Received: 7 December 2020 | Revised: 26 March 2021 | Accepted: 6 April 2021

DOI: 10.1002/pon.5700

WILEY

REVIEW

Quality of life, anxiety and depression patient-reported outcome measures in testicular cancer: A systematic review

Amine Nur Dincer[1] | Oliver Brunckhorst[1] | Oktay Genel[1] | Prokar Dasgupta[1] | Asif Muneer[2,3,4] | Kamran Ahmed[1,5]

[1]MRC Centre for Transplantation, Guy's Hospital Campus, King's College London, King's Health Partners, London, UK

[2]Division of Surgery and Interventional Science, University College London, London, UK

[3]Department of Urology, Institute of Andrology, University College London Hospitals NHS Foundation Trust, London, UK

[4]NIHR Biomedical Research Centre, University College London Hospital, London, UK

[5]Department of Urology, King's College Hospital, London, UK

Correspondence
Kamran Ahmed, MRC Centre for Transplantation, King's College London, 5th Floor Southwark Wing, Guys Hospital, London, SE1 9RT, UK.
Email: kamran.ahmed@kcl.ac.uk

Funding information
Malcolm Coptcoat Memorial Trust

Abstract

Objectives: Several patient-reported outcome measures (PROMs) are available for the assessment of quality of life (QoL), anxiety and depression for testicular cancer (TCa); however, these PROMs have uncertain validation of their psychometric properties for TCa-only cohorts. This systematic review aims to critically analyse and evaluate the psychometric properties of these QoL, anxiety and depression PROMs.

Methods: PubMed, EMBASE and PsycInfo were searched by two independent reviewers from inception to August 2020. Evaluative studies that assessed measurement properties of PROM(s) tools used for measuring QoL, anxiety and depression in TCa patients were included. The COnsensus-based Standards for the selection of health status Measurement Instruments (COSMIN) updated criteria for good measurement properties were used in the evaluation of PROM psychometric quality. This systematic review was registered on the PROSPERO database (CRD42020160232).

Results: Of 4,305 abstracts screened, a final eight full-text articles were included in this review. Five general and two TCa-specific PROMs were identified (depression, $n = 1$; anxiety and depression, $n = 2$; QoL, $n = 4$). All studies were incomplete in the validation of nine measurement properties and the modal methodological quality was 'indeterminate'. The European Organisation for Research and Treatment of Cancer Quality -Testicular Cancer 26 questionnaire and CAYA-T had the highest psychometric validation with three out of nine measurement properties being 'sufficient'.

Conclusion: This systematic review identifies a paucity of PROM-validation studies assessing anxiety, depression and QoL in TCa-only cohorts. We recommend further comprehensive and standardised psychometric validation studies of QoL, anxiety and depression PROMs in TCa-only study populations.

KEYWORDS
anxiety, cancer, depression, masculinity, mental health, oncology, patient reported outcome measures, psycho-oncology, quality of life, testicular neoplasms

1 | BACKGROUND

Testicular cancer (TCa) is the most common solid malignancy of men between the ages of 14 and 44[1] and represents 5% of urological cancers and 1%–2% of all male neoplasms. The global incidence of TCa is increasing, with 3–10 new cases occurring per 100,000 males/per year in European countries.[2,3] Earlier detection (e.g., by means of heightened public awareness relating to testicular self-examination), young age at presentation, high healing rate of testicular neoplasms and surgical technology have resulted in a high median five-year survival rate for TCa of ~95% and achievability of complete remission in metastatic TCa with high-dose cisplatin-based chemotherapy.[4–7]

However, despite the high survival rate, many TCa patients experience a complex set of psychological, social, and physical morbidity.[8] Chemotherapy-associated acute and long-term physical co-morbidities include nausea and vomiting, gastrointestinal mucositis, fatigue, neurotoxicity, Raynaud-like symptoms, ototoxicity (tinnitus and hearing impairment), thromboembolic events and leukaemia. Radiotherapy long-term toxicity includes radiation-induced secondary malignant neoplasms and decreased sexual satisfaction.[9] In addition, fertility concerns, body image issues and sexual dysfunction (including ejaculatory dysfunction) are intrinsic themes of TCa as the patient group affected are largely men within the fertile age range.[10] Several studies have shown that psychosocial and functional limitations of TCa compromise overall patient QoL and health-related QoL (HR-QoL). These studies have also shown higher rates of depressive symptoms (34%), depression and anxiety disorders (19% vs. 13.5%) compared to the general population.[11–14]

TCa patient-reported outcome measures (PROMs) assessing most prevalent mental health conditions in TCa (anxiety and depression), overall HR-QoL and overall quality of life (QoL) are an integral domain of clinical assessment in TCa person-centred care.[15,16] PROMs measure beyond biological functioning, morbidity, and mortality. They measure the physical, social, and emotional functions of a patient weighted by patient-perceived symptom burden and diagnosis and treatment related expectations.[17] Several validated PROMs of QoL, HR-QoL, anxiety and depression are reported in the clinical assessment of TCa. However, most of these PROMs have been developed as generic tools for a broad spectrum of disorders and domains (e.g., adherence, distress, health-behaviour, functioning) and not specifically for TCa. Importantly no systematic reviews have been identified to psychometrically validate the measurement properties (three domains of reliability, responsiveness, and validity) of these PROMs in a TCa population.

Current PROMs may therefore misrepresent the totality of psychosocial, depressive and anxiety symptoms that compose mental wellbeing and QoL of TCa patients. Clinical assessment performed with PROMs of unknown measurement property quality may further result in missed diagnoses and a greater cumulative psychosocial impact of co-morbidity on QoL in TCa patients and survivors. Therefore, establishing the psychometric properties of these PROMs utilised in TCa is of significant importance in screening, accurate assessment and management of mental health, specifically anxiety and depression, and global psychosocial well-being evaluated as QoL.

