OBSERVATIONAL STUDY

Can composite performance measures predict survival of patients with colorectal cancer?

Kuo-Piao Chung, Li-Ju Chen, Yao-Jen Chang, Yun-Jau Chang

Kuo-Piao Chung, Li-Ju Chen, Graduate Institute of Health Policy and Management, College of Public Health, National Taiwan University, Taipei 10055, Taiwan

Li-Ju Chen, Department of Ophthalmology, Heping Branch, Taipei City Hospital, Taipei 10065, Taiwan

Yao-Jen Chang, Department of General Surgery, Buddhist Xindian Tzu Chi General Hospital, Xindian City, Taipei 23142, Taiwan

Yun-Jau Chang, Department of General Surgery, National Taiwan University Hospital, Taipei 10002, Taiwan

Yun-Jau Chang, Department of General Surgery, Zhongxing Branch, Taipei City Hospital, Taipei 10341, Taiwan

Author contributions: Chen LJ drafted and revised the manuscript; Chang YJ conceived of and designed the study; Chang YJ analyzed the data, and participated in interpretation and discussion of results; Chung KP participated in data acquisition and reviewed the manuscript; all authors read and approved the final manuscript.

Supported by Bureau of Health Promotion, Department of Health, Taiwan No. DOH98-NH-1004

Correspondence to: Yun-Jau Chang, MD, PhD, Department of General Surgery, National Taiwan University Hospital, No. 7 Zhongshan South Road, Zhongzheng District, Taipei 10002, Taiwan. yunjauchang2003@gmail.com

Telephone: +886-2-25523234-6258 Fax: +886-2-23581541

Received: November 6, 2013   Revised: March 24, 2014
Accepted: May 19, 2014
Published online: November 14, 2014

Abstract

AIM: To assess the relationship between long-term colorectal patient survival and methods of calculating composite performance scores.

METHODS: The Taiwan Cancer Database was used to identify patients who underwent bowel resection for colorectal adenocarcinoma between 2003 and 2004. Patients were assigned to one of three cohorts based on tumor staging: cohort 1, colon cancer stage < Ⅲ; cohort 2, colon cancer stage Ⅲ; cohort 3, rectal cancer. A composite performance score (CPS) was calculated for each patient using five different aggregating methods, including all-or-none, 70% standard, equal weight, analytic hierarchy process (AHP), and principal component analysis (PCA) algorithms. The relationships between CPS and five-year overall, disease-free, and disease-specific survivals were evaluated by a Cox proportional hazards model. A goodness-of-fit analysis for all five methods was performed using Akaike's information criterion.

RESULTS: A total of 3272 colorectal cancer patients (cohort 1, 1164; cohort 2, 790; cohort 3, 1318 patients) with a mean age of 65 years were enrolled in the study. Bivariate correlation analysis showed that CPS values from the equal weight method were highly correlated with those from the AHP method in all cohorts (all $P < 0.05$). Multivariate Cox hazards analysis showed that CPS values derived from equal weight and AHP methods were significantly associated with five-year survivals of patients in cohorts 1 and 2 (all $P < 0.05$). In these cohorts, higher CPS values suggested a higher probability of five-year survival. However, CPS values derived from the all-or-none method did not show any significant process-outcome relationship in any cohort. Goodness-of-fit analyses showed that CPS values derived from the PCA method were the best fit to the Cox proportional hazards model, whereas the values from the all-or-none model showed the poorest fit.

CONCLUSION: CPS values may highlight process-outcome relationships for patients with colorectal cancer in addition to evaluating quality of care performance.

analysis

Core tip: Performance measures allow healthcare stake-holders to evaluate the quality of services provided and maintained by healthcare organizations. However, it is unknown whether the implementation of performance measurements may also improve patient outcome. This study aimed to investigate the association between composite performance scores calculated with five different algorithms and survival of surgical patients with non-stage IV colorectal cancer. Models with weighted schemes showed that higher performance scores correlated with increased five-year survivals. However, the most stringent criterion-based strategy, the all-or-none method, failed to show any association.

INTRODUCTION

In recent decades, performance measurements for medical care have been implemented to assess the quality of care provided by hospitals. These assessments consist of a standardized set of valid, reliable, and evidence-based measures applicable to the process of accreditation and for improvements in quality and performance[1]. In addition, the healthcare system can develop its own performance measures, and payer organizations can use individual measures as pay-for-performance criteria[2,3]. Although the goal of these assessments is to ultimately improve patient care and outcome, their effect has been largely untested[4]. A report from the Institute of Medicine (IOM) emphasized the importance of performance and quality assessment of cancer care[5]. However, studies of performance measures to date have been based on descriptions of given hospitals, geographic areas, or populations, and have primarily focused on breast cancer[6]. The IOM has recommended the use of summary measures (composite measures) to streamline assessments from various organizations and government agencies[7]. Such composite measures can eliminate the limitations associated with individual measures[8,9], and have increasingly been used to summarize healthcare performance and enhance provider and hospital accountability. However, no consensus has been established for the incorporation of assessments into a reliable composite measure[10].

Colorectal cancer is a major public health threat and the second leading cause of cancer deaths in United States, the third leading cause in Taiwan, and the fourth leading cause worldwide[11-15]. However, the advancement in available adjuvant therapies, including chemotherapeutic agents, target and radiation therapies[16,17], has not been paralleled by improvement in patient outcome. For this reason, an investigation of process-outcome correlation is crucial to ensure that the design of the performance measurement system is appropriate. This study compared five methods [all-or-none, 70% standard, equal weight, analytic hierarchy process (AHP), and principal component analysis (PCA)] for the calculation of a composite performance score (CPS) for the treatment of newly diagnosed colorectal cancer patients from hospitals participating in the Taiwan Cancer Database (TCDB). The aim of this study was to assess the relationship between the composite performance score and the long-term survival of patients.

