

BMJ Open Evaluation of guidelines regarding surgical treatment of breast cancer using the AGREE Instrument: a systematic review

Xin Lei,^{1,2,3} Fengtao Liu,^{1,2} Shuying Luo,^{2,4} Ya Sun,^{1,2} Liling Zhu,^{1,2} Fengxi Su,^{1,2} Kai Chen,^{1,2} Shunrong Li^{1,2}

To cite: Lei X, Liu F, Luo S, *et al.* Evaluation of guidelines regarding surgical treatment of breast cancer using the AGREE Instrument: a systematic review. *BMJ Open* 2017;**7**:e014883. doi:10.1136/bmjopen-2016-014883

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2016-014883>).

XL, FL and SL contributed equally.

Received 25 October 2016
Revised 31 July 2017
Accepted 11 August 2017



CrossMark

¹Breast Tumour Center, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China

²Guangdong Provincial Key Laboratory of Malignant Tumour Epigenetics and Gene Regulation, Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, Guangzhou, China

³Thyroid and Breast Surgery Department, The First Affiliated Hospital, Guangzhou, Guangdong, China

⁴Department of Pharmacy, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, Guangdong, China

Correspondence to

Dr. Kai Chen;
chenkai23@mail.sysu.edu.cn
and Dr Shunrong Li;
lishr@mail.sysu.edu.cn

ABSTRACT

Objectives Many clinical practice guidelines and consensus statements (CPGs/consensus statements) have been developed for the surgical treatments for breast cancer. This study aims to evaluate the quality of these CPGs/consensus statements.

Methods We systematically searched the PubMed and EMBASE databases, as well as four guideline repositories, to identify CPGs and consensus statements regarding surgical treatments for breast cancer between January 2009 and December 2016. We used the Appraisal of Guidelines for Research and Evaluation (AGREE) instrument to assess the quality of the CPGs and consensus statements included. The overall assessment scores from the AGREE instrument and radar maps were used to evaluate the overall quality. We also evaluated some factors that may affect the quality of CPGs and consensus statements using the Mann-Whitney U test or Kruskal-Wallis H test. All analyses were performed using SPSS V.19.0. This systematic review was conducted according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.

Results A total of 19 CPGs and four consensus statements were included. In general, the included CPGs/consensus statements (n=23) performed well in the 'Scope and Purpose' and 'Clarity and Presentation' domains, but performed poorly in the 'Applicability' domain. The American Society of Clinical Oncology (ASCO), National Institute for Health and Care Excellence (NICE), Scottish Intercollegiate Guidelines Network (SIGN), New Zealand Guidelines Group (NZGG) and Belgium Health Care Knowledge Centre (KCE) guidelines had the highest overall quality, whereas the Saskatchewan Cancer Agency, Spanish Society of Medical Oncology (SEOM), Japanese Breast Cancer Society (JBCS) guidelines and the D.A.C.H and European School of Oncology (ESO) consensus statements had the lowest overall quality. The updating frequency of CPGs/consensus statements varied, with the quality of consensus statements generally lower than that of CPGs. A total of six, eight and five CPGs were developed in the North American, European and Asian/Pacific regions, respectively. However, geographic region was not associated with overall quality.

Conclusions The ASCO, NICE, SIGN, NZGG and KCE guidelines had the best overall quality, and the quality of consensus statements was generally lower than that of

Strengths and limitations of this study

- This was a systematic review conducted following Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines, including descriptions of key methodological steps, results and discussion.
- This was the first study, to our knowledge, to systematically assess the methodological quality of CPGs and consensus statements regarding surgical treatments for breast cancer using the Appraisal of Guidelines for Research and Evaluation II instrument.
- We only searched two databases and four guideline repositories and only included literature published in English. Only CPGs and consensus statements published after January 2009 were included.

CPGs. More efforts are needed to identify barriers and facilitators for CPGs/consensus statement implementation and to improve their applicability.

INTRODUCTION

Surgical treatment is the major approach for patients with non-metastatic breast cancer.¹ The quality of surgical treatment of breast cancer depends on a variety of factors, including the surgeons' perspective as well as the patient's socioeconomic status and resources.² Among these, surgeons' perspective is an important factor that is associated with the services provided and is shaped by a variety of factors, including clinical practice guidelines or consensus statements (CPGs/consensus statements). CPGs/consensus statements have been developed to optimise and standardise the surgical management of breast cancer to improve the quality of care. They should provide clear, comprehensive and evidence-based recommendations to reduce the gap between research and clinical practice.³ However, when developed by different institutions and/or countries,

CPGs/consensus statements may provide equivocal or inconsistent recommendations due to different perspectives, local resources or updating frequency, among other factors. This can result in confusion among healthcare providers in clinical practice regarding which CPGs/consensus statements to follow and what to consider when applying the recommendations. Such confusion may affect healthcare providers' implementation and adherence to the CPGs/consensus statements, which may in turn affect long-term patient outcomes.⁴⁻⁶ For example, the National Comprehensive Cancer Network (NCCN) has incorporated the conclusions of the ACOSOG Z0011 study⁷; patients fitting the Z0011 criteria may be spared from axillary lymph node dissection (ALND) if their surgeons follow the NCCN recommendations. Some CPGs/consensus statements suggest that patients meeting the Z0011 criteria may be eligible to avoid ALND, a recommendation that may sometimes be ambiguous. Healthcare providers may not be able to find clear statements in these CPGs/consensus statements regarding additional considerations in patients fitting the Z0011 criteria to avoid ALND. Clarity and unambiguity of recommendations are important factors in the implementation of CPGs/consensus statements and reflect their methodological quality.

The methodological quality of CPGs/consensus statements is an important factor to guide surgeons regarding which CPGs/consensus statements they should follow and also aids CPGs/consensus developers in considering their strategy for developing and updating their CPGs/consensus statements.⁸ Several instruments have been developed to assess the methodological quality of CPGs/consensus statements. Among them, the Appraisal of Guidelines for Research and Evaluation (AGREE) II instrument is the most popular and has been validated internationally.⁹⁻¹¹ In this study, we conducted a systematic review of the CPGs/consensus statements regarding the surgical management of breast cancer and assessed their methodological quality using the AGREE II instrument. We also investigated potential factors that might be associated with quality.

METHODS

This review was performed following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines, thus providing a comprehensive framework for objectively assessing quality indicators and the risk of bias in the included CPGs/consensus statements.

