Author: Wouters, M.W.J.M
Title: Measuring and improving quality of care in surgical oncology
Issue Date: 2013-05-23

Chapter 8

Exemplary Care and Outcome (ECO): a composite measure for quality assessment in cancer surgery

M.W.J.M. Wouters[1,2], N.E. Kolfschoten[2,3] , R.A.E.M. Tollenaar[2], B.P.L. Wijnhoven[4], J.Kievit.[2,3]

1 Department of Surgical Oncology, Netherlands Cancer Institute / Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands.
2 Department of Surgery, Leiden University Medical Center, Leiden, The Netherlands.
3 Department of Medical Decision Making, Leiden University Medical Center, Leiden, The Netherlands.
4 Department of Surgery, Erasmus Medical Center, Rotterdam, The Netherlands.

ABSTRACT

Background:

There is on-going interest in measuring quality in clinical practice. Though quality of care is a multi-dimensional concept, it is often assessed using singular outcomes. The purpose of this study is to provide a multi-dimensional assessment of quality, using esophageal cancer surgery as an example.

Methods:

Two methods for multi-dimensional quality assessment were tested. A relevance-weighted quality score (RWQS) and a cumulative quality profile, in which relevant quality parameters are ordered by their relevance for long-term outcome. Subsequent higher levels in the profile represent progressively more strict quality-standards; the proportion of patients meeting all standards is called the Exemplary Care and Outcome (ECO) measure. The two methods were used both unadjusted and adjusted for case-mix. Both methods were tested on outcome data from 12 hospitals that performed 1439 esophagectomies between 1991 and 2004.

Results:

No hospital scored best on more than one Observed/Expected quality score. O/E scores varied between hospitals from 0.65 to 1.05 for hospital survival, from 0.1 to 1.69 for profile-ECO, and from 0.88 to 1.03 for RWQS. Both multidimensional quality scores differed significantly between high and low volume hospitals, and between academic and non-academic hospitals, while O/E-scores for single-dimension of hospital survival did not differ significantly.

Conclusion:

Quality of care can be measured by more than one parameter only. We designed two methods of combining multiple quality parameters and a top-quality measure of ECO. Both methods seemed feasible, and results suggest that these methods may better discriminate between higher and lower quality of care.

INTRODUCTION

Ever since the publication of the Harvard Medical Practice Study,[1] and of the Institute of Medicine report "To Err is Human",[2] public attention has focused upon quality and safety in health care, or on the lack of it. The Institute of Medicine (IoM) has defined quality as a multi-dimensional concept, encompassing the dimensions effectiveness, safety, timeliness and patient centeredness. The Agency for Healthcare Research and Quality (AHRQ) described quality health care as "doing the right thing, at the right time, in the right way, for the right person, and having the best possible results". In recent years, the IOM, AHRQ and other institutions such as the Institute of Health Care Improvement, the Leapfrog Group and the UK department of Health, have pioneered initiatives on quality assessment, improvement and transparency. However, although quality is conceptually clear, there is an on-going debate on how to measure quality in clinical practice. One approach has been to use simple and readily available outcomes, such as hospital mortality or duration of hospital admission, another approach to use procedural volume as a readily available quality-proxy[3;4]. However, neither of these simplifying approaches has adequate content validity, as it does no justice to the multi-dimensional concept of quality. High quality care is safe, effective, patient-centered and cost-effective, and its good outcomes are the result of high quality (infra)structure and process[5]. Thus, as quality is a multi-dimensional concept, it should be measured as such.

For esophageal cancer surgery, quality assessment has so far focused almost exclusively on in-hospital mortality. Although this mortality has declined in recent years,[6] marked differences between institutions still exist, ranging from 2 to 10%.[7;8] In a recently published study we were able to show that concentration of esophageal resections by outcome-based referral dramatically reduced hospital mortality[9]. However, high-quality esophageal cancer care encompasses more than in-hospital survival only. Anastomotic leakage (occurring in 10-25% of the patients) and other adverse outcomes may severely affect esophageal cancer patients' quality of life[10]. In addition, treatment effectiveness (radicality of cancer resection, long-term survival) is no less important than treatment safety.

There is a need for quality frameworks that encompasses and combines different dimensions or aspects of health care quality. If such frameworks are to be used to compare the quality of care between hospitals, they should take into account differences in patient, disease- and procedure-mix between hospitals. In the present study we designed and piloted multi-dimensional quality assessment that aims at providing a more valid insight into the quality of surgical oncological care than mortality alone. In general, such studies are hampered by the absence of a 'golden standard' for quality of care. In the absence of a clear reference standard, we used hospitals' academic status as a proxy. In addition, we investigated whether multidimensional quality assessment could have potential to discriminate better between hospitals with different levels of care than mortality alone. Thus, in this study, we addressed the following three research questions:

1. In cancer surgery, which parameters can be used to assess the quality of care provided to patients?
2. How can these parameters be combined into multidimensional assessment of quality, that could provide more insight than a singular quality measure (such as mortality) alone?
3. Is it plausible that multidimensional quality assessment, after correction for case-mix, better (or less good) discriminates between hospitals with higher and lower quality of care than a singular quality measure?