MATERIALS AND METHODS

Data sources and patient selection

The data in this study were obtained from the TCDB, a nation-wide cancer registry containing disease characteristics, treatment modalities, and updated follow-up status for patients with six major types of cancer, and were linked with the National Health Insurance Database (NHID), a population-based database containing claim data from all healthcare organizations in Taiwan[1]. Patients newly registered in the TCDB (with International Classification of Disease for Oncology ICD-O-3 C18.0 colon-C21.8 rectum) from 2003 to 2004 were identified. Patients who did not undergo curative surgery (*i.e.*, biopsy only, colostomy only, or no surgery of any kind) or transanal excision were excluded from the analysis. Curative surgery referred to total proctectomy, hemicolectomy, partial, subtotal, or total colectomy, or colectomy with resection of the diseased organ (operation type code 20-90). Patients with stage IV disease, pathological report other than adenocarcinoma, or no treatment record in the NHID were also excluded. Patients enrolled in this study were then linked to the 2002-2005 NHID to acquire comorbidity status and linked to a third database (2003-2009 death registry from the Ministry of the Interior) to recheck survival status and identify cause of death if any. Enrolled patients were divided into one of three cohorts according to disease stage. The variables available for this study included age at diagnosis, sex, comorbidity index, primary tumor site, stage, pathological grade (differentiation) and stage, surgical margin, tumor invasion depth, lymph node evaluation, chemotherapy, and radiotherapy. The primary endpoint was five-year survival, including overall survival (OS), disease-free survival (DFS), and disease-specific survival (DSS)[2].

Performance measures and AHP questionnaire

A set of performance measures for colorectal cancer has been developed by our group in collaboration with multidisciplinary professionals (proposed by the Taiwan Clinical Oncology Group from the National Health Research Institute)[1]. This multidisciplinary panel endorsed 11 measures covering three categories: two pre-treatment, six treatment, and three follow-up measures. For this

Category	Measure	Description	Global priority[1]	Weight[2]		
				Cohort 1	Cohort 2	Cohort 3
Pre-treatment (PT)	PT1	CRC patients with pre-operative chest X-ray and abdominal ultrasound, CT scan, or MRI	0.144	0.213	0.193	0.213
	PT2	Early stage of CRC	0.104	0.154	0.140	0.154
Treatment (T)	T1	CRC patients with history of surgical resection that were checked by colonoscopy or barium enema LGI series with sigmoidoscopy within six months peri-operatively	0.030	0.044	0.040	0.044
	T2	Patients with non-metastatic CRC offered curative resection or neoadjuvant therapy within six weeks of diagnosis	0.057	0.084	0.077	0.084
	T3	Patients with stage Ⅰ to Ⅲ CRC who underwent wide surgical resection with a "negative margin"	0.133	0.197	0.179	0.197
	T4	CRC patients who underwent surgery with pathology reports on tumor and node stage	0.116	0.172	0.156	0.172
	T5	Patients with stage Ⅰ to Ⅲ CRC with twelve or more lymph nodes examined in pathology reports	0.092	0.136	0.124	0.136
	T6	Patients (< 70 yr) with stage Ⅲ CRC who received chemotherapy within eight weeks after surgery	0.069		0.093	
Follow-up (F)[3]	F1	CRC patients (stages Ⅰ–Ⅲ) survived after five post-operative years	0.101			
	F2	RC patients (stages Ⅰ–Ⅲ) experienced no local recurrence after five post-operative years	0.082			
	F3	CRC patients expired within 30 d after surgery	0.073			
Sum of priorities (or weights)			1.001[4]	1.000	1.000	1.000

[1]Values from Chung *et al*[18], 2013; [2]Individual weight was derived proportionally from individual global priority; [3]Measures from follow-up category were not considered as process measures; [4]Value due to rounding. Cohort 1: Colon cancer patients, stage < Ⅲ; Cohort 2: Colon cancer patients, stage Ⅲ; Cohort 3: Rectal cancer patients. CRC: Colorectal cancer; RC: Rectal cancer; CT: Computed tomography; MRI: Magnetic resonance imaging; LGI: Lower gastrointestinal.

study, an emphasis was placed on pre-treatment and the first six treatment measures, which were modified to allow binary status assignments (yes/no, concordant/non-concordant) at the patient level[18]. Priorities or weights (Table 1) were used in subsequent CPS calculations in this study. The design of the study is shown in Figure 1.

All-or-none calculation of the CPS
The all-or-none score is a dichotomous variable that signifies patients for whom all of the audited aspects of care are met. This method has recently been advocated largely on the basis that acceptance of anything less than complete care is incompatible with the pursuit of excellence[6,19-22].

70% standard calculation of the CPS
The 70% standard method is a less stringent version of the all-or-none method that was recently adopted by Reeves *et al*[23]. The 70% standard score is a dichotomous variable with a lowered criterion such that only 70% of the audited aspects of care need to be met.

Equal weight calculation of the CPS
The equal weight score represents the mean percentage of audited aspects of care that are met for each patient method and assumes that all indicators contribute equally to the outcome[24]. The equal weight method has been used to evaluate changes in quality of care over a five-year period in a representative general practice sample in England[25] and to explore the trend of composite hospi-

tal measures following coronary artery bypass surgery[20].

AHP calculation of the CPS
The AHP method was developed by Thomas to solve complex problems involving multiple criteria and features an ordinal pair-wise comparison of attributes ranging from a score of 1 (equal importance) to 9 (extreme importance)[18]. The AHP technique results in priorities of attributes (weights of performance measures) that add up to 1. This method was used by Richman *et al*[26] for the selection of prostate cancer treatments. Among the five methods used in the present study, it is the only one that requires experts' choices in a questionnaire survey.

PCA calculation of the CPS
The PCA calculation is a simple, non-parametric method of extracting relevant information from complex datasets[27]. The goal of the PCA method is to reduce the dimensionality of a dataset containing a large number of interrelated variables, while retaining its variation to the greatest extent possible[10]. This method has previously been used to rank American hospitals[28].