Data sources and searches

Recent progress in scientific researches has led to advances in surgical treatments for breast cancer over the past decade, resulting in a need to update many CPGs and consensus statements. We therefore only searched studies published between January 2009 and December 2016. Two independent reviewers screened the PubMed and EMBASE databases for guidelines and consensus

statements on surgical treatments for breast cancer. The search strategy included terms related to breast cancer, surgical treatments, guideline and consensus. Online supplementary file 1 has the full PubMed search strategy, which was adapted to suit other databases. Additionally, four guideline repositories, the National Guideline Clearinghouse (USA), the National Library for Health (UK) on Guideline Finder, Canadian Medical Association Infobase (Canada) and the Guidelines International Network (G-I-N) International Guideline Library were manually searched. We also performed a search of the websites for the organisations that developed those CPGs/consensus statements.

Inclusion and exclusion criteria

According to the National Guideline Clearinghouse, we defined CPGs as statements that included recommendations intended to optimise and standardise patient care informed by a systematic review of evidence and assessment of the benefits and risks of alternative care options.^{12 13} Consensus statements based on comprehensive or systematic reviews and providing clinically relevant suggestions based on the collective opinion of an expert panel¹² were also included.

We included CPGs/consensus statements if they met the following criteria:

1. addressed issues about surgical management of breast cancer, including breast surgery and axillary surgery;
2. published in English;
3. full-text available.

We excluded:

1. CPGs/consensus statements focused on a specific topic that was irrelevant to the surgical management of breast cancer, for example, screening guidelines;
2. CPGs/consensus statements focused only on metastatic breast cancer, as surgical management in these patients is not the primary recommendation^{14 15};
3. CPGs/consensus statements focused on breast reconstruction surgery, such as prosthesis implantation, autologous reconstruction;
4. CPG/consensus statements 'for education and information purpose' or 'out of date' because the organisations declared that CPG/consensus statements may no longer be consistent with recent evidence;
5. draft or unpublished guidelines, discussion papers, personal opinions and obsolete guidelines replaced by updated recommendations from the same organisation.

Several additional principles were followed:

1. If multiple updated versions of a CPG/consensus statement were available, the most recent one was included.
2. If doubts existed regarding whether an article was a CPG/consensus statement, we verified its eligibility by checking the inclusion criteria of similar reports in the National Guideline Clearinghouse.

Two authors (XL and FTL) independently searched and identified eligible CPGs/consensus statements and

collected the full text of the CPGs/consensus statements and related supplementary materials if available. The authors met to gather and compile all available information to ensure that no relevant information was missed and to ensure that all three reviewers reviewed the same materials. Discrepancies or inconsistent findings were discussed together with the third author (SYL). Because this was a systematic review, the ethical approvals of Sun Yat-Sen Memorial Hospital and the First Affiliated Hospital, Sun-Yat Sen University, were waived.

Guideline quality assessment

The quality of each CPG/consensus statement was independently evaluated by three different reviewers (XL, KC and YS) using the AGREE II Instrument^{9 11} (updated: September 2009). The AGREE II¹¹ instrument evaluates 23 items categorised into six domains, including Scope and Purpose, Stakeholder Involvement, Rigour of Development, Clarity and Presentation, Applicability and Editorial Independence. Reviewers scored each item ranging from 1 (strongly disagree) to 7 (strongly agree). A score of 1 is assigned when no information concerning that item is available, while a score of 7 indicates that clear information is evident and the full criteria were met. The domain score is calculated by scaling the total obtained score of that domain as a percentage of the maximum possible score for that domain using the following formula:

$$\text{Domain score} = (\text{obtained score} - \text{minimum possible score}) / (\text{maximum possible score} - \text{minimum possible score})$$

For example, if three reviewers assessed the CPGs/consensus statements, with four items within a domain, the maximum possible score and minimum possible score was $7 \times 4 = 28$ and $1 \times 4 = 4$, respectively. If a total score of 30 was obtained, the domain score was $(30 - 4) / (28 - 4) = 0.25$.

If the actual total score of all items within one domain between any two of the three appraisers differed by $>30\%$ of the maximal total score of all items within that domain, disagreements were discussed by all reviewers, together with a fourth author (SXF), to ensure that all necessary information (supplementary files, website pages, full-text) was collected. After discussion, the three reviewers (XL, KC and YS) re-evaluated the CPGs/consensus statements and resubmitted their final domain scores. The reviewers could keep the previous score without any changes after discussion. Consistency among reviewers on AGREE II scores was assessed using the intraclass correlation coefficient (ICC).

In addition to the six domains, an overall assessment is included in the AGREE II instrument. Three reviewers (XL, KC and YS) scored the overall quality of the CPGs/consensus statements from 1 to 7; the overall assessment score was calculated using the same equation as that used for the domain scores. Additionally, we used a radar map to illustrate the domain scores of each CPG/consensus statement and calculated the total area of the radar map as a reflection of the overall quality of the

CPGs/consensus statements. The radar map areas were expressed as percentages of the maximal area. The association between radar map areas and overall assessment score was tested using linear regression analysis. The reviewers also categorised the CPGs/consensus statements into three groups: recommending the CPGs/consensus statements for use, recommending the CPGs/consensus statements for use with modifications and not recommending the CPGs/consensus statements for use.

Factors associated with guideline quality

Two authors (XL and LLZ) developed a data extraction plan to collect the main features of each CPG/consensus statement (eg, year of publication, country/region, year of publication, update frequency). The quality (radar map and overall assessment scores) of the CPGs/consensus statements according to these factors was compared using the Mann-Whitney U test or the Kruskal-Wallis H test as appropriate. $p < 0.05$ was considered statistically significant. All analyses were performed using SPSS V.19.0.

RESULTS

Search results and characteristics

A total of 19 guidelines^{16–34} and 4 consensus statements^{35–38} were identified for final evaluation (figure 1). Among the 19 CPGs included, six,^{18 23 28–30 33} eight^{16 17 24–26 31 32 34} and five^{19–22 27} CPGs were developed in North American, European and Asian/Pacific regions, respectively (table 1).

The update frequency of each CPGs/consensus statement varied (table 2). The National Institute for Health and Care Excellence (NICE),¹⁷ Malaysia,²¹ New Zealand Guidelines Group (NZGG),²⁰ Cancer Australia-National Breast and Ovarian Cancer Centre (CA-NBOCC)²² guidelines and the Biedenkopf³⁵ consensus statement have not been updated since 2011. The German Group for Gynaecological Oncology (AGO)³² and the National Comprehensive Cancer Network (NCCN)¹⁸ guidelines have been updated annually, and the St. Gallen³⁶ consensus statement has been updated every other year.

