Received 13 January 2011
Revised 21 September 2012
Accepted 21 September 2012

Advancing the use of performance evaluation in health care

Andreas Traberg and Peter Jacobsen
Department of Management Engineering, Technical University of Denmark, Lyngby, Denmark, and
Nadia Monique Duthiers
Radiology Department, Hospital of Southern Jutland, Kyngby, Denmark

Abstract

Purpose – The purpose of this paper is to develop a framework for health care performance evaluation that enables decision makers to identify areas indicative of corrective actions. The framework should provide information on strategic pro-/regress in an operational context that justifies the need for organizational adjustments.

Design/methodology/approach – The study adopts qualitative methods for constructing the framework, subsequently implementing the framework in a Danish magnetic resonance imaging (MRI) unit. Workshops and interviews form the basis of the qualitative construction phase, and two internal and five external databases are used for a quantitative data collection.

Findings – By aggregating performance outcomes, collective measures of performance are achieved. This enables easy and intuitive identification of areas not strategically aligned. In general, the framework has proven helpful in an MRI unit, where operational decision makers have been struggling with extensive amounts of performance information.

Research limitations/implications – The implementation of the framework in a single case in a public and highly political environment restricts the generalizing potential. The authors acknowledge that there may be more suitable approaches in organizations with different settings.

Practical implications – The strength of the framework lies in the identification of performance problems prior to decision making. The quality of decisions is directly related to the individual decision maker. The only function of the framework is to support these decisions.

Originality/value – The study demonstrates a more refined and transparent use of performance reporting by combining strategic weight assignment and performance aggregation in hierarchies. In this way, the framework accentuates performance as a function of strategic progress or regress, thus assisting decision makers in exerting operational effort in pursuit of strategic alignment.

Keywords Performance measurement, Decision support, Health services sector, Strategic alignment, Holistic performance

Paper type Research paper

Performance evaluation in health care

Managing modern health care is becoming increasingly complicated as institutions evolve into integrated health systems comprising hospitals, outpatient clinics and surgery centers, nursing homes, and home health services (Curtright *et al.*, 2000). In addition, increasing demands for individualized, high-performance services, intensified patient inflow and technological innovations all result in rising pressure on health expenditures (Strandberg-Larsen *et al.*, 2007; World Health Organization, 2008). This development has led to a growing need for more reliable performance evaluation tools in order to guide the increasingly complex decision-making processes (Swaminathan *et al.*, 2008). But health care service performance is often difficult to quantify (Basu *et al.*, 2010), and numerous methods have been suggested and debated as tools for assessing the performance and quality of health care services (Mohammadi *et al.*, 2007). Accurate diagnosis and treatment are no longer enough; stakeholders

Emerald

Journal of Health Organization and Management
Vol. 28 No. 3, 2014
pp. 422-436

1477-7266
DOI 10.1108/JHOM-01-2011-0004

require high performance in all facets of modern health care (Elleuch, 2008). This signifies that health care organizations need to move beyond a narrow medical view and embrace a holistic approach to the concept of health care performance. In the attempt to provide holistic performance information, measurement systems have consequently become more wide-ranging (Cheng and Thompson, 2006; Lega and Vendramini, 2008). Today, consensus exists about defining performance in relation to explicit goals that reflect the values and requirements of various stakeholders (such as patients, professions, regulators, etc.). Therefore, performance evaluation has evolved into a multi-faceted concept, which includes patient load analysis (Mital, 2010), work environment (Jones *et al.*, 2009), patient satisfaction (Kutney-Lee *et al.*, 2009), mortality rates (Barros, 2003), surgical performance (Treasure *et al.*, 2002), and incentive structures (Buetow, 2008) to name a few examples of the extensive work conducted within specific health care performance evaluation. Because of a common acceptance of the strength in both high-level and specific measures, many health care facilities are adopting both types within their internal evaluation procedures. Likewise, national and international agencies are evaluating the performance of health care services to an extent not previously seen. Performance indicators, quality audits, and accreditation standards are therefore gradually becoming fundamentals in the vocabulary of health care professionals around the world.

This variety of indicators to be measured has, however, left the health care industry without a unanimously accepted framework as a tool for measuring the quality and performance of health care services (Ondategui-Parra *et al.*, 2004). Disagreement arises from the fact that performance indicators in health care are inherently controversial, because they require an operational and clear-cut definition in order to be measured. As a result, the use and development of new measurement systems for evaluating health care performance has been rapidly evolving, and debated, in recent decades (Folan and Browne, 2005). In pursuit of accountability in all aspects, development and usage of comprehensive measurement systems are common practice, where variations of the balance scorecard (BSC) concept are by far the system most utilized (Bloomquist and Yeager, 2008; Curtright *et al.*, 2000; Neely, 2005). Most of these measurement concepts are well documented and well executed, but they pose one significant drawback. Even when arranged in comprehensive performance measurement systems, indicators are still stand-alones with no apparent relation to each other. The undying weakness of performance measurement systems is that assembly of a vast selection of self-contained indicators does not in itself provide an overview for individual decision makers (Loeb, 2004). Many organizations have turned to complex measurement systems, but surprisingly few have succeeded in presenting operational guidelines for how to analyze models with multiple performance measures (Matta and Patterson, 2007). Practitioners experience these fragmented structures to be heavy to administrate, and the feedback is often overwhelming and confusing (Kocakülâh and Austill, 2007). Thus, in many cases, the expansion of the administrative burden has not provided more operational value for health care organizations – just more work. Managers and a high percentage of operational employees spend huge portions of their time on administrative tasks related to reporting on performance and quality initiatives. Contradictory to the initial objective, the expanding quantity of registrations, reports, standards, budgets, etc., has limited the organizations' ability to make use of all the information at hand. Decision makers are constantly faced with a vast selection of indicators, which in some cases lead to administrative fatigue and information overload (Bovier and Perneger, 2003). Few employees are able to understand and grasp all the information produced at modern health care facilities. Unfortunately, the result is that many decisions are not based on quantitative data, even though the evidence is available. Instead, they are based on more subjective assessments and the risk is that they are not aligned with the organization's strategy (Ormrod, 1993). In order to comprehend all available information, more and more employees need to be responsible for sub-parts of the organizational decision-making hierarchy. This segregation of tasks into smaller areas of responsibility seems to be an obvious structural response (Evans and Weir, 1995), but such segregation demands considerably more from the managing processes within the organization (Walley *et al.*, 2006). The result is that decision making is moved away from the operational levels of health care organizations and into the strategic levels, thus prolonging the ability to make corrective adjustments and delaying necessary changes.