Statistical analysis
Statistical analyses were performed using SAS, version 8.2 (SAS Institute Inc., Cary, NC, United States) and SPSS 12.0 (SPSS Inc., Chicago, IL, United States) statistical software. Multivariate Cox proportional hazards models were constructed for survival analyses. All estimates were derived by controlling for age, sex, comorbidity, chemother-
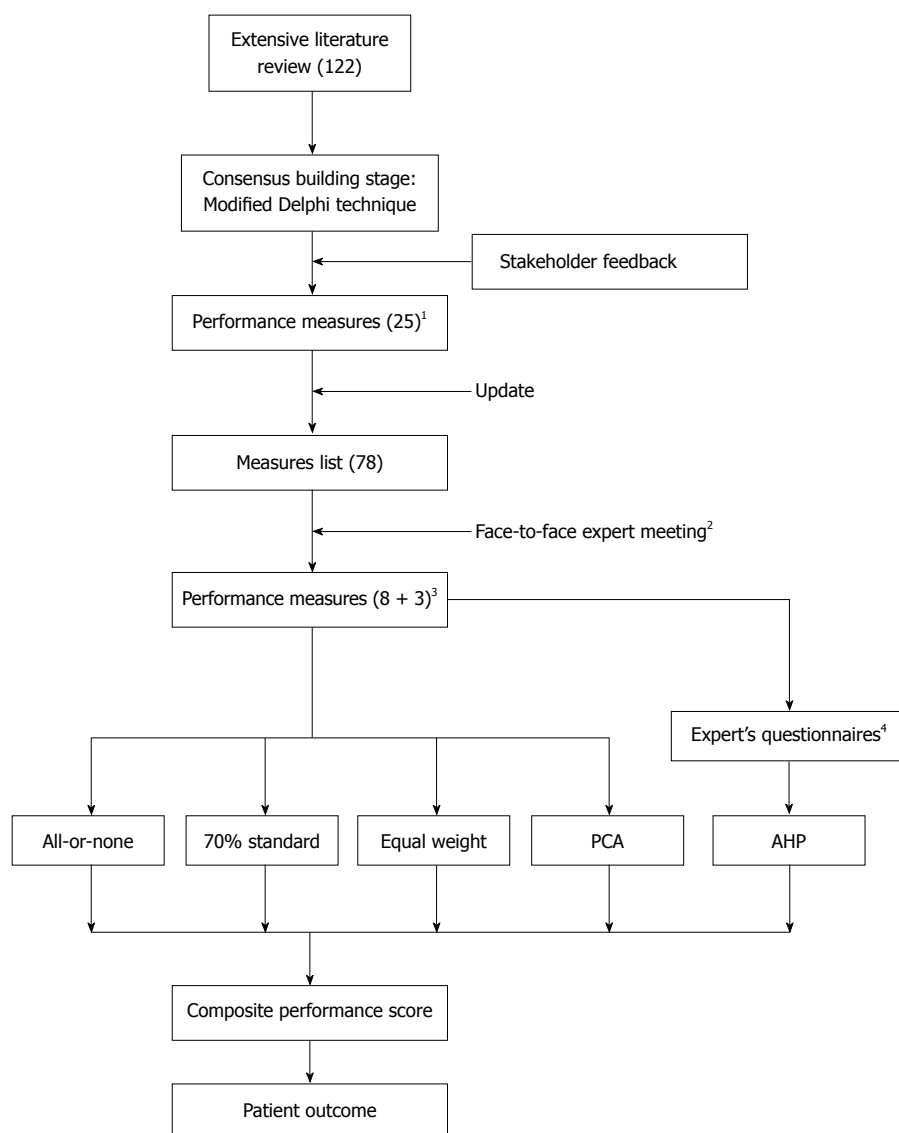
Figure 1 Flow chart of the study design (modified from Chung *et al*[1], 2010 and Chung *et al*[18], 2013). [1]Including 3 pre-treatment measures, 16 treatment measures, and 6 monitor measures (Chung, 2010); [2]Meetings were held on September 4 and October 2, 2009 at the College of Public Health, National Taiwan University; [3]Including 2 pre-treatment measures, 6 treatment measures, and 3 follow-up measures (Chung, 2013); [4]Questionnaire was mailed on October 25, 2009 and completed as of November 20, 2009 (Chung, 2013). PCA: Principal component analysis; AHP: Analytic hierarchy process.

apy, and tumor, lymph node, and metastasis (TNM) stage. Values are expressed as mean ± SD, and $P < 0.05$ was considered to be statistically significant. Akaike's information criterion (AIC) was used to represent the goodness-of-fit of models in the CPS. To identify the best model, AIC = 2k - 2 ln(L), where k refers to the number of parameters, and L refers to the maximized value of the likelihood function for the estimated model[29]. As the smallest value of AIC indicates the best model fit, data are reported as ΔAIC (difference from the best model). As result, the best model is defined by an AIC = 0.

RESULTS

Description of patient cohorts

A total of 8143 patients newly diagnosed with colorectal cancer from 2003 to 2004 were treated in 29 hospitals participating in the TCDB program (Figure 2). Of the 3272 patients who were eligible for this study, 1164 patients had stage < Ⅲ colon cancer (cohort 1), 790 patients had stage Ⅲ cancer (cohort 2), and 1318 patients had rectal cancer (cohort 3). The mean age of patients among the cohorts ranged from 64.3 to 66.8 years, and all cohorts were comprised of a greater proportion of male patients (Table 2).

