

Balanced scorecard-based performance evaluation of Chinese county hospitals in underdeveloped areas

Journal of International Medical Research

2018, Vol. 46(5) 1947–1962

© The Author(s) 2018

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/0300060518757606

journals.sagepub.com/home/imr



Hongda Gao^{1,*}, He Chen^{2,*}, Jun Feng^{1,*},
Xianjing Qin¹, Xuan Wang¹, Shenglin Liang¹,
Jinmin Zhao¹ and Qiming Feng¹

Abstract

Objective: Since the Guangxi government implemented public county hospital reform in 2009, there have been no studies of county hospitals in this underdeveloped area of China. This study aimed to establish an evaluation indicator system for Guangxi county hospitals and to generate recommendations for hospital development and policymaking.

Methods: A performance evaluation indicator system was developed based on balanced scorecard theory. Opinions were elicited from 25 experts from administrative units, universities and hospitals and the Delphi method was used to modify the performance indicators. The indicator system and the Topsis method were used to evaluate the performance of five county hospitals randomly selected from the same batch of 2015 Guangxi reform pilots.

Results: There were 4 first-level indicators, 9 second-level indicators and 36 third-level indicators in the final performance evaluation indicator system that showed good consistency, validity and reliability. The performance rank of the hospitals was $B > E > A > C > D$.

Conclusions: The performance evaluation indicator system established using the balanced scorecard is practical and scientific. Analysis of the results based on this indicator system identified several factors affecting hospital performance, such as resource utilisation efficiency, medical service price, personnel structure and doctor–patient relationships.

Keywords

County hospital, medically underserved area, balanced scorecard, performance evaluation, indicator system, China

Date received: 10 October 2017; accepted: 15 January 2018

¹Guangxi Medical University, Guangxi, China

²Guangxi University of Chinese Medicine, Guangxi, China

*These authors contributed equally to this work.

Corresponding author:

Qiming Feng, Guangxi Medical University, 22 Shuangyong Road, Nanning, Guangxi 530021, China.

Email: fengqm2013@163.com



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<http://www.creativecommons.org/licenses/by-nc/4.0/>)

which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

Introduction

In China, county-level public hospitals are the core providers of medical and health services in each county and form the top level of care in the rural three-tier health-care network. In addition, these institutions connect the medical and health systems of urban and rural areas. Public county hospitals are used for the treatment of common diseases, rehabilitation from serious diseases and the referral of difficult diseases. Public county hospitals also oversee training and guidance for grassroots medical institutions and the management of natural disasters and public health emergencies. In 2009, the Chinese State Council approved the *Opinions of the CPC Central Committee and the State Council on Deepening the Health Care System Reform*¹ and the *Implementation Plan for the Recent Priorities of the Health Care System Reform (2009–2011)*.² These contained five main tasks, one of which was to promote public hospital reform. Furthermore, county-level public hospital reform is a key component of public hospital reform, as it facilitates access to lower-cost medical services. In 2012, the General Office of the State Council issued the *Opinions of Pilot Projects for County-level Public Hospital Reform*,³ which focused on county-level public hospitals and prioritised their development.

Based on central reform guidelines and the local context, Guangxi Province implemented two batches of county-level public hospital reform pilots in 2012 and 2013, which involved 115 hospitals in 40 counties. At the end of 2015, the remaining 103 county-level public hospitals in 36 counties were reformed; thus, the pilots achieved full coverage and substantial advances were made toward the principle of 'ensure a foundation, strengthen the grassroots, construct the mechanism'. To further improve reform and identify problems affecting this process, we need to evaluate the

performance of county-level public hospitals. Since the implementation of county-level public hospital reform in 2012, research in different Chinese provinces has focused on how to establish a set of scientific and effective indicator systems to evaluate county-level public hospital performance. As Guangxi is an underdeveloped region that is home to the Zhuang ethnic minority group, it differs from other provinces in terms of its social customs. Therefore, a matched performance evaluation system for Guangxi county hospitals that closely reflects the social and cultural context is needed.

The balanced scorecard (BSC), introduced by Kaplan and Norton in 1992, is a popular performance management system that categorises organisational goals into four measurable and operable perspectives: Learning and Growth, Financial, Customer and Internal Business Process.⁴ The BSC has been successfully used worldwide in many institutions, such as government units, manufacturing companies, service organisations and non-profit companies.^{5–8} For example, researchers at Duke Children's Hospital in the USA worked with managers using the BSC. After 3 years' implementation of the system, they had turned the hospital's deficit into a profit, reduced costs and increased patient satisfaction.⁹ Early in 1994, the representatives of some Alberta and Ontario hospitals, the University of Toronto and government and policy groups explored the application of the BSC in hospital performance measurement in Canada.¹⁰ The BSC system has also been used in Europe. In the UK, the BSC has been successfully used for key government projects; both the Olympic Delivery Authority and the High Speed 2 railway project have used the BSC to summarise their procurement policies. In addition, the UK Department of Health has used the BSC to evaluate the performance of the National Health Service's

information technology strategy.¹¹ In Sweden, Bern University Hospital has designed a BSC system for the department of anesthesiology¹² and in 2002, the Netherlands launched a campaign to establish performance evaluation indicators for the national health system.¹³ In 2000, the BSC began to be used in healthcare in China and generated a wide range of research and applications.