To test the framework's feasibility, we used it to assess the (differences in) quality of care between hospitals that provided surgical treatment to patients with esophageal cancer between 1991 and 2004 in the Netherlands, and on which we reported one-dimensional outcome information in an earlier paper.[11]

METHODS

Patient data were derived from a database that was created to assess the quality of esophageal cancer surgery in 12 hospitals (2 academic hospitals, indicated by A1 and A2, and 10 general hospitals, G3 - G12) in the mid-western part of the Netherlands (1.7 million inhabitants). Through the 'Cancer Registry' of the Comprehensive Cancer Center Leiden we identified 1438 patients that were treated between 1990 and 2004.[11;12] Patient and disease characteristics, and information on treatment and outcome were extracted and analyzed from patient records and hospital information systems. In our earlier study, we reported on a few separate outcomes only and thus could not provide a broader assessment on the quality of care that patients experienced. The quality framework we now propose aims at addressing these limitations.

Constructing and testing multidimensional quality assessment

Design and testing of multidimensional quality assessment was done in 3 phases; being a) the choice of relevant quality dimensions and parameters, b) design of two methods of combining different parameters into multidimensional assessment (a cumulative quality profile that provides insight, and a relevance-weighted quality score that supports choice), and c) testing these assessments on the dataset described above, both without and with case-mix adjustment.

In the first phase, we used the four main quality dimensions (safety, effectiveness, efficiency and patient-centeredness) to guide the selection of appropriate quality parameters (from those that were available in our retrospective database).[13] To assess safety, 4 categories of adverse outcomes were selected from the Dutch surgical adverse outcome registry that is

Table 1. Patient, tumor, treatment, outcome and hospital characteristics

	Overall		Academic		General		
Characteristics	No. of patients	%	No. of patients	%	No. of patients	%	P value
Age							
median (yrs)	62.4		62.0		63.3		0.02
range (yrs)	28-89		28-89		32-87		
Gender							
Male	1099	76.4	775	77,9	324	72.9	0.04
Female	340	23.6	220	22.1	120	27.1	
Co-morbidity							
None	636	44.2	450	45,2	186	41.9	0.26
1 organ system	467	32.5	317	31.9	150	33.7	
2 organ systems	202	14.0	141	14.2	61	13.7	
≥ 3 organ systems	91	6.3	71	7.1	20	4.5	
Stage (pTNM)							
I	228	15.8	155	15.6	73	16.4	0.001
II	569	39.6	380	38.2	189	42.7	
III	463	32.2	313	31.4	150	33.6	
IV	179	12.4	147	14.8	32	7.3	
Esophagectomy							<0.001
Transthoracic	304	21.1	178	17.9	126	28.4	
Transhiatal	1101	76.5	787	79.1	314	70.7	
Adverse outcome							
None	509	35.4	400	40.2	109	24.5	<0.001
Grade 1	587	40.8	407	40.9	180	40.5	
Grade 2	172	12.0	97	9.7	75	16.9	
Grade 3	62	4.3	38	3.8	24	5.4	
Grade 4 (mortality)	102	7.1	53	5.3	49	11.0	
Length of stay							
mean (days)	22		20		28		<0.001
range (days)	5-273		5-173		9-273		
Radical resection	1054	73.2	729	73.3	325	73.2	0.98
1-year survival	1003	69.7	700	70.4	303	68.2	0.42
Hospital volume							
Low (<20/y)	566	39.3	122	12.3	444	100	<0.001
High (>20/y)	873	60.7	873	87.7	0	0.0	
Total no. of patients	1439		995		444		

Academic = academic hospitals; general= general hospitals; yrs = years; adverse outcome: Grade 1 = minor complications without re-intervention or permanent damage leading to hospital stay > 14 days; Grade 2 = complications needing re-intervention; Grade 3 = complications with permanent damage; Grade 4 = complications leading to in-hospital mortality

carried by most Dutch hospitals.[14] In decreasing order of severity these adverse outcomes were 'in-hospital death' (grade 4 adverse outcome), 'major complications with permanent/ long term morbidity' (grade 3 adverse outcome), 'major complications requiring re-operation' (grade 2 adverse outcome) and 'minor complications leading to delayed discharge (> 14 days) from hospital' (grade 1 adverse events). To assess effectiveness, 2 treatment goals were chosen, one short-term and one longer term: 'tumor free margins of cancer resection' (R0-resection) and '1-year survival'. No information was available on the dimensions

Table 2. The 6 quality parameters chosen and their correlation

	avoiding adverse outcomes				achieving goals	
	SurvHosp	No long-term AO	No Reop	Hosp <14 days	R0	Surv1YR
SurvHosp	1.000	-0.047	**-.089***	0.045	0.031	**.433***
No long-term AO	0.076 (ns)	1.000	**-.078***	**.138***	0.013	-0.033
No Reop	**0.001**	**0.003**	1.000	**.236***	**-.053***	0.005
Hosp<14 days	0.088 (ns)	**<0.001**	**<0.001**	1.000	0.019	**.086***
R0	0.244 (ns)	0.613 (ns)	**0.045**	0.475 (ns)	1.000	**.242***
Surv1YR	**<0.001**	0.205 (ns)	0.864 (ns)	**0.001**	**<0.001**	

pearson corr. (upper right table-half)

