

Benchmarking operating room departments in the Netherlands

Evaluation of a benchmarking collaborative between eight university medical centres

Benchmarking
OR
departments

1171

Elizabeth van Veen-Berkx

*Department of Operating Rooms, Erasmus University Medical Center,
Rotterdam, The Netherlands*

Dirk F. de Korne

*Singapore National Eye Centre, SingHealth, Singapore and
Institute for Health Policy and Management,
Erasmus University Rotterdam, Rotterdam, The Netherlands*

Olivier S. Olivier

Medworq-Connecting Excellence, Doesburg, The Netherlands

Roland A. Bal

*Institute for Health Policy and Management, Erasmus University Rotterdam,
Rotterdam, The Netherlands, and*

Geert Kazemier

*Department of Surgery, VU University Medical Centre Amsterdam,
Amsterdam, The Netherlands*

Received 25 April 2014
Revised 23 December 2014
Accepted 13 January 2015

Abstract

Purpose – Benchmarking is increasingly considered a useful management instrument to improve performance in healthcare. The purpose of this paper is to assess if a nationwide long-term benchmarking collaborative between operating room (OR) departments of university medical centres in the Netherlands leads to benefits in OR management and to evaluate if the initiative meets the requirements of the 4P-model.

Design/methodology/approach – The evaluation was based on the 4P-model (purposes, performance indicators, participating organisations, performance management system), developed in former studies. A mixed-methods design was applied, consisting of document study, observations, interviews as well as analysing OR performance data using SPSS statistics.

Findings – Collaborative benchmarking has benefits different from mainly performance improvement and identification of performance gaps. It is interesting that, since 2004, the OR benchmarking initiative still endures after already existing for ten years. A key benefit was pointed out by all respondents as “the purpose of networking”, on top of the purposes recognised in the 4P-model. The networking events were found to make it easier for participants to contact and also visit one another. Apparently, such informal contacts were helpful in spreading knowledge, sharing policy documents and initiating improvement. This benchmark largely met all key conditions of the 4P-model.

Research limitations/implications – The current study has the limitations accompanied with any qualitative research and particularly related to interviewing. Qualitative research findings must be viewed within the context of the conducted case study. The experiences in this university hospital context in the Netherlands might not be transferable to other (general) hospital settings or other countries. The number of conducted interviews is restricted; nevertheless, all other data sources are extensive.

Originality/value – A collaborative approach in benchmarking can be effective because participants use its knowledge-sharing infrastructure which enables operational, tactical and strategic learning.



Organisational learning is to the advantage of overall OR management. Benchmarking seems a useful instrument in enabling hospitals to learn from each other, to initiate performance improvements and catalyse knowledge-sharing.

Keywords Benchmarking, Health services, Performance indicators, Operating rooms, University medical centres, Performance management system, Collaborative benchmarking, Learning from others, Knowledge-sharing

Paper type Case study

Introduction

As in many countries in the world, the healthcare system in the Netherlands has been intensively reformed over the last decade (Schut and van de Ven, 2011; van de Bovenkamp *et al.*, 2013). The introduction of more competition in healthcare was one of the most important changes (Enthoven and Tollen, 2005; Scanlon *et al.*, 2008; Schut and van de Ven, 2005). This increasingly urges hospital administrators and clinicians to deliver transparent, high-quality care with strict financial budgets. The focus on performance improvement sparked the interest of healthcare providers to measure their performance and compare themselves with others in order to be enabled to perform more efficiently in their operational processes (Berg *et al.*, 2005; de Korne *et al.*, 2010; Kazemier and van Veen-Berkx, 2013). As indicated by Porter and Teisberg (2006) in their landmark “Redefining Health Care”, competition among healthcare providers should be focused on value (defined as “health care results per unit of costs”) and supported by widely available outcome data. Obtaining such data, however, requires appropriate management instruments that can disseminate business information and compare the performance of a single provider to others. Benchmarking, defined as “a process of continuous measuring and comparing an organization’s business against others” (APQC, 2008), is described as one of the approaches to obtain useful results (Zairi and Leonard, 1994; Zairi and Ahmed, 1999; Blank and Valdmanis, 2008; Dattakumar and Jagadesh, 2003).

To assess the application of benchmarking in hospitals, de Korne *et al.* (2010) have developed a “4P” conceptual framework (see Figure 1).