In cohorts 1 and 2, tumors were most often located in the sigmoid colon, followed by the ascending colon. Comorbidity indices of patients were more frequently ≥ 2, and ≥ 3 (71%) for cohort 2. The majority of tumors in all cohorts were moderately differentiated (about 81%). These tumors commonly penetrated through the muscularis propria of the colorectal wall into the subserosa (cohorts 1 and 3), and directly invaded other organs (cohort 2). TNM staging indicated that the majority of
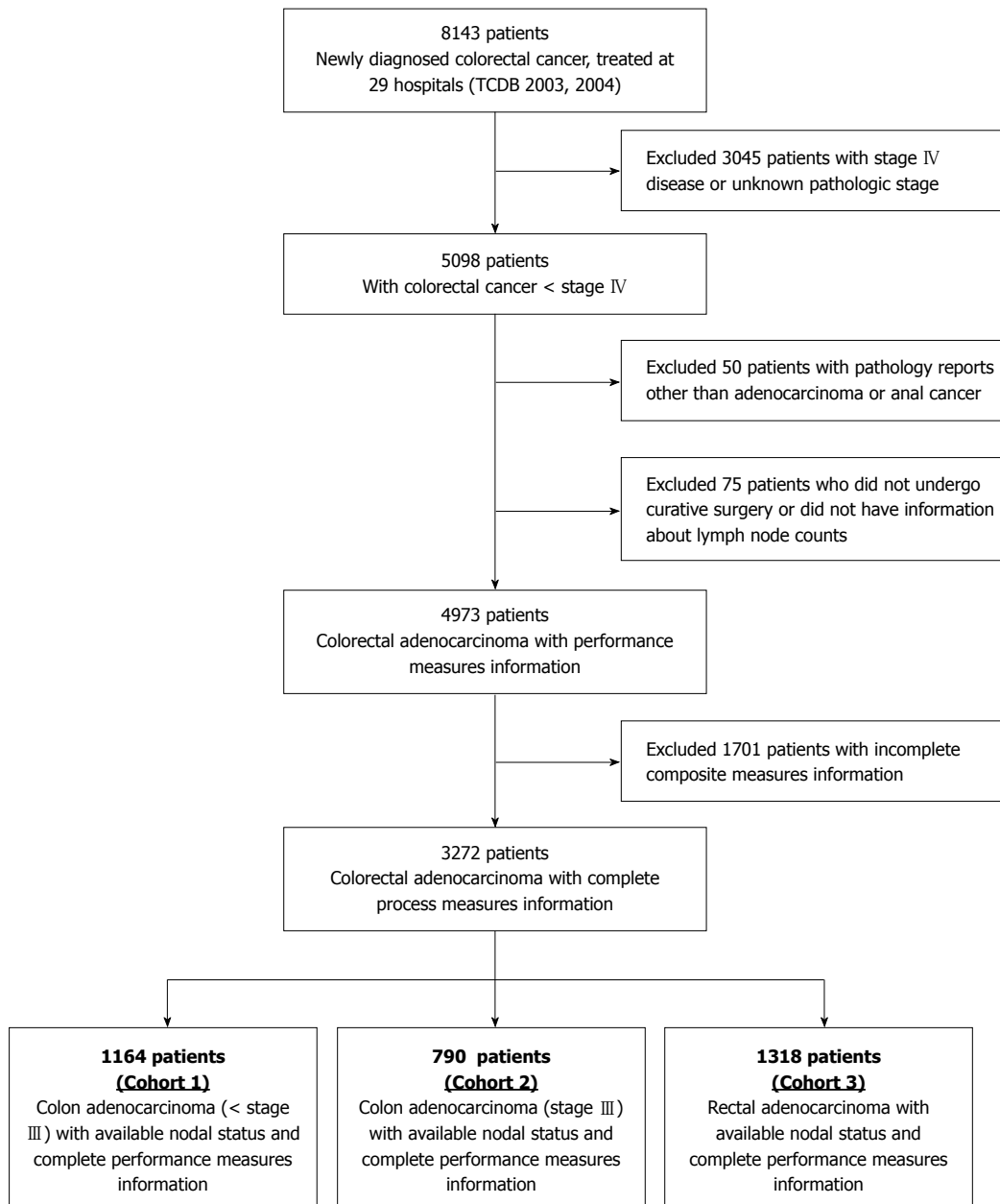
Figure 2 Enrollment of colorectal patients. TCDB: Taiwan cancer database.

cohort 3 patients (58.2%) presented with N0 status at the time of first treatment, whereas most patients of cohort 2 were of N1 staging (63.9% *vs* 35.4% in N2). The number of lymph nodes retrieved and examined was 17.7 ± 12.6, 19.3 ± 13.9, and 15.4 ± 10.0 for patients of cohorts 1, 2, and 3, respectively. Harvests of more than 25 lymph nodes were uncommon. Adjuvant chemotherapy was administered to 26.3, 75.9, and 49.6% patients, and radiotherapy to 2, 4.1, and 23.4% patients from cohorts 1, 2, and 3, respectively. The mean follow-up duration of all patients was 54.6 ± 21.5 mo (range, 0.0 to 84.0). Five-year overall survival rates were 87.1, 73.0, and 88.7% for cohorts 1, 2, and 3, respectively.

Association between composite performance score and survival

The CPS values calculated by each of the five methods are presented in Table 3, ranging from 0 to 1 for all methods except PCA. CPS values derived with the equal weight method were strongly correlated with values derived using the AHP and PCA methods in cohort 1, and with the AHP method only in cohorts 2 and 3 (Table 4). The relationships between CPS values derived from each of the five methods and patient survival are summarized in Table 5, including HR estimated after controlling for confounders. Five-year OS in cohort 1 was significantly associated with CPS values derived with the 70% stan-

