ScienceDirect
Contents lists available at sciencedirect.com
Journal homepage: www.elsevier.com/locate/jval

From Multiple Quality Indicators of Breast Cancer Care Toward Hospital Variation of a Summary Measure

Elvira L. Vos, MD, PhD, Linetta B. Koppert, MD, PhD,* Agnes Jager, MD, PhD, Marie-Jeanne T.F.D. Vrancken Peeters, MD, PhD, Sabine Siesling, PhD, Hester F. Lingsma, PhD

ABSTRACT

Objectives: To improve quality in breast cancer care, large numbers of quality indicators are collected per hospital, but benchmarking remains complex. We aimed to assess the validity of indicators, develop a textbook outcome summary measure, and compare case-mix adjusted hospital performance.

Methods: From a nationwide population-based registry, all 79 690 nonmetastatic breast cancer patients surgically treated between 2011 and 2016 in 91 hospitals in The Netherlands were included. Twenty-one indicators were calculated and their construct validity tested by Spearman's rho. Between-hospital variation was expressed by interquartile range (IQR), and all valid indicators were included in the summary measure. Standardized scores (observed/expected based on case mix) were calculated as above (>100) or below (<100) expected. The textbook outcome was presented as a continuous and all-or-none score.

Results: The size of between-hospital variation varied between indicators. Sixteen (76%) of 21 quality indicators showed construct validity, and 13 were included in the summary measure after excluding redundant indicators that showed collinearity with others owing to strong construct validity. The median all-or-none textbook outcome score was 49% (IQR 42%-54%) before and 49% (IQR 48%-51%) after case-mix adjustment. From the total of 91 hospitals, 3 hospitals were positive (3%) and 9 (10%) were negative outliers.

Conclusions: The textbook outcome summary measure showed discriminative ability when hospital performance was presented as an all-or-none score. Although indicator scores and outlier hospitals should always be interpreted cautiously, the summary measure presented here has the potential to improve Dutch breast cancer quality indicator efforts and could be implemented to further test its validity, feasibility, and usefulness.

Keywords: breast cancer, quality of care, quality indicators, hospital variation.

VALUE HEALTH. 2020; ■(■):■–■

Introduction

There is a growing demand for publishing information on quality of care to drive quality improvement.[1] Monitoring and publishing quality indicator information has previously led to quality improvement.[2–4] Further quality improvement is expected from comparing and ranking hospitals.[5,6]. Nevertheless, hospitals can score high on one indicator but low on another indicator, resulting in a complex web of information and challenges for benchmarking. The case mix may also influence variation seen in indicator scores between hospitals.[7–10] Interpretation and usability of hospital performance data are therefore difficult, and the impact on healthcare providers and consumers is limited.[11–15] Studies revealed that patients prefer a summary measure to gain insight into the performance of a hospital.[16–18] The same can be expected from other parties making use of performance data as healthcare providers, health insurance companies, and governmental agencies. We believe that a summary measure describing quality with one value has the potential to increase the understanding and impact of hospital performance data.

One important prerequisite for quality indicators is their validity. Validity means that indicators measure what they claim to measure.[19] One type of validity is construct validity, which evaluates the relation between indicators that measure the same underlying concept.[20] Indicators lacking construct validity or indicators that are redundant may be excluded from the summary measure. A commonly used approach to construct a summary measure is the textbook outcome, representing patients in whom optimal (ie, textbook) health outcomes are realized.[21] The textbook outcome can be presented as a continuous score (ie, the average number of indicators with a positive result) or an all-or-none score (ie, the number

* Address correspondence to: Linetta B. Koppert, Department of Surgery, Erasmus MC Cancer Institute, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands. Email: l.koppert@erasmusmc.nl

https://doi.org/10.1016/j.jval.2020.05.011

of patients with a positive result for all indicators) per hospital.[22]

Breast cancer care quality improvements efforts in The Netherlands are led by the National Breast cancer working group Netherlands (NABON) Breast Cancer Audit (NBCA). Since 2011, a set of structure, process, and outcome indicators are defined and regularly updated by multidisciplinary group consensus.[23] The NBCA provides regular feedback on the quality indicator scores to the individual hospitals. We aimed to develop a textbook outcome type of summary measure by construct validity testing of the NBCA set of breast cancer process and outcome quality indicators. Then, hospital performance was assessed by comparing and ranking them based on their case mix–adjusted textbook outcome standardized score and presented by both the continuous and all-or-none approach.

Methods

Data

From the NBCA, patient-level data were retrieved of all patients with primary invasive breast cancer or ductal carcinoma in situ (DCIS) who were surgically treated and diagnosed in all 91 hospitals between January 2011 and August 2016 in The Netherlands. In this period, 9 hospitals fused, resulting in a different total number of hospitals in each year. Hospitals choose to register the data themselves directly into a web-based system (20%-30% of all hospitals) provided by Dutch Institute for Clinical Auditing or have it registered by the Netherlands Comprehensive Cancer Organisation as part of the Netherlands Cancer Registry (NCR). All new malignancies have been registered in the NCR on a national level since 1989. All hospitals can review the data for inconsistencies. A third party anonymized all data before it was made available for this study. The data included the following: sex, age, World Health Organization performance status, method of tumor detection, palpability, type of surgery/surgeries, multifocality, histology, tumor size in millimeter, Bloom and Richardson differentiation grade, hormone and Her2neu receptor status, clinical and pathological tumor node metastasis staging,[24] radiation treatment, chemotherapy, and hormonal therapy.