This systematic review therefore aims to: (i) identify QoL, anxiety and depression PROMs that are psychometrically validated (reliability, validity, and responsiveness) in patients with TCa, (ii) systematically summarise and provide a comprehensive overview of psychometric validation studies and measurement properties of these PROM(s), and (iii) provide recommendations of the most appropriate validated generic and TCa-specific PROM(s).

2 | METHODS

2.1 | Protocol and registration

This systematic review was performed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols (PRISMA-P) guidelines[18] and the COSMIN methodology for systematic reviews of PROM(s). The review protocol was registered prospectively on PROSPERO, Registration Number: CRD42020160232.

2.2 | Study eligibility criteria

Full-text articles that evaluated one or more measurement properties, development of a PROM and/or interpretability of all disease-specific and generic PROMs assessing one of the following principal constructs: QoL, HR-QoL, anxiety and depression, in a study sample of men with TCa (all types, grades and treatment modalities) were included in this review. Measurement properties (quality of PROMs) were defined as per the COnsensus-based Standards for the selection of health status Measurement INstruments) (COSMIN) Taxonomy (Table S1).

The exclusion criteria included: low quality of evidence study designs (case reports, editorials and abstract-only articles); studies with clinician-assessed outcome measurement tools; studies using PROM(s) of interest as an outcome measurement instrument only; validation studies of PROM(s) measuring physical symptoms of well-being that is, fertility, erectile dysfunction for selectivity; full-text article not available in the English language.

2.3 | Information sources and search

A comprehensive systematic search was performed utilising three online databases: PubMed, EMBASE and PsycInfo were searched from date of inception to 31/08/2020 using Boolean operators to

combine three constructs (QoL, anxiety and depression), TCa and PROM instrument. The search terms used were (quality of life OR QoL OR health-related quality of life OR HRQoL OR depress* OR psychological OR anxiety OR symptom*) AND ([{testic* OR testis OR testes} AND {neoplasm OR cancer OR carcinoma OR malignancy}] OR [seminoma OR nonseminomatous germ cell tumour OR non-seminomatous germ cell tumour OR nsgct OR leydig cell OR sertoli cell]) AND (prom OR patient reported OR self-report OR measure OR questionnaire* OR survey* OR psychometric*). No language restrictions were used in the search strategy in order to report existence of all PROM-validation studies.

Reference screening of full-text articles included was performed. A search of grey literature was also conducted through ongoing unpublished research on ClinicalTrials.gov, US National Library of Medicine, to identify any additional relevant studies, with authors of any relevant studies contacted for any preliminary data available.

2.4 | Study selection

Titles and abstracts of all articles were assessed for inclusion or exclusion against pre-defined eligibility criteria in duplicate by two independent reviewers (Amine Nur Dincer and Oktay Genel) and disagreements were resolved by discussion between the two reviewers. Included abstracts were retrieved for full-text review and duplicates were identified and deducted. Full-text publications that satisfied the pre-defined study eligibility and exclusion criteria were included. A referencing software (Mendeley 2020, London) was used to manage all citations.

2.5 | Data collection and data items

Data were extracted on the characteristics of PROM(s), characteristics of the included studies, results on measurement properties (psychometric validation), interpretability and feasibility of PROM (s) scores and recorded onto a pre-defined extraction sheet by tworeview authors (Amine Nur Dincer and Oktay Genel).

Data extraction reflected parameters required for the evaluation of measurement property per PROM. Measurement properties were assessed in the following order: structural validity, internal consistency, cross-cultural validity/measurement invariance, reliability, measurement error, criterion validity, hypotheses testing for construct validity and responsiveness.

PROM quality was assessed by rating each measurement property against the updated criteria for good measurement properties as 'sufficient' (+), 'insufficient' (–), or 'indeterminate' (?) as per COSMIN updated criteria for good measurement properties (Table S2).[19]

Consistent results of each measurement property quality per PROM were then summarised to produce a total rating per measurement quality per PROM, also recorded as 'sufficient' (+), 'insufficient' (–), or 'indeterminate' (?) and presented in a Summary of Findings (SoF) table (Table 2).

Data extracted on study characteristics included: sample size, participant demographics (age, tumour histological type), country the PROM instrument was administered in, modes of administration (e.g., computer-based vs. paper), response rate (Table 1). Data extracted on PROM(s) included: PROM(s) name; domains/constructs assessed in the PROM(s); item types/response options (such as static or rating scales), scoring system and ranges (Table S3); measurement properties assessed per PROM (Table 2).

2.6 | Risk of bias assessment

A risk of bias assessment was performed assessing the methodological quality of included studies using the COSMIN Risk of Bias checklist.[20] The COSMIN Risk of Bias checklist was used to rate the overall quality of individual studies as 'very good', 'adequate', 'doubtful' or 'inadequate'. Heterogeneous terms and definitions for measurement properties in included articles were objectively re-defined to a measurement property defined by the COSMIN taxonomy.[19]

The Grading of Recommendations Assessment, Development and Evaluation (GRADE) principles were used to rate the overall quality of evidence of the review findings (Table 2; Table S4).

3 | RESULTS

3.1 | Study selection

A total of 4,305 studies were identified and a further 10 articles were identified through reference screening of full-text publications. Three hundred and four abstracts satisfied the eligibility criteria or were deemed indeterminate and full-text articles were retrieved and assessed for inclusion. After-deduplication, eight full-text articles were included in this systematic review (Figure 1).