2-sided chi-square (lower-left table-half)

SurvHosp = no complications leading to in-hospital mortality (grade IV adverse outcome); No Long-term AO = no complications with permanent damage (grade III adverse outcome) ; No Reop = no complications needing re-intervention (grade II adverse outcome); Hosp<14 days = no minor complications leading to hospital stay > 14 days (Grade I adverse outcome); R0 = microscopically radical resection; Surv1YR = patient alive one year after resection; ns = not significant. For Pearson correlation coefficients (upper right table-half) stars (*) indicate statistical significance (at alpha 0.05). In the lower-left table-half, significance of correlations is quantified by p-value of chi-square test

efficiency and patient-centeredness. For each of these (4 + 2 =) 6 quality parameters, a favorable quality standard was defined, being the absence of adverse outcome for the first four, and achieving the treatment goal for the latter two.

In the second phase we combined the 6 quality parameters into two methods of multidimensional quality assessment. The first method is a cumulative quality profile (CQP) in which each subsequent profile-level indicates whether the quality standards of the present and all preceding levels are met (scored 1) or not (scored 0). Higher levels represent progressively more strict standards for health care quality, their order being determined by their relevance for long term outcome (using Cox regression for survival). The second method is a relevance-weighted quality score (RWQS), being the sum of products of relevance weights and parameter outcomes. Relevance weights represent the extent to which each of the 6 parameters is considered relevant for choosing a high quality hospital. For the present study, these weights were obtained by questionnaire (shown in appendix) from 18 members of the Dutch Association of Surgical Oncologists who have special expertise in (the quality of) esophageal cancer care. At hospital level, the relevance weighted quality score is quantified as the average of scores for all patients.

For both multidimensional assessments (CQP and RWQS) 0 is the lowest score per patient and stands for (total) quality failure, while a score of 1 signifies that all predefined quality standards are met, i.e. that the patient has experienced exemplary care and outcome (ECO). Intermediate quality levels are represented on the quality profile by a score of 1 on lower profile only, and using the relevance weighted quality score by values between 0 and 1.

In the third phase we tested both these multidimensional quality assessment on the database of esophageal cancer patients, and compared them with the single-dimension quality measure of hospital survival. For all hospital comparisons, case-mix correction was applied by logistic regression for the covariates age, gender, co-morbidity and cancer stage to predict patient-specific outcomes for each parameter. For patient level-analysis, a case-mix corrected O/E-score is the quotient of the observed (0 or 1) and expected (between 0 and 1) patient-outcome. For hospital level-analysis, a case-mix corrected hospital score is the quotient of the observed and expected proportion of patients in whom the desirable outcome is achieved. The hospital ECO-score is the proportion of patients in whom all predefined quality standards are met. Hospital performance was analyzed both per hospital, and for specific categories of hospitals (high versus low volume, and academic versus non-academic), using O/E-scores for survival, ECO and RWQS respectively.

Statistics

Differences in patient, tumor and treatment characteristics and outcome measurements were assessed using Kruskal-Wallis and Mann Whitney U test for continuous variables, and chi-square testing for categorical variables. Spearman rank and two-sided chi-square test were used to test correlations between outcomes. Survival was analyzed using Cox regression for the period between the date of first surgery to either death or the last patient contact, with follow up monitoring being continued until December 31st 2006. Prediction of events was calculated using multivariate logistic regression. All analyses were conducted using SPSS software (version 18.0; SPSS inc., Chicago.IL), using an alpha of 0.05 as the significance-threshold.

RESULTS

Results are reported per phase (parameters, multidimensional assessment, and hospital comparisons. Table 1 provides the characteristics of all patients.
Correlations between patient-outcomes for the 6 quality parameters are shown in table 2. Interestingly, reoperation is not only associated with increased length of stay, but also with R0-resection, hospital survival and the absence of permanent or long term morbidity. 1 Year survival is not only (unsurprisingly) associated with hospital survival, but also with R0 resection and timely discharge. Table 3 shows the proportions and O/E-scores of patients meeting the 6 separate quality standards (i.e. having a favorable outcome on a quality parameter) in each of the 12 hospitals. No hospital scored best on more than 1 case-mix corrected parameter. That quality parameters are not always positively associated is illustrated by hospital G8, that scored best on (both absolute and case-mix corrected) hospital survival (97%, 1.05)

For the first multidimensional assessment, the cumulative quality profile, the order of the various parameters (on the basis of their relevance for long term outcome, by Cox regression) is shown in table 4. The most basic parameter is hospital survival, followed by 1-year survival, R0-resection, no reoperation, timely discharge, and no permanent/long term morbidity respectively. Case-mix corrected hospital ECO-scores varied from 0.24 (G10) to 1.63 (A2) and 1.69 (A1) for the two university hospitals. Figure 1 shows the quality profiles of the 12 hospital profiles graphically, both without (1a) and with (1b) case-mix correction.