Quality Indicators

The NBCA quality indicator set evolved over the years. Even though all indicators used between 2011 and 2017 were considered, only those generally believed to represent quality of care were studied.[23] For example, the indicator percentage of patients with magnetic resonance imaging (MRI) was not suitable, because a higher or lower score is not associated with better quality of care. For each patient, an indicator could be scored as positive or negative. For each hospital, indicator scores were calculated by dividing the numerator (ie, the number of patients with a positive score) by the denominator (ie, the number of eligible patients) as defined by the publicly available NBCA manual (Table 1).[25] Patients with distant metastasis at time of diagnosis were excluded from all denominators. Indicator scores presented here may deviate from the NBCA reports.

Statistical Analysis

Whether the patient, tumor, and treatment characteristics differed between hospitals was tested by Kruskal-Wallis test. Hospital variation referred to *between*-hospital variation, and *within*-hospital variation was not studied. Indicator scores and hospital variation were presented by median, interquartile range (IQR), and range. The effect of number of events on hospital variation was evaluated by presenting results from the total cohort (2011-2016) and from 1 year of data only (2015, since 2016 included data until August).

Construct Validity

Members of NABON were consulted for their expert opinions regarding indicators that may measure the same underlying construct. The direction of association (ie, negative or positive) was defined a priori. Construct validity was tested by Spearman correlation coefficient, and 95% confidence interval (CI) was obtained by bootstrapping (1000 random replicas). Construct validity was considered present if statistically significant in the expected direction of association. A Spearman's rho <0.40 was defined as a weak correlation, 0.40 to 0.59 was described as moderate, and >60 was strong.[26]

Textbook Outcome

Construct valid indicators were included in the summary measure textbook outcome. If a patient scored positive on all indicators, a textbook outcome was achieved. For most indicators, the denominator is different and therefore a patient did score positive if this patient was not included in the denominator of the indicator of interest. For example, patients with DCIS only were not included in the denominator of irradical breast-conserving surgery (BCS) in invasive disease, but did score positive. This enabled building a single textbook outcome summary measure, otherwise hospitals could not be allocated 1 summary score for all their different types of breast cancer patients with different treatments. Textbook outcome was presented as a continuous score (ie, the median number of indicators with a positive score) and as an all-or-none score (ie, the percentage of patients scoring positive on all indicators thus achieving textbook outcome).

Case Mix

Case-mix adjustment was performed by multivariable linear regression analysis for the continuous textbook outcome and by logistic regression analysis for the all-or-none textbook outcome. All general patient and tumor characteristics were used as case-mix factors: age, histology, pathological tumor and node stage, differentiation grade, multifocality, and estrogen and Her2neu receptor status. For each hospital, the standardized rate and 95% CI for textbook outcome was calculated by dividing the observed score by the expected score. The expected score was the mean from all hospitals for the unadjusted model and the predicted probability for an individual hospital for the case-mix adjusted model. A standardized rate larger than 100 means more textbook outcomes (ie, better-achieving hospital) and a standardized rate smaller than 100 means less textbook outcomes (ie, poorer-achieving hospital) compared to the average or expected. The standard error of the standardized rate was calculated by dividing the standardized rate by the root of the number of events.

Statistical analysis was performed by IBM SPSS Statistics version 22.0 (IBM, Armonk, New York) and R version 3.0.1 (R Foundation of Statistical Computing, Vienna, Austria). All P values were 2-sided, and P values of $<.05$ were considered statistically significant. For this type of study, approval from a medical ethical committee was not required.

Results

A total of 79 690 patients had invasive breast cancer or DCIS and were surgically treated in 91 different hospitals between January 2011 and August 2016 in The Netherlands (see Appendix Table A in Supplemental Materials found at https://doi.org/10.1

Table 1. Definition of NBCA Quality Indicators 2011-2017.

QI	Short name	Definition
	Radiology	
1	BI-RADS classification	Numerator: number of patients with BI-RADS category reported in diagnostic phase on mammography, ultrasound, or breast MRI. Denominator: number of patients surgically treated for invasive breast cancer or DCIS.
2	MRI in neoadjuvant chemotherapy	Numerator: number of patients with breast MRI prior to start of neoadjuvant chemotherapy. Denominator: number of patients with invasive breast cancer treated with neoadjuvant chemotherapy.
	Pathology	
3	Full pathology report as defined	Numerator: number of patients with standard pathology report including information about ER%, PR%, HER2, grade, tumor size, resection margin, and number of positive lymph nodes. Denominator: number of patients with a pathology report of invasive breast cancer of at least 1 cm without neoadjuvant therapy.
	Surgery/plastic surgery	
4	Irradical BCS in invasive disease	Numerator: number of patients with more than focally positive margins* after first BCS. Denominator: number of patients treated with BCS for invasive non-metastasized breast cancer and without neoadjuvant chemotherapy.
5	Irradical BCS in DCIS	Numerator: number of patients with positive margins after first BCS. Denominator: number of patients treated with BCS for DCIS.
6	Reexcision after BCS for invasive disease	Numerator: number of patients with reexcision. Denominator: number of patients with BCS for invasive non-metastasized breast cancer without neoadjuvant chemotherapy.
7	Reexcision after BCS for DCIS	Numerator: number of patients with reexcision. Denominator: number of patients with BCS for DCIS.
8	Breast contour–preserving treatment	Numerator: number of patients with (1) breast-conserving surgery including re-lumpectomies without neoadjuvant chemotherapy, (2) neoadjuvant chemotherapy, and (3) mastectomy with direct breast reconstruction. Denominator: number of patients with invasive non-metastasized breast cancer with and without neoadjuvant chemotherapy.
9	Immediate breast reconstruction in DCIS	Numerator: number of patients with immediate breast reconstruction. Denominator: number of patients with a primary mastectomy for DCIS.
10	Immediate breast reconstruction in invasive disease	Numerator: number of patients with immediate breast reconstruction. Denominator: number of patients with a primary mastectomy for invasive breast cancer.