The website links of the included CPG/consensus statements are listed in online supplementary table 1. The ICCs of the three reviewers for each guideline/consensus statement ranged between 0.90 and 0.99 (online supplementary table 2.1); the ICCs of the three reviewers for each domain of AGREE II ranged between 0.82 and 0.96 (online supplementary table 2.2), suggesting good agreement of rating scores among the three reviewers.

Guidelines appraisal

Overall quality assessment

The overall assessment scores and radar map areas were significantly correlated (online supplementary figure 1) ($R^2 = 0.835$, $p < 0.05$). The overall assessment scores suggested that the American Society of Clinical Oncology (ASCO),²⁸ NICE,¹⁷ NCCN,¹⁸ Scottish Intercollegiate Guidelines Network (SIGN),¹⁶ NZGG²⁰ and Belgium Health Care Knowledge Centre (KCE)²⁶

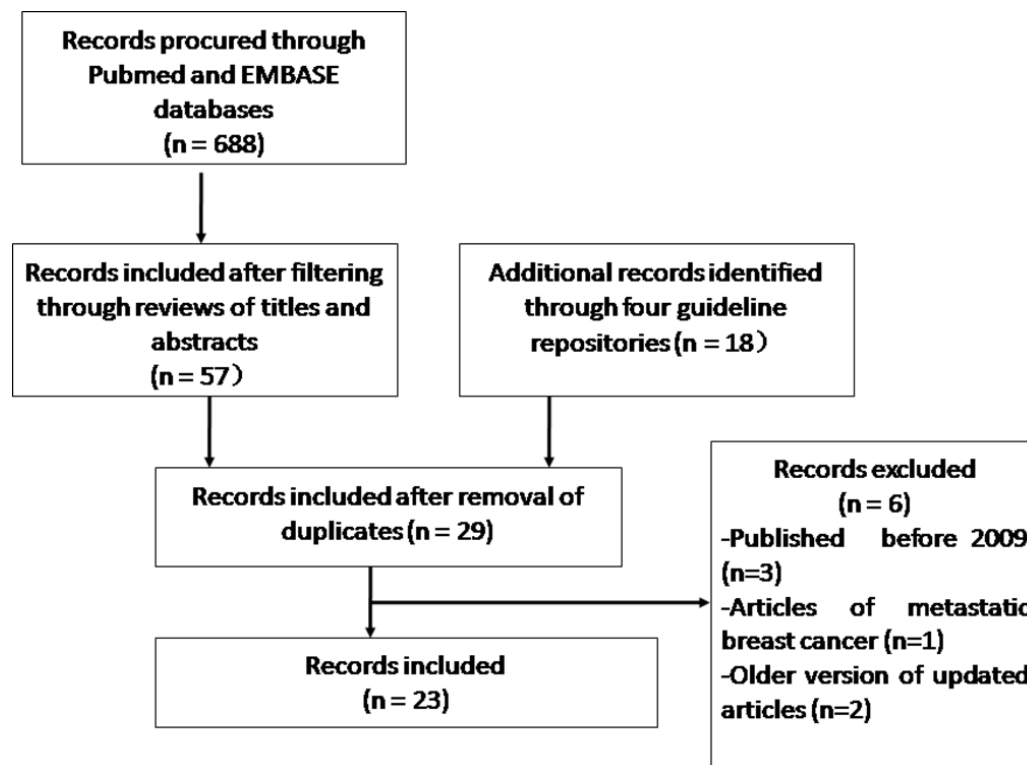


Figure 1 Flow chart of the systematic review.

guidelines had the best overall quality, whereas the Saskatchewan Cancer Agency (SASK),²³ Spanish Society of Medical Oncology (SEOM),³¹ Japanese Breast Cancer Society (JBCS)¹⁹ guidelines, and the D.A.C.H.³⁷ and European School of Oncology (ESO)³⁸ consensus statements had the poorest overall quality (table 3). Radar map areas suggested that the ASCO,²⁸ SIGN,¹⁶ NICE,¹⁷ NZGG²⁰ and KCE²⁶ guidelines had the best overall quality, whereas the SASK,²³ SEOM³¹ and JBCS¹⁹ guidelines and the D.A.C.H.³⁷ and ESO³⁸ consensus statements had the poorest overall quality (table 3, figure 2). All three reviewers categorised the SIGN,¹⁶ KCE,²⁶ ASCO²⁸ and Malaysia²¹ guidelines as ‘recommend for use’, whereas all three reviewers categorised the SASK²³ guideline as ‘not recommend for use’ in the overall assessment (table 3).

Domain assessment

In general, the median domain scores (range) of the Scope and Purpose, Stakeholder Involvement, Rigour of Development, Clarity and Presentation, Applicability and Editorial Independence domains were 74.1% (57.4%–92.6%), 38.9% (13.0%–94.4%), 54.9% (14.6%–89.6%), 79.6% (40.7%–92.6%), 23.6% (4.2%–73.6%) and 63.9% (8.3%–91.7%), respectively. All included CPGs/consensus statements scored >50% in the Scope and Purpose domain. In contrast, only five CPGs/consensus statements^{16–18 20 28} scored >50% in the Applicability domain. Five of the CPGs/consensus statements^{16–18 20 28} had all domain scores >50%, while in SEOM,³¹ all but the Scope and Purpose domain scored <50%. The domain scores of each CPGs/consensus statement are listed in table 2.

Factors associated with quality

CPGs versus consensus statements

In total, 4 consensus statements and 19 CPGs were included in this study. In general, consensus statements had lower overall quality than CPGs. The median (range) of the radar map area was 19% (15.1%–22.6%) and 34.6% (5.8%–67.7%) for consensus statements and CPGs, respectively (p=0.10). The median (range) of overall assessment scores ranged between 72.2% (22.2%–94.4%) and 41.7% (33.3%–44.4%) for consensus statements and CPGs, respectively (p=0.01). As shown in table 2, consensus statements had lower average domain scores than CPGs in Stakeholder Involvement (consensus statements 30.1% vs CPGs 52.7%, p=0.133), Rigour of Development (consensus statements 30.1% vs CPGs 61.3%, p=0.062) and Applicability (consensus statements 14.6% vs CPGs 33.9%, p=0.088) domains, none of which were statistically significant.

Geographic regions

The median (range) of the radar map areas were 26.4% (7.6%–67.5%), 41.4% (5.8%–67.7%) and 47.6% (9.4%–59.6%), and the median (range) of overall assessment scores were 63.9% (22.2%–88.9%), 72.2% (44.4%–94.4%) and 72.2% (38.9%–77.8%) for CPGs published in Europe, North America and Asian/Pacific regions, respectively. However, we did not observe any statistically significant differences in the radar map areas or the overall assessment scores among CPGs developed in different geographic regions (p>0.05).