Table 3. Hospital performance on each of the 6 separate quality parameters, both uncorrected (expressed as the proportion of patients meeting the quality standard), and with correction for case-mix (expressed as observed/expected ratio for each parameter)

Hospitals		Outcome measures											
		SurvHosp		No Long-term AO		No Reop		Hosp <14 days		R0		Surv1YR	
Id	Volume	%	O/E	%	O/E	%	O/E	%	O/E	%	O/E	%	O/E
A-1	873*	94%	1.02	97%	1.01	91%	1.03	53%	1.16	72%	1.00	69%	1.01
A-2	122	93%	1.01	91%	0.95	85%	0.97	54%*	1.17*	80%	1.04	79%*	1.07
G-3	108	87%	0.94	96%	1.00	79%	0.90	48%	1.03	76%	1.01	76%	1.05
G-4	88	83%	0.90	94%	0.99	82%	0.93	27%	0.61	68%	0.94	56%	0.80
G-5	54	93%	1.01	94%	0.99	78%	0.88	15%	0.32	81%	1.03	70%	0.95
G-6	39	95%	1.02	100%*	1.04	92%	1.05	33%	0.72	67%	0.94	74%	1.08*
G-7	37	95%	1.03	92%	0.96	86%	0.98	3%	0.06	62%	0.86	70%	1.04
G-8	33	97%*	1.05*	94%	0.98	82%	0.93	36%	0.80	61%	0.83	73%	1.04
G-9	28	82%	0.89	100%*	1.05	100%*	1.11*	25%	0.52	75%	0.94	68%	0.91
G-10	25	92%	1.03	76%	0.80	80%	0.90	12%	0.30	84%	1.15	64%	0.98
G-11	21	90%	0.97	100%*	1.04	81%	0.91	19%	0.42	95%*	1.25*	71%	0.97
G-12	10	60%	0.65	100%*	1.05*	90%	1.01	40%	0.94	80%	1.05	50%	0.71

Id = hospital identification G = general hospital; A = academic hospital; SurvHosp = no complications leading to in-hospital mortality (grade IV adverse outcome); No Long-term AO = no complications with permanent damage (grade III adverse outcome) ; No Reop = no complications needing re-intervention (grade II adverse outcome); Hosp <14 days = no minor complications leading to hospital stay > 14 days (Grade 1 adverse outcome); R0 = microscopically radical resection; Surv1YR = patient alive one year after resection; * = highest score for separate measures

Table 4. Hospital performance on the cumulative quality profile, both uncorrected (expressed as the percentage of patients that meets each of the 6 progressively stricter cumulative quality levels), and with correction for case-mix (expressed as observed/expected ratio for each level).

Hospitals		Cumulative standards satisfied (in O/E)											
		SurvHosp		+ Surv1YR		+ R0		+ NoReop		+ No Long-term AO		+ Hosp <14days (= ECO)	
Id	Volume	%	O/E	%	O/E	%	O/E	%	O/E	%	O/E	%	O/E
A-1	873*	94%	1.02	69%	1.08	56%	1.21	51%	1.25	49%	1.25	31%	1.69*
A-2	122	93%	1.01	79%*	1.19*	65%	1.26	56%	1.23	50%	1.14	32%*	1.63
G-3	108	87%	0.94	76%	1.14	59%	1.21	44%	1.00	43%	1.03	27%	1.37
G-4	88	83%	0.90	56%	0.90	45%	0.97	33%	0.76	31%	0.75	10%	0.60
G-5	54	93%	1.01	70%	1.04	59%	1.12	43%	0.94	37%	0.82	9%	0.39
G-6	39	95%	1.02	74%	1.17	54%	1.18	51%	1.18	51%	1.23	21%	1.11
G-7	37	95%	1.03	70%	1.12	49%	1.09	41%	1.04	41%	1.09	3%	0.10
G-8	33	97%*	1.05*	73%	1.12	52%	0.96	39%	0.78	33%	0.70	12%	0.49
G-9	28	82%	0.89	68%	0.96	54%	0.88	54%	0.98	54%	1.03	4%	0.14
G-10	25	92%	1.02	64%	1.08	56%	1.26	44%	1.16	28%	0.80	4%	0.24
G-11	21	90%	0.96	71%	1.06	71%*	1.43*	62%*	1.37*	62%*	1.43*	14%	0.74
G-12	10	60%	0.66	50%	0.75	50%	0.96	40%	0.85	40%	0.90	10%	0.56

O/E = observed / expected ratio; Id = hospital identification G = general hospital; A = academic hospital; SurvHosp = no complications leading to in-hospital mortality (grade IV adverse outcome); No Long-term AO = no complications with permanent damage (grade III adverse outcome) ; No Reop = no complications needing re-intervention (grade II adverse outcome); Hosp <14 days = no minor complications leading to hospital stay > 14 days (Grade 1 adverse outcome); R0 = microscopically radical resection; Surv1YR = patient alive one year after resection; * = highest score for separate measures without and with casemix correction.