Table 2 Characteristics of patients undergoing colectomy for colorectal adenocarcinoma *n* (%)			
Parameter	Cohort 1	Cohort 2	Cohort 3
	(*n* = 1164)	(*n* = 790)	(*n* = 1318)
Age (yr)	66.8 ± 13.1	65.4 ± 13.6	64.3 ± 12.8
Sex (M/F)	656 (56.4)/	431 (54.6)/	781 (59.3)/
	508 (43.6)	359 (45.4)	537 (40.7)
Location			
Rectum	-	-	1318 (100)
Rectosigmoid junction	124 (10.7)	88 (11.1)	-
Sigmoid	440 (37.8)	313 (39.6)	-
Descending	124 (10.7)	72 (9.1)	-
Splenic flexure	35 (3.0)	17 (2.2)	-
Transverse	107 (9.2)	70 (8.9)	-
Hepatic flexure	47 (4.0)	38 (4.8)	-
Ascending	173 (14.9)	118 (14.9)	-
Cecum and appendix	77 (6.6)	44 (5.6)	-
Overlapping	12 (1.0)	7 (0.9)	-
Unspecified	25 (2.1)	23 (2.9)	-
Comorbidity			
CCI (0 or 1)	63 (5.4)	26 (3.3)	73 (5.5)
CCI (2)	659 (56.6)	103 (25.7)	531 (40.3)
(> 3)	442 (37.9)	661 (71.0)	714 (54.2)
Grade			
Well differentiated	102 (8.8)	46 (5.8)	101 (7.7)
Moderately differentiated	937 (80.5)	640 (81.0)	1072 (81.3)
Poorly differentiated	61 (5.2)	87 (11.0)	89 (6.8)
Un-differentiated	1 (0.1)	-	1 (0.1)
Unspecified	63 (5.4)	17 (2.2)	55 (4.2)
Staging (p)			
Stage I	295 (25.3)	-	311 (23.6)
Stage II	869 (74.7)	-	459 (34.8)
Stage III	-	790 (100.0)	548 (41.6)
pT1	133 (11.4)	15 (1.9)	107 (8.1)
pT2	164 (14.1)	37 (4.7)	287 (21.8)
pT3	694 (59.6)	574 (72.7)	719 (54.6)
pT4	162 (13.9)	158 (20.0)	198 (15.0)
Unspecified	11 (0.9)	6 (0.8)	7 (0.5)
pN0	1164 (100)	-	767 (58.2)
pN1	-	505 (63.9)	299 (22.7)
pN2	-	280 (35.4)	246 (18.7)
Unspecified	-	5 (0.6)	6 (0.5)
Lymph node examination	17.7 ± 12.6	19.3 ± 13.9	15.4 ± 10.0
Chemotherapy (yes)	306 (26.3)	600 (75.9)	654 (49.6)
Radiotherapy (yes)	23 (2.0)	32 (4.1)	308 (23.4)
Follow-up months (range)	57.1 ± 20.1	51.3 ± 23.2	54.7 ± 21.3
	(0-84.0)	(0.1-84.0)	(0-83.8)
Five-year survival rate	87.1	73	88.7

Pathological staging and T and N classifications as defined by the American Joint Committee on Cancer, 6th edition. CCI: Charlson comorbidity index; M/F: Male/Female; p: Pathological.

dard, equal weight, AHP, and PCA methods (HR = 0.76, 95%CI: 0.59-0.98; HR = 0.32, 95%CI: 0.14-0.74; HR = 0.32, 95%CI: 0.14-0.73; and HR = 0.85, 95%CI: 0.78-0.92, respectively, all *P* < 0.05). Similar results were found for the five-year DSS and DFS (except marginal significance for DFS by the 70% standard method). These data indicate that a patient with a higher CPS in cohort 1 had an increased probability for five-year survival. For patients in cohort 2, the CPS values derived with the AHP method were significantly associated with five-year OS or DSS, and values derived with the equal weight method were only associated with the DSS. There were no significant

associations between five-year survivals and CPS values for patients in cohort 3. Moreover, CPS values derived with the all-or-none method gave no prediction for five-year survival. The goodness-of-fit model analysis using AIC showed that the best method for predicting five-year survivals, with the lowest AIC values, was the PCA method, followed next by the AHP and equal weight methods (Table 6).

DISCUSSION

The complexity of colorectal cancer and its treatment limit the availability of solid evidence concerning the quality of care and its relation to patient outcome. This study is the first, to our knowledge, to assess such process-outcome relationships by use of composite measures in colorectal cancer therapy. The most important finding of this study is that almost all of the aggregating methods used to generate a CPS, the use of which have been controversial[20], correlated with patient survival.

The results of this study demonstrate that the all-or-none method of generating a composite measure of treatment quality was not associated with patient outcome for any of the cohorts examined. This is in contrast to results obtained by Cheng *et al*[6], who demonstrated a positive relationship between all-or-none performance measures and five-year OS and DFS for patients with breast cancer within a single hospital. This discordance could be attributed to the fact that the level of performance measures for breast cancer is higher than for colorectal cancer. Additionally, the heterogeneity of population-based data may play a role, as patients from 29 hospitals were included in the present study. Despite the differences in methodologies, both studies reveal the possibility that CPS shows prognostic significance in the field of cancer care. The current study also substantiates the role for AHP in the performance measurement of colorectal cancer. The AHP method incorporates multi-criteria selection, which is useful for deciding priorities in performance measures, and valuable expert clinical input to facilitate clinicians' adherence to the chosen performance measures. Unlike the equal weight method, the AHP provides prioritized ranking of target measures for subsequent discretion, which may prove most useful for the design of pay-for-performance programs[18].

The selection of an appropriate aggregating method for calculating a CPS is not always stereotyped. The "absolute score" methods (all-or-none and 70% standard) represent the degree to which a provider has achieved a predefined threshold of quality of care for each patient. The judgment of 70% is arbitrary, aimed to strike a balance between strictness and leniency[23]. The scores derived from the remaining aggregating methods, however, are modified by weighting each measure before aggregation. These methods, namely the AHP, PCA, and equal weight, differ on the relative contribution that each measure makes to the composite score.

The use of a weighting scheme raises complex questions, such as how to select a suitable criterion on which

Table 3 Composite performance scores (mean, range) calculated by different methods

Cohort	All-or-none	70% standard	Equal weight	AHP	PCA
1	0.07 ± 0.26	0.78 ± 0.42	0.74 ± 0.14	0.78 ± 0.14	0 ± 1.00
	(0.00-1.00)	(0.00-1.00)	(0.29-1.00)	(0.22-1.00)	(-6.44-1.04)
2	NA	0.68 ± 0.47	0.73 ± 0.12	0.73 ± 0.12	0 ± 1.00
	NA	(0.00-1.00)	(0.25-0.88)	(0.27-0.86)	(-2.03-1.23)
3	0.05 ± 0.21	0.76 ± 0.43	0.73 ± 0.14	0.77 ± 0.13	0 ± 1.00
	(0.00-1.00)	(0.00-1.00)	(0.14-1.00)	(0.19-1.00)	(-13.13-0.59)

Cohort 1: Colon cancer patients stage < Ⅲ; Cohort 2: Colon cancer patients stage Ⅲ; Cohort 3: Rectal cancer patients. AHP: Analytic hierarchy process; NA: Not applicable; PCA: Principal component analysis.