continued on next page

Table 1. Continued

QI	Short name	Definition
	Radiotherapy	
11	Seen by radiation oncologist prior to neoadjuvant chemotherapy	Numerator: number of patients seen by radiation oncologist prior to neoadjuvant chemotherapy. Denominator: number of patients with invasive breast cancer treated with neoadjuvant chemotherapy, surgery, and postoperative radiotherapy.
12	Radiotherapy for locally advanced	Numerator: number of patients treated with radiotherapy. Denominator: number of patients with invasive non-metastasized locally advanced[†] breast cancer and treated with mastectomy.
	General	
13	Preoperative MDT meeting	Numerator: number of patients for whom the information in the registry is complete and discussed in a preoperative MDT meeting. Denominator: number of surgically treated patients with primary invasive breast cancer or DCIS.
14	Postoperative MDT meeting	Numerator: number of patients for whom the information in the registry is complete and discussed in a postoperative MDT meeting. Denominator: number of surgically treated patients with primary invasive breast cancer or DCIS.
15	Neoadjuvant chemotherapy within 5 weeks of diagnosis	Numerator: number of patients receiving neoadjuvant chemotherapy within ≤5 weeks after diagnosis. Denominator: number of patients with neoadjuvant chemotherapy for invasive non-metastasized breast cancer.
16	Surgery within 5 weeks (without reconstruction) of diagnosis	Numerator: number of patients receiving surgery within 5 weeks of diagnosis. Denominator: number of patients with primary surgery without immediate breast reconstruction for invasive non-metastasized breast cancer or DCIS and without neoadjuvant chemotherapy.
17	Surgery with breast reconstruction within 5 weeks of diagnosis	Numerator: number of patients receiving surgery within 5 weeks of diagnosis. Denominator: number of patients with primary surgery with breast reconstruction[‡] for invasive non-metastasized breast cancer or DCIS and without neoadjuvant chemotherapy.
18	Radiotherapy within 5 weeks of final operation	Numerator: number of patients receiving radiotherapy within ≤5 weeks after surgery. Denominator: number of patients with invasive non-metastasized breast cancer or DCIS treated with surgery and radiotherapy (without chemotherapy between the 2 treatments).
19	Radiotherapy within 5 weeks of last chemotherapy	Numerator: number of patients receiving radiotherapy within ≤5 weeks after chemotherapy. Denominator: number of patients with invasive non-metastasized breast cancer with chemotherapy and radiotherapy.

continued on next page

Table 1. Continued

QI	Short name	Definition
20	Chemotherapy within 5 weeks of final operation	Numerator: number of patients receiving chemotherapy within ≤5 weeks after surgery. Denominator: number of patients with invasive non-metastasized breast cancer with surgery and chemotherapy (without radiotherapy between the 2 treatments).
21	Chemotherapy within 5 weeks of last radiotherapy	Numerator: number of patients receiving chemotherapy within ≤5 weeks after radiotherapy. Denominator: number of patients with invasive non-metastasized breast cancer with radiotherapy and chemotherapy.

BCS indicates breast-conserving surgery; BI-RADS, Breast Imaging-Reporting and Data System; DCIS, ductal carcinoma in situ; MDT, multidisciplinary team; MRI, magnetic resonance imaging; QI, quality indicator.
*More than focally positive margins is defined as tumor touching the inked margin over a length of 4 mm or more.
†Clinical T3, T4, any N, M0 and T, N2-3, M0 with ≥cT3, or ≥pT2 (except for pT3N0).
‡Including both primary and secondary mastectomies and both prosthesis and autologous breast reconstruction.

016/j.jval.2020.05.011). The subcohort of 2015 consisted of 15 101 patients who were treated in 82 different hospitals. Variation between hospitals was present for all 21 quality indicators (Fig. 1; Appendix Table B in Supplemental Materials found at https://doi.org/10.1016/j.jval.2020.05.011). Nevertheless, the magnitude of the variation was influenced by the number of patients in the following indicators: the IQR of between-hospital variation for irradical BCS in invasive disease was 2.1 to 4.2 in the total cohort and 0.8 to 4.6 in the subcohort; for irradical BCS in DCIS it was 16 to 24 and 11 to 29, respectively, for reexcision after BCS for invasive disease 5.0 to 8.9 and 4.2 to 9.0, for reexcision after BCS for DCIS 11 to 19 and 5 to 22, and for chemotherapy within 5 weeks of final operation 54 to 83 and 50 to 100, respectively.