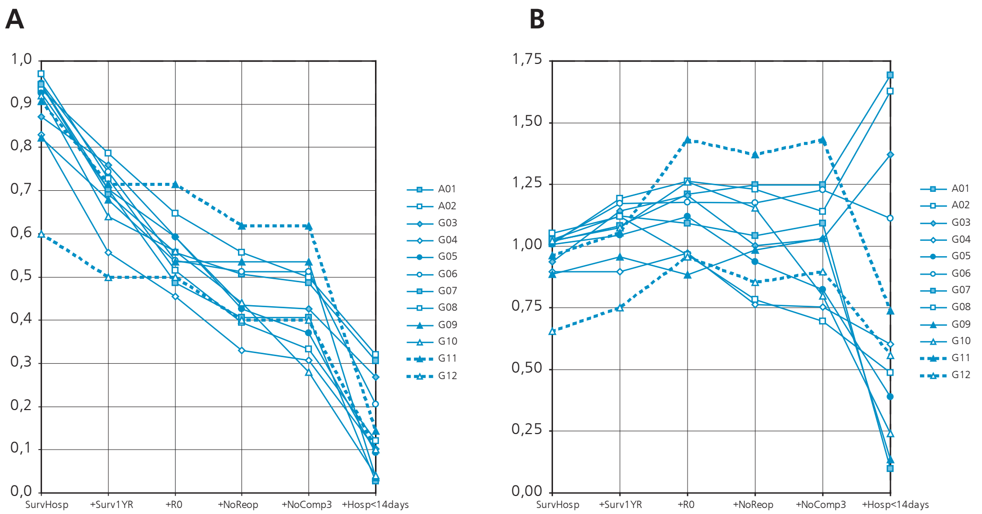

Figure 1. Cumulative outcome profile for 12 hospitals performing esophagectomy for cancer: **A.** shows the percentage of patients meeting the present and all preceding quality standards. **B.** shows the observed/ expected percentage, based on gender, age, cancer stage and number of co-morbidities. G = general hospital; A = academic hospital; SurvHosp = no in-hospital mortality; Surv1YR = 1 year survival; R0 = microscopically radical resection;NoGrade2 = no complications needing re-intervention; NoGrade3= no complications with permanent damage; NoGrade1= no minor complications leading to hospital stay > 14 days.

For the second assessment, the multidimensional weighted quality score, we used relevance weights that were obtained from the 18 surgeons-experts, and that are shown in table 5. Hospital quality scores are shown in table 6, varying from 0.88 for G12 to 1.03 for G11.

Correlations between case-mix corrected hospital quality scores on single dimension hospital survival, and on multidimensional ECO-score and RWQS were compared, and yielded modest to low correlations, varying from 0.022 (Pearson correlation between hospital scores on O/E-survival and O/E-ECO), to 0.526 (for O/E-hospital survival and O/E-RWQS) and 0.276 (for O/E-ECO and O/E-RWQS). This is not surprising, as these scores deal with failure (mortality/survival), perfection (ECO) and the whole quality range (RWQS) respectively.

Table 5. Results of assessment of relevance for quality of care by 18 surgeon-experts.

Outcome parameter	SurvHosp	R0	No Long-term AO	Surv1YR	No Reop	Hosp <14days
Mean	26.7	24.8	19.2	14.7	11.5	3.2
Stdev	7.9	8.4	5.2	8.9	7.9	2.9
Max	45	45	30	35	25	5
Min	10	10	10	0	0	0

SurvHosp = no complications leading to in-hospital mortality (grade IV adverse outcome); No Long-term AO = no complications with permanent damage (grade III adverse outcome) ; No Reop = no complications needing re-intervention (grade II adverse outcome); Hosp <14 days = no minor complications leading to hospital stay > 14 days (Grade 1 adverse outcome); R0 = microscopically radical resection; Surv1YR = patient alive one year after resection; stdev = standard deviation; max = maximum; min = minimum

Table 6. Hospital performance on the relevance-weighted quality score, both uncorrected (expressed as hospital averages for absolute quality scores), and with correction for case-mix (expressed as averages of observed/expected ratios for quality).

Hospitals		Outcome measures													
Relevance weights		26.7		24.8		19.2		14.7		11.5		3.2		100	
Id	Volume	SurvHosp		R0		No Long-term AO		Surv1YR		No Reop		Hosp <14days		Weighted total	
		O	O/E	O	O/E	O	O/E	O	O/E	O	O/E	O	O/E	O	O/E
A-1	873*	94%	1.02	72%	1.00	97%	1.01	69%	1.01	91%	1.03	53%	1.16	84%	1.02
A-2	122	93%	1.01	80%	1.04	91%	0.95	79%*	1.07	85%	0.97	54%*	1.17*	85%	1.01
G-3	108	87%	0.94	76%	1.01	96%	1.00	76%	1.05	79%	0.90	48%	1.03	82%	0.98
G-4	88	83%	0.90	68%	0.94	94%	0.99	56%	0.80	82%	0.93	27%	0.61	76%	0.91
G-5	54	93%	1.01	81%	1.03	94%	0.99	70%	0.95	78%	0.88	15%	0.32	83%	0.97
G-6	39	95%	1.02	67%	0.94	100%*	1.04	74%	1.08*	92%	1.05	33%	0.72	84%	1.01
G-7	37	95%	1.03	62%	0.86	92%	0.96	70%	1.04	86%	0.98	3%	0.06	79%	0.96
G-8	33	97%*	1.05*	61%	0.83	94%	0.98	73%	1.04	82%	0.93	36%	0.80	80%	0.97
G-9	28	82%	0.89	75%	0.94	100%*	1.05	68%	0.91	100%*	1.11*	25%	0.52	82%	0.96
G-10	25	92%	1.03	84%	1.15	76%	0.80	64%	0.98	80%	0.90	12%	0.30	79%	0.97
G-11	21	90%	0.97	95%*	1.25*	100%*	1.04	71%	0.97	81%	0.91	19%	0.42	87%*	1.03*
G-12	10	60%	0.65	80%	1.05	100%*	1.05*	50%	0.71	90%	1.01	40%	0.94	74%	0.88