Table 1 Guidelines and consensus included

Organisation	Title	Year of publication	Country	Published in journal
ESMO	Primary breast cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up	2015	Europe	<i>Annals of Oncology</i>
NCCN	Clinical guidelines in oncology: breast cancer	2016	USA	
JBCS	Clinical practice guideline for surgical treatment of breast cancer	2016	Japan	<i>Breast Cancer</i>
ASCO	Sentinel lymph node biopsy for patients with early-stage breast cancer	2016	USA	<i>Journal of Clinical Oncology</i>
AGO	Recommendations for the diagnosis and treatment of patients with early breast cancer	2015	German	
NICE	Early and locally advanced breast cancer diagnosis and treatment	2009	UK	
CCO	Locoregional therapy of locally advanced breast cancer	2015	Canada	
SIGN	Treatment of primary breast cancer	2013	Scotland, UK	
ACR	Ductal carcinoma in situ	2015	USA	<i>Journal of American College of Radiology</i>
KCE	Breast cancer in women: diagnosis, treatment and follow-up	2013	Belgium	
SSO-ASTRO	Consensus guideline on margins for breast-conserving surgery with whole breast irradiation in stage I and II invasive breast cancer	2014	USA	<i>Journal of Clinical Oncology</i>
SEOM	Clinical guidelines in early stage breast cancer	2015	Spain	<i>Clinical and Translational Oncology</i>
EUSOMA	Recommendations for the management of young women with breast cancer	2012	Europe	<i>European Journal of Cancer</i>
CA-NBOCC	Recommendations for staging and managing the axilla in early breast cancer	2011	Australia	
CA-BRCA	Recommendations for the management of early breast cancer in women with identified BRCA1 or BRCA2 gene mutation or at high risk of a gene mutation	2014	Australia	
SASK	Breast cancer treatment guideline	2012	Canada	
Malaysia	Management of breast cancer	2010	Malaysia	
SIOG	Management of elderly patients with breast cancer	2012	Europe	<i>Lancet Oncology</i>
NZGG	Management of early breast cancer	2009	New Zealand	
D.A.C.H	Diagnosis and local treatment of axilla in breast cancer	2013	International	<i>European Journal of Cancer</i>
St. Gallen	Tailoring therapy-improving the management of early breast cancer	2015	International	<i>Annals of Oncology</i>
ESO	First international consensus guidelines for breast cancer in young women	2014	International	<i>The Breast</i>
Biedenkopf	Locoregional treatment of primary breast cancer	2010	International	<i>Cancer</i>

AGO, German Group for Gynecological Oncology; ESMO, European Society for Medical Oncology; NCCN, National Comprehensive Cancer Network; JBCS, Japanese Breast Cancer Society; ASCO, American Society of Clinical Oncology; NCI, National Institute for Health and Care Excellence; CCO, Cancer Care Ontario; SIGN, Scottish Intercollegiate Guidelines Network; ACR, American College of Radiology; KCE, Belgium Health Care Knowledge Center; SSO-ASTRO, Society of Surgical Oncology-American Society for Radiation oncology; SEOM, Spanish Society of Medical Oncology; EUSOMA, European Society of Breast Cancer Specialists; CA-NBOCC, Cancer Australia-National Breast and Ovarian Cancer Center; SASK, Saskatchewan Cancer Agency; Malaysia Academy of Medicine of Malaysia; SIOG, Society of Geriatric Oncology; NZGG, New Zealand Guidelines Group; D.A.C.H, German, Australia, Swiss Societies of Senology, St Gallen St Gallen Consensus; ESO, European School of Oncology, Biedenkopf the Biedenkopf expert panel members.

Table 2 Update frequency of the included CPGs/consensus statements

Guidelines	First year of publication	2009	2010	2011	2012	2013	2014	2015	2016
SIGN	2005					★			
NICE	2006	★							
KCE	2007		★			★			
SIOG	2007				★				
AGO	2012				★	★	★	★	
EUSOMA	2012				★				
ESMO	2005		★			★		★	
SEOM	2010		★					★	
ASCO	2005						★		★
CCO	2015							★	
SSO-ASTRO	2014						★		
NCCN	1995	★	★	★	★	★	★	★	★
ACR	1996				★			★	
SASK	2012				★				
Malaysia	2010		★						
NZGG	2009	★							
CA-BRCA	2001						★		
CA-NBOCC	2001			★					
JBCS	2014						★	★	★
St Gallen	1987	★		★		★		★	
ESO	2014						★		
D.A.C.H.	2013					★			
Biedenkopf	2010		★						

AGO, German Group for Gynecological Oncology; ESMO, European Society for Medical Oncology; NCCN, National Comprehensive Cancer Network; JBCS, Japanese Breast Cancer Society; ASCO, American Society of Clinical Oncology; NCIE, National Institute for Health and Care Excellence; CCO, Cancer Care Ontario; SIGN, Scottish Intercollegiate Guidelines Network; ACR, American College of Radiology; KCE, Belgium Health Care Knowledge Center; SSO-ASTRO, Society of Surgical Oncology-American Society for Radiation oncology; SEOM, Spanish Society of Medical Oncology; EUSOMA, European Society of Breast Cancer Specialists; CA-NBOCC, Cancer Australia-National Breast and Ovarian Cancer Center; SASK, Saskatchewan Cancer Agency, Malaysia Academy of Medicine of Malaysia; SIOG, Society of Geriatric Oncology; NZGG, New Zealand Guidelines Group; D.A.C.H., German, Australia, Swiss Societies of Senology, St Gallen St Gallen Consensus; ESO, European School of Oncology, Biedenkopf the Biedenkopf expert panel members.

DISCUSSION

Importance of CPGs/consensus statements and our major findings

The practice of breast cancer surgery varies in clinical practice. The underlying reasons for this variation may be multifactorial such as patient preferences, local resources and surgeons' perspectives. CPGs/consensus statements with clear structure and presentation may help reduce the disparity in clinical practice and potentially increase the quality of care.