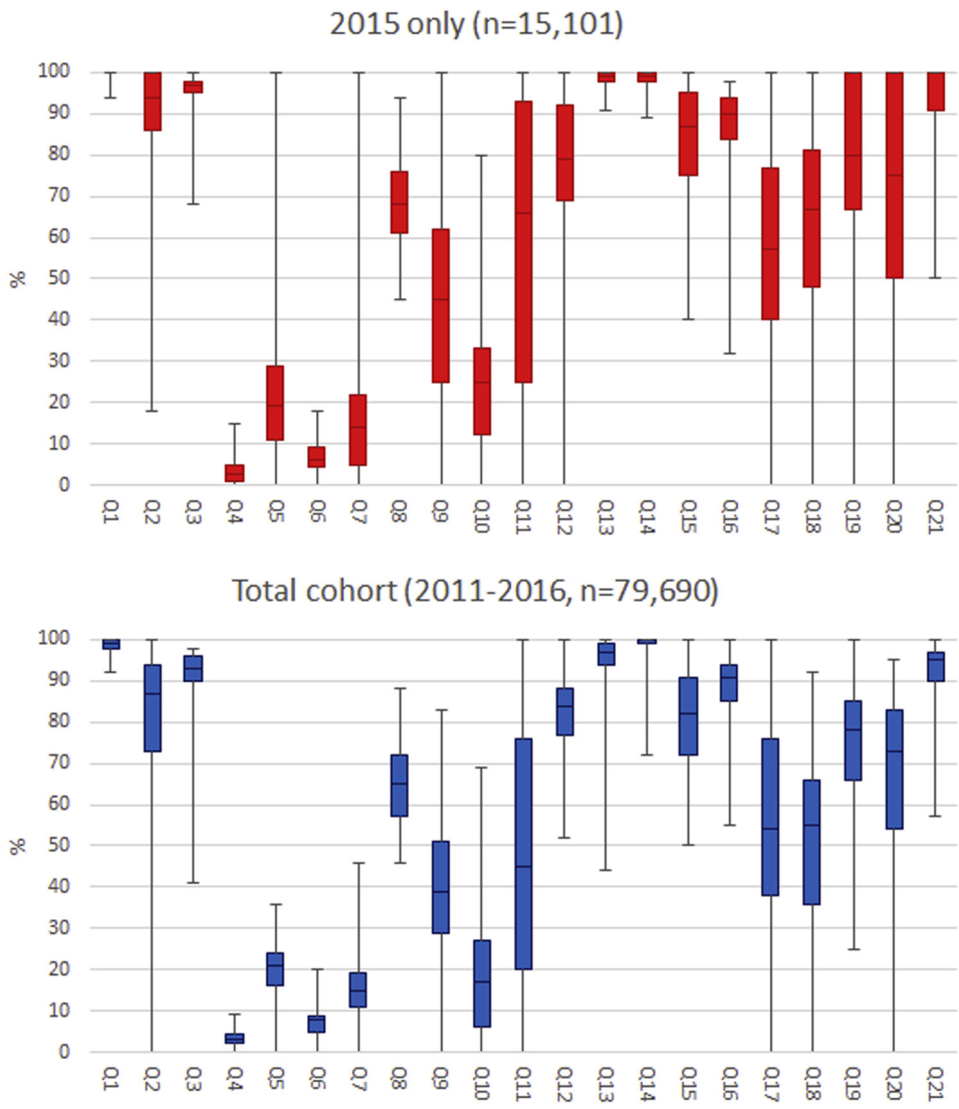
Construct Validity

A total of 14 correlations were hypothesized to measure the same underlying construct in the set of 21 indicators, and some indicators were used in more than 1 correlation. Nine correlations were found to be significant in the subcohort of 2015 and 11 correlations in the total cohort of 2011 to 2016 representing 16 (76%) of the indicators (Fig. 2; Appendix Table C in Supplemental Materials found at https://doi.org/10.1016/j.jval.2020.05.011). A strong correlation was found between irradical BCS in invasive disease and reexcision after BCS for invasive disease (Spearman's rho 0.60, P <.001), breast contour–preserving treatment and immediate breast reconstruction in invasive disease (Spearman's rho 0.62, P<.001), and immediate breast reconstruction in invasive disease and immediate breast reconstruction in DCIS (Spearman's rho 0.77, P <.001). A moderate correlation was found between preoperative multidisciplinary team (MDT) meeting and postoperative MDT meeting (Spearman's rho 0.48, P<.001), irradical BCS in invasive disease and irradical BCS in DCIS (Spearman's rho 0.44, P<.001), irradical BCS in DCIS and reexcision after BCS in DCIS (Spearman's rho 0.52, P<.001), and surgery (without reconstruction) within 5 weeks of diagnosis and surgery with breast reconstruction within 5 weeks of diagnosis (Spearman's rho 0.51, P<.001). A weak correlation was found between the full pathology report and Breast Imaging-Reporting and Data System classification (Spearman's rho 0.31, P=.002), being seen by a radiation oncologist prior to neoadjuvant chemotherapy, and MRI in neoadjuvant chemotherapy (Spearman's rho 0.28, P=.007), reexcision after BCS in invasive disease and reexcision after BCS in DCIS (Spearman's rho 0.37, P<.001), neoadjuvant chemotherapy within 5 weeks of diagnosis and surgery within 5 weeks (without reconstruction) of diagnosis (Spearman's rho 0.22, P=.040). No correlation was found between irradicality in invasive disease and breast contour–preserving treatment, radiotherapy within 5 weeks of final operation and radiotherapy within 5 weeks of last chemotherapy, and chemotherapy within 5 weeks of final operation and chemotherapy within 5 weeks of last radiotherapy.

Textbook Outcome

All indicators showing weak or moderate construct validity were included in the summary measure. From the indicators with strong construct validity, 1 of the 2 was included. Indicators lacking construct validity were excluded. The following 13 indicators were included in the summary measure textbook outcome: Breast Imaging-Reporting and Data System classification, MRI in neoadjuvant chemotherapy, full pathology report, irradical BCS in invasive disease, irradical BCS in DCIS, breast contour–preserving treatment, seen by radiation oncologist prior to neoadjuvant chemotherapy, radiotherapy for locally advanced, preoperative MDT meeting, postoperative MDT meeting, neoadjuvant chemotherapy within 5 weeks of diagnosis, surgery (without reconstruction) within 5 weeks of diagnosis, and surgery with breast reconstruction within 5 weeks of diagnosis (see Appendix Table D in Supplemental Materials found at https://doi.org/10.1016/j.jval.2020.05.011). Textbook outcome was achieved in 38 956 (48.9%) patients; 29 142 (36.6%) patients scored positive on 12 indicators, 9313 (11.7%) on 11, 1979 (2.5%) on 10 276 (0.3%) on 9, 21 patients on 8 indicators, and 3 patients scored positive on 7 indicators. Hospital volume was not correlated with achieving textbook outcome (Pearson correlation of 0.05 and P=.404). The median (IQR) of the continuous textbook outcome score was 12.3 (12.2-12.4) before and 12.3 (12.3-12.3) after case-mix adjustment. The median (IQR) all-or-none textbook outcome score was 49% (42%-54%) before and 49% (48%-51%) after case-mix adjustment. Besides a reduction in hospital variation, the individual hospital score increased or decreased after case-mix adjustment (Fig. 3). For the continuous textbook outcome, the standardized rate for the individual hospital scores ranged between 94 and 102. After case-mix adjustment, the score increased with a range between 0.1 and 2.1 points in 13 hospitals, the score did not change in 14 hospitals, and the score decreased with a range between 0.1 and