Id = hospital identification G = general hospital; A = academic hospital; SurvHosp = no complications leading to in-hospital mortality (grade IV adverse outcome);No Long-term AO = no complications with permanent damage (grade III adverse outcome) ; No Reop = no complications needing re-intervention (grade II adverse outcome; Hosp <14 days = no minor complications leading to hospital stay > 14 days (Grade 1 adverse outcome); R0 = microscopically radical resection; Surv1YR = patient alive one year after resection; * = highest score for separate measures without and with casemix correction;

Table 7. Comparisons of different categories of hospitals (high versus low volume, academic versus non-academic) by single dimension patient O/E-scores for hospital survival, and by multidimensional patient O/E-scores for ECO and RWQS

Hospital categories	n (patients)	HospSurv		ECO		RWQS	
		average	stdev	average	stdev	average	stdev
High volume	873	1.020	0.255	1.692	2.970	1.021	0.247
Low Volume	565	0.968	0.335	0.911	2.131	0.970	0.242
p (Mann Whitney U)		0.114		< 0.001*		< 0.001*	
Academic	995	1.019	0.257	0.168	2.928	1.021	0.243
Non-academic	443	0.957	0.350	0.713	1.933	0.956	0.248
p (Mann Whitney U)		0.095		< 0.001*		< 0.001*	

Comparison of the quality provided to patients treated in different categories of hospitals (high versus low volume, academic versus non-academic) is shown in table 7, and shows that O/E-hospital survival scores did not differ (p=0.114 for high vs. low volume, and p=0.085 for academic vs. non-academic by Mann Whitney U-test), while O/E-ECO and O/E-RWQS did, highly significantly (p<0.001).

DISCUSSION

Our study addresses the fact that quality of care should be measured by more than one outcome only, and illustrates this point by introducing multidimensional assessment for the quality of esophageal cancer surgery. We introduce 6 quality parameters, including the traditional measure of hospital mortality, and combine these parameters into a cumulative quality profile and a relevance-weighted quality score. Both the finding that the separate quality parameters may be negatively correlated (table 2) and that no hospital scored best on all parameters, supports the relevance of such a multidimensional approach.

For both frameworks we introduce the concept of ECO, exemplary care and outcome. Per patient, this signifies the provision of care that meets all of the predefined quality standards, while at hospital level, the ECO score is the proportion of patients treated with ECO-quality. We tested the feasibility of multidimensional quality assessment on the performance of 12 hospitals providing esophageal cancer surgery to patients treated in the west of the Netherlands between 1991 and 2004. Although no hospital performed best on more than 1 parameter, overall scores of the two university hospitals (A1 and A2) and one low-volume hospital (G11) tended to have higher scores in many assessments. Finally, we clustered hospitals in categories expected to provide different levels of esophageal cancer care (high versus low volume hospitals, and academic versus non-academic) and assessed quality between these categories. The results are shown in table 7, and demonstrate that differences in case-mix corrected O/E-ECO and O/E-RWQS between hospital categories are highly significant, whereas O/E-hospital survival does not differ significantly. This demonstrates that both multidimensional methods can provide a broader assessment of quality (suggesting better construct validity) but could also be more sensitive to differences in the quality of care provided (i.e. better criterion validity)

In recent years, a plethora of articles describing variation in outcome between different institutions has been published. Most studies are population-based and assess differences in outcome between large groups of hospitals, without evaluating quality of care on the level of individual institutions. Adopting Donabedian's paradigm of structure, process and outcome, Birkmeyer et al. mentioned the relative merits of different approaches to measuring the quality of surgical care.[3] More recently, Porter made a strong case for multidimensional outcome measurement, and for taking the point of view of the care-seeker, not the care-provider.

However, comparison of health outcomes between hospitals is not without its problems, the most obvious problems being data-quality and case-mix correction[11]. In addition, for rare outcomes, there is 'the problem of small sample size'. Dimick et al. investigated the minimum hospital caseloads necessary to detect a doubling of the mortality rate for different procedures, and found that an annual volume of 77 esophageal resections is necessary to detect significant mortality differences.[15] He concluded that most operations are not performed frequently enough to use surgical mortality as an indicator of hospital quality.