Table 4 Correlation matrix of composite performance scores calculated by different methods

Cohort	Method	All-or-none	70% standard	Equal weight	AHP	PCA
1	All-or-none	1				
	70% standard	0.149[b]	1			
	Equal weight	0.500[b]	0.784[b]	1		
	AHP	0.476[b]	0.710[b]	0.918[b]	1	
	PCA	0.288[b]	0.664[b]	0.819[b]	0.853[b]	1
2	All-or-none	NA				
	70% standard	NA	1			
	Equal weight	NA	0.843[b]	1		
	AHP	NA	0.753[b]	0.891[b]	1	
	PCA	NA	0.436[b]	0.502[b]	0.823[b]	1
3	All-or-none	1				
	70% standard	0.125[b]	1			
	Equal weight	0.429[b]	0.794[b]	1		
	AHP	0.416[b]	0.700[b]	0.904[b]	1	
	PCA	0.071[b]	0.407[b]	0.484[b]	0.419[b]	1

Cohort 1: Colon cancer patients stage < Ⅲ; Cohort 2: Colon cancer patients stage Ⅲ; Cohort 3: Rectal cancer patients. [b]$P < 0.01$ *vs* control. AHP: Analytic hierarchy process; NA: Not applicable; PCA: Principal component analysis.

Table 5 Associations between composite performance score and five-year survival by multivariate Cox proportional hazards model

Cohort	Survival	Method														
		All-or-none			70% standard			Equal weight			AHP			PCA		
		HR	95%CI	*P* value	HR	95%CI	*P* value	HR	95%CI	*P* value	HR	95%CI	*P* value	HR	95%CI	*P* value
1	OS	0.88	0.45-1.73	NS	0.76	0.59-0.98	0.032	0.32	0.14-0.74	0.007	0.32	0.14-0.73	0.007	0.85	0.78-0.92	< 0.001
	DSS	0.93	0.40-2.20	NS	0.69	0.52-0.91	0.009	0.30	0.12-0.79	0.014	0.28	0.11-0.70	0.006	0.85	0.77-0.95	0.002
	DFS	0.93	0.49-1.80	NS	0.78	0.61-1.01	0.057	0.40	0.17-0.92	0.031	0.36	0.16-0.80	0.013	0.87	0.79-0.95	0.002
2	OS		NA		0.92	0.73-1.17	NS	0.42	0.16-1.08	NS	0.39	0.15-0.98	0.045	0.92	0.83-1.03	NS
	DSS		NA		0.88	0.69-1.13	NS	0.35	0.13-0.94	0.038	0.31	0.12-0.81	0.017	0.90	0.81-1.01	NS
	DFS		NA		0.96	0.76-1.21	NS	0.48	0.19-1.23	NS	0.47	0.19-1.18	NS	0.95	0.83-1.03	NS
3	OS	1.07	0.54-2.13	NS	0.93	0.75-1.16	NS	0.07	0.34-1.46	NS	0.72	0.35-1.49	NS	0.98	0.89-1.08	NS
	DSS	1.33	0.60-2.92	NS	0.9	0.81-1.01	NS	0.61	0.28-1.34	NS	0.66	0.30-1.43	NS	1.00	0.89-1.11	NS
	DFS	1.13	0.58-2.19	NS	0.95	0.75-1.16	NS	0.73	0.35-1.50	NS	0.76	0.37-1.55	NS	0.99	0.90-1.08	NS

Cohort 1: Colon cancer patients stage < Ⅲ; Cohort 2: Colon cancer patients stage Ⅲ; Cohort 3: Rectal cancer patients. AHP: Analytic hierarchy process; DFS: Disease-free survival; DSS: Disease-specific survival; NA: Not applicable; OS: Overall survival; PCA: Principal component analysis.

to base the weights, how to determine the value each weight should take, and how to determine the sensitivity of the score to the weights. The equal weight method is the simplest and most straightforward, ensuring that each performance measure makes an equal contribution, regardless of how discrepant their priorities are. The PCA method, based on statistical models, and the AHP method, based on experts' preferences for measures, assign different weights to each measure. The PCA method can eliminate the double counting of performance measures, though the set of weights used in the composite score will change as a result of its sensitivity to data modification, revision, and updating[24]. The AHP method recruits participation of expert medical professionals *via* a questionnaire survey at the outset, which can increase the transparency of the composite score. This can encourage cooperation between clinical practitioners and academic researchers, as well as other stakeholders, which may be

Table 6 Model fit analysis of Akaike's information criterion values calculated by different methods

Cohort	Survival	Method				
		All-or-none	70% standard	Equal weight	AHP	PCA
1	OS	12.819	8.544	6.002	5.895	0
	DSS	8.823	2.412	3.047	1.691	0
	DFS	9.406	6.023	4.921	3.434	0
2	OS	NA	14.655	11.914	11.205	0
	DSS	NA	4.593	1.363	0	2.520
	DFS	NA	2.385	0.246	0	1.605
3	OS	0.843	0.483	0	0.105	0.750
	DSS	1.053	0.380	0	0.421	1.507
	DFS	0.617	0.543	0	0.177	0.659

Cohort 1: Colon cancer patients stage < Ⅲ; Cohort 2: Colon cancer patients stage Ⅲ; Cohort 3: Rectal cancer patients. Values represent the ΔAIC, where the best model is defined as AIC = 0. AHP: Analytic hierarchy process; AIC: Akaike's information criterion; DFS: Disease-free survival; DSS: Disease-specific survival; OS: Overall survival; PCA: Principal component analysis.

more likely to result in the incorporation of quality improvement programs into daily practice, thus paving the way for enhancement of patient outcome.

As corresponding academic associations develop performance measures of cancer care quality independently, discrepancies between the specification of measures used in this study and other works should not be overlooked. Two of the performance measures (T3, T4) in this study contained combinations of two to three measures, as in the study by Malin *et al*[30]. However, unlike their study, we recommended that chest X-rays, ultrasound, and computed tomography scans be ordered for pre-operative checks for patients with colorectal cancer instead of ultrasound alone, as was suggested for patients with rectal cancer by their study. Another discrepancy was the commencement of chemotherapy for stage Ⅲ colon cancer within two months of diagnosis, in contrast to Desch *et al*[31] who advocated treatment within four months. However, the minimal requirement of examining 12 lymph nodes remained the same in our study.