Figure 1. Between-hospital variation in quality indicator scores.



0.6 points in 64 hospitals. For the all-or-none textbook outcome, the standardized rate for the individual hospital scores ranged between 47 and 124. After case-mix adjustment, the score increased with a range between 0.1 and 2.1 points in 5 hospitals, the score did not change in 84 hospitals, and the score decreased with 0.1 point in 2 hospitals.

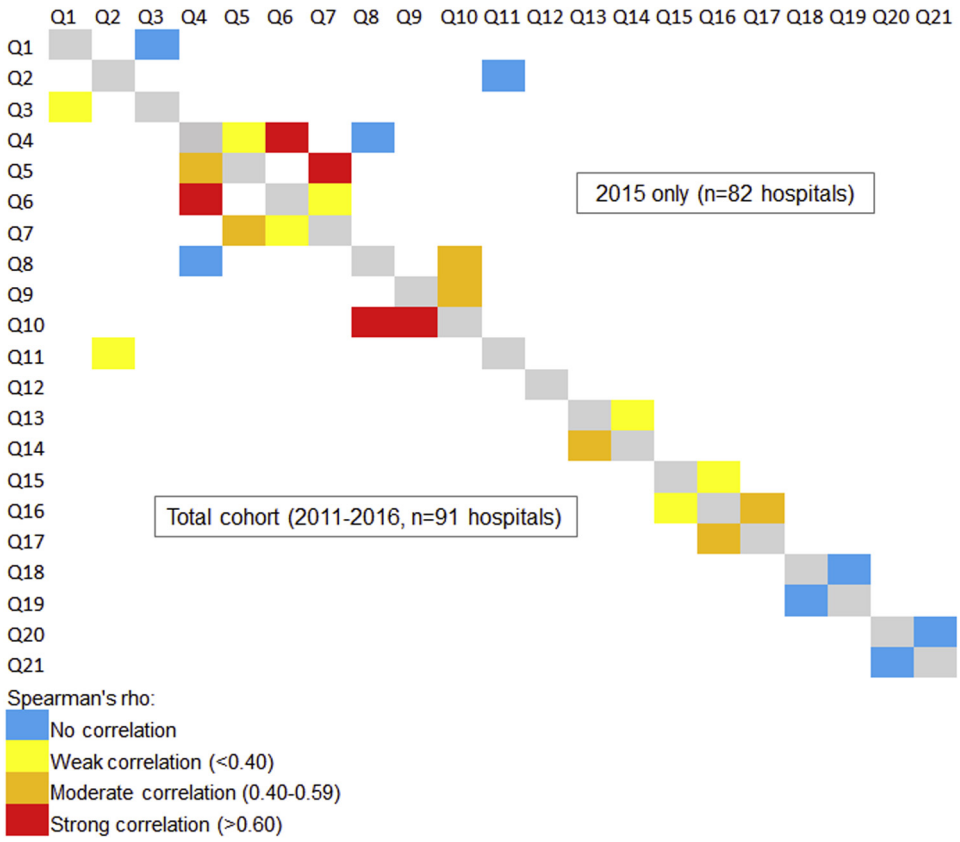
After ranking hospitals based on their continuous textbook outcome score, no outliers were identified before and after case-mix adjustment (Fig. 4). By the all-or-none textbook outcome approach, 3 (3.3%) hospitals were identified as positive outliers and 9 (9.9%) hospitals were identified as negative outliers both before and after case-mix adjustment, meaning that these hospitals had a statistically significantly higher and lower adjusted rate of textbook outcomes compared to the expected average.

Discussion

Sixteen (76%) of 21 Dutch breast cancer process and outcome quality indicators showed construct validity. Thirteen indicators were included in the textbook outcome summary measure that was achieved in 38 956 (49%) patients. Case mix–adjusted hospital variation in textbook outcome among 91 hospitals was present, and 9 (9.8%) negative and 3 (3.3%) positive outlier hospitals could be identified by the all-or-none method.

