CrossMark

ORIGINAL ARTICLE

Cancer-related fatigue: appraising evidence-based guidelines for screening, assessment and management

Elizabeth J. M. Pearson[1] · Meg E. Morris[2] · Carol E. McKinstry[1]

Abstract

Purpose There is inconsistent management of cancer-related fatigue (CRF) by health professionals worldwide. This research aims to identify the most appropriate guidelines for the management of cancer-related fatigue.

Methods A systematic search of international literature identified evidence-based clinical practice guidelines for CRF. Four reviewers independently appraised the highest quality guidelines using the AGREE-II instrument and National Heath and Medical Research Council (NHMRC) guideline standards.

Results Five guidelines met the inclusion criteria. Of these, the 2015 Canadian Association of Psychosocial Oncology (CAPO) CRF guidelines and the 2014 American Society of Clinical Oncology (ASCO) fatigue guidelines for cancer survivors were selected for in-depth appraisal. The CAPO guideline scored higher than the ASCO for five domains of the AGREE-II. For one domain, the differences were statistically significant ($p \leq 0.05$). The CAPO guideline met 37 of 47 NHMRC mandatory guideline standards and the ASCO guideline met 20. The difference in the proportion of standards met was statistically significant for one domain ($p \leq 0.05$). Both guidelines had low scores for applicability and implementation.

✉ Elizabeth J. M. Pearson
  elizabeth.jm.pearson@gmail.com

1 Rural Health School, La Trobe University, Bendigo, Victoria, Australia

2 School of Allied Health, La Trobe University, Melbourne, Victoria, Australia

Conclusions Currently, the CAPO guideline for cancer-related fatigue has the strongest evidence for use. To enhance implementation, further strategies for guideline dissemination and application are needed.

Severe fatigue can be a disabling experience for people with cancer. Fatigue is recognised as one of the most prevalent and debilitating symptom of cancer and affects up to one third of cancer survivors [1, 2]. The US National Comprehensive Cancer Network (NCCN) defined cancer-related fatigue (CRF) as 'a distressing, persistent, subjective sense of physical, emotional, and/or cognitive tiredness or exhaustion related to cancer and/or cancer treatment that is not proportional to recent activity and interferes with usual functioning' [3].

Implementation of robust evidence-based guidelines is arguably needed for the effective management of CRF [4, 5]. Several guidelines exist although they appear to be under-utilised by oncology health professionals [4, 6].

The research question for the current study was 'Which clinical practice guideline for cancer-related fatigue is the most suitable for application?' A guideline is defined as 'a rule or instruction that shows or tells how something should be done' [7]. Clinical practice guidelines are 'statements that include recommendations intended to optimise patient care that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options' [8] p.4. The purpose of clinical guidelines is to assist practitioners and patients choose the most appropriate therapeutic interventions [8, 9]. To determine validity of clinical guidelines, a number of factors can be considered. These include validity; development procedures; stakeholder involvement; peer

review; the level, quality, and completeness of evidence, and the clarity of published recommendations [8, 10, 11].
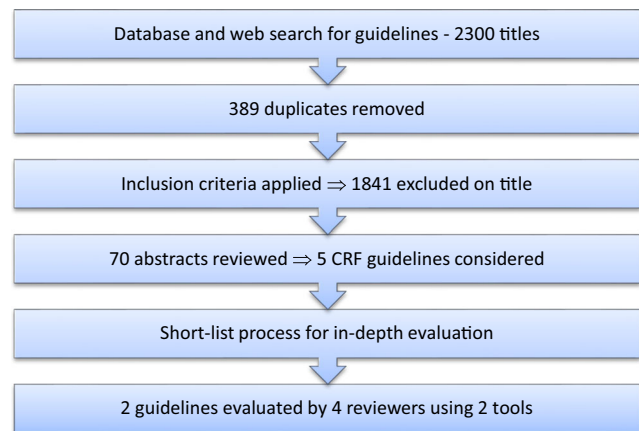
Methods

Guideline search and short listing

We conducted a systematic search for published guidelines for screening, assessment and treatment of CRF using terms described by Howell et al. [12] (Fig. 1). The databases MEDLINE®, PsychINFO, EMBASE®, and CINAHL® were searched in July 2015. The search for CRF guidelines extended to websites of four guideline portals[1] and Google search engine. Websites of identified guideline developers were checked for recent updates. Reference lists of guidelines and reviews describing guidelines were examined.