3.2 | Study characteristics and result synthesis

Seven PROM instruments were identified in this review; two of seven identified PROMs have been developed for TCa patients (European Organisation for Research and Treatment of Cancer Quality - Testicular Cancer 26 [EORTC QLQ-TC26], CAYA-T; Table 1). Results were synthesised into construct(s) measured and sub-categorised into each measurement property per PROM. Comparative analysis and total PROM score were generated for each measurement property (Table 2). Consistent results identified were qualitatively summarised to determine overall grade of the measurement properties per PROM. Meta-analyses were not performed because of insufficient quantitative or qualitative data available on measurement properties per PROM(s).

TABLE 1 Overview of the characteristics of included studies

PROM(Ref)	Construct(s) measured	*N*/sample size	Mean age (SD), years	TCS or TCP	Tumour histological type, %	Disease Duration (SD), months	Recall period, weeks	Questionnaire completion rate, %	Country	COSMIN RoB score
EORTC QLQ-TC26 (Sztanky)	QoL	313	38.6 (9.5)	TCS and TCP	*S* = 62.5 NS = 37.5	5.4–6.7 (3.4–17.7)	2, 4	91.2–97.6	Austria, Italy, The Netherlands, Poland, Serbia, Spain, UK	*V*
CES-D (Wang and Hoyt)	Depressive symptoms	171	25.2 (3.32)	TCS	N/A	Up to 36 (for 69%)	N/A	N/A	USA	*I*
CES-D (Tuinman)	Depressive symptoms	93	29.4 (7.5)	TCP	NS = 100	N/A	12, 48	90	The Netherlands	*I*
FACT-G (Hoyt)	QoL	171	25.2 (3.32)	TCS	N/A	30.1 (14.4)	4	N/A	USA	*I*
HADS (Skaali)	Anxiety and depressive symptoms	135	34.8 (8.9)	TCP	*S* = 53 NS = 47	1.2	N/A	N/A	Norway	*I*
HADS (Fossa)	Anxiety and depressive symptoms	820	44	TCS	*S* = 49 NS = 51	N/A	N/A	81	Norway	*I*
DASS21 (Smith)	Anxiety and depressive symptoms	244	38.3 (10.3)	TCS	N/A	N/A	N/A	N/A	Australia	*I*
SF-36v2 (Smith)	HR-QoL	244	38.3 (10.3)	TCS	N/A	N/A	N/A	N/A	Australia	*I*
CAYA-T (Hoyt)	HR-QoL	171	25.2 (3.32)	TCS	N/A	32.4 (19.3)	4	N/A	USA	*G*

Abbreviations: CAYA-T, cancer assessment for young adults-testicular; CES-D, centre for epidemiologic studies depression scale; DASS-21, depression anxiety stress scales-21; EORTC QLQ-TC26, European organisation for research and treatment of cancer quality-testicular cancer 26; FACT-G, functional assessment of cancer therapy general; G, good; HADS-A, hospital anxiety and depression score-anxiety; HADS-D, hospital anxiety and depression score-depression; I, inadequate; NS, non-seminoma; PROM, patient-reported outcome measure; RoB, risk of bias; S, seminoma; SD, standard deviation; SF-36V2, short form-36 version 2; TCP, testicular cancer patient; TCS, testicular cancer survivor; V, very good.

TABLE 2 Summary of Findings (SoF) table for results of studies on measurement properties of identified PROM(s)

	PROM Instrument (*n* of validation studies)	CAYA-T (1)	EORTC-QLQ TC26 (1)	CES-D (2)	FACT-G (1)	HADS (2)	DASS21 (1)	SF-36 V2 (1)
Updated COSMIN criteria for good measurement properties summary (overall rating)	Structural validity	?	+	?	?	?	?	?
	Internal consistency	+	–	?	?	?	?	?
	Reliability	+	–	?	?	?	?	?
	Measurement error	?	?	?	?	?	?	?
	Hypotheses testing for construct validity	+	+	?	?	?	?	?
	Cross-cultural validity	?	?	?	?	?	?	?
	Criterion validity	?	?	?	?	?	?	?
	Responsiveness	+	+	?	?	?	?	?
Total PROM score		+/?	+/?	?	?	?	?	?
GRADE quality of Evidence		High	High	Low	Low	Low	Low	Low

Abbreviations: CAYA-T, cancer assessment for young adults-testicular; CES-D, centre for epidemiologic studies depression scale; COSMIN, consensus-based standards for the selection of health measurement instruments; DASS-21, depression anxiety stress scales-21; EORTC, QLQ-TC26; European, organisation for research and treatment of cancer quality-testicular cancer 26; FACT-G, functional assessment of cancer therapy general; GRADE, grading of recommendations assessment development and evaluation; HADS-A, hospital anxiety and depression score-anxiety; HADS-D, hospital anxiety and depression score-depression; PROM, patient-reported outcome measure; SF-36V2, short form-36 version 2.

3.3 | Quality of life and health-related quality of life

3.3.1 | European Organisation for Research and Treatment of Cancer Quality -Testicular Cancer 26-item

The EORTC QLQ-TC26 is a supplementary module to the European Organisation for Research and Treatment of Cancer Quality- Core 30 (a gender-neutral questionnaire assessing biopsychosocial and physical health). The EORTC QLQ-TC26 assesses TCa-specific symptoms modulating HR-QoL: treatment side effects (hair loss, olfactory/gustatory/auditory disturbances, gastrointestinal symptoms (stomach pain, acid reflux), peripheral neuropathy, skin problems), patient satisfaction, future perspective (uncertainty, anxiety), job/education problems, physical limitations, family disruption, fertility concerns, masculinity, sexual functioning, satisfaction with testicular implant.