Methods

Data source

To establish a performance evaluation indicator system for county hospitals, we consulted professional persons in healthcare and studied research on hospital performance evaluation in China and other countries. We generated an indicator framework based on the BSC. Then, we used the Delphi method to modify and improve the framework and produce a final indicator system. We used the indicator system in a case study of five county hospitals randomly selected from the third-batch county hospital reform pilots in Guangxi. To evaluate the hospitals, we used data from questionnaires distributed by the Guangxi Zhuang Autonomous Region Health and Family Planning Commission. The questionnaires were completed by medical staff in the relevant departments and collected by each hospital liaison. Trained investigators obtained patient satisfaction data using one-to-one questionnaire interviews at each hospital. The Topsis method was used with the indicator system to evaluate these hospitals' performance. Microsoft Office Excel 2007 (Microsoft, Redmond, WA, USA) and IBM SPSS Statistics, version 19 (IBM Corp., Armonk, NY, USA) were used for all calculations. The study protocol was approved by the Medical Ethics Committee of Guangxi Medical University. All staff and patients provided

verbal informed consent before the study began.

Establishment of the performance evaluation indicator system

The indicator system framework we constructed was based on BSC theory. The framework was generated by consulting professional persons in healthcare and reviewing research on hospital performance evaluation from China and other countries. Figure 1 shows the performance evaluation indicator framework used in this study. The Delphi method was used to filter the indexes and grade the importance of the indicators. The relative weights of each indicator were determined using the analytic hierarchy process method. Finally, the reliability and validity of the indicator system were tested.

The Delphi method

The importance of each index was categorised according to five levels: very important, important, normal, unimportant and very unimportant. We selected 25 experts from administrative units, universities and hospitals, choosing individuals with a good knowledge of county hospital reform. We administered self-designed questionnaires to the experts, who provided suggestions for modifying the indicator framework and graded the importance of the indicators. This feedback was used to revise the indicator system. Table 1 shows basic demographic information about the experts who participated in the Delphi process.

Reliability of the expert suggestions. We used Cr to test the reliability of the expert suggestions (Cr = the average of Ck , Ca and Cs ; Ck = the knowledge level of experts, Ca = the experts' judgement basis and Cs = the experts' familiarity with each indicator). Larger values of Cr indicated greater expert reliability. Values of $Cr > 0.7$

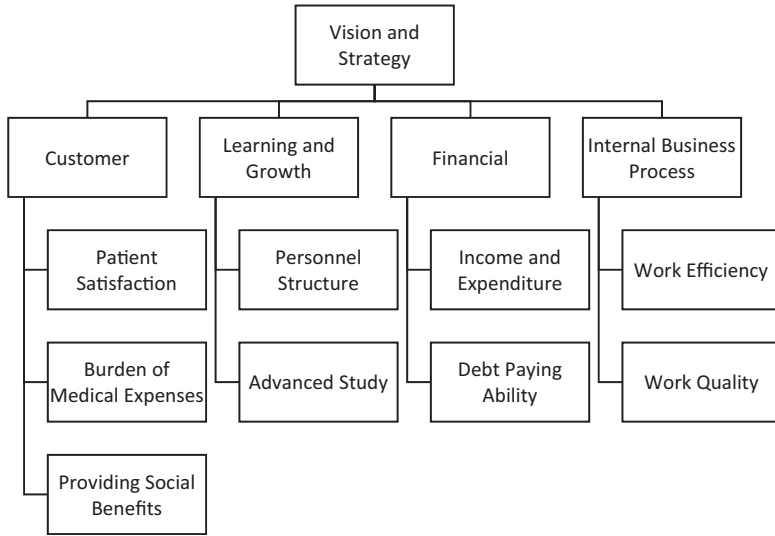


Figure 1. Evaluation indicator framework based on the balanced scorecard.

Table 1. Basic information of experts who participated in the Delphi process

Item	Category	Number	Proportion
Sex	Male	16	64%
	Female	9	36%
Age	<40	1	4%
	40–50	7	28%
	>50	17	68%
Working Years	15–20	3	12%
	>20–30	11	44%
	>30–40	9	36%
	>40	2	8%
Education	Bachelor	11	44%
	Master	8	32%
	Doctor	6	24%
Professional Title	Intermediate Title	3	12%
	Vice-Senior Title	9	36%
	Senior Title	13	52%

titles as 0.7. *Ca* values were based on types of judgement basis: theoretical analysis was scored as 0.8, practical experience as 0.6, knowledge from peers as 0.4 and intuition as 0.2. *Cs* values were based on expert familiarity with each indicator: very familiar was scored as 1.0, familiar as 0.75, generally familiar as 0.50, unfamiliar as 0.25 and very unfamiliar as 0.00.