This is a key concern, because it can lead to a descending performance spiral (administrative tasks vs operational productivity), where operational employees spend their time on administrative tasks instead of value-adding activities. Hence, if performance information is to be used as proactive decision support without increasing the administrative burden, the representation of organizational performance needs to be changed. A more intuitive and holistic representation is needed in order to ease the identification of performance problems throughout the organization.

Motivation and methodology

It is assumed that comprehensible performance information, along with an intuitive representation, is a necessity for modern health care organizations to reach their strategic objectives. This paper therefore focusses on the question of constructing holistic, aggregated performance information, capable of portraying strategic change based on operational performance measures. The aggregated measures must provide information about current operational performance compared to past performance, in order to represent strategic pro-/regress. The framework needs to justify whether operations are under statistical control or not according to the strategic objectives of an organization. This signifies that the framework detects large variations in lower level indicators, even though upper level indicators are assessed stabile, as extreme fluctuations are expected to have impact on general operations. The advancement of performance measurement systems is considered a key step toward improving the productivity/performance of the health care sector. The motivation for this study has been a wish to contribute to this advancement by focussing on holistic performance measurement, taking strategic objectives as the point of departure and transforming this into decision-support information for operational management.

The study is performed as a single case study (Morgan and Morgan, 2009; Voss *et al.*, 2002), performed by external researchers in close collaboration with the staff at a Danish radiology department. The empirical focus of this study is on the magnetic resonance imaging (MRI) modality. The framework has been developed in two steps: a qualitative development phase and a quantitative test phase. The qualitative development of the framework was performed in collaboration with staff members, who participated in workshops and interviews (Winter and Munn-Giddings, 2001). The justification for using workshops is that they provide the possibility of reflection at a higher level than during interviews. In preparation for these workshops, interviews were used as a promoter of discussion topics. All internal process-related data were collected from the Radiology Information System/Picture Archiving and Communication System (RIS/PACS), and Human Resources (HR) data were collected from the Hospital Information System (HIS). External data were collected from five federal units and government agencies:

(1) the Unit of Patient-Perceived Quality's survey of patients' experiences in Danish hospitals (The Unit of Patient-Perceived Quality's web site, 2009);

(2) the Danish Quality framework (The Danish Institute for Quality and Accreditation in healthcare web site, 2010);

(3) the National Indicator Project (the National Indicator Project's web site, 2009);

(4) patient safety records created by the National Board of Health (The National Board of Health's web site, 2009); and

(5) the Danish Patient-Safety Database (The Danish Patient Safety Database's web site, 2010).

It is important to note that all external data are publicly available and validated by the federal units and governmental agencies issuing them. The data collection has striven to base the performance measures on already collected, validated, and published data, with an eye to enhancing the credibility and validity of the outcomes.

Structural outline

In the construction of any decision support framework two key issues have to be properly addressed; the indicators to be measured (Flapper *et al.*, 1996), and how to evaluate these indicators (Dummer, 2007). In order to deal with these two key issues, this particular framework is divided into four successive steps:

(1) Selection and placement of indicators in hierarchies. The selection of suitable indicators is regarded as being of critical importance, because it establishes the organization's goals and priorities (Neely *et al.*, 1994).

(2) Creating a structure for mutually weighing the indicators, assigning mathematical weight in accordance with strategic significance. If the weights are not strategically aligned, the usefulness of the information as decision support is assumed limited.

(3) Normalization of outcomes. Normalization assigns a dimensionless quantity to indicators, thereby making them comparable regardless of initial value.

(4) The aggregation procedure calculates an aggregated performance index. This procedure provides information on change according to past performance.

Combined, these four steps constitute the guiding structure which allows for the interpretation of performance in relation to strategic objectives. The specifics in each of these steps are adapted from state-of-the-art performance measurement proposals and fitted to the particular settings constituting modern health care.