The strength of association between indicators showing construct validity was variable. For indicators with a strong association, monitoring and reporting both indicators is superfluous. For example, irradical BCS and reexcision after BCS were strongly associated and monitoring both indicators is minimally informative. Redundancy is under discussion for indicators with a moderate or weak association; they may provide complementary information.[27] Nevertheless, irradical BCS in invasive disease and irradical BCS in DCIS were weakly to moderately associated, but it can also be concluded that the lack of a strong correlation means that achieving radical margins cannot be explained by surgical performance only and other patient- or tumor-related factors are of influence. This may be a reason to exclude these indicators from external hospital performance reports. The composite indicator breast contour–sparing surgery was designed to replace the indicator irradical BCS that is believed to falsely stimulate performing mastectomy to keep irradicality rates low. Nevertheless,

Figure 2. Construct validity of hypothesized correlations between indicators.



testing construct validity showed that breast contour–sparing surgery is driven by immediate breast reconstruction rates and merely influenced by irradical BCS rates. A lack of construct validity could result from low numbers of events. For example, the correlation between being seen by a radiation oncologist prior to neoadjuvant chemotherapy and MRI prior to neoadjuvant chemotherapy was not significant in the subcohort of 2015, but was significant in the total cohort of 2011 to 2016. A lack of construct validity could also imply that the indicators are simply not valid and do not measure what they intend to measure. We suggest using these indicators for internal purposes only. In contrast to our findings, construct validity for quality indicators in hip replacement and colorectal cancer were limited.[27,28]

Our textbook outcome summary measure was achieved in 49% of patients by the all-or-none approach and a median 12.3 of 13 indicators were achieved by the continuous approach. The all-or-

Figure 3. The effect of case-mix adjustment on the standardized rate of textbook outcome in 91 hospitals (2011-2016) on the x-axis: unadjusted standardized rate and on the y-axis: case mix–adjusted standardized rate. (A) Continuous textbook outcome. (B) All-or-none textbook outcome.

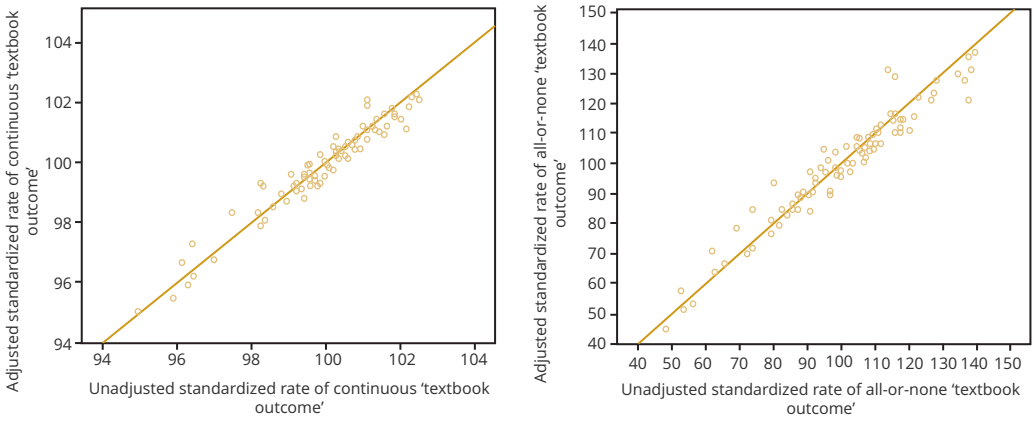
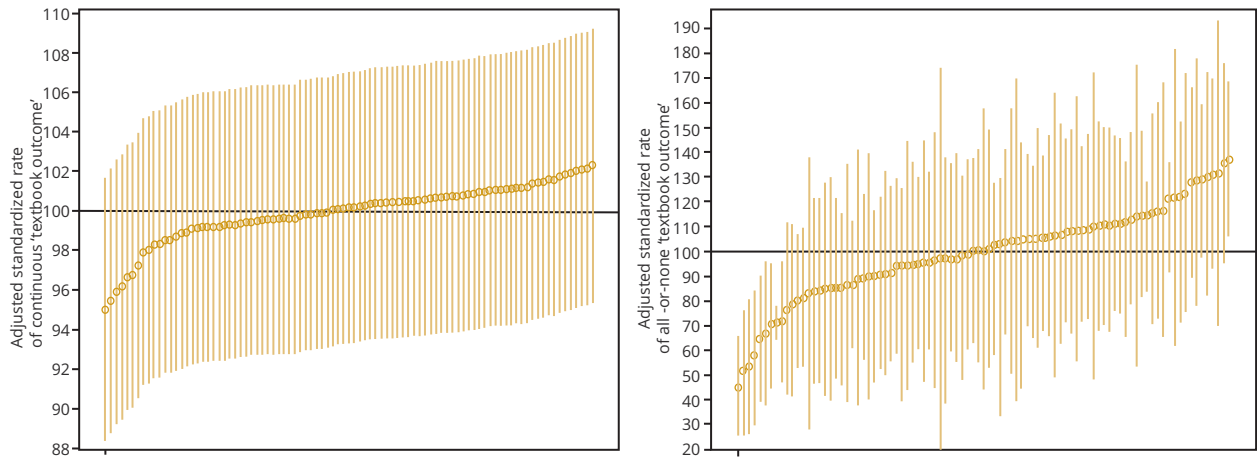


Figure 4. The case mix–adjusted standardized rate of textbook outcomes (on the y-axis) with a reference line at the expected mean of 100. The hospitals are sorted on the x-axis in order of textbook outcome score with 95% confidence interval. (A) Continuous textbook outcome. (B) All-or-none textbook outcome.



none method sets a higher benchmark and has better discriminative ability,[22,29] but results in lower scores.[18,22] Textbook outcome, using the all-or-none approach, has also been studied in esophageal cancer, gastric cancer, and cholangiocarcinoma, and was achieved in 23% to 45% of patients.[30–32] The average all-or-none textbook outcome score for colon cancer patients in the Dutch Surgical Colorectal Audit was found to be 49%, likewise here, and they identified 8 negative outlier hospitals as compared to 9 here.[21] In contrast to our study, these textbook outcomes included only 6 to 10 individual quality indicators, which were applicable to all patients. We found that this is not achievable for a breast cancer study population because of the more individualized treatments. Case-mix adjustment reduced between-hospital variation from an interquartile range of 42% to 54% to 48% to 51% without changing the median all-or-none textbook outcome score of 49%. More importantly, the ranking position of individual hospitals changed, emphasizing the importance of case-mix adjustment. The hospital ranking by continuous score was comparable to the hospital ranking by all-or-none score, but outlier hospitals could only be identified by the all-or-none score. The larger between-hospital variation and its ability to identify outlier hospitals makes the all-or-none approach the preferable method for quality improvement efforts.