The inclusion and exclusion criteria were applied by one researcher (EP). The target population included adults in the post-treatment phase with a diagnosis of any cancer type or disease stage. Guidelines for assessment or treatment of CRF and written in English were considered if they detailed the development methodology. To ensure recommendations were evidence-based, care plans or algorithms without explicit links to the evidence were excluded. Older versions of guidelines, including those developed more than 5 years ago without an update, were eliminated due to the rapidly changing evidence base. Five CRF guidelines met the inclusion criteria. Guideline developers were the American Society of Clinical Oncology (ASCO) [13], the National Comprehensive Cancer Network (NCCN) [3, 14], the Oncology Nursing Society (ONS) [15], and the Canadian Association of Psychosocial Oncology (CAPO) [12, 16].

Previous quality appraisals of included guidelines using the AGREE-II instrument were used to short list the most rigorously developed guidelines for further appraisal. The AGREE-II instrument [17] is a valid guideline quality appraisal tool used by guideline developers [18] and statutory bodies [11]. The AGREE-II has 23 items within six domains of scope and purpose, stakeholder involvement, rigour of development, clarity of presentation, applicability, and editorial independence. Each item is rated on a Likert scale from 1 (strongly disagree) to 7 (strongly agree). Domain scores are reported as a percentage of the maximum possible score [18]. Initially, quality scores for 'rigour of development' (domain 3) were compared. Guideline scope and other domain scores assisted the overall decision. Recent guidelines without published scores were also considered for appraisal.

---

[1] National Guideline Clearinghouse www.guideline.gov; Clinical Practice Guideline Portal www.clinicalguidelines.gov.au; NICE Guidance www.nice.org.uk/guidance; Guidelines International Network www.g-i-n.net

Fig. 1 Summary of methods used in the appraisal

Appraisal methodology

Two instruments were used to appraise the guidelines: the AGREE-II instrument [17] and a checklist of Australian National Health and Medical Research Council (NHMRC) guideline standards [11]. AGREE-II [17] has been endorsed as the most comprehensive appraisal tool for local, national, and international clinical practice guidelines [19, 20]. Evaluation using AGREE-II is limited to the appraisal of the quality of guideline development processes and documentation [19]. The second tool was used to extend the evaluation.

For endorsement and approval for practice, clinical practice guidelines need to meet scientific standards. The NHMRC standard aligns with and expands upon most of the 20 US Institutes of Medicine clinical practice guideline standards [8]. Therefore, this appraisal was considered to be internationally relevant. For NHMRC approval, there are 54 mandatory criteria to be met, and additional 33 desirable criteria are listed [11]. The NHMRC domains are governance and stakeholder involvement, scope and purpose, evidence review, guideline recommendations, structure and style, public consultation and dissemination, and implementation.

Seven standards were considered not applicable to international guidelines due to specific references to NHMRC processes or indigenous people. The selected CRF guidelines were evaluated against a checklist of 47 guideline requirements [11]. Categorical responses were 'met', 'not met', or 'not applicable'. Reviewers recorded qualifying statements and the location of the evidence in the document. Guidelines, technical reports, administrative reports, and developer websites were the key sources of information used in the appraisals.

The AGREE Research Trust [21] recommended that at least two, and ideally four reviewers, should independently appraise each guideline (www.agreetrust.org). Four reviewers including one consumer were purposively recruited by direct invitation. Inclusion criteria were determined to ensure an informed

multidisciplinary review panel. The inclusion criteria were a relevant qualification in medicine, nursing, or occupational therapy AND expertise in clinical practice or research in the field of cancer supportive care; OR a consumer of health care with sufficient knowledge in guideline evaluation to complete the appraisal. Details of the reviewers' professional discipline, qualifications, age, gender, and location were recorded. Reviewers were offered payment for 8 h at a senior postdoctoral rate.