Step 1: selection and placement of indicators in hierarchies

To fulfill the intention to present performance by a few key measures characterizing overall performance, indicators are structured in a hierarchy (see Figure 1). The aggregated performance index is thereby represented as a common denominator for all indicators included in the framework. The justification for adopting the concept of aggregated measures is that it can incorporate a vast array of information and at the same time decrease complexity (Jollands *et al.*, 2003). This approach has previously been used by Nakajima (1986), who introduced the use of aggregated indicators in his

Overall-Equipment-Efficiency (OEE) indicator. The OEE included availability, performance, and quality combined into a single measure. The aggregation of indicators provides an index of performance which is a representation of all lower level indicators included.

The hierarchy consists of four levels: a superior aggregated indicator, strategic dimensions, indicator clusters, and performance indicators. The selection of suitable indicators for a given department is individual (Evans, 2004); e.g. an intensive ward is likely to choose other indicators than the radiology department. Thus, the decision makers who are to use the framework need to be an integrated part of the construction of hierarchies. In this process, they guide the selection of dimensions, clusters, and indicators related to operational users and stakeholders within the department (Matta and Patterson, 2007; Moullin, 2004). This particular construction is useful in tracing poor performance backwards through the hierarchy, since poor performance outcomes will be reflected in the upper levels of the hierarchy. By enhancing transparency, the root cause of any overall performance problems is more easily identified.

Step 2: weighting of indicators

When incorporating several indicators into a performance information system, the indicators will inevitably be of different strategic importance. Without individually assigned weights, indicators in "large" clusters will mathematically seen have less weight than indicators in "small" clusters as long as the comparison is made by simple average values. This arrangement constitutes a problem, because some indicators merely support a decision, while others govern which decision is made. To compensate for this, the framework adapts the concept of analytical hierarchy process (AHP) (Saaty, 1982) to make a quantitative distinction between the indicators, the clusters, and the dimensions within the framework. AHP is a multiple criteria, decision-making approach that assigns mathematical weights based on either qualitative assumptions or quantitatively underpinned arguments. The application of this approach allows subjective as well as objective factors to be considered in a decision-making process (Dey *et al.*, 2008).

Once the hierarchy of suitable indicators is constructed, the decision makers conduct a systematic pair-wise comparison of the incorporated indicators, assigning values of relative intensity to each individual indicator within a cluster (see Figure 2). If, e.g. indicator P_1 is assessed to be of extreme importance (score of 9) in relation to P_2, then the intensity (I) of 9 is placed on the I_{12} position and the reciprocal value (score of 1/9) on the I_{21} position. The mathematical construction of the matrix signifies that each indicator's comparative intensity within the cluster is stated as the sum of each column.

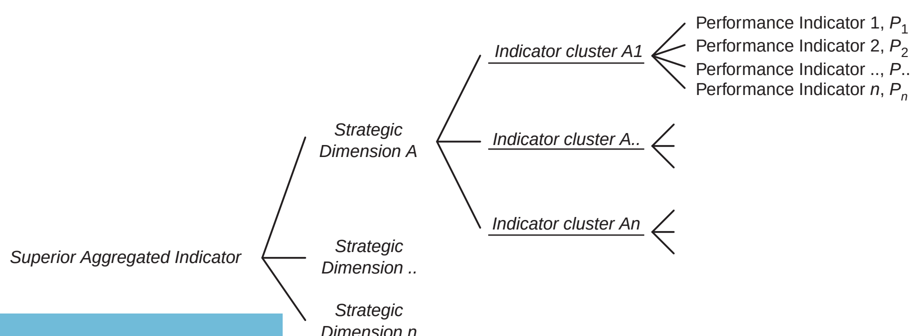

Figure 1.
Structural principle of indicator hierarchy

Scale for pair wise comparison		
Intensity	*Definition*	*Explanation*
1	Equal importance	Two elements contribute equally
3	Moderate importance	Sight favor of one indicator over another
5	Strong importance	Strong favor for one indicator over another
7	Very strong importance	Very strong favor for one indicator over another, demonstrated in practise
9	Extreme importance	Favor for one indicated with highest possible importance
Intensities of 2,3,6, and 8 can be used to express intermediate values		
Reciprocal of above	If performance indicator *n* has of the above values assigned to it when compared with performance indicator *m*, then *m* has the reciprocal value when compared *n*	

Matrix for pair wise comparison				
	P_1	P_2	$P_{..}$	P_{1m}
P_1	1	I_{12}	$I_{1..}$	I_{1n}
P_2	I_{21}	1	$I_{2..}$	I_{2n}
$P_{..}$	$I_{..1}$	$I_{..2}$	1	$I_{..m}$
P_{n1}	I_{n1}	I_{n2}	$I_{n..}$	1
Sum	$\sum_{j=1}^{n} I_{j1}$	$\sum_{j=1}^{n} I_{j2}$	$\sum_{j=1}^{n} I_{j..}$	$\sum_{j=1}^{n} I_{jm}$

Source: Adopted from Saaty (2008)

Figure 2.
Scale and matrix for indicator comparison

To normalize the individual intensities, intensity is divided by the corresponding sum. Then, the relative weight of each indicator within the cluster is calculated as the sum of each row (Figure 3).

The procedure repeats itself when determining the relative weights in comparing clusters to clusters and likewise dimensions to dimensions. The procedure is independent of the amount of indicators, clusters, or dimensions compared. This constitutes strength in relation to the practical usage, where some organizations prefer to implement more measures than others. Furthermore, this also signifies that expansion of a hierarchy does not in any way present a problem.