A disadvantage of the proposed summary measure is that not all indicators apply to all breast cancer patients. For example, only 8 of 13 indicators applied to a patient with purely DCIS, and positive scores were given for the remaining 5 indicators. As a result, a high score does not necessarily mean better quality of care. For example, a hospital performing more mastectomies over BCS does receive points for radical BCS in invasive disease and radical BCS in DCIS and can still achieve a high textbook outcome score. Presenting a separate summary score for each subgroup of breast cancer patients is not possible, because almost each indicator has a different denominator. Also, the influence of each indicator on quality of care is not equal and weights should be added. Nevertheless, the simplicity of unweighted measures encourages implementation.[18]

An important strength of this study was the large nationwide study population with almost 80 000 patients treated in more than 90 different hospitals. The data are relatively complete, accurate, consistent, and reproducible with clear indicator definitions.[25,33,34] Each hospital recorded 99% of its total number of patients.[23] As far as we know, construct validity testing and constructing a textbook outcome summary measure has not been done before with quality indicators for breast cancer. The addition of case-mix adjustment increased the validity of our hospital comparisons and rankings.[7]

Next Steps

To enhance implementation of our findings, we suggest studying within-hospital variation of the textbook outcome to increase our understanding of the summary measure. A potential threat of validity is random variation owing to small sample sizes. Smaller hospitals can have extreme indicator scores simply owing to chance. Future research should focus on using random effect models, which can handle random variation, to estimate the performance of individual hospitals. Further, the validity of the summary indicator should be tested by relating it to clinically and patient relevant outcomes. We suggest implementing the summary measure in several hospitals to further test its validity, feasibility, and usefulness. For example, the case mix–adjusted textbook outcome could be presented in a comparative format followed by the individual indicators. Moreover, we advise excluding invalid indicators to lower the registration burden. On the other hand, the increasing focus on quality of life in breast cancer patients prompts the addition of patient-reported outcome measures.

Implications

Advantages of a summary measure include reducing the visible size of the quality indicator set without losing information and improving the communication of hospital performance to stakeholders. Stakeholders include patients, and the limited impact of comparative information on patient hospital choice can be improved.[11–18,35] Caregivers and hospital management can readily see how they perform compared to the expected or average and investigate individual indicators in more detail if necessary. Regulators, such as the NABON, can focus their quality improvement

efforts on negative outlier hospitals whereby the positive outliers can function as examples of best practices.[1]

Conclusions

Most (16 of 21) breast cancer process and outcome quality indicators used in The Netherlands showed construct validity, and suggestions were made for how to prioritize indicators. We constructed a textbook outcome summary measure including 13 indicators that was achieved in 49% of patients. Case-mix adjustment reduced the variation among 91 hospitals substantially. In contrast to the continuous approach of presenting textbook outcome performance, the all-or-none approach was superior in its discriminative ability, and 9 (10%) negative and 3 (3%) positive outlier hospitals could be identified. The summary measure presented here has the potential to improve the Dutch breast cancer quality indicator efforts and could be implemented in several hospitals to further test its validity, feasibility, and usefulness.

Supplemental Material

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jval.2020.05.011.

Article and Author Information

Published Online: xxxx

doi: https://doi.org/10.1016/j.jval.2020.05.011

Authors' Affiliations: Department of Surgery, Erasmus MC Cancer Institute, Rotterdam, The Netherlands; Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands; Department of Surgery, Netherlands Cancer Institute, Amsterdam, The Netherlands; Department of Research, Netherlands Comprehensive Cancer Organisation, Utrecht, The Netherlands; Department of Health Technology and Services Research, Technical Medical Centre, University of Twente, Enschede, The Netherlands; Department of Public Health, Erasmus MC, Rotterdam, The Netherlands

Acknowledgments: The authors thank the registration teams of the Netherlands Comprehensive Cancer Organisation for the collection of data for the NCR. The authors also thank the registrars of individual hospitals for entering patient data into the NBCA database, as well as the scientific staff of the NBCA. Because of the sensitive nature of the data collected for this study, requests to access the dataset from qualified researchers trained in human participant confidentiality protocols should be agreed upon by the Netherlands Cancer Registry and may be sent to L.B. Koppert at l.koppert@erasmusmc.nl.

This study was financially supported by the Dutch Cancer Society (Grant ID: EMCR 2015-7784).