Because a growing number of institutions, working groups and/or governmental agencies have developed CPGs/consensus statements regarding surgical treatment, it would be helpful to know which CPGs/consensus statements are the most reliable. Therefore, assessing the quality of CPGs/consensus statements for breast cancer surgical

treatments is important and informative. In this study, we found that the ASCO,²⁸ NICE,¹⁷ SIGN,¹⁶ NZGG²⁰ and KCE²⁶ guidelines had the best overall quality, whereas the SEOM,³¹ SASK,²³ JBCS¹⁹ guidelines and the D.A.C.H.³⁷ and ESO³⁸ consensus statements had the poorest overall quality. These results were similar to those reported by Gandhi *et al*,³⁹ which was done for CPGs/consensus statements for early breast cancer systemic therapy. They found that the NICE, ASCO and NZGG guidelines had the highest overall assessment scores, whereas the SASK, SEOM guidelines and the St. Gallen consensus statement had the lowest overall assessment scores. The SASK²³ guideline had the poorest quality in both our and S. Gandhi's studies; it scored poorly in the 'Applicability' and 'Editorial Independence' domains. Low scores in the Applicability domain might suggest poor guideline implementation.

Table 3 Standardised scores (%) on the AGREE instrument assigned to included CPGs/consensus

Guidelines	Domain (%)					AGREE overall assessment		Radar map area (%)
	Scope and purpose	Stakeholder involvement	Rigour of development	Clarity and presentation	Applicability	Editorial independence	Score (%)	
Guidelines from Europe								
SIGN	81.5	94.4	89.6	92.6	66.7	66.7	77.8	67.5
NICE	87	87	87.5	85.2	73.6	63.9	88.9	65.3
KCE	87.1	87	85.4	88.9	41.7	75	72.2	58.3
SIOG	64.9	50	56.9	87	34.7	47.2	72.2	31.3
AGO	64.8	20.4	49.3	53.7	23.6	75.1	55.6	21.5
EUSOMA	68.5	40.7	26.4	72.2	23.6	44.4	50.0	20.4
ESMO	62.9	31.5	43.8	79.6	12.5	58.3	55.6	19.3
SEOM	70.4	27.8	30.6	40.7	6.9	33.3	22.2	7.6
Guidelines from North America								
ASCO	92.6	77.8	88.2	83.3	61.1	91.7	94.4	67.7
CCO	92.6	61.1	86.9	83.3	34.7	88.9	72.2	54
NCCN	61.1	75.9	79.9	81.5	51.4	66.7	72.2	48.2
SSO-ASTRO	90.7	31.5	76.4	87	12.5	75	77.8	34.6
ACR	61.1	31.4	65.9	51.9	15.3	50	55.6	20
SASK	57.4	13	14.6	68.5	13.9	8.3	44.4	5.8
Guidelines from Asian and Pacific region								
Malaysia	92.6	74.1	68.1	85.2	31.9	75	72.2	59.6
NZGG	87	66.7	77.8	88.9	56.9	88.9	77.8	49.6
CA- BRCA	81.5	68.5	54.9	77.8	47.2	88.9	77.8	47.9
CA-NBOCC	81.5	38.9	54.9	75.9	31.9	8.3	55.6	21.4
JBCS	57.4	24.1	27.1	55.6	4.2	30.6	38.9	9.4
Consensus statements								
Biedenkopf	74.1	35.2	31.3	75.9	19.4	63.7	44.4	22.6
St. Gallen	66.7	31.5	27.8	70.4	18.1	66.7	38.9	19.8
ESO	74.1	27.8	31.3	83.2	12.5	50	33.3	18.2
D.A.C.H.	77.8	25.9	36.1	57.4	8.3	41.7	44.4	15.1

●, recommend; ★, recommend with modification; ▲, not recommended.

AGO, German Group for Gynecological Oncology; ESMO, European Society for Medical Oncology; NCCN, National Comprehensive Cancer Network; JBCS, Japanese Breast Cancer Society; ASCO, American Society of Clinical Oncology; NCI, National Institute for Health and Care Excellence; CCO, Cancer Care Ontario; SIGN, Scottish Intercollegiate Guidelines Network; ACR, American College of Radiology; KCE, Belgium Health Care Knowledge Center; SSO-ASTRO, Society of Surgical Oncology-American Society for Radiation oncology; SEOM, Spanish Society of Medical Oncology; EUSOMA, European Society of Breast Cancer Specialists; CA-NBOCC, Cancer Australia-National Breast and Ovarian Cancer Center; SASK, Saskatchewan Cancer Agency; Malaysia Academy of Medicine of Malaysia; SIOG, Society of Geriatric Oncology; NZGG, New Zealand Guidelines Group; D.A.C.H., German, Australia, Swiss Societies of Senology, St Gallen St Gallen Consensus; ESO, European School of Oncology, Biedenkopf the Biedenkopf expert panel members.