The four reviewers included an oncology nurse coordinator, a medical oncologist, an occupational therapist, and a consumer representative. All reviewers were female with tertiary qualifications at bachelor level and average age of 42.5 years (SD 12).

After written consent was obtained, reviewers were sent relevant guideline documentation, website links, and electronic versions of the NHMRC checklist and AGREE-II rating forms. A link to online training for AGREE-II and the user manual were provided. Reviewers were instructed to read the guideline documentation in detail and then rate their level of agreement with statements in the AGREE-II instrument and whether each NHMRC standard was met. The completed forms were returned to the research team.

The La Trobe University Human Ethics Sub-Committee of the College of Science, Health and Engineering approved all procedures in this study, reference number FHEC14/270.

The results from the four reviewers for each appraisal tool were tabulated into spreadsheets and analysed using Statistical Package for Social Sciences (SPSS®) software version 22 (IBM®). Non-parametric tests were used to determine the statistical significance of differences in domain scores between the two guidelines because of the small sample size [22].

The null hypotheses tested in the analyses were that the median of differences between guidelines of AGREE-II domain scores *and* of the proportion of NHMRC standards met in each domain equals zero ($p < 0.05$).

Data handling

Raw data for each instrument were modified to enable comparisons. The raw data obtained for each standard in the NHMRC appraisal was a categorical variable for each reviewer. Two researchers independently adjusted any ambiguous ratings to either 'met or not met', using the reviewers' notes. 'Unsure', 'N/A' or 'partly met' were adjusted to 'not met'. 'Mostly met' was adjusted to 'met'. Adjustment discrepancies were resolved by mutual agreement between the researchers.

The research team defined compliance for each NHMRC standard a priori as being positively endorsed by at least three of the four reviewers. Using this definition, the data were further adapted to an overall rating of standard 'met' or 'not met'. If one or two reviewers rated a standard as 'met', reviewer notes were used to determine whether the overall rating should be changed. The proportion of standards met in each domain was used as the unit for comparative analysis.

The AGREE-II unit of analysis is the domain score and is expressed as a percentage of the maximum possible score [17]. Domain scores were calculated using the formula specified by The AGREE Research Trust [21]:

$$\frac{(\text{Obtained score} - \text{Minimum possible score}) \times 100}{\text{Maximum possible score} - \text{Minimum possible score}}.$$

Individual reviewer scores, median, and overall scores were determined for each AGREE-II domain. Changes in reviewer ratings for each domain were plotted using Minitab® statistical software.

Analysis

Inter-rater reliability was calculated in SPSS® using the Kappa statistic for individual adjusted NHMRC data sets and the intra-class coefficient (ICC) for AGREE-II scores. The Kappa statistic was used to determine inter-rater agreement for independent evaluation of all cases by the same reviewers, using categorical variables with same number of categories [23]. For the NHMRC ratings, Kappa was calculated for each pair of raters (6 pairs) and the results were averaged as described by Light [24]. Proportions of agreement for adjusted NHMRC ratings were conducted for each of 6 pairs of reviewers using the tool at http://vassarstats.net. A two-way mixed, consistency average-measure ICC was calculated to assess the degree of reviewer consistency in rating of AGREE-II domains. This approach reflected the non-random sample of reviewers rating the guidelines, the unit of analysis as the 'average rating', and consistency of response as appropriate for Likert scales [23]. The AGREE Rating Concordance Calculator (available from Guidelines Resource Centre at www.cancerview.ca) was used to determine whether additional reviewers were required. Decision rules were based on the standard deviations of raw scores for both of the guidelines.

McNemar's exact test was performed using SPSS® to evaluate the statistical significance of differences in the proportion of NHMRC standards met in each domain. McNemar's test evaluates the significance of differences in pairs of dichotomous variables using 2×2 contingency table [25]. A related-sample Wilcoxon signed rank test [22] was performed in SPSS® to determine the significance of the difference of AGREE-II domain scores between the two guidelines. Significance level was set at $p < 0.05$ for all analyses.

