrijck, M. (2013). Global incidence and outcome of testicular cancer. *Clinical Epidemiology*, *5*, 417.
- Sim, J., & Wright, C. C. (2005). The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Physical Therapy*, *85*(3), 257–268.
- Smith, A. B., Butow, P., Olver, I., Luckett, T., Grimison, P., Toner, G. C., Stockler, M. R., Hovey, E., Stubbs, J., Turner, S., Hruby, G., Gurney, H., Alam, M., Cox, K., & King, M. T. (2016). The prevalence, severity, and correlates of psychological distress and impaired health-related quality of life following treatment for testicular cancer: A survivorship study. *Journal of Cancer Survivorship*, *10*(2), 223–233.
- Smith, A. B., King, M., Butow, P., Luckett, T., Grimison, P., Toner, G. C., Stockler, M., Hovey, E., Stubbs, J., Hruby, G., & Gurney, H. (2013). The prevalence and correlates of supportive care needs in testicular cancer survivors: A cross-sectional study. *Psycho-Oncology*, *22*(11), 2557–2564.
- Sztankay, M., Aaronson, N. K., Arraras, J. I., Basso, U., Bumbasirevic, U., Efficace, F., Giesinger, J. M., Johnson, C. D., van Leeuwen, M., Oberguggenberger, A. S., & Sosnowski, R. (2018). International phase IV validation study of an EORTC quality of life questionnaire for testicular cancer patients: The EORTC QLQ-TC26. *BMC Cancer*, *18*(1), 1104.
- Taft, C., Karlsson, J., & Sullivan, M. (2001). Do SF-36 summary component scores accurately summarize subscale scores?. *Quality of Life Research*, *10*(5), 395–404.
- Vogt, C. A., Andereck, K. L., & Pham, K. (2020). Designing for quality of life and sustainability. *Annals of Tourism Research*, *83*, 102963.
- World Health Organization. (2014, March 11). *WHO | WHOQOL: Measuring quality of life*. <https://www.who.int/healthinfo/survey/whoqol-qualityoflife/en/>
- Zhu, Y. X., Li, T., Fan, S. R., Liu, X. P., Liang, Y. H., & Liu, P. (2016). Health-related quality of life as measured with the Short-Form 36 (SF-36) questionnaire in patients with recurrent vulvovaginal candidiasis. *Health and Quality of Life Outcomes*, *14*(1), 65.