There were several limitations of the present study. First, although TCDB covered about 60% of all newly diagnosed colorectal cancer patients, the results cannot be considered as universal. Second, fewer than seven measures in each criterion are permitted in the AHP framework. This restriction inevitably precludes the consideration of some measures that may be meaningful, though a recent study shows that the tendency to reduce the number of performance measures is in pursuit of a useful and meaningful end[31]. Third, the choice of weighting schemes was constrained by the concern for case-based adherence, rather than for comparison between the organizational or nationwide levels. Fourth, the failure to demonstrate a process-outcome relationship in patients with rectal cancer indicates that there is still a need for revisal of these measures.

In conclusion, the results suggest that the implementation of performance measurement may improve patient outcomes. For clinicians, adherence to composite measures may increase colorectal cancer patients' five-year survivals. For policymakers, the all-or-none method should not be used to evaluate colorectal cancer care, as this method failed to correlate with patient outcome. For healthcare researchers, performance measurement is still not predictive for survival in patients with rectal cancer, indicating that research concerning the quality of cancer care should be expanded by comprehensive survival studies or clinical trials. Future research may uncover additional performance measures to allow for extension of this process-outcome assessment to other cancer domains and to achieve both an authentic improvement in the quality of care as well as better patient prognosis.

COMMENTS

Background
Performance measures have been developed in recent decades in an effort to ensure cancer care quality for several types of malignancies. For these measures to be meaningful, they should relate to patient outcome. A recent study reported a relationship between individual measures and patient outcome in breast cancer using data from a single hospital. However, there have been no large-scale studies investigating the relationship between performance measures and patient outcome for colorectal cancer.

Research frontiers
The selection of methods for calculating composite performance scores is arbitrary. There are "absolute score" methods, such as the all-or-none and 70% standard methods, as well as weighted methods, including the equal weight, analytic hierarchy process, and principal component analysis methods. By comparing these five aggregating methods, the influences of each method on process-outcome linkage in the field of colorectal cancer care can be evaluated in-depth.

Innovations and breakthroughs
There is a large gap between clinicians and other healthcare stakeholders concerning the development and application of cancer care performance measurements. These differences may be resolved with the introduction of the analytic hierarchy process method for construction and determination of performance measures, which will subsequently set a standard for clinical daily practices.

Applications
The results of this study support the existence of a process-outcome linkage in quality cancer care. Higher composite performance scores can predict increased five-year survival in patients with colorectal cancer, and can be used to effectively evaluate patient care.

Terminology
The composite performance score is acquired by aggregating scores of individual performance measures. Several levels of composite performance scores can be calculated, including national, hospital, and individual levels. The analytic hierarchy process is a weighting scheme that incorporates the priority rankings from experts' responses to a questionnaire. The principal component

analysis is a method to weight individual performance measures to reduce the dimensionality of dataset.

Peer review

This is a population-based study analyzing the predictive capability of composite performance scores on colorectal cancer patient survival. This is the first study evaluating various strategies to calculate composite performance scores, and demonstrates that weighted analyses of performance measurements may provide the best method to evaluate quality of colorectal cancer care.

REFERENCES

1 **Chung KP**, Chang YJ, Lai MS, Kuo RN, Cheng SH, Chen LT, Tang R, Liu TW, Shieh MJ. Is quality of colorectal cancer care good enough? Core measures development and its application for comparing hospitals in Taiwan. *BMC Health Serv Res* 2010; **10**: 27 [PMID: 20105287 DOI: 10.1186/1472-6963-10-27]

2 **Chang YJ**, Chen LJ, Chang YJ, Chung KP, Lai MS. Application of propensity score model to examine the prognostic significance of lymph node number as a care quality indicator. *Surg Oncol* 2012; **21**: e75-e85 [PMID: 22221938 DOI: 10.1016/j.suronc.2011.12.003]

3 **Chang YJ**, Chang YJ, Chen LJ, Chung KP, Lai MS. Evaluation of lymph nodes in patients with colon cancer undergoing colon resection: a population-based study. *World J Surg* 2012; **36**: 1906-1914 [PMID: 22484567 DOI: 10.1007/s00268-012-1568-2]

4 **Higashi T**, Shekelle PG, Adams JL, Kamberg CJ, Roth CP, Solomon DH, Reuben DB, Chiang L, MacLean CH, Chang JT, Young RT, Saliba DM, Wenger NS. Quality of care is associated with survival in vulnerable older patients. *Ann Intern Med* 2005; **143**: 274-281 [PMID: 16103471 DOI: 10.7326/0003-4819-143-4-200508160-00008]

5 **Hewitt M**, Simone JV. Ensuring quality cancer care. Washington DC: National Academy Press, 1999

6 **Cheng SH**, Wang CJ, Lin JL, Horng CF, Lu MC, Asch SM, Hilborne LH, Liu MC, Chen CM, Huang AT. Adherence to quality indicators and survival in patients with breast cancer. *Med Care* 2009; **47**: 217-225 [PMID: 19169123 DOI: 10.1097/MLR.0b013e3181893c4a]

7 **Hurtado MP**, Swift EK, Corrigan JM. Envisioning the national healthcare quality report: designing the national healthcare quality report. Washington DC: National Academy Press, 2001

8 **Shwartz M**, Ren J, Peköz EA, Wang X, Cohen AB, Restuccia JD. Estimating a composite measure of hospital quality from the Hospital Compare database: differences when using a Bayesian hierarchical latent variable model versus denominator-based weights. *Med Care* 2008; **46**: 778-785 [PMID: 18665057 DOI: 10.1097/MLR.0b013e31817893dc]

9 **Staiger DO**, Dimick JB, Baser O, Fan Z, Birkmeyer JD. Empirically derived composite measures of surgical performance. *Med Care* 2009; **47**: 226-233 [PMID: 19169124 DOI: 10.1097/MLR.0b013e3181847574]

10 **Singh RK**, Murty HR, Gupta SK, Dikshit AK. Development of composite sustainability performance index for steel industry. *Ecological Indicators* 2007; **7**: 565-588