The key conditions, based on literature study, are: purposes (learning from others, identifying performance gaps, implementing best practices); performance indicators (SMART indicators, comparable indicator information, reliable data gathering and sharing); participating organisation similarities (in structure, process, outcomes; no competition between participants, voluntary and involved participation); and performance management system (cyclical, internal). The model has been validated in international and US domestic benchmark initiatives between eye hospitals (de Korne *et al.*, 2010, 2012) but has not been applied in other settings.

Therefore we have studied an on-going collaborative benchmarking initiative between the operating room (OR) departments of eight university medical centres (UMCs). In the Netherlands, OR departments of all eight UMCs established a nationwide benchmarking collaborative in 2004. The objective of the benchmark is to compare the utilisation of OR resources and the economic aspects of OR performance between the UMCs, with the aim to improve this performance. Each UMC provides their surgical case records to a central OR benchmark database. This extensive database, presently comprising more than one million surgical case records, is used to calculate key performance indicators related to the utilisation of OR capacity. The results from benchmarking – by name of UMC – are only accessible to the participants. However, the database is also used for multicentre research on OR

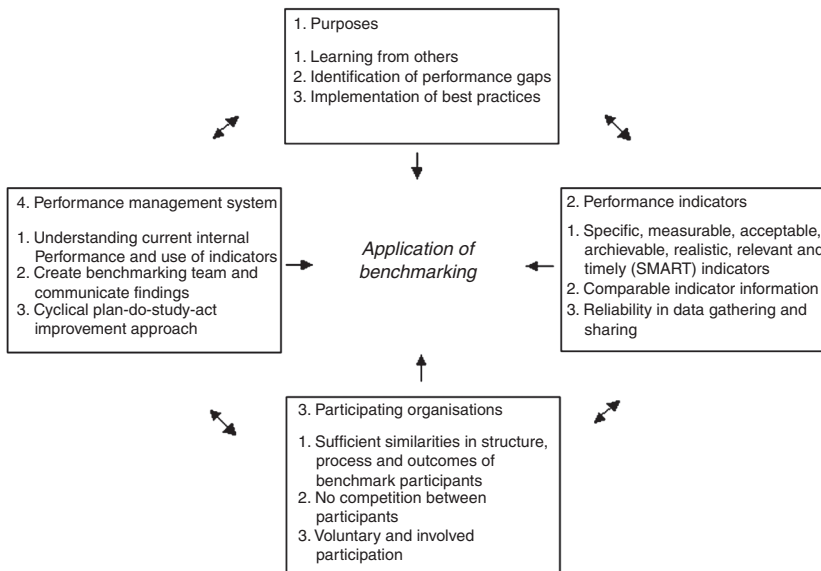


Figure 1.
The “4P model” with
key conditions for
the application of
benchmarking

Source: De Korne *et al.* (2010)

scheduling topics and OR efficiency, and therefore results from benchmarking are published anonymously (Kazemier and van Veen-Berkx, 2013; van Veen-Berkx *et al.*, 2014b, c).

The aim of this study is to assess if the collaborative, long-term approach of the Dutch OR benchmarking initiative leads to benefits in OR management and to evaluate if the initiative meets the requirements of the 4P-model. Based on the findings we discuss the applicability of the 4P-model and present key findings useful for benchmarking in (university) hospital settings.

Literature

Literature identifies several types of benchmarking: internal and external (Camp and Tweet, 1994; Joint Commission, 2012). While internal benchmarking focuses on performance measurement and comparing within one organisation over time, external benchmarking can be categorised in competitive, functional, generic and collaborative benchmarking. Competitive, functional and generic benchmarking are commonly conducted independently, while the collaborative approach to traditional benchmarking is performed by groups of organisations that work jointly to achieve the same goals. Collaborative benchmarking entails more than merely comparing performance: organisations share their ideas, approaches, process designs and interventions (Joint Commission, 2012; Mosel and Gift, 1994). This approach implicates the formation of a voluntary network of healthcare organisations that cooperate in carrying out the benchmarking study and commit to this long-term (Mosel and Gift, 1994; Gift *et al.*, 1994).

Although benchmarking was developed for the business industry, it is increasingly being observed in the public sector (Booth *et al.*, 2005; Guven-Uslu, 2005). However, empirical research on the use and function of benchmarking in health organisations is

still scarce (Blank, 2008; Blank and Valdmanis, 2008). Sower *et al.* (2008) have described practical benchmarking experiences at the Bronson Methodist Hospital, the Columbus Children's Hospital, and the North Mississippi Health Services. They concluded that benchmarking could help close the widening gap between hospitals that deliver exemplary patient service and those that provide lower levels of care.