Table 1 Characteristics of included guidelines

Guideline developers, location	American Society of Clinical Oncology (ASCO), USA and Canada	National Comprehensive Cancer Network (NCCN), USA	National Comprehensive Cancer Network (NCCN), USA	Oncology Nursing Society (ONS), USA	Canadian Association of Psychosocial Oncology (CAPO), Canada
Citation	Bower et al. [13]	NCCN [14]	NCCN [3]	Mitchell et al. [15]	Howell et al. [16]
Version last update	Original April 2014	Survivorship version 1.2015 February 2015	Fatigue version 2.2015 January 2015	Version 3 August 2014	Version 2 2015 April 2015
Literature search	March 2013	September 2014	July 2014	March 2014	November 2014
Development methodology	ADAPTE [18]	Expert consensus supported by evidence	Expert consensus supported by evidence	Systematic review	Systematic review and ADAPTE [18]
Population					
Age	Adult-onset	Adult-onset	Adults and children	Adult	Adult
Treatment phase	After treatment	After treatment	During and after	During and after	During and after
Disease stage	Remission	Remission	All including end of life	All including end of life	All including end of life
With CRF	Yes	–	–	–	Yes
Categories of evidence used	*NCCN and ONS categories* [3, 26]	*NCCN categories of evidence and consensus* [3] Cat 1: based upon high-level evidence, there is uniform NCCN consensus that the intervention is appropriate Cat 2A: based upon lower-level evidence, there is uniform NCCN consensus that the intervention is appropriate Cat 2B: based upon lower-level evidence, there is NCCN consensus that the intervention is appropriate Cat 3: based upon any level of evidence, there is major NCCN disagreement that the intervention is appropriate		*ONS weight-of-evidence categories* [26] Effectiveness established Likely to be effective Benefits balanced with harms Effectiveness not established Supported by expert opinion	*GRADE ratings* [27] Strong or weak with high-, moderate-, or low-quality evidence Also 'Expert panel consensus informed by guideline evidence' for some recommendations

Table 2 Published AGREE-II domain scores for fatigue guidelines

Guideline/Version	Review source	Domain 1	Domain 2	Domain 3	Domain 4	Domain 5	Domain 6
ASCO 2014	Howell et al. [16]	100	100	62	100	31	96
NCCN Fatigue 2.2009	Howell et al. [12]	58	42	24	86	44	79
NCCN Fatigue 1.2010	Harris et al. [27]	55.7	53.7	28.5	87.0	40.3	61.1
NCCN Fatigue 1.2013[a]	Bower et al. [13]	75	64	47	86	35	96
NCCN Fatigue 1.2014	Howell et al. [16]	94	94	79	100	21	58
NCCN Survivorship 2013[a]	Bower et al. [13]	92	67	44	86	17	28
ONS 2009	Harris et al. [27]	77.8	44.2	58.3	77.8	6.3	58.3
ONS 2007	Howell et al. [12]	53	31	50	72	29	29
CAPO 2011	Howell et al. [16]	83	89	67	69	25	100
CAPO 2011	Jacobs et al. [26]	-	-	75	-	-	38
CAPO 2011	Bower et al. [13]	-	-	86.5	-	-	-

[a] Data provided by ASCO on request

Results

Of the included guidelines, the CAPO, ONS, and NCCN Fatigue guidelines have application in all stages of cancer. The ASCO and NCCN Survivorship guidelines are specific to disease free survivors of adult–onset cancer. Guideline development methodology, target populations, and evidence categories varied between the guidelines as summarised in Table 1. Four guidelines included screening, assessment, and treatment of fatigue. The Oncology Nursing Society (ONS) guideline [15] focused on assessment and treatment of CRF only. Guideline recommendations were relatively consistent, but some differences were apparent, particularly evidence level. The guideline recommendations are summarised in Appendix 1.

Five publications were identified that reported AGREE-II results for one or more fatigue management guideline [12, 13, 16, 28, 29]. The domain scores for each pair of reviewers are shown in Table 2. Domain 3 represents 'rigor of development'.