Multidimensional quality assessment aims at better construct validity (by taking a greater proportion of relevant quality aspects into account) and (hopefully) at better criterion validity (being better able to discriminate between hospitals with better and lower overall quality). However, it also adds a new challenge to existing ones; the question of how to combine different quality parameters. O'Brien, analyzing quality measurement of adult cardiac surgery on behalf of the Quality Measurement Task Force of the Society of Thoracic Surgeons, tested four methods of composite scoring (1 – an opportunity-based approach, 2 - [weighted or un-weighted] averaging of item-specific estimates, 3 - all or none scoring, and 4 - latent trait analysis), and concluded that none is without flaws or limitations.[16]
In the present study, we use two ways of composite scoring. One, the relevance-weighted quality score, is a specific example of O'Brien's second method. The parameters weights we used were obtained from 16 oncological surgeons-experts. They were asked to image that they were esophageal cancer patients choosing the most appropriate hospital, or that they were asked for such advice by a family member. Their relevance weights would thus represent both their professional insight, and their "as-if"-patient perspective.
The other way, the cumulative quality profile, is new and combines advantages of transparency and integration. In another study from our institution, we found that such a summary measure is very well received by patients.[17]

Our study has several limitations. The first was mentioned in the introduction and is hampering quality assessment in general: the absence of a 'golden standard' for quality of care. We selected six parameters available in our database and gave them weights representing the extent to which each of the 6 parameters is considered relevant for choosing a high quality hospital. We used substitutes for patient preferences, obtained from surgeons-experts by an "as-if" questionnaire. However interesting and valid their stated preferences may be, they are no substitute for real patients' preferences. The models presented in this study are therefore no more than a 'proof of principle', and the appropriateness of relative importance, either by sequence or by weights, requires more in-depth study in real patients.
The second limitation is the low number of institutions evaluated and there is the obvious problem of low patient numbers per institution and the ensuing lack of statistical power for rare outcomes. The fact that the Dutch Health Care Inspectorate has recently banned esophageal surgery in low volume hospitals, will provide part of the solution to this problem.[4] Another part of the solution may be our finding that multidimensional measures of quality may be more sensitive to quality differences, than single dimensional outcomes that rarely occur. In addition, following examples in our neighboring countries, nation-wide clinical audits for oncological and other care have recently been started in the Netherlands (www.clinicalaudit.nl).[18] In 2011, the Dutch Upper Gastro-intestinal Cancer Audit has been initiated in which all patients in who an esophageal or gastric resection for cancer has been performed are registered. The results of the present study suggest that multidimensional quality assessment may be a valid tool to analyze differences in quality of care in these clinical audits. It may provide broader and more relevant information on quality to patients,

doctors and society, but also may have better power to identify those hospitals that provide care of appropriate quality. The data from these much larger and detailed clinical databases may give us, or other researchers, the opportunity to refine and validate multi-dimensional frameworks for the assessment of quality of care, like the Exemplary Care and Outcome (ECO) concept presented in this study.

REFERENCES

1. Brennan TA, Leape LL, Laird NM, Hebert L, Localio AR, Lawthers AG et al. Incidence of adverse events and negligence in hospitalized patients: results of the Harvard Medical Practice Study I. 1991. *Qual Saf Health Care* 2004; 13(2):145-151.
2. Kohn L, Corrigan J, Donaldson MS. Institute of Medicine committee on Quality of Health Care in America. To err is human: building a safer health system. 1999. Washington, United States, National Academies Press.
3. Birkmeyer JD, Dimick JB, Birkmeyer NJ. Measuring the quality of surgical care: structure, process, or outcomes? *J Am Coll Surg* 2004; 198(4):626-632.
4. Wouters MW, Krijnen P, Le Cessie S, Gooiker GA, Guicherit OR, Marinelli AW et al. Volume- or outcome-based referral to improve quality of care for esophageal cancer surgery in The Netherlands. *J Surg Oncol* 2009; 99(8):481-487.
5. Donabedian A. Evaluating the quality of medical care. *Milbank Mem Fund Q* 1966; 44(3):Suppl-206.
6. Whooley BP, Law S, Murthy SC, Alexandrou A, Wong J. Analysis of reduced death and complication rates after esophageal resection. *Ann Surg* 2001; 233(3):338-344.
7. Birkmeyer JD, Siewers AE, Finlayson EV, Stukel TA, Lucas FL, Batista I et al. Hospital volume and surgical mortality in the United States. *N Engl J Med* 2002; 346(15):1128-1137.
8. Varghese TK, Jr., Wood DE, Farjah F, Oelschlager BK, Symons RG, Macleod KE et al. Variation in esophagectomy outcomes in hospitals meeting leapfrog volume outcome standards. *Ann Thorac Surg* 2011; 91(4):1003-1010.
9. Wouters MW, Karim-Kos HE, Le CS, Wijnhoven BP, Stassen LP, Steup WH et al. Centralization of esophageal cancer surgery: does it improve clinical outcome? *Ann Surg Oncol* 2009; 16(7):1789-1798.
10. Schuchert MJ, Abbas G, Nason KS, Pennathur A, Awais O, Santana M et al. Impact of anastomotic leak on outcomes after transhiatal esophagectomy. *Surgery* 2010; 148(4):831-838.
11. Wouters MW, Wijnhoven BP, Karim-Kos HE, Blaauwgeers HG, Stassen LP, Steup WH et al. High-Volume versus Low-Volume for Esophageal Resections for Cancer: The Essential Role of Case-Mix Adjustments based on Clinical Data. *Ann Surg Oncol* 2007.
12. Schouten LJ, Jager JJ, van den Brandt PA. Quality of cancer registry data: a comparison of data provided by clinicians with those of registration personnel. *Br J Cancer* 1993; 68(5):974-977.
13. Institute of Medicine committee on Quality of Health Care in America. Crossing the Quality Chasm: A New Health System for the 21st Century. 2001. National Academy Press, Washington, D.C.
14. Kievit J, Krukerink M, Marang-van de Mheen PJ. Surgical adverse outcome reporting as part of routine clinical care. *Qual Saf Health Care* 2010; 19(6):e20.
15. Dimick JB, Welch HG, Birkmeyer JD. Surgical mortality as an indicator of hospital quality: the problem with small sample size. *JAMA* 2004; 292(7):847-851.
16. O'Brien SM, Shahian DM, DeLong ER, Normand SL, Edwards FH, Ferraris VA et al. Quality measurement in adult cardiac surgery: part 2–Statistical considerations in composite measure scoring and provider rating. *Ann Thorac Surg* 2007; 83(4 Suppl):S13-S26.
17. Dijs-Elsinga J, Otten W, Versluijs MM, Smeets HJ, Kievit J, Vree R et al. Choosing a hospital for surgery: the importance of information on quality of care. *Med Decis Making* 2010; 30(5):544-555.
18. van Gijn W, Wouters MW, Peeters KC, van de Velde CJ. Nationwide outcome registrations to improve quality of care in rectal surgery. An initiative of the European Society of Surgical Oncology. *J Surg Oncol* 2009; 99(8):491-496.