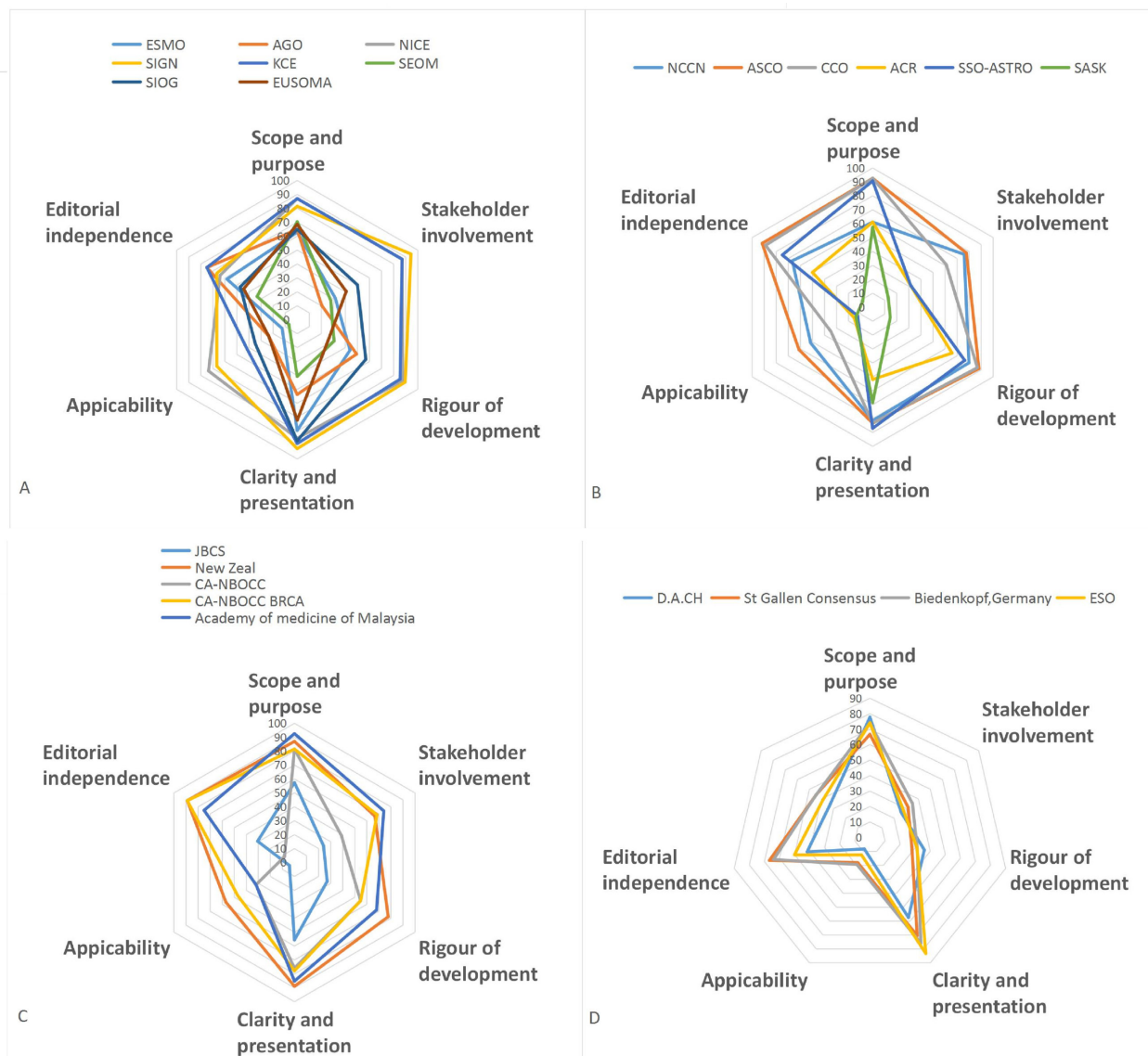


Figure 2 Radar map to show the six dimensions (domain) of the quality of CPGs developed in Europe (A), North America (B), Asian/Pacific (C) regions and of consensus statements (D). ACR, American College of Radiology; AGO, German Group for Gynaecological Oncology; ASCO, American Society of Clinical Oncology; ; CA-NBOCC, Cancer Australia-National Breast and Ovarian Cancer Centre; CCO, Cancer Care Ontario; ESMO, European Society for Medical Oncology; EUSOMA, European Society of Breast Cancer Specialists; JBCS, Japanese Breast Cancer Society; KCE, Belgium Health Care Knowledge Centre; Malaysia, Malaysia Academy of Medicine of Malaysia; NCCN, National Comprehensive Cancer Network; NICE, National Institute for Health and Care Excellence; NZGG, New Zealand Guidelines Group; SASK, Saskatchewan Cancer Agency; SEOM, Spanish Society of Medical Oncology; SIGN, Scottish Intercollegiate Guidelines Network; SIOG, Society of Geriatric Oncology; SSO-ASTRO Society of Surgical Oncology-American Society for Radiation Oncology; ESO, European School of Oncology; St. Gallen, St. Gallen Consensus; D.A.C.H, German, Australia, Swiss Societies of Senology. Biedenkopf, the Biedenkopf expert panel members.

In addition, we did not find any statement in the SASK guideline regarding conflicts of interest of the guideline development group members, which led to its low score in the Editorial Independence domain. Healthcare providers should therefore use caution when choosing which CPG/consensus statement to follow.

Update frequency

The Rigour of Development domain of the AGREE II instrument assesses whether the CPGs/consensus statements provide a procedure for updating the guideline. However, there is no recommended optimal schedule for updating CPGs/consensus statements. We found that the update frequency of CPGs/consensus statements varied. Timely updates based on newly published studies could facilitate the acceptance and implementation of these

CPGs/consensus statements. For example, the Z0011 study⁷ was published in 2011, when some controversies existed. However, the NCCN guideline incorporated the results of the Z0011 in 2012. Meanwhile, the ALND rate significantly decreased in the USA, from 71% in the pre-Z0011 era (January 2007–April 2011) to 7% in the post-Z0011 era (April 2011–February 2014).⁴⁰ The reduction in the ALND rate was also observed in studies from different countries.^{40–42} Therefore, the timely updating of the NCCN guidelines may accelerate the change of clinical practices. In contrast, the Malaysia²¹ and the NZGG²⁰ guidelines did not include any recommendations about the Z0011 trials as they have not been updated since 2011. Therefore, these two CPGs should be considered to be out of date. Physicians should use caution when adhering to these CPGs, despite them having higher scores using the AGREE II instrument.

CPGs versus consensus statements

We found that, although without statistical significance, the overall methodological quality of CPGs was better than that of consensus statements, which was consistent with Jacobs' findings.⁴³ In their study, they found that the score of the Rigour of Development domain for consensus statements was 32% lower than that of CPGs ($p < 0.0001$). The score of the Editorial Independence domain was 15% lower for consensus statements than for CPGs ($p = 0.0003$). The differences between CPGs and consensus statements may be multifactorial. First, systematic reviews are performed more frequently for CPGs than for consensus statements. Some consensus statements are based on comprehensive literature searches rather than systematic reviews. Second, most consensus statements are developed by one round of voting of panel members, whereas for CPGs, several rounds of drafting, revision and discussion, voting and peer reviews are used. Third, the authors of consensus statements may not necessarily comply with all domains of the AGREE instrument. However, despite less rigorous development of consensus statements, they are still valuable resources if they are developed in response to a recently identified issue or newly recognised gap in healthcare based on high-quality evidence, such as the optimal negative margin for DCIS patients who will receive BCS. Therefore, physicians should weigh the advantages and disadvantages of consensus statements when they apply their recommendations in clinical practice.

Domain assessment

The median scores of the Scope and Purpose and Clarity of Presentation domains for all CPGs/consensus statements were $>70\%$, suggesting that most of them had clear purposes and provided clear recommendations. The most poorly performing domain was the Applicability domain, which refers to the facilitators and barriers to guideline implementation strategies used to improve uptake and resource availability.¹¹ Poor performance of CPGs in the Applicability domain is a common problem,^{12 39 44}

reflecting that the implementation of guidelines and its barriers were not well addressed globally. To facilitate CPGs/consensus statement implementation, pilot studies and/or barrier analysis⁴⁵ may identify facilitators and barriers to implementation.^{44 46 47} Feedback from stakeholders and users could also be informative and help to improve the incorporation of CPGs/consensus statements. Furthermore, widely accepted resource-stratified CPGs/consensus statements would be helpful. In some low-income and middle-income countries where certain diagnostic tests and treatments are unavailable, CPGs/consensus statements should be able to differentiate which services are basic standard of care from those services that could provide major improvements in disease outcomes but are cost prohibitive. Although this may be difficult for some reasons, such as considerations of patient values and preferences in each country/region, costs and resource-use implications, it is possible. The NCCN Framework for Resource Stratification stratified treatment pathways into four levels based on available resources—Basic, Core, Enhanced and NCCN guidelines^{18 48}—and provided a tool to optimise treatment options given specific resource constraints. Additionally, ongoing efforts in healthcare quality improvement policy, such as the establishment of National Quality Strategy^{49–52} and the Institute for Health Improvement (<http://www.ihl.org/Pages/default.aspx>), should be recognised.