A single validation study of EORTC QLQ-TC26 performed by Sztankay et al.[21] was identified. Pooled data from Sztankay et al.[21] of a total of 313 patients (mean age 38.6, SD 9.5), 200 testicular cancer survivors (TCSs) and 113 TCa patients, exhibited completion rates of 91.2% for sexuality items (percentage of missing responses 8.8%–1.1%) and 97.6% for all other items. Completion time required for the questionnaire was 8.1 min (SD 4.2) and 6% of patients required assistance to complete the questionnaire. A single international phase four validation study for EORTC-QLQ TC26 questionnaire was identified.[21]

Sztankay et al.[21] assessed the structural validity measurement property of the validity domain. Confirmatory factor analysis (CFA) was conducted and the Comparative Fit Index (CFI) = 0.974, Tucker-Lewis Index (TLI) = 0.963 and the Root Mean Square Error of Approximation (RMSEA) = 0.046 indicated 'sufficient' structural validity. Internal consistency was 'sufficient' (Cronbach alpha = 0.79–0.90) for 10 scales, but 'insufficient' for Communication (alpha = 0.67) and Sexual Functioning (alpha = 0.62). A Cronbach alpha less than 0.70 for each subscale/unidimensional score were graded 'insufficient'. The reliability domain was graded using internal consistency and reliability measurement properties. Test-retest reliability was 'sufficient' in 8 of 12 scales (intraclass correlation [ICC]: 0.71–0.91). Four subscales: physical limitations (0.67), family problems (0.65), treatment satisfaction (0.48) and testicular implant satisfaction (0.69) were classified 'insufficient'. An ICC below 0.70 was rated as 'insufficient'. Responsiveness to change over time was also 'sufficient'. Measurement error and criterion validity were not assessed and graded 'indeterminate'.

3.3.2 | Functional Assessment of Cancer Therapy General

The Functional Assessment of Cancer Therapy General (FACT-G) questionnaire is used to assess the health-related QoL of cancer patients. The 27-item questionnaire uses a five-point response scale

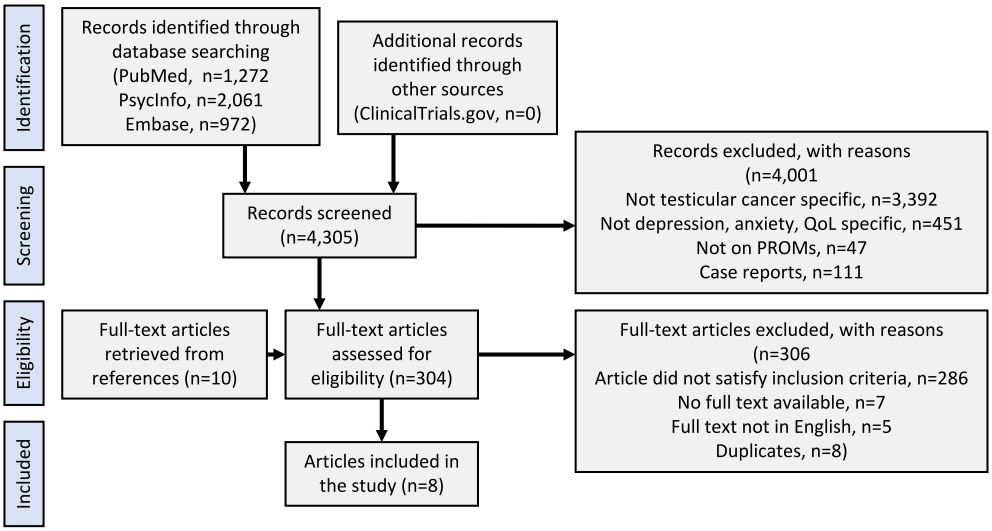

FIGURE 1 Study screening and selection; adapted from Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)

to subjectively measure patient symptomatic burden and physical well-being over the past 7 days, (e.g., 'I am forced to spend time in bed'; 'I feel close to my friends'). The FACT-G questionnaire assesses four domains: physical (seven items reporting lethargy, nausea, role limitations, pain, treatment-related side effects, perception of illness); social/family (eight items including sexual functioning) emotional well-being (six items) and functional well-being (seven items). One study validating the FACT-G questionnaire was identified.[22] The internal consistency of the seven-item Physical Well-Being domain had a Cronbach's alpha of 0.86, based on the study by Hoyt et al.[22] however, structural validity was not assessed, and therefore internal consistency was graded 'indeterminate'. No other measurement property was assessed and were rated as 'indeterminate'. Internal consistency was reported as a single score and Cronbach's alpha was not calculated for each unidimensional scale. The study included TCa survivors exclusively of adequate sample size (n = 171) and the age bracket of the study population was a true representative of a TCa cohort (mean age 25.2, SD = 3.32).

3.3.3 | Short Form 36-item version 2

The Short Form-36 version 2 (SF-36V2) measures eight multi-scale domains: physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional and mental health (symptoms of depression and anxiety, nervousness, fatigue, pain). A single study assessing psychometric properties of SF-36 V2 was identified.[23] Smith et al. [23] reports Cronbach's alpha as a range above >0.70, alpha = 0.82–0.93, however structural validity was not reported. The internal consistency of the SF-36 V2 questionnaire was 'indeterminate'. No other measurement property was assessed and thus these measurement properties were rated as 'indeterminate'. The study population included only TCa survivors, previously managed with surgical intervention and with no evidence of TCa recurrence. This study population may not be representative of TCa patients who have been treated with different treatment modalities, including active surveillance or adjuvant chemotherapy or radiotherapy.