Concordance of the expert suggestions. Once a consensus of expert opinion is reached, the Delphi process should be concluded. To test the concordance of expert suggestion, we calculated Kendall’s coefficient of concordance (*W*) using Equations (1) and (2). *m* represents the number of experts, *n* represents indicators graded by experts, *R_i* represents the summation of rank assigned to the *i*th indicator.

$$W = \frac{12S}{m^2n(n^2 - 1)} \tag{1}$$

$$S = \sum_{i=1}^n \left[R_i - \frac{m(n+1)}{2} \right]^2 \tag{2}$$

indicated good reliability of expert suggestions.¹⁴ Different criteria were used to assign *Ck*, *Ca* and *Cs* values. *Ck* values were based on each expert’s professional title: senior titles were scored as 1.0, vice-senior titles as 0.9 and intermediate

The analytic hierarchy process method

We then transformed the importance scores of the indicators into index-weighted scores using the analytic hierarchy process method. This method was proposed by T. L. Saaty in 1970 and is a popular multi-criteria decision-making method that combines quantitative and qualitative analysis.¹⁵ It has been widely used to calculate indicator weights in many studies on hospital management, environmental protection and other areas.¹⁶⁻¹⁸ The calculation process is as follows:

1. Based on Saaty’s scale of pairwise comparisons, we translated the importance to value a_{ij} using pairwise comparison between two indicators from the same level.¹⁹ A judgement matrix was then produced: $A = \{a_{ij}\}$.
2. We first calculated the initial weight coefficient W'_i using Equation (3). In Equation (3), m represents the number of indicators in the same level, a_{ij} represents the scale value obtained by pairwise comparison between two indicators. The weight W_i was calculated using Equation (4):

$$W'_i = \sqrt[m]{a_{i1} * a_{i2} * \dots * a_{ij} * \dots * a_{im}} \tag{3}$$

$$W_i = W'_i / \sum_{i=1}^m W'_i \tag{4}$$

3. After obtaining the indicator weights, we needed to determine the degree of consistency to check the logicity of the indicator system. The consistency ratio was calculated (CR , $CR = CI/RI$). Generally, if $CR \leq 0.1$, matrix A is considered acceptable. Otherwise, the matrix needs to be adjusted.²⁰

In Equation (5), CI =the consistency index calculated using Equations (5) to (7), RI =the random index, with values

assigned using Saaty’s scale of pairwise comparisons²¹ and λ_{max} represents the largest eigenvalue. A good consistency is generally assumed if m is no larger than 2; if m is larger than 2, the consistency is acceptable only if CR is less than 0.10.²²

$$CI = (\lambda_{max} - m) / (m - 1) \tag{5}$$

$$\lambda_{max} = \left(\sum_{i=1}^m \lambda_i \right) / m \tag{6}$$

$$\lambda_i = \left[\sum_{j=1}^m (a_{ij} * W_j) \right] / W_i \tag{7}$$

Reliability and validity

After establishing the performance evaluation indicator system, we needed to check its reliability and validity. Reliability was measured using Cronbach’s coefficient alpha: an alpha larger than 0.6 indicated that the factors were reliable.²³ We measured both content validity and construct validity. Construct validity was measured using the Kaiser–Meyer–Olkin (KMO) and Bartlett’s tests. Content validity was assessed according to the source of the information used to develop the system.

Performance evaluation using the Topsis method.

The Topsis method was used to evaluate the performance based on the established indicator system. Topsis (a Technique For Order Preference By Similarity To An Ideal Solution) is an effective multiobjective decision method. Its advantage is that it has no special data requirements and preserves the original data information.²⁴ In addition, the results can be presented in the form of ranks, which is very intuitive. Its calculation steps are as follows:

1. Normalise all data to allow comparisons across criteria. For efficiency indicators,

larger values represent a more positive result, such as the indicator of cure rate. For cost indicators, larger values represent a more negative result, such as the indicator of outpatient expense.²⁵ Negative indicators must be transformed into positive indicators using the reciprocal method or the difference method.

2. Process the data using the normalisation method shown in Equation (8).

$$Z_{ij} = \frac{X_{ij}}{\sqrt{\sum_{i=1}^n X_{ij}^2}}, \quad j = 1, 2, \dots, m \quad (8)$$

3. Find the optimal vector Z^+ and worst vector Z^- and calculate the difference (D^+) between Z_{ij} and Z^+ using Equation (9), and the difference (D^-) between Z_{ij} and Z^- using Equation (10); m represents the number of indicators, n presents the number of hospitals evaluated and a_j represents the weight of each indicator.

$$D_i^+ = \sqrt{\sum_{j=1}^m [a_j(Z_{ij} - Z_j^+)]^2}, \quad i = 1, 2, \dots, n \quad (9)$$

$$D_i^- = \sqrt{\sum_{j=1}^m [a_j(Z_{ij} - Z_j^-)]^2}, \quad i = 1, 2, \dots, n \quad (10)$$

4. Calculate the relative similarity (C_i) between Z_{ij} and the best solution using Equation (11).

$$C_i = \frac{D^-}{D_i^+ + D_i^-}, \quad i = 1, 2, \dots, n \quad (11)$$

Results

The performance evaluation indicator system and weights

All 25 invited experts responded (response rate: 100%). The Crs were 0.84, 0.80, 0.83 and 0.84 for the perspectives of Learning and Growth, Financial, Customer and Internal Business Process, respectively. The average Cr was larger than 0.7, which indicated that the expert suggestions had good reliability. Kendall's coefficient of concordance (W) was 0.277 ($\chi^2 = 235.458$, $P = 0.000 < 0.001$), indicating that the expert opinions were consistent. Based on the Delphi expert opinions, we repeatedly modified the indicators and eventually developed a performance evaluation system with remarkable consistency ($CR < 0.10$). The performance evaluation indicator system contained 4 first-grade indicators, 9 second-grade indicators and 36 third-grade indicators. Table 2 shows the performance evaluation indicator system of Guangxi county-level public hospitals and the weights W_i .