REFERENCES

1. Hassett MJ. Quality improvement in the era of big data. *J Clin Oncol*. 2017;35:3178–3180.
2. Neuss MN, Malin JL, Chan S, et al. Measuring the improving quality of outpatient care in medical oncology practices in the United States. *J Clin Oncol*. 2013;31:1471–1477.
3. Jacobson JO, Neuss MN, McNiff KK, et al. Improvement in oncology practice performance through voluntary participation in the Quality Oncology Practice Initiative. *J Clin Oncol*. 2008;26:1893–1898.
4. Campanella P, Vukovic V, Parente P, et al. The impact of Public Reporting on clinical outcomes: a systematic review and meta-analysis. *BMC Health Serv Res*. 2016;16:296.
5. Fung V, Schmittdiel JA, Fireman B, et al. Meaningful variation in performance: a systematic literature review. *Med Care*. 2010;48:140–148.
6. Selby JV, Schmittdiel JA, Lee J, et al. Meaningful variation in performance: what does variation in quality tell us about improving quality? *Med Care*. 2010;48:133–139.
7. Fischer C, Lingsma H, Hardwick R, et al. Risk adjustment models for short-term outcomes after surgical resection for oesophagogastric cancer. *Br J Surg*. 2016;103:105–116.
8. Talsma AK, Reedijk AM, Damhuis RA, et al. Re-resection rates after breast-conserving surgery as a performance indicator: introduction of a case-mix model to allow comparison between Dutch hospitals. *Eur J Surg Oncol*. 2011;37:357–363.
9. van Dishoeck AM, Koek MB, Steyerberg EW, et al. Use of surgical-site infection rates to rank hospital performance across several types of surgery. *Br J Surg*. 2013;100:628–636. discussion 637.
10. Gooiker GA, Tollenaar RA. Performance indicators. How valid are the indicators? *BMJ*. 2010;341:c4284.
11. Marshall MN, Shekelle PG, Leatherman S, et al. The public release of performance data: what do we expect to gain? A review of the evidence. *JAMA*. 2000;283:1866–1874.
12. Prang KH, Canaway R, Bismark M, et al. Public performance reporting and hospital choice: a cross-sectional study of patients undergoing cancer surgery in the Australian private healthcare sector. *BMJ Open*. 2018;8:e020644.
13. Faber M, Bosch M, Wollersheim H, et al. Public reporting in health care: how do consumers use quality-of-care information? A systematic review. *Med Care*. 2009;47:1–8.
14. de Groot IB, Otten W, Smeets HJ, et al. Is the impact of hospital performance data greater in patients who have compared hospitals? *BMC Health Serv Res*. 2011;11:214.
15. Turnpenny A, Beadle-Brown J. Use of quality information in decision-making about health and social care services – a systematic review. *Health Soc Care Community*. 2015;23:349–361.
16. Dijs-Elsinga J, Otten W, Versluijs MM, et al. Choosing a hospital for surgery: the importance of information on quality of care. *Med Decis Making*. 2010;30:544–555.
17. Marang-van de Mheen PJ, Dijs-Elsinga J, Otten W, et al. The relative importance of quality of care information when choosing a hospital for surgical treatment: a hospital choice experiment. *Med Decis Making*. 2011;31:816–827.
18. Nolan T, Berwick DM. All-or-none measurement raises the bar on performance. *JAMA*. 2006;295:1168–1170.
19. Cronbach LJ, Meehl PE. Construct validity in psychological tests. *Psychol Bull*. 1955;52:281–302.
20. Bland JM, Altman DG. Statistics Notes: validating scales and indexes. *BMJ*. 2002;324:606–607.
21. Kolfschoten NE, Kievit J, Gooiker GA, et al. Focusing on desired outcomes of care after colon cancer resections; hospital variations in 'textbook outcome'. *Eur J Surg Oncol*. 2013;39:156–163.
22. Reeves D, Campbell SM, Adams J, et al. Combining multiple indicators of clinical quality: an evaluation of different analytic approaches. *Med Care*. 2007;45:489–496.
23. van Bommel AC, Spronk PE, Vrancken Peeters MT, et al. Clinical auditing as an instrument for quality improvement in breast cancer care in the Netherlands: the national NABON Breast Cancer Audit. *J Surg Oncol*. 2017;115:243–249.
24. Singletary SE, Allred C, Ashley P, et al. Staging system for breast cancer: revisions for the 6th edition of the AJCC Cancer Staging Manual. *Surg Clin North Am*. 2003;83:803–819.
25. Anema HA, van der Veer SN, Kievit J, et al. Influences of definition ambiguity on hospital performance indicator scores: examples from The Netherlands. *Eur J Public Health*. 2014;24:73–78.
26. Evans JD. *Straightforward Statistics for the Behavioral Sciences*. Pacific Grove, CA: Brooks/Cole Publishers; 1996.
27. Gooiker GA, Kolfschoten NE, Bastiaannet E, et al. Evaluating the validity of quality indicators for colorectal cancer care. *J Surg Oncol*. 2013;108:465–471.
28. Fischer C, Lingsma HF, Anema HA, et al. Testing the construct validity of hospital care quality indicators: a case study on hip replacement. *BMC Health Serv Res*. 2016;16:551.
29. O'Brien SM, Shahian DM, DeLong ER, et al. Quality measurement in adult cardiac surgery: part 2—statistical considerations in composite measure scoring and provider rating. *Ann Thorac Surg*. 2007;83:S13–S26.
30. van der Kaaij RT, de Rooij MV, van Coevorden F, et al. Using textbook outcome as a measure of quality of care in oesophagogastric cancer surgery. *Br J Surg*. 2018;105:561–569.
31. Levy J, Gupta V, Amirazodi E, et al. Gastrectomy case volume and textbook outcome: an analysis of the Population Registry of Esophageal and Stomach Tumours of Ontario (PRESTO). *Gastric Cancer*. 2020;23(3):391–402.
32. Merath K, Chen Q, Bagante F, et al. A multi-institutional international analysis of textbook outcomes among patients undergoing curative-intent resection of intrahepatic cholangiocarcinoma. *JAMA Surg*. 2019;154:e190571.
33. Anema HA, Kievit J, Fischer C, et al. Influences of hospital information systems, indicator data collection and computation on reported Dutch hospital performance indicator scores. *BMC Health Serv Res*. 2013;13:212.
34. Pringle M, Wilson T, Grol R. Measuring "goodness" in individuals and healthcare systems. *BMJ*. 2002;325:704–707.
35. Joint Research Centre, European Commission. *Handbook on Constructing Composite Indicators: Methodology and User Guide*. Paris, France: OECD Publishing; 2008.