The results of the appraisal by Bower et al. [13] suggested that the Canadian Association of Psychosocial Oncology (CAPO) fatigue guideline [12] was developed with substantial rigour, compared with the 2013 NCCN fatigue and NCCN survivorship guidelines. Two reviews compared the NCCN fatigue and ONS guidelines. Both rated the NCCN guideline's rigour very low, at approximately half the ONS rigour scores [12, 29]. Domain scores for the 2014 NCCN fatigue guideline

in the review by Howell et al. [16] were markedly higher than other reviews. This was considered to be inconsistent with previously reported scores, perhaps due to individual rater marking styles. The two NCCN guidelines and ONS guideline were then eliminated from further consideration in this study due to lower methodological rigour and lack of screening recommendations in the ONS guideline.

Based upon two independent reviews [13, 28], the CAPO fatigue guideline was selected and the 2015 version was appraised [16]. The ASCO fatigue guideline for survivors [13] was also selected due to promising scores in several domains in its only review [16].

Appraisal (1): NHMRC guideline standards

Inter-rater agreement across both guidelines using Light's kappa [30] was 0.48 with a standard error of 0.09, indicating a moderate agreement between reviewers ($0.41 \leq \kappa \geq 0.6$) [24] (see Table 3). This was consistent with Kappa values calculated for each guideline. The mean observed proportions of agreement was 0.76 with 95 % CI 0.56 to 0.93 (data on request).

The number and proportion of standards per domain meeting the 47 NHMRC guideline standards are shown in Table 4. The CAPO guideline met 37 standards and the ASCO guideline met 20 standards. The proportion of standards met by each guideline differed for four of seven domains. The difference

Table 3 Inter-rater agreement for NHMRC standards for both guidelines using Cohen's κ

Reviewer pairs	1 + 2	1 + 3	1 + 4	2 + 3	2 + 4	3 + 4	Light's κ (average)
Cohen's Kappa	0.560	0.318	0.727	0.284	0.604	0.369	0.477
95 % CI	0.389, 0.731	0.101, 0.535	0.583, 0.871	0.078, 0.490	0.440, 0.769	0.157, 0.581	–
Asymptotic standard error	0.086	0.096	0.073	0.087	0.083	0.095	0.087
Approximate p value	.000	.001	.000	.001	.000	.000	–

Table 4 Results of NHMRC evaluation

Domain	Number of items[a]	ASCO		CAPO		McNemar's exact test (Pan-Canadian–ASCO)	
		Standards met				Difference in proportion met	Two-tailed p value
		N	Proportion	N	Proportion		
A. Governance/ Stakeholder involvement	7	4	0.57	5	0.71	0.143	1
B. Scope and purpose	4	4	1.00	4	1.00	0	1
C. Evidence review	8	3	0.38	8	1.00	0.625	.063
D. Recommendations	12	2	0.17	10	0.84	0.667	.008
E. Structure and style	12	6	0.50	10	0.83	0.333	.219
F. Public consultation	2	0	0.00	0	0	0	1
G. Dissemination/ Implementation	2	0	0.00	0	0	0	1

[a] Seven items were not applicable to non-Australian guidelines

was statistically significant with a moderate effect size for domain D 'Recommendations' ($p = 0.008$), see Table 4.

Appraisal (2): AGREE-II instrument

The inter-rater reliability ICC for all domain scores was 0.86 (95 % CI 0.66 to 0.95) for absolute agreement and 0.89 (95 % CI 0.73 to 0.96) for consistency. Separate ICCs calculated for each guideline did not differ substantially from the combined ICC. These figures were in the excellent range as defined by Cicchetti [31]. According to AGREE Trust decision rules, additional reviewers were not required [21].

Domain scores (mean), median, and range using the AGREE-II are reported in Table 5. A related-sample Wilcoxon signed rank test was performed in SPSS to test the null hypothesis. The sum of median differences was 19 ($p = 0.046$) in favour of the CAPO guideline, which was not consistent with the null hypothesis. Further analysis considered the comparison of median scores for separate domains, using

the results at the reviewer level (data on request). The largest median difference in individual domains was for Editorial Independence; this comparison was statistically significant with higher scores for CAPO ($p < 0.05$). This difference represented a small effect size according to Cohen's criteria [32].