Limitations

Several limitations of this study should be addressed.

First, lack of content appraisal is one of the major limitations of our study. To comprehensively evaluate CPGs/consensus statements, we need to assess not only the strength of their development processes, structure and presentation but also the content and strength of the evidence. Therefore, gathering a panel of experts or using an instrument, such as the Grade Approach⁵³ developed by the National Guideline Clearinghouse, to evaluate the content and strength of evidence of CPGs/consensus statements should be considered in the future.

Second, the AGREE II instrument has a manual to guide reviewers on how to appraise CPGs/consensus statements, and reviewers score each item based on how much information is provided related to that item. However, reviewers cannot evaluate how much information is provided quantitatively, and scoring each item is therefore a subjective process.

Third, we only included CPGs/consensus statements published in English, so relevant non-English CPGs/consensus statements may have been missed.

Fourth, we included CPGs/consensus statements with different scopes, which may have used different approaches for development and presentation and therefore may have affected the methodological quality.

SUMMARY

Our study showed that the ASCO, NICE, SIGN, NZGG and KCE had the highest overall quality, whereas SASK, SEOM, JBCS, D.A.C.H. and ESO had the lowest overall quality. All of the CPGs/consensus statements generally had lower scores in the Applicability domain. The consensus statements generally had lower quality than CPGs. The geographic regions in which the CPGs/consensus statements were developed were not associated with methodological quality. To comprehensively assess CPGs/consensus in the future, more efforts are needed to appraise content and the frequency of updates. Additional resource-stratified CPGs/consensus statements with more applicability for implementation in clinical practice are necessary.

Contributors Each author certifies that he/she has made a direct and substantial contribution to the conception and design of the study, development of the search strategy, establishment of the inclusion and exclusion criteria, data extraction, analysis and interpretation. XL was involved in the literature search, data collection and analysis, quality appraisal and writing. FTL was involved in the literature search and writing. SYL was involved in the data collection and writing. YS conducted the quality appraisal. LLZ extracted and analysed the data. FS provided critical revision of the paper. KC was involved in the design of this study, conducted the quality appraisal and provided critical revision of the paper. SL was involved in the design of this study and provided critical revision of the paper. All authors read and provided final approval of the version to be published.

Funding This study was supported by the National Natural Science Foundation of China (grant # 81402201/81372817), National Natural Science Foundation of Guangdong Province (grant # 2014A030310070) and grant [2013] 163 from Key Laboratory of Malignant Tumour Molecular Mechanism of Guangzhou Bureau of Science and Information Technology. We appreciate the statistical advice provided by Yilong Education.

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement All the tables and figures can be accessed on BMJ Open, and all the supplementary materials can be accessed upon request via email to the corresponding authors of this study.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES

- Jemal A, Bray F, Center MM, *et al.* *Global cancer statistics*. CA: A Cancer Journal for Clinicians, 2011;61:69–90.
- Margenthaler JA, Ollila DW. Breast conservation therapy versus mastectomy: Shared decision-making strategies and overcoming decisional conflicts in your patients. *Ann Surg Oncol* 2016;23:3133–7.
- McAlister FA, van Diepen S, Padwal RS, *et al.* How evidence-based are the recommendations in evidence-based guidelines? *PLoS Med* 2007;4:e250.
- Wöckel A, Kurzeder C, Geyer V, *et al.* Effects of guideline adherence in primary breast cancer—a 5-year multi-center cohort study of 3976 patients. *Breast* 2010;19:120–7.
- DeSnyder SM, Hunt KK, Smith BD, *et al.* Assessment of Practice Patterns Following Publication of the SSO-ASTRO Consensus Guideline on Margins for Breast-Conserving Therapy in Stage I and II Invasive Breast Cancer. *Ann Surg Oncol* 2015;22:3250–6.
- Tsao MW, Cornacchi SD, Hodgson N, *et al.* A Population-Based Study of the Effects of a Regional Guideline for Completion Axillary Lymph Node Dissection on Axillary Surgery in Patients with Breast Cancer. *Ann Surg Oncol* 2016;23:3354–64.
- Giuliano AE, Hunt KK, Ballman KV, *et al.* Axillary dissection vs no axillary dissection in women with invasive breast cancer and sentinel node metastasis: a randomized clinical trial. *JAMA* 2011;305:569–75.
- Turner T, Misso M, Harris C, *et al.* Development of evidence-based clinical practice guidelines (CPGs): comparing approaches. *Implement Sci* 2008;3:45.
- AGREE Collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. *Qual Saf Health Care* 2003;12:18–23.
- Vlayen J, Aertgeerts B, Hannes K, *et al.* A systematic review of appraisal tools for clinical practice guidelines: multiple similarities and one common deficit. *Int J Qual Health Care* 2005;17:235–42.
- Brouwers MC, Kho ME, Browman GP, *et al.* AGREE II: advancing guideline development, reporting and evaluation in health care. *J Clin Epidemiol* 2010;63:1308–11.
- Nagler EV, Vanmassenhove J, van der Veer SN, *et al.* Diagnosis and treatment of hyponatremia: a systematic review of clinical practice guidelines and consensus statements. *BMC Med* 2014;12:1.
- JiWon Jane S, Lohr KN. Introducing the National Guideline Clearinghouse Revised Inclusion Criteria. <https://guideline.gov/expert/expert-commentary/46924/>
- Khan SA. Primary tumor resection in stage IV breast cancer: consistent benefit, or consistent bias? *Ann Surg Oncol* 2007;14:3285–7.
- NICE guideline. Advanced breast cancer overview [web page]. 2014 <http://pathways.nice.org.uk/pathways/advanced-breast-cancer#>
- Heys SD AA, *et al.* Treatment of primary breast cancer Scotland United Kingdom: Scottish intercollegiate guideline network. 2013 www.sign.ac.uk
- Yarnold J. Early and locally advanced breast cancer: diagnosis and treatment National Institute for Health and Clinical Excellence guideline 2009. *Clin Oncol* 2009;21:159–60.
- Gradishar W, Salerno KE. NCCN Guidelines Update: Breast Cancer. *J Natl Compr Canc Netw* 2016;14(5 Suppl):641–4.
- Komoike Y, Inokuchi M, Itoh T, *et al.* Japan Breast Cancer Society clinical practice guideline for surgical treatment of breast cancer. *Breast Cancer* 2015;22:37–48.
- Management of early breast cancer. 2009. www.nzgg.org.nz
- Har YC. Management of breast cancer. 2010. www.acadmed.org.my
- Recommendations for staging and managing the axilla in early breast cancer, 2011. www.canceraustralia.gov.au
- Breast cancer treatment guideline, 2012. www.saskcancer.ca
- Biganzoli L, Wildiers H, Oakman C, *et al.* Management of elderly patients with breast cancer: updated recommendations of the International Society of Geriatric Oncology (SIOG) and European Society of Breast Cancer Specialists (EUSOMA). *Lancet Oncol* 2012;13:e148–60.
- Cardoso F, Loibl S, Pagani O, *et al.* The European Society of Breast Cancer Specialists recommendations for the management of young women with breast cancer. *Eur J Cancer* 2012;48:3355–77.
- Wildiers H SS. Breast cancer in women: diagnosis, treatment and follow-up: Belgian Health Care Knowledge Center, 2013. www.kce.fgov.be/content/
- Kirk JEJ, *et al.* Recommendations for the management of early breast cancer in women with identified BRCA1 or BRCA2 gene mutation or at high risk of a gene mutation, 2014. www.canceraustralia.gov.au
- Lyman GH, Temin S, Edge SB, *et al.* Sentinel lymph node biopsy for patients with early-stage breast cancer: American Society of Clinical Oncology clinical practice guideline update. *J Clin Oncol* 2014;32:1365–83.
- Moran MS, Schnitt SJ, Giuliano AE, *et al.* Society of Surgical Oncology-American Society for Radiation Oncology consensus guideline on margins for breast-conserving surgery with whole-breast irradiation in stages I and II invasive breast cancer. *J Clin Oncol* 2014;32:1507–15.
- Brackstone M, Fletcher GG, Dayes IS, *et al.* Locoregional therapy of locally advanced breast cancer: a clinical practice guideline. *Curr Oncol* 2015;22:54–66.
- Garcia-Saenz JA, Bermejo B, Estevez LG, *et al.* SEOM clinical guidelines in early-stage breast cancer 2015. *Clin Transl Oncol* 2015;17:939–45.
- Hanf V, Schütz F, Liedtke C, *et al.* AGO Recommendations for the Diagnosis and Treatment of Patients with Early Breast Cancer: Update 2015. *Breast Care* 2015;10:189–97.
- Kaufman SA, Harris EE, Bailey L, *et al.* ACR Appropriateness Criteria® Ductal Carcinoma in Situ. *Oncology* 2015;29:44660–1.