3.3.4 | Cancer assessment for young adults-testicular

The Cancer Assessment for Young Adults-Testicular (CAYA-T) is a 90-item questionnaire that includes the seven biopsychosocial domains assessing concentration and memory, self-care, education and work, social functioning, sexual relationships, emotional functioning (fear, preoccupation with illness, anxiety and depressive symptoms). One study, with the study population including TCa survivors, psychometrically validating the CAYA-T questionnaire was identified.[24] Hoyt et al.[24] used the Rasch model (person separation index 0.34–0.82) and overall Cronbach's alpha for internal consistency which was 'sufficient', alpha = 0.85 (0.70–0.91). Reliability was assessed using intraclass correlation coefficient (ICC). ICC was >0.80 and the reliability of the CAYA-T questionnaire was 'sufficient' (ICC more than or equal to 0.70). No other measurement property was assessed. Internal consistency (Cronbach's alpha) and reliability (ICC) was also 'sufficient' for each unidimensional scale except for the Financial Maintenance (ICC = 0.62), Recreational Pursuit (ICC = 0.66), Spiritual Stability (ICC = 0.58) subscales. The CAYA-T questionnaire may have increased specificity as it utilises a three-point Likert scoring scale. Subscale response specificity is often reduced by value cross-over in five-/six-point Likert scales in assessing symptom duration and severity.

The lengthiness of the CAYA-T questionnaire may adversely affect response rates, resulting in decreased diagnostic accuracy and reliability.

3.4 | Depressive symptoms

3.4.1 | Centre for Epidemiologic Studies Depression Scale

The Centre for Epidemiologic Studies Depression Scale (CES-D) is a 20-item self-report questionnaire on depressive symptoms. Two studies[25,26] were identified however all only assessed the internal consistency of the CES-D questionnaire. Tuinman et al.[26] included TCa patients and the study by Wang and Hoyt[25] reported on TCa survivors. All studies used the Cronbach's alpha statistical analysis criteria. Tuinman et al.[26] (alpha = 0.88–0.91), Wang and Hoyt[25] (a = 0.94) measured internal consistency Cronbach's alpha above 0.70. Hoyt et al.[25] further reported 'good reliability' (a = 0.94), but intraclass correlation (ICC) and/or weighted Kappa was not recorded for reliability. Structural validity was graded 'indeterminate', and the measurement property internal consistency was 'indeterminate'. No hypotheses were defined by the review team and the responsiveness domain was 'indeterminate'. No other measurement property was assessed.

3.5 | Anxiety and depressive symptoms

3.5.1 | Hospital Anxiety and Depression Scale

Two studies were identified that validate measurement properties of the Hospital Anxiety and Depression Scale (HADS) measures 14 items divided into HADS-Anxiety (HADS-A) and HADS-depression (HADS-D) with seven items per domain. Two studies validating HADS were identified.[27,28] No criteria were reported for structural validity in either study. Internal consistency for HADS-A was measured using Cronbach's alpha in both studies: a = 0.78 and a = 0.82 respectively. Internal consistency for HADS-D was also above 0.70 (a = 0.85 and a = 0.86). No other measurement property was assessed in either study. Hypothesis was defined by the review team and the responsiveness domain scored 'indeterminate'. Both studies were evaluated in Norway and one of the studies by Fossa et al.[27] exclusively validated the HADS questionnaire in TCa survivors. Skaali et al.[28] validated the PROM in TCa patients.

3.5.2 | Depression Anxiety Stress Scales 21-item

The Depression Anxiety Stress Scales 21 is a short-form version of the 42-item Depression Anxiety Stress Scales (DASS) questionnaire measuring psychological distress in three multi-item domains: anxiety, depression and stress/tension, experience over the past 1 week. The depression domain assessed self-reported hopelessness, self-depreciation, interest, anhedonia, inertia and dysphoria. The anxiety domain evaluated experience of anxiety and relevant physical effects. The stress/tension domain aimed to measure agitation, irritability, impatience, difficulty relaxing (e.g., 'I found myself getting agitated', 'I found it difficult to relax'). A single validation study on TCa survivors was identified for the DASS-21 questionnaire. Smith et al.[23] measures internal consistency using Cronbach's alpha per domain: depression items (a = 0.93), anxiety items (a = 0.83) and stress/tension items (a = 0.91). Information or model fit on structural validity was not reported and was rated as 'indeterminate'. Internal consistency was also 'indeterminate' as criteria for evidence of 'sufficient' structural validity was not met.

3.6 | Risk of bias

The COSMIN Risk of Bias checklist was used modularly to rate the overall quality of measurement properties in individual studies as 'very good', 'adequate', 'doubtful' or 'inadequate'. Internal consistency was 'very good' for all included studies as an internal consistency statistic was calculated for each unidimensional scale with preferred statistical standards (reported using Cronbach's alpha) in all identified PROM-validation studies. Overall quality was calculated as 'very good' for the EORTC QLQ-TC26 validation study and 'good' for the CAYA-T validation study, with all other PROM validation studies scoring 'inadequate' in the assessment of risk of bias.

4 | DISCUSSION

A standardised assessment of QoL/HR-QoL, and screening of anxiety and depression in patients diagnosed with TCa is important for patient-centred delivery of healthcare.[10,21] There is a clear paucity of research in relation to validation and assessment of psychometric properties of PROMs measuring these prevalent constructs in a TCa-only cohort and limited studies have psychometrically validated the measurement properties of these PROMs in a TCa population. To the authors' knowledge this is the first systematic review to identify and comprehensively evaluate the methodological quality and measurement properties of QoL/HR-QoL, anxiety and depression PROM-validation studies in TCa cohorts, using the updated COSMIN criteria for good measurement properties, and in accordance with COSMIN guidelines.