	P_1	P_2	$P_{..}$	P_{1m}	Weight
P_1	$\frac{1}{\sum_{j=1}^{n} I_{j1}}$	$\frac{I_{12}}{\sum_{j=1}^{n} I_{j2}}$	$\frac{I_{1..}}{\sum_{j=1}^{n} I_{j..}}$	$\frac{I_{1m}}{\sum_{j=1}^{n} I_{jm}}$	$W_{E1}=\sum_{i=1}^{m} E_{1i}$
P_2	$\frac{I_{21}}{\sum_{j=1}^{n} I_{j1}}$	$\frac{1}{\sum_{j=1}^{n} I_{j2}}$	$\frac{I_{2..}}{\sum_{j=1}^{n} I_{j..}}$	$\frac{I_{2m}}{\sum_{j=1}^{n} I_{jm}}$	$W_{E2}=\sum_{i=1}^{m} E_{2i}$
$P_{..}$	$\frac{I_{..1}}{\sum_{j=1}^{n} I_{j1}}$	$\frac{I_{..2}}{\sum_{j=1}^{n} I_{j2}}$	$\frac{1}{\sum_{j=1}^{n} I_{j..}}$	$\frac{I_{..m}}{\sum_{j=1}^{n} I_{jm}}$	$W_{En}=\sum_{i=1}^{m} E_{..i}$
P_{n1}	$\frac{I_{n1}}{\sum_{j=1}^{n} I_{j1}}$	$\frac{I_{n2}}{\sum_{j=1}^{n} I}$	$\frac{I_{n..}}{\sum_{j=1}^{n} I_{j..}}$	$\frac{1}{\sum_{j=1}^{n} I_{jm}}$	$W_{En}=\sum_{i=1}^{m} E_{ni}$

Note: $\sum_{j=1}^{n} w_j = 1$

Figure 3.
Normalizing matrix

Step 3: normalization procedure

If aggregation of multiple, dissimilar indicators are to be possible, outcomes need to be normalized in order to make them comparable. As data normalization method the standard score, more commonly referred to as the *z*-score, is chosen (see Equation (1)):

$$z\text{-score}(\sigma) = \frac{(\text{Data point} - \text{Mean value})}{\text{SD}}$$

The *z*-score corresponds to a data point in a normal distribution, and converts all data into a common scale, making them comparable regardless of initial units (Stapenhurst, 2009). A positive *z*-value indicates performance above mean for a given period of time. The magnitude, positive or negative, indicates how much the value differs from the mean with regard to the standard deviation. When normalizing all performance data according to *z*-scores, the outcome becomes an index in relation to past performance. The number of data points can be adjusted to the character of the individual indicator, depending on the time span necessary to provide reasonable comparative values. The normalization procedure is performed for all indicators ($P_1 - P_n$) in a dimension or cluster (see Figure 4).

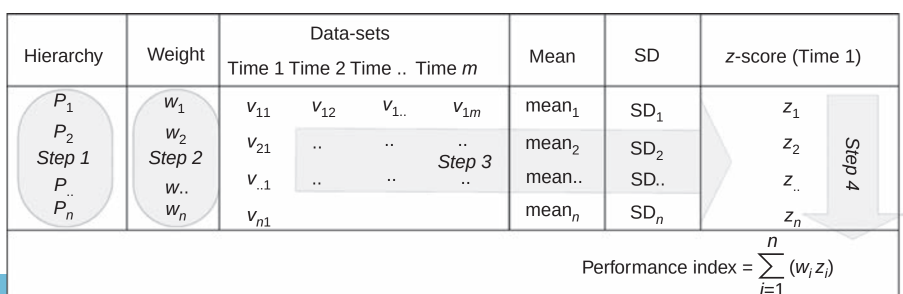

Figure 4.
Schematic outline of evaluation framework

The mean value and the standard deviation are calculated for each data set containing value points (v_{11}-v_{1m}) and are then transformed into a corresponding *z*-score.

The justification for choosing the *z*-score is that health care facilities commonly wish to reduce variation in delivered service (Lim *et al.*, 1999). It is considered more desirable to perform acceptably in all aspects than to be perfect in some and poor in others. By adopting the *z*-score, the framework encourages mean scores over a high variation, with the aim of securing high overall performance. This premise is further supported in Danish health care by the vision of the Danish Quality Framework, which states: "All patients have the right to the *same homogeneous high quality service*, no matter where they are treated" (Danish Institute for Quality and Accreditation in healthcare, 2008).

Step 4: aggregation procedure

Finally, all *z*-scores are aggregated through weighted averages, thereby presenting a measure relative to performance history, represented by step 4 in Figure 3.

As a consequence of this approach, the amount of incorporated indicators is case dependent, because each indicator acts as a contributor to a given cluster with the assigned weight. The number of data points can be adjusted to the nature of the individual indicator. In some cases, for example, it is necessary to incorporate retrospective data for several years, whereas in other cases a few months' data are suitable. The performance index, which is determined as the weighted average, now represents performance outcomes as a representation of organizational importance. Lowly prioritized indicators will not have as much impact as highly prioritized indicators. This allows for monitoring performance as a function of strategic importance.