11 **Siegel R**, Ward E, Brawley O, Jemal A. Cancer statistics, 2011: the impact of eliminating socioeconomic and racial disparities on premature cancer deaths. *CA Cancer J Clin* 2011; **61**: 212-236 [PMID: 21685461 DOI: 10.3322/caac.20121]

12 Causes of death in Taiwan, DOH 2012 [cited 2014 March 22]. Available from: URL: http://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_full.pdf

13 Global Burden of Disease Estimates. WHO 2008 [cited 2014 March 22]. Available from: URL: http://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_part2.pdf

14 **Jemal A**, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin* 2011; **61**: 69-90 [PMID: 21296855 DOI: 10.3322/caac.20107]

15 **Chang YJ**, Chen LJ, Chang YJ, Chung KP, Lai MS. Risk groups defined by Recursive Partitioning Analysis of patients with colorectal adenocarcinoma treated with colorectal resection. *BMC Med Res Methodol* 2012; **12**: 2 [PMID: 22214198 DOI: 10.1186/1471-2288-12-2]

16 **Yuan Y**, Li MD, Hu HG, Dong CX, Chen JQ, Li XF, Li JJ, Shen H. Prognostic and survival analysis of 837 Chinese colorectal cancer patients. *World J Gastroenterol* 2013; **19**: 2650-2659 [PMID: 23674872 DOI: 10.3748/wjg.v19.i17.2650]

17 **Trakarnsanga A**, Ithimakin S, Weiser MR. Treatment of locally advanced rectal cancer: controversies and questions. *World J Gastroenterol* 2012; **18**: 5521-5532 [PMID: 23112544 DOI: 10.3748/wjg.v18.i39.5521]

18 **Chung KP**, Chen LJ, Chang YJ, Chang YJ, Lai MS. Application of the analytic hierarchy process in the performance measurement of colorectal cancer care for the design of a pay-for-performance program in Taiwan. *Int J Qual Health Care* 2013; **25**: 81-91 [PMID: 23175531 DOI: 10.1093/intqhc/mzs070]

19 **Auerbach AD**, Hilton JF, Maselli J, Pekow PS, Rothberg MB, Lindenauer PK. Shop for quality or volume? Volume, quality, and outcomes of coronary artery bypass surgery. *Ann Intern Med* 2009; **150**: 696-704 [PMID: 19451576]

20 **O'Brien SM**, Shahian DM, DeLong ER, Normand SL, Edwards FH, Ferraris VA, Haan CK, Rich JB, Shewan CM, Dokholyan RS, Anderson RP, Peterson ED. Quality measurement in adult cardiac surgery: part 2--Statistical considerations in composite measure scoring and provider rating. *Ann Thorac Surg* 2007; **83**: S13-S26 [PMID: 17383406]

21 **Nolan T**, Berwick DM. All-or-none measurement raises the bar on performance. *JAMA* 2006; **295**: 1168-1170 [PMID: 16522838]

22 **Couralet M**, Guérin S, Le Vaillant M, Loirat P, Minvielle E. Constructing a composite quality score for the care of acute myocardial infarction patients at discharge: impact on hospital ranking. *Med Care* 2011; **49**: 569-576 [PMID: 21532510 DOI: 10.1097/MLR.0b013e31820fc386]

23 **Reeves D**, Campbell SM, Adams J, Shekelle PG, Kontopantelis E, Roland MO. Combining multiple indicators of clinical quality: an evaluation of different analytic approaches. *Med Care* 2007; **45**: 489-496 [PMID: 17515775]

24 **Nardo M**, Saisana M, Saltelli A, Tarantola S, Hoffman A. Handbook on constructing composite indicators: methodology and user guide. Paris: OECD Statistics Working Papers, Organisation for Economic Cooperation and Development, 2005

25 **Campbell SM**, Roland MO, Middleton E, Reeves D. Improvements in quality of clinical care in English general practice 1998-2003: longitudinal observational study. *BMJ* 2005; **331**: 1121 [PMID: 16257992]

26 **Richman MB**, Forman EH, Bayazit Y, Einstein DB, Resnick MI, Stovsky MD. A novel computer based expert decision making model for prostate cancer disease management. *J Urol* 2005; **174**: 2310-2318 [PMID: 16280831]

27 **Margel D**, Harel A, Yossepowitch O, Baniel J. A novel algorithm to improve pathologic stage prediction of clinically organ-confined muscle-invasive bladder cancer. *Cancer* 2009; **115**: 1459-1464 [PMID: 19152435 DOI: 10.1002/cncr.24138]

28 **McFarlane E**, Murphy J, Olmsted MG, Drozd EM, Hill G. America's Best Hospitals 2008 Methodology. US News & World Report 2008 [cited 2014 March 22]. Available from: URL: http://www.usnews.com/usnews/health/best-hospitals/methodology/ABH_Methodology_2008.pdf

29 **Kawaoka T**, Aikata H, Takaki S, Hashimoto Y, Katamura Y, Hiramatsu A, Waki K, Takahashi S, Kamada K, Kitamoto M, Nakanishi T, Ishikawa M, Hieda M, Kakizawa H, Tanaka J, Chayama K. Transcatheter chemoembolization for unresectable hepatocellular carcinoma and comparison of five staging systems. *Hepatol Res* 2010; **40**: 1082-1091 [PMID: 20880059

DOI: 10.1111/j.1872-034X.2010.00714.x]

30   **Malin JL**, Schneider EC, Epstein AM, Adams J, Emanuel EJ, Kahn KL. Results of the National Initiative for Cancer Care Quality: how can we improve the quality of cancer care in the United States? *J Clin Oncol* 2006; **24**: 626-634 [PMID: 16401682]

31   **Desch CE**, McNiff KK, Schneider EC, Schrag D, McClure J, Lepisto E, Donaldson MS, Kahn KL, Weeks JC, Ko CY, Stewart AK, Edge SB. American Society of Clinical Oncology/National Comprehensive Cancer Network Quality Measures. *J Clin Oncol* 2008; **26**: 3631-3637 [PMID: 18640941 DOI: 10.1200/JCO.2008.16.5068]

Baishideng®