We identified eight general and two TCa-specific PROM questionnaires and supplementary modules. Remarkably, the overall methodological quality of studies reporting psychometric properties were on average rated 'inadequate' and the modal value for measurement property rating was 'indeterminate', as a result of absent or incomplete validation and reporting. Our findings demonstrated that the CAYA-T and QRQTC-36 questionnaires are the most psychometrically validated PROM questionnaires currently available for QoL assessment in TCa patients and/or survivors. Internal

consistency (with majority of studies applying Cronbach's alpha, based on a reflective model) was the most analysed measurement property in the included studies with all studies reporting an average internal consistency of alpha >0.70. Reliability (reported with intraclass correlation coefficient), responsiveness (i.e., the result being in accordance with the hypothesis) and structural validity were only partially (in less than 50% of studies) reported. Structural validity is the most important measurement property and all included studies, except for the phase four validation study for the EORTC QLQ-TC26 questionnaire, failed to report on this using either classical test theory or item response theory. The EORTC QLQ-TC26 questionnaire is appropriate for the assessment of HR-QoL in both TCa patients and early post-treatment survivors (defined as up to 1 year following treatment completion). The remaining PROMs identified shared insufficient reporting and 'indeterminate' psychometric properties. In comparison, majority of these PROMs, including the HADS, SF-36, FACT-G and CES-D tools, have been validated more extensively in other cancers such as breast and prostate cancer,[29–34] further highlighting the paucity of psychometric validation studies of PROMs assessing anxiety, depression and QoL in TCa research.

Importantly, only two of the seven identified PROMs were developed specifically for TCa patients. TCa patients and survivors experience several treatment-related side effects (e.g., tinnitus and chronic fatigue), social, emotional and physical morbidity that are generic for oncological patients and it may be plausible to assess social/emotional and physical well-being using validated generic PROMs. However, there is a spectrum of other symptoms and themes specific to TCa including testicular prosthesis satisfaction, body image issues, fertility concerns and loss of masculinity that are not addressed and omitted from assessment using generic PROMs. The EORTC-QLQ TC26 and CAYA-T questionnaires have been developed for TCa incorporating several subscales on sexual functioning. Particularly, the lowest completion rates were reported for the sexuality items in EORTC-QLQ TC26 suggesting that sexual health and TCa-specific symptoms are often not assessed and/or reported adequately. Reasons for omitting the sexuality items may include patient perceived inappropriateness or irrelevance of the subscale or associated feelings of self-consciousness and distress in answering these subscales. Importantly, the reasons for the incomplete reporting of sexuality items have not been addressed in either study. Given that fertility and sexual functioning is an important consideration for majority of TCa patients and survivors, additional or updated subscales on sexual functioning should be developed through focus groups and thematic analysis, and Delphi studies with clinical professionals and service users, for a comprehensive evaluation of QoL.

PROM(s) utilised in TCa patients have very limited validation of their psychometric properties. The clinical or research-based use of PROM(s) of unknown (graded 'indeterminate) or poor (graded 'insufficient') measurement properties, reliability and feasibility is unreliable. Fractional validation of identified PROM tools' measurement properties results in a poor overall assessment of PROM psychometric quality. Additionally, the partial assessment of PROM psychometric quality may be inadequate and insufficient, limiting recommendations on the most suitable PROM tool for QoL/HR-QoL, anxiety and depression in TCa. We therefore recommend future high-quality validation studies analysing all nine psychometric properties with preferred statistical methods of QoL/HR-QoL, anxiety and depression PROMs in a TCa-only cohort.

4.1 | Study limitations

Limitations of this review include the small number of eligible and included original PROM-validation studies (n = 10) assessing psychometric properties of QoL, anxiety and depression tools in TCa. Studies identified utilising PROMs in an evaluative manner, that is, to assess intervention effects, were not included. Studies graded 'inadequate' were not excluded from data analysis, per COSMIN guidelines, hence the overall reliability and accuracy of results may have been downgraded. Furthermore, there was substantial heterogeneity in reporting of the results between the studies and thus quantitative pooled summary or meta-analyses was not performed. Finally, despite a comprehensive search strategy performed by two independent authors, limitations may be present in the data set and relevant studies within the literature may have been missed in view of the variety of scientific literature sources available.

4.2 | Clinical implications

PROMs offer immense potential for improving the quality of care and clinical outcomes and promoting patient choice in TCa. TCa patients experience a broad and unique spectrum of physical and psychosocial symptoms that affect both functional and clinical outcomes. This reflects the importance of a comprehensive evaluation of health-related QoL, anxiety, and depressive symptoms in this patient population. This review supports utilisation of the EORTC-QLQ TC26 and CAYA-T questionnaire in clinical practice as the best currently available tools. Psychometric parameters and values reported in this review can be useful in choosing the most appropriate PROM questionnaire in both clinical and research settings.

In consultation, psychometrically validated PROM questionnaires assessing QoL and mental health aids holistic conversations exploring patient-reported biopsychosocial and physical symptoms. A diagnosis of TCa has a significant impact on mental health and in like manner, identification of a mental health disorder affects the treatment plan. TCa patients have a higher suicide mortality rate, depression and anxiety prevalence when compared to the general population estimates.[11–14] Currently, there is no recommended universal PROM questionnaire for screening of these constructs in TCa, only guidance that selection of the most appropriate PROM depends on the clinical setting (i.e., treatment modality, patient characteristics).[35] The use of psychometrically validated PROMs for screening these prevalent mental health disorders and QoL limitations results in better detection of mental health and/or psychological needs,

earlier intervention by referral to psychological or psychiatric support or other appropriate services and subsequent better outcomes overall for TCa patients.[36,37] Furthermore, accurate estimation of depression and anxiety prevalence is essential and patient-reported questionnaires report a higher prevalence rate in comparison to clinician-reported questionnaires.[38] However, patient-reported questionnaires are limited in their diagnostic accuracy compared to a standardised diagnostic interview and are therefore best utilised as screening tools.[39] This review also highlights the paucity of primary PROM-validation studies and therefore recommends further validation studies of PROMs assessing anxiety, depression and QoL in TCa patients and survivors that need to be conducted in primary or tertiary clinical settings.