Framework tryout at an MRI unit

To test whether the four-step procedure is useful as a tool for health care decision makers, the framework was applied in an MRI unit in a Danish hospital. The case is a radiology department at a Danish non-profit hospital with four individual sites, the result of a fusion between four formerly independent hospitals. The hospitals were merged at management level, but the four sites still act as separate operational parts in the new hospital. The radiology department employs 128 staff members in total, and examines and treats approximately 145 patients on a daily basis. The quantitative test is performed exclusively at the MRI unit on one site, which examines approximately 3,500 patients per year, distributed on about 70 different types of MRI scans. The unit receives both acute and planned patients. The unit employs both full-time and part-time employees and students.

The MRI hierarchy

The dimensions used in the construction of the hierarchy were based on the strategic plan of the hospital, where patients, employees, and operations constitute the backbone. The selection and placement of indicators were determined through a series of workshops with radiologists, technicians, a project manager, and the head of the department. It was desirable to include a large amount of indicators in order to provide as complete a performance picture as possible (Curtright *et al.*, 2000). Based on the indicator selection and the clustering, the full hierarchy was constructed as shown in Table I.

The hierarchy includes 26 performance indicators distributed in nine clusters. The indicators chosen show overall coherence with other scientific work dealing with

	Dimension	Cluster	Indicator	Formula
Superior indicator	Patients	Safety	Adverse advents	# Adverse advents / # Total procedures
			Incorrect treatment	# Incorrect treatments / # Total procedures
			Re-called patients	# Re – called patients / # Total procedures
		Information	Received written info	# Patients without written info / # Total procedures
			Satisfaction (written info)	Adopted from external survey
			Satisfaction (oral info)	Adopted from external survey
		Satisfaction	Satisfaction survey	Adopted from external survey
			Waiting time for treatment	Adopted from external Database
			Complaints	# Complaints / # Total procedures
	Employees	Occupation profile	Part-time employees	# Part – time employees / # Full – time employees
			Available posts	# Open posts / # Full – time employees
			Educational positions	# Students / # Full – time employees
		Work environment	Overtime	# Overwork hours / # Scheduled hours
			Sick leave	Sick leave [days] / # Scheduled [days]
			Turnover rate	# Employees who left / # Full – time employees
			Satisfaction survey	Adopted from external survey
		Risk	Reported work hazards	# Reported hazards / # Full – time employees
			Long-term sickness absence	Sick leave > 15 days / Sick leave
	Operations	Planning	Acute load	# Acute patients / # Total procedures
			Non-Attending patients	# Non – attending patients / # Total procedures
			Cancelled examinations	# Cancelled examinations / # Total procedures
		Efficiency	Operational time	∑ Equipment producing [min] / # Total procedures
			% procedures (7–15)	# Examinations (between 7 – 15) / # Total procedures
			Throughput	∑ ([Description end] – exam start) / # Total procedures
		Utilization	Employee utilization rate	(# Exams × Operational time) – Overwork / ∑ Working hours
			Equipment utilization rate	(# Exams × Operational time) / (60 × 24 × 365) – Downtime

Table I.
Indicator hierarchy
for the MRI unit

health care performance measurement (e.g. Byrne, 2006; De Toni *et al.*, 2007; Kollberg *et al.*, 2005). The allocation of weights was also conducted in workshops, where clinicians discussed the relative importance of indicators. The data collection was performed by the researchers, and data validity was continuously confirmed by the workshop attendees. After applying quantitative data from the MRI unit, the aggregated performance profile was calculated (see Table II). Due to the normalization method, positive values indicate that the MRI unit is performing above the average of retrospective data, and negative values indicate that current performance is lower than past performance.

As the aggregated result indicates, overall performance is under control, with a small increase of 0.13 in total aggregated performance. The underlying reason for

		Dimension (weight)	*z*-score	Cluster (weight)	*z*-score	Indicator (weight)	*z*-score
						Adverse advents *(0.630)*	0.00
				Safety *(0.503)*	0.09	Incorrect treatment *(0.250)*	0.31
						Re-called patients *(0.120)*	0.12
						Received written info *(0.463)*	0.45
		Patient *(0.573)*	0.28	Information *(0.348)*	0.18	Satisfaction (written info) *(0.329)*	0.76
						Satisfaction (oral info) *(0.208)*	–1.33
						Satisfaction survey *(0.586)*	1.88
				Satisfaction *(0.148)*	1.13	Waiting time for treatment *(0.224)*	0.69
						Complaints *(0.190)*	0.00
						Part-time employees *(0.595)*	0.45
				Occupation profile *(0.570)*	0.30	Available posts *(0.277)*	0.13
						Educational positions *(0.129)*	0.00
						Overtime *(0.438)*	–0.95
Result	0.13	Employee *(0.320)*	0.07	Work environment *(0.259)*	–0.36	Sick leave *(0.240)*	0.45
						Turnover rate *(0.202)*	0.20
						Satisfaction survey *(0.120)*	–0.77
				Risk *(0.171)*	–0.08	Reported work hazards *(0.833)*	0.11
						Long-term sickness absence *(0.167)*	–1.00
						Acute load *(0.387)*	–1.06
				Planning *(0.684)*	–0.62	Non-Attending patients *(0.443)*	–0.48
						Cancelled examinations *(0.170)*	0.00
						Operational time *(0.657)*	–0.53
		Operation *(0.107)*	–0.51	Efficiency *(0.244)*	0.06	% procedures (7-15) *(0.207)*	2.14
						Throughput *(0.136)*	–0.27
				Utilization *(0.072)*	–1.38	Employee utilization rate *(0.875)*	–1.37
						Equipment utilization rate *(0.125)*	–1.46