Extensive research exists regarding hospital benchmarking studies using data envelopment analysis's (DEA). Benchmarking studies applying the method of DEA are predominantly initiated by the government or regulatory offices (e.g. the Dutch Healthcare Authority) and carried out by a separate academic statistical unit. Datasets are obtained from the Ministries of Health, in other words, "external data collectors". DEA is known as a nonparametric mathematical programming approach for determining a best practice of resource usage and service delivery, and is also focused on explaining variations in cost efficiency due to a hospital's environment. Cost-efficiency scores measured by the DEA approach are principally used by policy makers who are interested in budget allocation for hospitals (Blank, 2008; Blank and Valdmanis, 2008, 2010; Kacak *et al.*, 2014). The character of benchmarks using DEAs, however, is essentially different from the character of benchmarking collaboratives initiated by hospitals themselves and not by a third external party. Although this type of collaborative benchmarking is increasingly used in hospitals, well-described experiences and systematic empirical research are scarce (van Lent *et al.*, 2010).

Several studies though have assessed the efficacy of performance reports in stimulating hospital quality improvement. Hibbard *et al.* (2005) found in a large study in Wisconsin that disclosure of performance, in private and public reports, resulted in improvement in the clinical area reported upon. Devers *et al.* (2004) indicated different mechanisms that drive hospital quality improvement: regulation, professionalism and market forces; benchmarking and reporting performances is thought to be a key strategy for influencing market forces and, to a lesser extent, professionalism. Also in more government-driven systems there is evidence for positive effects of performance reporting. Levay and Waks (2009) analysed national quality registries in Sweden and describe how professional groups are actively engaged in transparency technologies and found them meaningful, despite initial resistance, and continued discontent with specific aspects of the monitoring systems. In the UK's National Health Service, the use of targets and disclosure of performance have resulted in increasing performance due to forces of reputation management (Bevan and Hood, 2006). Hibbard *et al.* (2005) argue that the feedback inherent to both public and private reports will be sufficient to stimulate efforts to improve quality, simply because of professional norms around maintaining standards and self-governance. Therefore, benchmarking has been suggested to be applied broader in hospital care. A recent study by Welborn and Bullington (2013) indicated that of all process improvement techniques available, benchmarking was found to be the most popular and widely used among a group of award winning US healthcare organisations (Welborn and Bullington, 2013).

Benchmarking has often been approached as a competitive activity resulting in rankings and with a focus on creating competition between participants as driver for improvement. Since benchmarking was initiated in Japan (Zairi, 1992) in order to improve competitiveness and since Xerox in 1979 discovered benchmarking as an advantageous management instrument (Camp and Tweet, 1994), it is not unexpected that benchmarking initiatives are typically associated with "competition" instead of "collaboration" between organisations. In literature and previous research, as referred

to by Wolfram Cox *et al.* (1997), it is generally agreed that the motivation behind benchmarking is to improve and to reduce the performance gap compared with the superior competitor (Camp, 1989; Spendolini, 1992; Watson, 1993). Due to this competitive nature, most benchmarking studies performed by organisations have been conducted individually.

Already in 1994, however, Mosel and Gift as well as Gift *et al.* (1994) refer to the need for healthcare providers to consider an alternative to the individual method, which is found in the collaborative approach of benchmarking. Wolfram Cox *et al.* (1997) clearly contrast the two approaches:

- (1) the collaborative one is characterised by “learning with and from others as aim”, “partnership as relationship between participants”, “a joint action” and “the visual picture is horizontal and visiting (sharing knowledge from the kitchen)”; and
- (2) the competitive one is characterised by “superiority or learning to gain position over the other organisation”, “a relationship of rivalry”, “a unilateral action to gain position on the ladder of success” and “the visual picture is vertical ranking”.