5 | CONCLUSIONS

The use of PROMs enables better healthcare delivery via synergistic integration of physical (assessed via oncological and clinician-reported functional outcomes) and psychological health (measured using validated QoL, anxiety and depression PROMs) in TCa. Identification of biopsychosocial and physical symptoms and limitations through PROMs can guide service and healthcare delivery for TCa patients. We demonstrate that the EORTC QLQ-TC26 questionnaires are at present the most psychometrically validated QoL PROMs for use in TCS and TCa patients and recommend utilisation of these QoL PROMs in the standard care pathway and clinical trials. Utilisation of the CAYA-T questionnaire is recommended in TCSs. Nonetheless, we identify a consistent lack of PROM-validation studies for QoL, anxiety and depression in TCa specific cohorts. Thus, future psychometric validation studies are recommended to identify the most appropriate PROM tool for QoL, anxiety and depression in TCa.

ACKNOWLEDGEMENTS

Amine Nur Dincer, Prokar Dasgupta and Kamran Ahmed acknowledge research support from the Malcolm Coptcoat Trust. Oliver Brunckhorst, Prokar Dasgupta and Kamran Ahmed acknowledge research support from the Medical Research Council (MRC) Centre for Transplantation and funding from the King's Medical Research Trust (KMRT). Prokar Dasgupta further acknowledges funding from the King's College London-Vattikuti Institute of Robotic Surgery and the Guy's and St. Thomas' (GSTT) Charity. Asif Muneer is supported by the NIHR Biomedical Research Centre University College London Hospital. Kamran Ahmed further acknowledges funding from the Royal College of Surgeons of England, The Urology Foundation, and the Pelican Foundation. This work was funded by the generous contributions of the Malcolm Coptcoat Memorial Trust.

CONFLICT OF INTERESTS

The authors of this systematic review have no conflicts of interest to disclose.

ETHICS STATEMENT

Ethics approval was not required for this systematic review.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Oliver Brunckhorst https://orcid.org/0000-0002-2073-922X
Kamran Ahmed https://orcid.org/0000-0002-5135-9211

REFERENCES

1. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Canc*. 2015;136(5):86:E359.
2. Park JS, Kim J, Elghiaty A, Ham WS. Recent global trends in testicular cancer incidence and mortality. *Med Baltim*. 2018;97(37):e12390.
3. Bergström R, Adami H-O, Möhner M, et al. Increase in testicular cancer incidence in six European countries: a birth cohort phenomenon. *J Natl Cancer Inst*. 1996;88(11):727-733.
4. Cheng L, Albers P, Berney DM, et al. Testicular cancer. *Nat Rev Dis Primers*. 2018;4(1):29.
5. Ozturk C, Fleer J, Hoekstra HJ, Hoekstra-Weebers JE. Delay in diagnosis of testicular cancer; a need for awareness programs. *PLoS One*. 2015;10(11):e0141244.
6. Cheng L, Albers P, Berney DM, et al. Testicular cancer. *Nat Rev Dis Primers*. 2018;4(1):29. https://doi.org/10.1038/s41572-018-0029-0. PMID:30291251
7. EAU Guidelines. *Edn. presented at the EAU Annual Congress Amsterdam*. 2020. ISBN 978-94-92671-07-3.
8. Schepisi G, De Padova S, De Lisi D, et al. Psychosocial issues in long-term survivors of testicular cancer. *Front Endocrinol*. 2019;10:113.
9. Rutskij R, Gaarden T, Bremnes R, et al. A study of coping in long-term testicular cancer survivors. *Psychol Health & Med*. 2010;15(2):146-158.
10. Ljungman L, Eriksson LE, Flynn KE, et al. Sexual dysfunction and reproductive concerns in young men diagnosed with testicular cancer: an observational study. *J Sex Med*. 2019;16(7):1049-1059.
11. Dahl AA, Østby-Deglum M, Oldenburg J, et al. Aspects of posttraumatic stress disorder in long-term testicular cancer survivors: cross-sectional and longitudinal findings. *J Cancer Surviv*. 2016;10(5):842-849.
12. Orre IJ, Fosså SD, Murison R, et al. Chronic cancer-related fatigue in long-term survivors of testicular cancer. *J Psychosomatic Res*. 2008;64(4):363-371.
13. Kreiberg M, Bandak M, Lauritsen J, et al. Cohort profile: the Danish testicular cancer late treatment effects cohort (DaTeCa-LATE). *Front Oncol*. 2018;8:37.
14. Kenney LB, Antal Z, Ginsberg JP, et al. Improving male reproductive health after childhood, adolescent, and young adult cancer: progress and future directions for survivorship research. *J Clin Oncol*. 2018;36(21):2160-2168.
15. Santana MJ, Manalili K, Jolley RJ, Zelinsky S, Quan H, Lu M. How to practice person-centred care: a conceptual framework. *Health Expect* 2018;21(2):429-440.
16. Montague T, Gogovor A, Aylen J, et al. Patient-centred care in Canada: key components and the path forward. *Healthc Q*. 2017;20(1):50-56.
17. Yang LY, Manhas DS, Howard AF, Olson RA. Patient-reported outcome use in oncology: a systematic review of the impact on