Discussion

The trustworthiness of the study was enhanced using four independent reviewers to rate two guidelines with two instruments. The study rigour was increased through the use of an internationally recognised valid guideline appraisal instrument, the AGREE-II [17, 19] together with national guideline standards [11]. Few statistically significant differences in domain scores were found in this analysis. It is noted that for domains with very few items, small p values and consequent rejection of the null hypothesis were not expected due to the influence of measurement error [33]. The AGREE-II domain of Editorial Independence has only two items, and the clinical

Table 5 Results of AGREE-II evaluations

Domain	Number of items	ASCO guideline %				Pan-Canadian guideline %				Wilcoxon signed rank test	
		Mean score	SD	Median	Range	Mean score	SD	Median	Range	Mean difference[a]	p value
1. Scope and purpose	3	89	0.96	88.9	78–100	99	0.17	100	94–100	−10	.190
2. Stakeholder involvement	3	74	1.16	77.8	50–89	71	0.93	72.2	56–83	4	.803
3. Rigour of development	8	66	1.80	64.3	23–73	86	0.64	75.9	68–77	−20	.170
4. Clarity of presentation	3	71	1.33	75.0	44–89	85	0.85	80.6	78–100	−14	.264
5. Applicability	4	23	1.16	20.8	4–49	41	1.95	33.3	13–83	−18	.351
6. Editorial independence	2	67	1.26	66.7	50–83	94	0.00	95.8	83–100	−24	.020

[a] Difference in mean domain scores of the four raters on each guideline

relevance of this finding was uncertain. In contrast, the significant finding of a moderate effect size for the 12-item NHMRC Recommendations domain indicated important differences in the methodological quality of the recommendations between the guidelines. The significant AGREE-II overall result suggested that the quality of development and reporting for the CAPO fatigue guideline was superior to that of the ASCO guideline. Additionally, the CAPO fatigue guideline met 17 more NHMRC guideline standards than did the ASCO. This suggested that the CAPO is the more suitable guideline for clinical use.

This study had several limitations. Although short-listing of guidelines using AGREE-II was recommended [18], the use of previously published scores to guide selection is novel. Benchmark AGREE-II quality scores have not been published, and relative scores were used to rank quality domains. Readers should not consider these scores as absolute but rather in the context of other ratings by the same pairs of reviewers. Comparing different iterations of guidelines could be misleading if development methodology and evidence changes. The use of NHMRC standards as a tool by which to compare guideline properties is also novel. Because the domain constructs of the NHMRC Standards remain untested, their validity may be questioned [34].

The AGREE-II and NHMRC guideline standard evaluations were dependent on obtaining accurate and complete guideline documentation. It is possible that materials were overlooked or incomplete, which could result in lower ratings [35]. Dichotomous scoring in the NHMRC appraisal may have also reduced scores if a standard were partially met [35]. To address this, we adjusted the ambiguous ratings based on the notes of all four reviewers. The overall results were unchanged by this procedure.

Both of the tools used appraised documentation and methodological quality [19]. The guidelines both scored poorly for applicability or implementation. Neither the validity nor clinical appropriateness of guideline recommendations was evaluated in this appraisal. Additional evaluation methods are required to determine the acceptability, feasibility, and effectiveness of the guidelines. Further evaluation could include pre-implementation studies using end-user feedback [36], tools such as the GuideLine Implementability Appraisal [37], and pilot clinical studies [38].

Conclusions

The 2015 CAPO guideline for cancer-related fatigue appears to be appropriate for clinical use worldwide. Further enhancement of the guidelines is needed to enable application to local contexts. It is recommended that guideline developers make the application of evidence-based guidelines easier to enhance their implementation.