Table II.
Aggregated performance result for the MRI unit

this small progress is found in the positive result of 0.28 within the patients' dimension. The relatively large positive patient dimension constitutes a significant element in the overall outcome. This is of course due to the significant weight assigned to this element. The operations dimension shows a negative score with negative values in both planning and utilization. The reason that the operations dimension value does not bring down the overall aggregated result is the relatively low weight assigned to this dimension. The negative value in utilization points out another curiosity. Surprisingly, equipment utilization has yielded a very poor result in a period with a fair amount of overtime. The explanation is that the department did not have an increased number of planned patients in this period and had fewer acute patients. Throughput time increased during the same period. In a period with an increase in overtime and an increase in throughput time it is not irrelevant that technicians have been working overtime and perhaps thereby created additional work for clinicians without extra resources being allocated to reporting. Data collected at the radiology department have shown that a high amount of non-attending patients use private sector MRI for diagnostics, due to prolonged waiting times in public sector MRI. This relation is enhanced by the increase in the use of private insurance, which allows the use of private health care facilities in Denmark. The correlation is also evident in the hierarchy shown by a decrease in waiting times and a decrease in non-attending patients. In a period with shorter waiting times the result is thus fewer non-attending patients.

Discussion

In light of the research objective, the authors believe that the framework possesses the desired strengths, particularly by providing a still-life of current performance compared to past performance. The distinctiveness of this framework lies in its combination of normalization according to past performance, and the use of the AHP concept as a method for prioritization, analyzing, and representing the outcome in few key measures. This allows for the monitoring of performance progress and regress as a function of strategic importance. The framework has the potential for including large amounts of information, and at the same time it can target this information for use in strategic decision support. The hierarchical construction calls for strategic alignment, since indicator structures are deduced through the organization from the strategic objectives. The traditional target-setting procedure is replaced by the normalization procedure, which portrays each indicator as a reflection of past performance. Hence, the target becomes progress instead of a fixed target value, and the weight profile indicates which areas should be the center of attention. The framework thus provides a clear indication of whether an organization is progressing satisfactorily toward its predetermined strategic objectives. Weak areas would be visible, as well as their significance in relation to overall performance. The framework deals with the reality of facing massive amounts of information and the request for targeted information. Information is presented in a holistic manner, and at the same time allows focus on clustered and target measures. This enables decision makers to trace poor performance through the hierarchy, easing the identification of root causes. Output represents operational performance with a positive or negative value according to strategic progress or regress. The strength of specific measures is still apparent, enabling operational decision makers to correct unacceptable performance.

However, a very structured and methodical development phase is a prerequisite if the above-mentioned benefits are to be obtained. The hierarchy must be arranged to reflect the desired aspects at the unit being analyzed. In this case, the hospital's strategic plan places patients and employees at the center of all activities. The choice of "operations" as a third element was based on a radiological focus. As a result, the highest weights in the hierarchy were assigned to the patient and employee dimensions. Obviously, sloppy selection and placement would result in an unusable reporting system. Assigning weights is also a difficult process, because a hospital's strategic plan usually does not prioritize dimensions, but states areas in focus. Weight assignment would therefore be a result of subjective assessments made among workshop participants. The authors do not ignore the difficulty of assigning strategic weights to indicators, which can be further complicated by unclear strategic objectives (Neely and Al Najjar, 2006). At the same time, the weight assignment needs to be justifiable within the highly political environment in which health care is found. In spite of this, the weighting process is made easier due to the clustering of indicators in the framework. Subsequently to the implementation of the framework, the Radiology Department has continuously evaluated the weighting of indicators, clusters, and dimensions so that it reflects the current focus. In order to maintain overall strategic stability, however, it is kept in mind to keep priority changes at a minimum. In this context, it is still important to note that the framework itself does not secure the quality of decisions. Quality of decisions is closely related to the individual decision maker. Even while high quality information is provided, the decision process may be influenced by internal pressure, politicians, or even poor judgment. The aim of the study has not been to validate decisions, but only to provide valid and transparent information to be used for decision support.

There are two main issues regarding the validity of the mathematical outcome. First, the validity of external input data is secured by the agencies issuing them, and internal input data by the employees at the department. Second, the mathematical construction of the framework does not affect data validity. The normalization and aggregation of data does not in itself change the reliability of the data, but it naturally affects the interpretation of outcomes.

In order to test whether the presentation of results corresponds to perceived reality, a method of recognition was used. In order to test whether the quantitative results reflect "reality," a blind test was conducted with the staff members who had participated in the construction of the hierarchies. The blind test was performed as the quantitative result in the hierarchy was blinded and the participant should "rank" dimensions and clusters according to best achieved performance. The test validated the data saying that the qualitative construction of the hierarchies was in concordance with a perceived reality, because the quantitative result reflected staff members' perception of performance. The framework showed progress and regress in the organizational areas that were expected. The notion of a "perceived reality" is important to emphasize here, because there are no absolute values for good or bad performance. The weight profiles are subjective, since the assignment is conducted based on the interviewees' perception of mutual importance. Thus, the interpretation of performance is biased to represent the "perceived reality" of those who have participated in the construction of the hierarchy and designed the weight profile. The advantage, though, is that the priorities are explicitly formulated, instead of being implicit.

Implementation of the framework in a single case in a public and highly political environment restricts the generalizing potential. The authors acknowledge that there may be more suitable approaches in organizations with different settings. Despite this consideration we share the clear conviction that the framework can easily be adapted to different settings.