The average Cronbach's alpha was 0.837, which is larger than 0.6 and so indicates good reliability. The average KMO was 0.704, indicating that the data were suitable for factor analysis. Bartlett's test was less than 0.001, indicating that the variables were correlated sufficiently for factor analysis to be performed. The factor analysis showed that the construct validity was acceptable. Furthermore, the development of the indicator system (from the framework construction to the calculation of the weights) had been approved by experts; therefore, the content validity was also appropriate. These tests suggested that our evaluation indicator system could provide reasonable results.

Table 2. Performance evaluation indicator system and weights (W_i)

First-Grade Indicators (Weight W_i)	Second-Grade Indicators (Weight W_i)	Third-Grade Indicators (Weight W_i)	Synthetic Weight
Financial (0.460)	Income and Expenditure (0.529)	% of Government grants in total income (0.398)	0.097
		% of Staff expenses in business expenditure (0.281)	0.068
		% of Drug income in business income (0.213)	0.052
		% of Examination income in medical income (0.054)	0.013
		% of Management expenses in business expenditure (0.054)	0.013
	Debt Paying Ability (0.471)	Asset-liability ratio (0.545)	0.118
		Current ratio (0.233)	0.051
		Quick ratio (0.139)	0.03
		Business income from per 100 RMB fixed assets (0.084)	0.018
Internal Business Process (0.303)	Work Efficiency (0.485)	Rate of bed utilisation (0.303)	0.045
		Average hospitalization days (0.303)	0.045
		Turnover rate of hospital beds (0.165)	0.024
		Physician burden of medical treatment per day (0.165)	0.024
		Physician burden of hospitalization duration per day (0.065)	0.01
	Work Quality (0.516)	Coincidence rate of admission and discharge diagnosis (0.368)	0.058
		Coincidence rate of admission and clinic diagnosis (0.207)	0.032
		Cure rate (0.207)	0.032
		Improvement rate (0.109)	0.017
		Successful recovery rate of inpatients (0.109)	0.017
Learning and Growth (0.094)	Personnel Structure (0.514)	Ratio of doctors to nurses (0.331)	0.016
		Ratio of beds to nurses (0.331)	0.016
		% of Vice-senior titles or above in health technical professionals (0.146)	0.007
		% of Health technical professionals in all employees (0.096)	0.005
		% of Junior college education or above in all employees (0.096)	0.005
	Advanced Study (0.486)	Frequency per medical worker of further study in upper-level hospitals (0.511)	0.023
		Frequency per medical worker of external short-term training (0.490)	0.022

(continued)

Table 2. Continued

First-Grade Indicators (Weight W_i)	Second-Grade Indicators (Weight W_i)	Third-Grade Indicators (Weight W_i)	Synthetic Weight
Customer (0.143)	Patient Satisfaction (0.570)	Inpatient satisfaction (0.582)	0.047
		Outpatient satisfaction (0.348)	0.028
	Burden of Medical Expenses (0.333)	Number of medical disputes per 1000 discharged patients (0.070)	0.006
		Expenses per inpatient (0.400)	0.019
		Hospitalization expenses per day (0.400)	0.019
		Expenses per outpatient (0.200)	0.01
	Providing Social Benefits (0.097)	% of Public welfare expenses in total expenditure (0.420)	0.006
		Frequency per 100 medical workers of training basic medical unit staff (0.269)	0.004
		Frequency per 100 medical workers of undertaking sudden public health events and emergency medical rescue (0.190)	0.003
		Frequency per 100 medical workers of providing counterpart assistance to basic medical units (0.121)	0.002

Table 3. Consistency index (C_i) and ranks for four balanced scorecard perspectives

Hospital (A–E)	Financial		Internal Business Process		Learning and Growth		Customer		Total Performance	
	C_i	Rank	C_i	Rank	C_i	Rank	C_i	Rank	C_i	Rank
A	0.52	3	0.68	1	0.81	1	0.61	3	0.47	3
B	0.83	1	0.55	2	0.75	2	0.06	5	0.67	1
C	0.33	4	0.46	3	0.06	5	0.18	4	0.42	4
D	0.02	5	0.30	4	0.12	4	0.73	2	0.18	5
E	0.76	2	0.12	5	0.44	3	0.80	1	0.61	2

Performance evaluation results calculated using the Topsis method

Tables 3 to 6 show the initial data from the five county hospitals according to the four BSC perspectives. Table 7 shows the C_i and performance ranks of the five county hospitals from the four BSC perspectives and shows the total ranks. For example, hospital B performs the best and hospital D the worst; hospital A is the best in Internal Business Process and Learning and

Growth. We discussed these results with the experts and confirmed their agreement of the interpretation.