References

1. Prue G, Rankin J, Allen J, Gracey J, Cramp F (2006) Cancer-related fatigue: a critical appraisal. Eur J Cancer 42(7):846–863. doi:10.1016/j.ejca.2005.11.026

2. Wu H-S, Harden JK (2015) Symptom burden and quality of life in survivorship: a review of the literature. Cancer Nurs 38(1):E29–E54. doi:10.1097/NCC.0000000000000135

3. NCCN (2015) Cancer-related fatigue version 2.2015. http://www.nccn.org/professionals/physician_gls/pdf/fatigue.pdf. Accessed 16/7/2015

4. Berger AM, Mitchell SA, Jacobsen PB, Pirl WF (2015) Screening, evaluation, and management of cancer-related fatigue: ready for implementation to practice? CA Cancer J Clin. doi:10.3322/caac.21268

5. Silver JK, Gilchrist LS (2011) Cancer rehabilitation with a focus on evidence-based outpatient physical and occupational therapy interventions. Am J Phys Med Rehabil 90(5 Suppl 1):S5–15. doi:10.1097/PHM.0b013e31820be4ae

6. Pearson EJM, Morris ME, McKinstry CE (2015) Cancer related fatigue: a survey of health practitioner knowledge and practice. Support Care Cancer 23(12):3521–3529. doi:10.1007/s00520-015-2723-8

7. Merriam-Webster (2014) Merriam-Webster online dictionary. doi:http://www.merriam-webster.com/

8. Institutes of Medicine (2011) Clinical practice guidelines we can trust. The National Academies Press, Washington, DC doi:http://nap.edu/catalog/13058.htm

9. Woolf S, Schünemann HJ, Eccles MP, Grimshaw JM, Shekelle P (2012) Developing clinical practice guidelines: types of evidence and outcomes; values and economics, synthesis, grading, and presentation and deriving recommendations. Implement Sci 7:61. doi:10.1186/1748-5908-7-61

10. Clubb AB, Dahm P (2011) How to critically appraise a clinical practice guideline. Indian J Urol 27(4):498–502. doi:10.4103/0970-1591.91441

11. National Health and Medical Research Council (2011) Procedures and requirements for meeting the 2011 NHMRC standard for clinical practice guidelines. National Health and Medical Research Council, Melbourne doi:https://www.nhmrc.gov.au/guidelines/publications/cp133-and-cp133a

12. Howell D, Keller-Olaman S, Oliver TK, Hack TF, Broadfield L, Biggs K, Chung J, Gravelle D, Green E, Hamel M, Harth T, Johnston P, McLeod D, Swinton N, Syme A, Olson K (2013) A pan-Canadian practice guideline and algorithm: screening, assessment, and supportive care of adults with cancer-related fatigue. Curr Oncol 20(3):e233–e246. doi:10.3747/co.20.1302

13. Bower JE, Bak K, Berger AM, Breitbart W, Escalante CP, Ganz PA, Schnipper HH, Lacchetti C, Ligibel JA, Lyman GH, Ogaily MS, Pirl WF, Jacobsen PB (2014) Screening, assessment, and management of fatigue in adult survivors of cancer: an American Society of Clinical Oncology clinical practice guideline adaptation. J Clin Oncol 32(17):1840–1850. doi:10.1200/JCO.2013.53.4495

14. NCCN (2014) Survivorship version 1.2015. National Comprehensive Cancer Network. http://www.nccn.org/professionals/physician_gls/pdf/survivorship.pdf. Accessed 17/7/2015

15. Mitchell SA, Hoffman AJ, Clark JC, DeGennaro RM, Poirier P, Robinson CB, Weisbrod BL (2014) Putting evidence into practice: an update of evidence-based interventions for cancer-related fatigue during and following treatment. Clin J Oncol Nurs 18(Suppl):38–58. doi:10.1188/14.cjon.s3.38-58

16. Howell D, Keshavarz H, Broadfield L, Hack T, Hamel M, Harth T, Jones J, McLeod D, Olson K, Phan S, Sawka A, Swinton N, Ali M, on behalf of the Cancer Journey Advisory Group of the Canadian Partnership Against Cancer (2015) A pan Canadian practice guideline for screening, assessment, and management of cancer-related fatigue in adults Version 2-2015. Canadian Association of Psychosocial Oncology, Toronto doi:http://www.capo.ca/pdf/CRF_Guideline.pdf

17. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, Fervers B, Graham ID, Grimshaw J, Hanna SE, Littlejohns P, Makarski J, Zitzelsberger L (2010) AGREE II: advancing guideline development, reporting and evaluation in health care. Can Med Assoc J 182(18):E839–E842. doi:10.1503/cmaj.090449