Conclusion

The main objective has been to cluster large amounts of performance information in order to make it understandable for operational decision makers. The paper proposes that performance indicators be built into a hierarchical construction, prioritized according to strategic significance, normalized according to past performance, and finally aggregated to present an expression to present an evidence of overall performance. The advantage of using such a combination is that it allows for the monitoring of performance progress and regress as a function of strategic objectives. Placing indicators in a hierarchy provides the possibility to trace performance from a strategic level through to tactical and operational indicators. However, in designing the hierarchy, the thoroughness of the prioritization according to strategic objectives becomes a key issue with regard to reliability as it determines the end result. With meticulous prioritization and construction according to strategic objectives, the construction of the performance measurement approach confronts the issue of increasing informational complexity. What otherwise would have been a qualitative assessment of strategic importance is now quantified by representing performance as weighted, aggregated measures. The study demonstrates a more refined and transparent use of performance reporting, as it portrays operational performance as a function of strategic progress and regress in a unified manner.

References

Barros, P.P. (2003), "Random output and hospital performance", *Health Care Management Science*, Vol. 6 No. 4, pp. 219-227.

Basu, A., Howell, R. and Gopinath, D. (2010), "Clinical performance indicators: intolerance for variety?", *International Journal of Health Care Quality Assurance*, Vol. 23 No. 4, pp. 436-449.

Bloomquist, P. and Yeager, J. (2008), "Using balanced scorecards to align organizational strategies", *Healthcare Executive*, Vol. 23 No. 1, pp. 24-26, 28.

Bovier, P.A. and Perneger, T.V. (2003), "Predictors of work satisfaction among physicians", *The European Journal of Public Health*, Vol. 13 No. 4, pp. 299-305.

Buetow, S. (2008), "Pay-for-performance in New Zealand primary health care", *Journal of Health Organisation and Management*, Vol. 22 No. 1, pp. 36-47.

Byrne, M. (2006), "Implementing performance management in the Irish health sector", *The Health Care Manager*, Vol. 25 No. 2, pp. 114-121.

Cheng, S.M. and Thompson, L.J. (2006), "Cancer care Ontario and integrated cancer programs: portrait of a performance management system and lessons learned", *Journal of Health Organisation and Management*, Vol. 20 No. 4, pp. 335-343.

Curtright, J.W., Stolp-Smith, S.C. and Edell, E.S. (2000), "Strategic performance management: development of a performance measurement system at the mayo clinic", *Journal of Healthcare Management*, Vol. 45 No. 1, pp. 58-68.

Danish Institute for Quality and Accreditation in healthcare (2008), *Den Danske Kvalitetsmodel*, Danish Institute for Quality and Accreditation in Healthcare, AArhus.

De Toni, A.F., Andrea, F. and Mattia, M. (2007), "A performance measurement system for facility management: the case study of a medical service authority", *International Journal of Productivity & Performance Management*, Vol. 56 No. 5, pp. 417-435.

Dey, P.K., Hariharan, S. and Despic, O. (2008), "Managing healthcare performance in analytical framework", *Benchmarking: An International Journal*, Vol. 15 No. 4, pp. 444-468.

Dummer, J. (2007), "Health care performance and accountability", *International Journal of Health Care Quality Assurance*, Vol. 20 No. 1, pp. 34-39.

Elleuch, A. (2008), "Patient satisfaction in Japan", *International Journal of Health Care Quality Assurance*, Vol. 21 No. 7, pp. 692-705.

Evans, J.R. (2004), "An exploratory study of performance measurement systems and relationships with performance results", *Journal of Operations Management*, Vol. 22 No. 3, pp. 219-232.

Evans, J. and Weir, C. (1995), "Decision processes, monitoring, incentives and large firm performance in the UK", *Management Decision*, Vol. 33 No. 6, pp. 32-38.

Flapper, S.D., Fortuin, L. and Stoop, P.P.M. (1996), "Towards consistent performance management systems", *International Journal of Operations and Production Management*, Vol. 16 No. 7, pp. 27-37.

Folan, P. and Browne, J. (2005), "A review of performance measurement: towards performance management", *Computers in Industry*, Vol. 56 No. 7, pp. 663-680.

Jollands, N., Lermit, J. and Patterson, M. (2003), "The usefulness of aggregate indicators in policy making and evaluation: a discussion with application to eco-efficiency indicators in New Zealand", *Ecological Economics, The ISEE Journal*.

Jones, M.K., Jones, R.J., Latreille, P.L. and Sloane, P.J. (2009), "Training, job satisfaction, and workplace performance in Britain: evidence from WERS 2004", *Labour*, Vol. 23 No. S1, pp. 139-175.

Kocakülâh, M.C. and Austill, A.D. (2007), "Balanced scorecard application in the health care industry: a case study", *Journal of Health Care Finance*, Vol. 34 No. 1, pp. 72-99.

Kollberg, B., Elg, M. and Lindmark, J. (2005), "Design and implementation of a performance measurement system in Swedish health care services: a multiple case study of 6 development teams", *Quality Management in Health Care*, Vol. 14 No. 2, pp. 95-111.

Kutney-Lee, A., McHugh, M.D. and Sloane, D.M. (2009), "Nursing: a key to patient satisfaction", *Health Affairs*, Vol. 28 No. 4, pp. w669-77.