Discussion

Many methods are currently used to evaluate performance, such as the key performance indicator method, the target management method and the data envelope analysis method.^{26–30} However, many of

Table 4. Financial indicator data for hospitals A–E

Hospital (A–E)	Income and Expenditure				Debt Paying Ability				Business Income from per 100 RMB Fixed Assets (RMB)
	% of Government Grants in Total Income (%)	% of Staff Expenses in Business Expenditure (%)	% of Drug income in Business Income (%)	% of Examination Income in Medical income (%)	% of Management Expenses in Business Expenditure (%)	Asset-liability Ratio (%)	Current Ratio (%)	Quick Ratio (%)	
A	8.48	29.25	40.47	4.77	8.39	43.00	136.00	126.00	222.52
B	4.93	37.77	21.13	9.47	15.68	30.81	177.00	159.00	110.38
C	9.21	27.31	41.59	6.86	2.98	19.44	123.30	113.20	134.24
D	11.71	23.60	34.38	8.63	13.98	38.71	60.23	45.54	126.98
E	6.80	46.34	22.74	11.71	1.66	28.80	131.43	86.42	81.78

Table 5. Internal business process indicator data for hospitals A–E

Hospital (A–E)	Work Efficiency				Work Quality				
	Rate of Bed Utilisation (%)	Average Hospitalization Days (%)	Turnover Rate of Hospital Beds (%)	Physician Burden of Medical Treatment per Day (Nos.)	Physician urden of Hospitalization Duration per Day (Nos.)	Coincidence Rate of Admission and Discharge Diagnosis (%)	Coincidence Rate of Admission and Clinic Diagnosis (%)	Improvement Rate (%)	Successful Recovery Rate of Inpatients (%)
A	104.50	8.50	44.90	11.10	1.20	98.90	98.10	57.70	93.60
B	91.20	7.30	45.10	9.40	4.30	99.50	96.00	36.80	89.10
C	74.30	6.20	44.70	4.60	2.90	92.70	92.90	57.50	94.30
D	86.40	6.90	45.60	4.40	2.90	99.20	98.10	77.60	94.50
E	72.65	8.30	36.84	4.00	3.10	96.70	97.50	49.36	93.10

Table 6. Learning and growth indicator data for hospitals A–E

Hospital (A–E)	Personnel Structure					Advanced Study	
	Ratio of Doctors to Nurses (%)	Ratio of Beds to Nurses (%)	% of Vice-senior Titles or Above in Health Professionals (%)	% of Health Technical Professionals in All Employees (%)	% of Junior College Education or Above in All Employees (%)	Frequency per Medical Worker of Further Study in Upper-Level Hospitals (%)	Frequency per Medical Worker of External Short-term Training (%)
A	0.46	201.88	2.33	77.90	82.61	3.52	36.70
B	0.52	171.96	3.45	80.67	83.67	1.36	32.28
C	0.64	147.52	3.70	89.24	74.23	1.05	28.42
D	0.50	189.00	7.13	78.66	79.54	2.71	36.36
E	0.84	111.29	1.36	92.45	97.28	2.31	30.51

Table 7. Customer indicator data for hospitals A–E

Hospital (A–E)	Patient Satisfaction		Burden of Medical Expenses			Providing Social Benefits			Frequency per 100 Medical Workers of Undertaking Sudden Public Health Events and Emergency Medical Rescue (Nos.)
	Inpatient Satisfaction (Score)	Outpatient Satisfaction (Score)	Number of Medical Disputes per 1000 Discharged Patients (Nos.)	Expenses per Inpatient (RMB)	Hospitalization Expenses per Day (RMB)	Expenses per Outpatient (RMB)	% of Public Welfare Expenses in Total Expenditure (%)	Frequency per 100 Medical Workers of Providing Counterpart Assistance to Basic Medical Units (Nos.)	
A	85.71	86.36	2.00	4072.50	508.70	76.49	0.37	8.00	
B	76.19	56.00	9.00	4581.00	452.70	103.00	0.07	6.00	
C	83.33	54.17	6.00	3580.00	582.00	149.80	0.38	13.00	
D	95.00	80.00	4.00	4000.00	500.00	75.00	0.50	9.00	
E	86.96	80.95	2.00	2540.27	450.00	82.15	0.21	4.00	

these methods have shortcomings. For example, some performance evaluation methods focus on economic indicators and ignore the growth and development of medical staff, patient satisfaction and internal processes. Some methods place too much emphasis on objective indicators or, conversely, only use subjective surveys and thus lack an objective perspective. In addition, the theoretical foundation of some evaluation indicator systems is not comprehensive and relies on personal experience or judgement instead of consultation with relevant stakeholders. Although their performance evaluation goal is the same, indicator systems vary across different provinces. In view of the shortcomings of previous methods, this study used the BSC to establish an indicator system framework from four perspectives. The Delphi method was used to modify and expand the framework based on expert opinions. This study is the first to combine the BSC with performance evaluation for Guangxi county hospitals; as such, the results may be very useful for Guangxi hospital reform. The results indicated that the level of expert authority was high and the expert opinions tended to be consistent, suggesting that the reliability of the expert suggestions can be trusted. The indicator system was developed based on these expert opinions. As the system showed good reliability and validity, the results of the performance evaluation can be assumed to be accurate.