patient-clinician communication. *Support Care Canc*. 2018;26(1): 41-60.
18. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev*. 2015;4(1):1.
19. Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res*. 2018;27(5):1171-1179.
20. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res*. 2018;27(5):1147-1157.
21. Sztankay M, Aaronson NK, Arraras JI, et al. International phase IV validation study of an EORTC quality of life questionnaire for testicular cancer patients: the EORTC QLQ-TC26. *BMC Canc*. 2018; 18(1):1104.
22. Hoyt MA, Nelson CJ, Darabos K, Marín-Chollom A, Stanton AL. Mechanisms of navigating goals after testicular cancer: meaning and emotion regulation. *Psycho-Oncol*. 2017;26(6):747-754.
23. Smith AB, King M, Butow P, et al. The prevalence and correlates of supportive care needs in testicular cancer survivors: a cross-sectional study. *Psycho-Oncol*. 2013;22(11):2557-2564.
24. Hoyt MA, Cano SJ, Saigal CS, Stanton AL. Health-related quality of life in young men with testicular cancer: validation of the Cancer Assessment for Young Adults (CAYA). *J Cancer Surviv*. 2013;7(4): 630-640.
25. Wang AW, Hoyt MA. Cancer-related masculinity threat in young adults with testicular cancer: the moderating role of benefit finding. *Anxiety Stress Coping*. 2020;33(2):207-215. https://doi.org/10.1080/10615806.2020.1713447
26. Tuinman MA, Hoekstra HJ, Vidrine DJ, et al. Sexual function, depressive symptoms and marital status innonseminomatesticular cancer patients: a longitudinal study. *Psycho-Oncol*. 2010;19(3): 238-247.
27. Skaali T, Fosså SD, Andersson S, Langberg CW, Lehne G, Dahl AA. Is psychological distress in men recently diagnosed with testicular cancer associated with their neuropsychological test performance? *Psycho-Oncol*. 2011;20(4):369-377.
28. Fosså SD, Dahl AA, Loge JH. Fatigue, anxiety, and depression in long-term survivors of testicular cancer. *J Clin Oncol*. 2003;21(7): 1249-1254.
29. Okamoto T, Shimozuma K, Katsumata N, et al. Measuring quality of life in patients with breast cancer: a systematic review of reliable and valid instruments available in Japan. *Breast Cancer*. 2003;10(3):204-213. https://doi.org/10.1007/BF02966719.PMID: 12955032
30. Treanor C, Donnelly M. A methodological review of the Short Form Health Survey 36 (SF-36) and its derivatives among breast cancer survivors. *Qual Life Res*. 2015;24(2):339-362.Epub 2014 Aug 20. PMID: 25139502. https://doi.org/10.1007/s11136-014-0785-6
31. Harrington S, Lee J, Colon G, Alappattu M. Oncology section EDGE task force on prostate cancer. *Rehabil Oncol*. 2016;34(1):27-35. PMID: 27134804; PMCID: PMC4851243.
32. Knight SJ, Chmiel JS, Sharp LK, et al. Spouse ratings of quality of life in patients with metastatic prostate cancer of lower socioeconomic status: an assessment of feasibility, reliability, and validity. *Urology*. 2001;57(2):275-280.PMID: 11182336. https://doi.org/10.1016/s0090-4295(00)00934-1
33. Hamoen EHJ, De Rooij M, Witjes JA, Barentsz JO, Rovers MM. Measuring health-related quality of life in men with prostate cancer: a systematic review of the most used questionnaires and their validity. *Urologic Oncol Seminars Orig Investigations*. 2015;33(2):e19-69. Epub 2014 Jan 13. PMID: 24433753. https://doi.org/10.1016/j.urolonc.2013.10.005
34. Rnic K, Linden W, Tudor I, Pullmer R, Vodermaier A. Measuring symptoms in localized prostate cancer: a systematic review of assessment instruments. *Prostate Cancer Prostatic Dis*. 2013;16(2): 111-122.Epub 2013 Feb 5. PMID: 23381695. https://doi.org/10.1038/pcan.2013.1
35. Niedzwiedz CL, Knifton L, Robb KA, Katikireddi SV, Smith DJ. Depression and anxiety among people living with and beyond cancer: a growing clinical and research priority. *BMC Canc*. 2019;19:943.
36. Recklitis CJ, Syrjala KL. Provision of integrated psychosocial services for cancer survivors post-treatment. *Lancet Oncol*. 2017;18: e39-e50.
37. Galway K, Black A, Cantwell M, Cardwell CR, Mills M, Donnelly M. Psychosocial interventions to improve quality of life and emotional wellbeing for recently diagnosed cancer patients. *Cochrane Database Syst Rev*. 2012;11:CD007064.
38. Levis B, Yan XW, He C, Sun Y, Benedetti A, Thombs BD. Comparison of depression prevalence estimates in meta-analyses based on screening tools and rating scales versus diagnostic interviews: a meta-research review. *BMC Med*. 2019;17:65.
39. Hartung TJ, Friedrich M, Johansen C, et al. The hospital anxiety and depression scale (HADS) and the 9-item patient health questionnaire (PHQ-9) as screening instruments for depression in patients with cancer. *Cancer*. 2017;123(21):4236-4243. Epub 2017 Jun 27. PMID: 28654189. https://doi.org/10.1002/cncr.30846

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Dincer AN, Brunckhorst O, Genel O, Dasgupta P, Muneer A, Ahmed K. Quality of life, anxiety and depression patient-reported outcome measures in testicular cancer: A systematic review. *Psycho-Oncology*. 2021;1–10. https://doi.org/10.1002/pon.5700