APPENDIX

Evaluation of quality of care for esophageal resections for cancer: a survey *(this survey is translated from the original Dutch version)*

Publications regarding the inverse relationship between hospital volume and postoperative mortality have initiated an extensive debate on the quality of care for patients undergoing esophagectomy for cancer. However, in-hospital mortality is not the only factor that's important in the evaluation of quality of care for this patient group. Preferably, a more extensive set of measures would be used for the assessment of quality of care in individual institutions. Though, the data needed for this quality assessment are not always available and are usually limited to data regarding the outcome after surgery, like postoperative complications, re-operations, radicality of the resection and survival. In addition, it's not always clear how these quality aspects have to be weighed. Through this survey we would like to investigate how medical specialists treating patients with esophageal cancer value different outcomes of esophageal cancer surgery, their selves.

"Suppose that you, or a member of your family should be treated for esophageal cancer and you had to make a choice from a number of hospitals. From every hospital only a limited set of data regarding 6 quality aspects is available, collected by an independent authority. The differences in 'case-mix' of these hospitals do not have to be taken into consideration. The 6 quality aspects are:

	Quality aspect	Definition
1	In-hospital mortality	The percentage of patients that dies postoperatively during the same hospital admission.
2	Complications	The percentage of patients that has complications with permanent damage concerning functional loss or handicaps (for example: hoarseness, dysphagia, dependence on feeding tubes after failure of gastric tube reconstruction, cardiac failure after myocardial infarction)
3	Re-operations	The percentage of patients in who one or more re-operations have been performed due to complications after esophagectomy.
4	Length of stay	Percentage of patients with a length of stay longer than 2 weeks.
5	Radicality	Percentage of patients with a radical resection (R0).
6	1-year survival	Percentage of patients alive 1 year after the operation

"If there was a hospital that scores perfect on every quality aspect, you would obviously choose for that hospital. Unfortunately, there's no perfect hospital and hospitals score better or worse on different quality aspects. Therefore it's necessary to make a choice based on a combination of different quality aspects. Would you be so kind to answer the 2 questions best as you can:

Question 1

Please select which quality aspect is most important for you, and which aspects gradually would be less important.

Quality aspect	Ranking*
- In-hospital mortality	
- Complications	
- Re-operations	
- Length of stay	
- Radicality	
- 1-year survival	

* give ranking from "1" = most important to "6" = least important.

Question 2

You get 100 weight points. Can you divide these between the 6 quality aspects in such a way that the number of points represents the weight that you assign to the different aspects.

Quality aspect	Points*
- In-hospital mortality	
- Complications	
- Re-operations	
- Length of stay	
- Radicality	
- 1-year survival	
Total	100

* divide 100 points between 6 quality aspects

Examples

- *Answers in the two examples beneath are intentionally chosen at random* -

Answer 1

Quality aspect	Ranking*
- In-hospital mortality	4
- Complications	6
- Re-operations	2
- Length of stay	1
- Radicality	5
- 1-year survival	3

* Rank from "1" = most important to "6" = least important.

Answer 2

Quality aspect	Points*
- In-hospital mortality	8
- Complications	7
- Re-operations	20
- Length of stay	30
- Radicality	10
- 1-year survival	25
Total	100

* divide 100 points between 6 quality aspects