Lega, F. and Vendramini, E. (2008), "Budgeting and performance management in the Italian national health system (INHS): assessment and constructive criticism", *Journal of Health Organisation and Management*, Vol. 22 No. 1, pp. 11-22.

Lim, P.C., Tang, N.K.H. and Jackson, P.M. (1999), "An innovative framework for health care performance measurement", *Managing Service Quality*, Vol. 9 No. 6, pp. 423-433.

Loeb, J.M. (2004), "The current state of performance measurement in health care", *International Journal for Quality in Health Care*, Vol. 16 No. S1, pp. i5-i9.

Matta, M.E. and Patterson, S.S. (2007), "Evaluating multiple performance measures across several dimensions at a multi-facility outpatient center", *Health Care Management Science*, Vol. 10 No. 2, pp. 173-194.

Mital, K.M. (2010), "Queuing analysis for outpatient and inpatient services: a case study", *Management Decision*, Vol. 48 No. 3, pp. 419-439.

Mohammadi, S.M., Mohammadi, S.F. and Hedges, J.R. (2007), "Conceptualizing a quality plan for healthcare", *Health Care Analysis*, Vol. 15 No. 4, pp. 337-361.

Morgan, D. and Morgan, R. (2009), *Single-Case Research Methods for the Behavioral and Health Sciences*, 1st ed., Sage, Thousand Oaks, CA.

Moullin, M. (2004), "Eight essentials of performance measurement", *International Journal of Health Care Quality Assurance*, Vol. 17 No. 3, pp. 110-112.

Nakajima, S. (1986), "TPM – Challenge to the improvement of productivity by small group activities", *Maintenance Management International*, Vol. 6 No. 2, pp. 73-83.

Neely, A. (2005), "The evolution of performance measurement research: developments in the last decade and a research agenda for the next", *International Journal of Operations & Production Management*, Vol. 25 No. 12, pp. 1264-1277.

Neely, A. and Al Najjar, M. (2006), "Management learning not management control: the true role of performance measurement", *California Management Review*, Vol. 48 No. 3, pp. 99-114.

Neely, A., Mills, J., Platts, K., Gregory, M. and Richards, H. (1994), "Realizing strategy through measurement", *International Journal of Operations & Production Management*, Vol. 14 No. 3, pp. 140-152.

Ondategui-Parra, S., Bhagwat, J.G., Gill, I.E., Nathanson, E., Seltzer, S. and Ros, P.R. (2004), "Essential practice performance measurement", *Journal of the American College of Radiology*, Vol. 1 No. 8, pp. 559-566.

Ormrod, J. (1993), "Decision making in health service managers", *Management Decision*, Vol. 31 No. 7, p. 8.

Saaty, T.L. (1982), "Priority setting in complex problems", *IEEE Transactions on Engineering Management*, Vol. EM-30 No. 3, pp. 140-155.

Saaty, T.L. (2008), "Relative measurement and its generalization in decision making why pairwise comparisons are central in mathematics for the measurement of intangible factors", *Estadística e Investigación Operativa/Statistics and Operations Research*, Vol. 102 No. 2, pp. 251-318.

Stapenhurst, T. (2009), "Normalization: how to compare apples with pears", *The Benchmarking Book: A How-to-Guide to Best Practice for Managers and Practitioners*, 1st ed., Oxford, pp. 112-137.

Strandberg-Larsen, M., Nielsen, M.B., Vallgårda, S., Krasnik, A., Vrangbæk, K. and Mossialos, E. (2007), "Health systems in transition", *European Observatory on Health Systems and Policies*, Vol. 9 No. 6, pp. 1-164.

Swaminathan, S., Chernew, M. and Scanlon, D.P. (2008), "Persistence of HMO performance measures", *Health Services Research*, Vol. 43 No. 6, pp. 2033-2049.

The Danish Institute for Quality and Accreditation in Healthcare web site (2010), 1-3-2010, Ref Type: Internet Communication.

The Danish Patient Safety Database's web site (2010), 10-4-2010, Ref Type: Internet Communication.

The National Board of Health's web site (2009), 10-12-2009, Ref Type: Internet Communication, December 12.

The National Indicator Project's web site (2009), 10-12-2009, Ref Type: Internet Communication, December 12.

The Unit of Patient-Perceived Quality's web site (2009), 3-2-2010, Ref Type: Internet Communication, February 2.

Treasure, T., Valencia, O., Sherlaw-Johnson, C. and Gallivan, S. (2002), "Surgical performance measurement", *Health Care Management Science*, Vol. 5 No. 4, pp. 243-248.

Voss, C., Tsikriktsis, N. and Frohlich, M. (2002), "Case research in operations management", *International Journal of Operations & Production Management*, Vol. 22 No. 2, pp. 195-219.

Walley, P., Silvester, K. and Mountford, S. (2006), "Health-care process improvement decisions: a systems perspective", *International Journal of Health Care Quality Assurance*, Vol. 19 No. 1, pp. 93-104.

Winter, R. and Munn-Giddings, C. (2001), *A Handbook for Action Research in Health and Social Care*, 1st ed., Routledge, New York, NY.

World Health Organization (2008), *The World Health Report 2008: Primary Health Care Now More Than Ever*, WHO Press, Geneva.

Corresponding author
Dr Andreas Traberg can be contacted at: andreas.traberg@siemens.com