Analysis of performance evaluation indicator system

The weightings of the first-level indicators showed the following relationship: Financial > Internal Business Process > Customer > Learning and Growth. Each indicator had a different weight at different levels and further analysis of the indicators is discussed below.

Financial perspective

A government policy to cancel drug price increases has meant that all drugs must be sold at their purchase price. Because of this, hospitals have lost some of their income. To balance the income gap, the government has introduced measures such as adjusting the price of medical services, increasing government subsidies, strengthening hospital accounting and saving on running costs. However, these measures have had some negative effects such as inadequate compensation in some areas and inconsistent adjustment of medical service prices, which can make hospitals appear to be operating poorly.^{31,32} To meet growing medical demands, county hospitals purchase large medical devices, introduce medical expertise and develop advanced medical technology, all of which increases hospital debt. To prevent the reappearance of these problems in the new health care reforms, attention must be paid to good management of funds and efficient medical service price adjustments. Improper use of funds wastes health resources and affects the development of county hospitals. Therefore, public subsidies need to be used properly, medical service prices adjusted on a scientific basis and assets and liabilities controlled properly. The effective management of hospital finances would have a substantial effect on the development of county hospitals.

Internal business process and customer perspective. Finance was identified as the primary problem, but other issues are also important. Both Internal Business Process and Customer indicators are correlated with Finance. As mentioned above, the cancellation of drug price increases has substantially reduced hospital income (Finance). This is likely to reduce the salaries of medical staff and so decrease their enthusiasm for work, which affects work

efficiency (Internal Business Process). The Internal Business Process indicator measures work efficiency and work quality status in county hospitals. The Customer indicator measures patient satisfaction with medical services. These two indicators reflect the patient-oriented approach of county hospitals, which are public welfare institutions. Internal Business Process had a greater weighting than Customer because the primary task of county hospitals is to guarantee the quality of medical services and work efficiency. Customer satisfaction is affected by many subjective factors like medical service quality, the service attitude of medical staff and media orientation. Regarding the scientific basis and reliability of performance evaluation, objective indicators have more stability and accuracy than subjective indicators, which may explain why Internal Business Process has a higher weight index than Customer.

Learning and growth perspective. Learning and Growth was ranked last of the four indicators for the following reasons. According to Chinese healthcare system reform policy, the goal of county hospitals is to treat common diseases, transfer patients suffering from difficult and complicated diseases, provide rehabilitation for patients with serious diseases, provide medical guidance and training to personnel in rural areas and oversee public services such as infectious disease control, natural disasters and emergency rescue. The central work of county hospitals focuses on regional medical treatment and public health, which require more practical work than teaching or scientific research. This explains why those indicators have a lower weight. However, county hospitals require a certain number of physicians, nurses and psychiatric beds to ensure medical quality and efficiency, which explains the higher weight for personnel structure. However, there is a lack of high-level talent in most county hospitals

in China, (and little difference among county hospitals on this factor); therefore, it is meaningless to try to evaluate this indicator.³³ Furthermore, the flow of talented personnel is affected by regional economy and policy, which county hospitals cannot control. Counties in Guangxi Province are characterised by poor economy, education, living environment and access to cities; therefore, county hospitals will continue to experience problems in attracting talented personnel until the government implements policies to relieve these problems. Therefore, the personnel structure of the hospitals did not reflect a full range of talent and so this indicator was assigned a small weight.

Analysis of performance evaluation results

Hospital B was ranked first on performance. Hospital B scored highest on Finance, indicating that it would be relatively easy for this hospital to improve technology or to employ good staff. Moreover, the ratio of hospital B drug income was the lowest and the examination income ratio was similar to the best, which indicates that hospital B performed well in cancelling drug price increases and adjusting the examination price. Hospital B was ranked second on physician burden of medical treatment per day, which shows a good performance in treating common diseases of local residents. However, hospital B was ranked lowest on patient satisfaction; this result could be attributed to the large burden of medical staff. Excessive workloads can lead to staff being less patient and having a poor attitude to patients.

Hospital D was ranked last on performance. From a Finance perspective, the financial structure of hospital B was unscientific; government grants formed the main part of hospital income and management expenses were the main outgoing. From an Internal Business Process perspective,

the physician burden of medical treatment per day was small and the turnover rate of hospital beds was low, which indicated that there were few patients and some beds were superfluous. From the Learning and Growth perspective, hospital D had a high ratio of beds to nurses and the staff structure was problematic: the ratio of health technical staff was low whereas the ratio of executives was high. However, hospital B scored highest on the Customer perspective, because it undertook more social welfare services and public health events than the other four hospitals. Because of its involvement in public services, hospital D received less revenue from medical services, which partly explains its poor medical performance.

Finally, from the Learning and Growth and Internal Business Process perspectives, hospital A performed well on medical quality, with a high utilisation ratio and many patients, which meant that hospital A scored well on treating common diseases of residents in county areas. Hospital B scored less than hospital A on patient expenses and drug income proportion, which is beneficial for patients. That is to say, hospital B performed better on solving the problem of expensive medical treatment. More importantly, hospital B had a higher score on the Financial perspective, and (because Finance was assigned the largest weight) therefore the overall performance score of hospital B was higher than that of hospital A.