34. Senkus E, Kyriakides S, Ohno S, *et al.* Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2015;26(Suppl 5):v8–30.
35. Kaufmann M, Morrow M, von Minckwitz G, *et al.* Locoregional treatment of primary breast cancer: consensus recommendations from an International Expert Panel. *Cancer* 2010;116:1184–91.
36. Coates AS, Winer EP, Goldhirsch A, *et al.* Tailoring therapies—improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. *Ann Oncol* 2015;26:1533–46.
37. Hoffmann J, Souchon R, Lebeau A, *et al.* German, Austrian and Swiss consensus conference on the diagnosis and local treatment of the axilla in breast cancer. *Eur J Cancer* 2013;49:2277–83.
38. Partridge AH, Pagani O, Abulkhair O, *et al.* First international consensus guidelines for breast cancer in young women (BCY1). *Breast* 2014;23:209–20.
39. Gandhi S, Verma S, Ethier JL, *et al.* A systematic review and quality appraisal of international guidelines for early breast cancer systemic therapy: Are recommendations sensitive to different global resources? *Breast* 2015;24:309–17.
40. Kenny TC, Dove J, Shabahang M, *et al.* Widespread Implications of ACOSOG Z0011: Effect on Total Mastectomy Patients. *Am Surg* 2016;82:53–8.
41. Gondos A, Jansen L, Heil J, *et al.* Time trends in axilla management among early breast cancer patients: Persisting major variation in clinical practice across European centers. *Acta Oncol* 2016;55:712–9.
42. Joyce DP, Lowery AJ, McGrath-Soo LB, *et al.* Management of the axilla: has Z0011 had an impact? *Ir J Med Sci* 2016;185:145–9.
43. Jacobs C, Graham ID, Makarski J, *et al.* Clinical practice guidelines and consensus statements in oncology—an assessment of their methodological quality. *PLoS One* 2014;9:e110469.
44. Wammes BM, Blom CA, Koelen M, *et al.* Implementation research for 'evidence-based' guideline development by dietitians: a pilot study to test an instrument. *J Hum Nutr Diet* 2002;15:243–54.
45. Gifford WA, Graham ID, Davies BL. Multi-level barriers analysis to promote guideline based nursing care: a leadership strategy from home health care. *J Nurs Manag* 2013;21:762–70.
46. Pulkki K, Suvisaari J, Collinson P, *et al.* A pilot survey of the use and implementation of cardiac markers in acute coronary syndrome and heart failure across Europe. The CARdiac MARker Guideline Uptake in Europe (CARMAGUE) study. *Clin Chem Lab Med* 2009;47:227–34.
47. Jagt-van Kampen CT, Kremer LC, Verhagen AA, *et al.* Impact of a multifaceted education program on implementing a pediatric palliative care guideline: a pilot study. *BMC Med Educ* 2015;15:194.
48. al GWe. NCCN Framework for Resource Stratification of NCCN Guidelines. 2016 <https://www.nccn.org/framework/>
49. McKinney M, . First, do no harm. HHS' National Quality Strategy uses broad approach to communicate expectations to providers. *Mod Healthc* 2011;41:6–7.
50. Schroeder SD. Feds implement priorities of National Quality Strategy. *S D Med* 2011;64:261.
51. Kennedy R, Murphy J, Murphy DW. An Overview of the National Quality Strategy: Where Do Nurses Fit? *Online J Issues Nurs* 2013;18:5.
52. O'Hare AM, Armistead N, Schrag WL, *et al.* Patient-centered care: an opportunity to accomplish the "Three Aims" of the National Quality Strategy in the Medicare ESRD program. *Clin J Am Soc Nephrol* 2014;9:2189–94.
53. Calonge N. New Promise for Uniform Evidence-based Guideline Development: The GRADE Approach. <https://guideline.gov/expert/expert-commentary/16440/>