In the Netherlands, healthcare reform and the introduction of more competition has been a driver for hospitals to compare themselves to others in the challenge to deliver safe, high quality, transparent, accountable and efficient care. Since reforms in the healthcare system, vertical ranking is increasingly used in order to provide (hospital) performance information and help patients to choose (Quartz *et al.*, 2013). However, these rankings generally tend to compare apples and oranges, because they show observed differences in outcome measures between organisations while outcome measures are bounded by methodological difficulties (Lingsma *et al.*, 2010; Mant, 2001). Variation in outcome between organisations is subject to patient case mix, differences in measurement (registration reliability), statistical uncertainty (chance), and real differences in quality of care (structure/process) (Lingsma *et al.*, 2010; Mant, 2001). For example, a Dutch hospital can score a high rank in one league table yet at the same time score a low rank in another league table. Therefore, ranking can provide inadequate information and the current public reporting can easily be wrongly interpreted by patients (Giard, 2006; Lingsma *et al.*, 2010; van Dishoeck *et al.*, 2011). Van Dishoeck *et al.* (2011) even claim that current outcome indicators, used by the Dutch Healthcare Inspectorate, are not suitable for ranking hospitals because of the influence of random variation.

Origin of the Dutch OR benchmarking collaborative

In 2004, the OR departments of all eight UMCs in the Netherlands established a benchmarking collaborative (Kazemier and van Veen-Berkx, 2013; van Veen-Berkx *et al.*, 2014b). This is a joint initiative of the eight Dutch UMCs. Each UMC provides surgical case records extracted directly from the hospital’s self-reported OR data management system to a central OR benchmark database. This central database is used to calculate key performance indicators of the utilisation of OR capacity, e.g. first-case tardiness, turnover time and raw utilisation. These performance measurements are shared and benchmarked between the UMCs, which enables the identification of areas of improvement by comparing one’s own performance to that of other, similar organisations.

ORs are of paramount importance to a hospital, given the fact that more than 60 per cent of patients admitted to a hospital are treated in the OR (Eijkemans *et al.*, 2010). Efficient use of OR capacity is pivotal since it is considered a high-cost

environment and a limited hospital resource (Marjamaa *et al.*, 2008). Due to the aging population and various developments in surgery, demands for OR facilities are likely to increase. Moreover, due to shortages of qualified OR staff, optimal utilisation of ORs is an ever increasing challenge. Benchmarking can be applied to identify improvement potential (Watson, 1993) and measure the effectiveness of interventions in an OR environment.

The initiators knowingly decided to develop a “self-led” collaboration with its own budget (paid for by the eight UMCs themselves) and management, independently from external consultancy organisations and external funding. Independence from external companies as well as external financial resources allows the collaborative to make its own decisions regarding the choice of performance indicators and builds more trust concerning knowledge sharing in a safe learning environment.

The collaborative consists of an organisation containing of a steering committee (head of Department of Surgery or Anesthesiology) and a project committee (OR managers) in which all eight UMCs are represented. One full time project manager is hired for planning, organising, securing and managing resources. This project management position is financed by the subscription that is annually paid by the UMCs. The project manager cooperates directly with the members of the project and steering committee on a frequent basis.

Methods

A mixed-methods design was applied (Bowling, 2009; Creswell, 2003; Yin, 2003). Based on a literature study, the 4P-model (purposes, performance indicators, participating organisations, performance management system) (de Korne *et al.*, 2010) was used to evaluate the collaborative in a case study of the OR departments of all eight UMCs in the Netherlands.

We analysed the OR performance data using SPSS statistics version 21. Data were abstracted from the central OR benchmark database. Regarding the OR performance indicators, all elective, inpatient surgical cases were included. If an OR complex of a single UMC was divided into a main location and sub locations such as a Cancer Centre, Children’s Hospital and Thorax Centre, merely the main (largest) inpatient OR location was included.

The interquartile range (IQR) is a measure of statistical dispersion, which contains the middle 50 per cent of the data (the top and bottom 25 per cent of data are left out) (Field, 2013). It is calculated as the difference between the upper and lower quartiles: $IQR = Q_3 - Q_1$.

We conducted ten semi-structured face-to-face interviews with key stakeholders from OR management ($n = 3$), surgical planning ($n = 2$), OR nursing ($n = 1$), data management ($n = 2$), policy consultant ($n = 1$) and CEO ($n = 1$). These key stakeholders came from five (out of eight) UMCs that represented different parts of the country. The interviews were a maximum of 1.5 hours, transcribed verbatim, and subsequently analysed. The semi-structured face-to-face interviews were guided by a topic list based on the 4P-model. Questions involved the purposes of benchmarking, the performance indicators, the reliability in data gathering and sharing, the participating organisations and their characteristics and environment, the involvement of participants, the performance management system and the cyclical plan-do-study-act (PDSA) improvement approach.