Suggestions for the development of county hospitals

In terms of basic investment, the government should strictly control hospital construction criteria, bed numbers and the purchase of large equipment. Furthermore, it should forbid construction or the buying of large equipment if a hospital is in debt. County hospitals should

adjust the number of beds according to county resident numbers. Once it reaches the standard scale set out in the national plan, a hospital should be barred from further expansion. Hospitals that exceed the standard or begin construction while in debt should be held accountable.

To reduce patient burden, county hospitals should set a reasonable price for medical services. The Guangxi government has implemented a zero margin drug profit policy and has claimed that county hospitals could address the income loss by adjusting medical service prices, saving costs and obtaining more government grants. However, price adjustments must reflect the labour value of medical staff while considering factors such as county economic development, medical insurance payment capacity and the medical cost burden of residents. County hospitals could obtain extra revenue by providing high-quality or distinctive services and reducing the cost of medical consumables and large medical equipment.

Addressing the shortage of qualified professional personnel is the most important issue for county hospital performance. To solve this problem and attract professionals from higher-level hospitals, a mechanism is needed to increase the personnel flow between urban and rural hospitals. County hospitals could introduce high-quality professionals using project employment, task employment or skills cooperation. However, to attract speciality or scarce personnel, or to address urgent staff shortages, hospitals should increase recruitment by reducing some requirements, such as education and age, and simplify the recruitment procedure. Furthermore, county hospitals should provide focal training to medical staff in key business positions and train core doctors while encouraging them to obtain in-service education.

Improving patient satisfaction and creating good relationships between doctors and

patients is also beneficial for performance. Further education in the humanities is first needed for medical staff to strengthen their understanding of medical ethics and retain professionalism. Then, the media needs to strengthen publicity and guide public opinion to encourage people to respect and value health workers. County hospitals should perfect their patient complaint mechanisms and ethics committees should be established to investigate complaints about improper medical behaviour and improve communication channels. If necessary, local government should establish a medical dispute resolution body to ensure the appropriate regulation of medical services. To guarantee the lawful rights and interests of doctors and patients, medical violence must be strictly prohibited. Finally, it is necessary to develop medical accident insurance and medical liability insurance, and to establish a mechanism for sharing medical risk between doctors and patients.

Future research prospects

This study has some limitations. Using the BSC, we evaluated the performance of Guangxi county hospitals from an academic perspective and provided some recommendations for hospital reform. The large number of indicators makes this performance evaluation system problematic to implement in terms of cost and efficiency; further refinement of the system is needed before it can be fully implemented. Because of funding and personnel limitations, we only selected five county-level public hospitals for this case study. Therefore, the system needs to be tested further on a larger sample of hospitals. In addition, the applicability of the performance evaluation system for other types of county-level hospitals, such as Chinese medicine hospitals, needs further investigation.

In future research, we plan to apply this performance evaluation system to additional county hospitals. We are also aiming to expand the range of this case study and explore the use of the indicator system in other types of hospitals, such as county-level Chinese medicine hospitals and maternal and child health care hospitals. In addition, to verify the evaluation results, we aim to compare the suitability of different methods to evaluate performance, such as the comprehensive index method and the rank sum ratio method.

We are also planning further studies using this system to evaluate the performance of hospital departments. These performance results will be combined with management data to provide more comprehensive recommendations for hospital development and decision making.

Author contributions

Hongda Gao generated the initial idea for the study, analysed the data and wrote the manuscript. He Chen and Jun Feng revised the manuscript and modified the English language. Qiming Feng, the corresponding author, designed the study project and provided the funding sources. Jinmin Zhao, a co-corresponding author, participated in designing the study and carried out the study project. Xuan Wang, Shenglin Liang and Xianjing Qin participated in data collection and cleaning. All authors read and approved the final manuscript.

Acknowledgments

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The authors would like to thank the Guangxi Zhuang Autonomous Region Health and Family Planning Commission for research coordination and data preparation, and thank all participants in this study for their cooperation.

Declaration of conflicting interest

The authors declare that there is no conflict of interest.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