18. The ADAPTE Collaboration (2009) The ADAPTE process: resource toolkit for guideline adaptation version 2.0. http://www.g-i-n.net. Accessed 13 Nov 2014

19. Cronin P, Rawson JV, Heilbrun ME, Lee JM, Kelly AM, Sanelli PC, Bresnahan BW, Paladin AM (2014) How to critically appraise the clinical literature. Acad Radiol 21(9):1117–1128. doi:10.1016/j.acra.2014.05.004

20. Siering U, Eikermann M, Hausner E, Hoffmann-Esser W, Neugebauer EA (2013) Appraisal tools for clinical practice guidelines: a systematic review. PLoS One 8(12):e82915. doi:10.1371/journal.pone.0082915

21. The AGREE Research Trust (2013) Appraisal of guidelines for REsearch and evaluation II. AGREE II Instrument. www.agreetrust.org. Accessed 6 August 2015

22. Pallant JF (2013) SPSS survival manual - a step by step guide to data analysis using IBM SPSS, 5 edn. Allen & Unwin, Crows Nest

23. Hallgren KA (2012) Computing inter-rater reliability for observational data: an overview and tutorial. Tutor Quant Methods Psychol 8(1):23–34 doi:nihms37295.pdf

24. Light RJ (1971) Measures of response agreement for qualitative data: some generalizations and alternatives. Psychol Bulletin 76(5):365–377. doi:10.1037/h0031643

25. Liddell FDK (1983) Simplified exact analysis of case-referent studies: matched pairs; dichotomous exposure. J Epidemiol Community Health 37(1):82–84. doi:10.1136/jech.37.1.82

26. Mitchell SA, Friese CR (2012) ONS PEP (Putting Evidence into Practice) weight of evidence classification schema: decision rules

for summative evaluation of a body of evidence. https://www.ons.org/practice-resources/pep/evaluation-process. Accessed July 3 2015

27. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ (2008) GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 336(7650):924–926. doi:10.1136/bmj.39489.470347.AD

28. Jacobs C, Graham ID, Makarski J, Chasse M, Fergusson D, Hutton B, Clemons M (2014) Clinical practice guidelines and consensus statements in oncology - an assessment of their methodological quality. PLoS One 9(10). doi:10.1371/journal.pone.0110469

29. Harris SR, Schmitz KH, Campbell KL, McNeely ML (2012) Clinical practice guidelines for breast cancer rehabilitation. Cancer 118(S8):2312–2324. doi:10.1002/cncr.27461

30. Landis JR, Koch GG (1977) An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. Biometrics 33(2):363–374. doi:10.2307/2529786

31. Cicchetti DV (1994) Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess 6(4):284–290. doi:10.1037/1040-3590.6.4.284

32. Cohen J (1988) Statistical power analysis for the behavioral sciences, 2 edn. Academic Press Inc, Hillsdale, NJ

33. Nunnally JC (1994) Psychometric theory. McGraw-Hill series in psychology, 3rd edn. McGraw-Hill, New York

34. Eisinga R, Grotenhuis M, Pelzer B (2013) The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? Int J Public Health 58(4):637–642. doi:10.1007/s00038-012-0416-3

35. Reames BN, Krell RW, Ponto SN, Wong SL (2013) Critical evaluation of oncology clinical practice guidelines. J Clin Oncol 31(20):2563–2568. doi:10.1200/JCO.2012.46.8371

36. Kastner M, Estey E, Hayden L, Chatterjee A, Grudniewicz A, Graham ID, Bhattacharyya O (2014) The development of a guideline implementability tool (GUIDE-IT): a qualitative study of family physician perspectives. BMC Fam Pract 15:19. doi:10.1186/1471-2296-15-19

37. Shiffman RN, Dixon J, Brandt C, Essaihi A, Hsiao A, Michel G, O'Connell R (2005) The GuideLine implementability appraisal (GLIA): development of an instrument to identify obstacles to guideline implementation. BMC Med Inform Decis Mak 5:23. doi:10.1186/1472-6947-5-23

38. National Institute of Clinical Studies (2006) Assessing the implementability of guidelines. NICS, Melbourne doi:http://www.nhmrc.gov.au/