- Opinions of the CPC Central Committee and the State Council on Deepening the Health Care System Reform. The Central People's Government of the People's Republic of China Website. http://www.gov.cn/test/2009-04/08/content_1280069.htm. Published March 17, 2009. Accessed February 5, 2017.
- Implementation Plan for the Recent Priorities of the Health Care System Reform (2009–2011). The Central People's Government of the People's Republic of China Website. http://www.gov.cn/zwgk/2009-04/07/content_1279256.htm. Published March 18, 2009. Accessed February 5, 2017.
- Opinions of Pilot Projects for County-level Public Hospital Reform. The Central People's Government of the People's Republic of China Website. http://www.gov.cn/zwgk/2012-06/14/content_2161153.htm. Published June 7, 2012. Accessed February 5, 2017.
- Kaplan RS and Norton DP. The balanced scorecard—measures that drive performance. *Harv Bus Rev* 1992; 70: 70–79.
- Madsen D and Slåtten K. The role of the management fashion arena in the cross-national diffusion of management concepts: the case of the balanced scorecard in the Scandinavian countries. *Adm Sci* 2013; 3: 110–142.
- Stas D, Lenort R, Wicher P, et al. Green Transport Balanced Scorecard Model with Analytic Network Process Support. *Sustainability* 2015; 7: 15243–15261.
- Lin Z, Zengbiao Y and Zhang L. Performance outcomes of balanced scorecard application in hospital administration in China. *China Agricultural Economic Review* 2014; 30: 1–15.
- Falle S, Rauter R, Engert S, et al. Sustainability management with the sustainability balanced scorecard in SMEs: findings from an Austrian case study. *Sustainability* 2016; 8: 545.
- Meliones JN, Ballard R, Liekweg R, et al. No mission no margin: it's that simple. *J Health Care Finance* 2001; 27: 21–29.
- Baker GR and Pink GH. A balanced scorecard for Canadian hospitals. *Healthcare Management Forum* 1995; 8: 7–13.
- Procuring Growth Balanced Scorecard. The UK Government Website. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/560247/Balanced_Scorecard_paper.pdf. Published October 14, 2016. Accessed February 10, 2017.
- Voelker KE, Rakich JS and French GR. The balanced scorecard in healthcare organizations: a performance measurement and strategic planning methodology. *Hospital Topics* 2001; 79: 13–24.
- Ten AA, Arah OA, Geelhoed J, et al. Developing a national performance indicator framework for the Dutch health system. *Int J Qual Health Care* 2004; 16: 155–171.
- Jianli Z. *Public hospital performance evaluation indicator system and evaluation system design*. Huazhong University of Science & Technology, 2012.
- Zhang S, Wei Z, Liu W, et al. Indicators for environment health risk assessment in the Jiangsu Province of China. *Int. J. Environ. Res. Public Health* 2015; 12: 11012–11024.
- Xu S, Wang T and Hu S. Dynamic assessment of water quality based on a variable fuzzy pattern recognition model. *Int. J. Environ. Res. Public Health* 2015; 12: 2230–2248.
- Li C and Yu C. Performance evaluation of public non-profit hospitals using a bp artificial neural network: the case of Hubei Province in China. *Int. J. Environ. Res. Public Health* 2013; 10: 3619–3633.
- Yang X and Ma H. Natural environment suitability of China and its relationship with population distributions. *Int. J. Environ. Res. Public Health* 2009; 6: 3025–3039.
- Anvari AR, Mojahed M, Zulkifli N, et al. A group ahp-based tool to evaluate effective factors toward leanness in automotive

- industries. *Journal of Applied Sciences* 2011; 11: 3142–3151.
20. Wang TK, Zhang Q, Chong HY, et al. Integrated supplier selection framework in a resilient construction supply chain: an approach via analytic hierarchy process (AHP) and grey relational analysis (GRA). *Sustainability* 2017; 9: 289.
 21. Lin C, Kou G and Ergu D. A statistical approach to measure the consistency level of the pairwise comparison matrix. *J. Oper. Res. Soc* 2014; 65: 1380–1386.
 22. Sang S, Wang Z and Yu C. Evaluation of health care system reform in Hubei Province, China. *Int. J. Environ. Res. Public Health* 2014; 11: 2262–2277.
 23. Wang M, Fang H, Bishwajit G, et al. Evaluation of rural primary health care in Western China: a cross-sectional study. *Int. J. Environ. Res. Public Health* 2015; 12: 13843–13860.
 24. Lu C, Xue B, Lu C, et al. Sustainability investigation of resource-based cities in Northeastern China. *Sustainability* 2016; 8: 1058.
 25. Pavlič I, Portolan A and Puh B. (Un)supported current tourism development in UNESCO protected site: the case of Old City of Dubrovnik. *Economies* 2017; 5: 9.
 26. Hübnerbloder G and Ammenwerth E. Key performance indicators to benchmark hospital information systems - a Delphi study. *Methods Inf Med* 2009; 48: 508–518.
 27. Yan X, Zhuoyun W and Dian Z. Study on the construction of management by objective system of clinical departments in the comprehensive hospital. *Chinese Hospital Management* 2014; 34: 21–23.
 28. Prakash V and Annapoorni D. Performance evaluation of public hospitals in Tamil Nadu: DEA approach. *J Health Manag* 2015; 17: 417–424.
 29. Farzianpour F, Aghababa S, Delgoshaei B, et al. Performance evaluation a teaching hospital affiliated to Tehran University of medical sciences based on Baldrige excellence model. *American Journal of Economics & Business Administration* 2014; 3: 272–276.
 30. Thomas E, Dyson M and Clerc M. An analysis of performance evaluation for motor-imagery based BCI. *J Neural Eng* 2013; 10: 031001.
 31. Minmin W. Discussion on causes, risks and countermeasures of developing in debt for public hospitals. *Modern Economic Information* 2015: 173–173.
 32. Ronggui C. Totally improve county level professional capacity by activity promoting integrated reform. *Chinese Hospitals* 2011; 15: 1–4.
 33. Xiaoyan Z, Zhenwei L and Xiao F. Thinking on problem of talents in the reform of county public hospitals. *Chin J Hosp Admin* 2014; 30: 171–173.