We performed document analyses to reveal information from management reports, policy documents and performance indicator reports. We analysed the minutes of

40 benchmarking meetings with representatives from all eight hospitals involved and performed observations during two benchmarking focus group study meetings. The interview data, documents and transcripts were analysed by using labels from the 4P-model as well as open labels. The labels were used to code and categorise the transcripts and identify recurrent themes, relying on the theoretical proposition from the 4P-model (as suggested by Yin, 2003). All interviews and observations were conducted by the third author. The preliminary comparative analyses were done separately by the second and third author. To increase the construct validity, the first, fourth and last author reviewed the drafts of the analyses. Data triangulation was used when comparing data gathered from different sources (Bowling, 2009).

Data triangulation applies multiple sources of information and data to investigate complex situations and to increase the validity of the study. This also means continuously comparing different findings with findings from other sources in order to contribute to a broader and deeper description and understanding of the case (Bowling, 2009). In this study there were four empirical sources: first, the quantitative OR data from the central OR Benchmark database; second, the transcripts of the conducted interviews; third, a review of relevant documents and minutes of meetings; and fourth, the field notes of the observations performed during benchmarking focus group study meetings.

Findings

Assessment of the four key conditions

Purposes. In accordance with the conditions found in the literature, all respondents ($n = 10$) had high expectations and indicated that the benchmark collaborative was focused on learning, sharing knowledge and improving of performance. The third purpose of the 4P-model “implementation of best practices” was mentioned in documents, however, not literally mentioned by respondents. Respondents focused more on improving performance; one way to achieve improvement could be the implementation of best practices. The term “performance” was used by respondents as a collective noun for “efficiency”, “productivity”, “patient safety”, “patient satisfaction” and/or “quality of care”.

The partnership, signed by the chairman of the board of every UMC, described three purposes of benchmarking at the start of the collaborative:

- (1) to compare the utilisation of OR resources as well as the economic aspects of OR performance and learn from similar organisations, with the aim to improve this performance, as indicated by this respondent:

I believe it is important to compare your own performance with other organisations. Especially in Germany, hospital chains share OR data and are able to improve their performance (UMC8 manager).

- (2) to avoid comparing apples to oranges, information and knowledge about the underlying organisational characteristics (see Table I) and methods/processes is therefore also gathered and shared; and
- (3) to learn about the application of benchmarking in university hospitals.

Additionally, two respondents mentioned:

It is always inspiring to have the opportunity to have a look in someone else’s backyard (UMC8 manager and UMC2 OR scheduler).

Table I.
Organisational characteristics of the participating UMCs

		UMC1		UMC2		UMC3		UMC4		UMC5		UMC6		UMC7		UMC8										
		n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD							
1	Total number of annual surgical cases performed (elective as well as non-elective)	99,659	95,716	132,404	102,789	76,663	73,495	148,209	84,664	27,716	47,357	50,381	29,470	23,459	82,113	50,618										
	Inpatient cases																									
	Outpatient cases																									
	Cancer Centre (with separate OR location)																									
	Children's hospital (with separate OR location)																									
2	Effective or non-elective/emergency cases performed (%)	72/26	69/31	71/29	73/27	77/23	80/20	83/17	71/29	71/29	71/29	80/20	83/17	71/29	71/29											
	Included are all inpatient surgical cases																									
	Excluded are cases performed at separate OR locations																									
3	Anesthesia-controlled time (in minutes): anesthesia-controlled time is defined as the sum of the time starting when the patient enters the OR to the time when positioning or skin preparation can begin plus the time starting when the surgical dressing is completed and ending when the patient leaves the OR, in other words, ACT is the sum of anesthesia induction time plus anesthesia emergence time	68,213	33	21	67,550	42	25	95,492	44	26	77,125	31	19	60,574	36	21	48,485	33	22	122,625	30	21	65,957	31	19	
	Included are all elective, inpatient surgical cases																									
	Excluded are cases performed at separate OR locations																									
4	Surgeon-controlled time (in minutes): the time starting when patient positioning and/or skin preparation can begin to when surgical dressing is completed, included are all elective, inpatient surgical cases	68,213	124	102	67,550	136	112	95,492	131	110	77,125	112	108	60,574	124	107	48,485	131	125	122,625	95	92	65,957	136	109	
	Included are all elective, inpatient surgical cases																									
	Excluded are cases performed at separate OR locations																									
5	Total case duration (in minutes): patient in to patient out of the OR room. In other words, anesthesia-controlled time plus surgeon-controlled time, included are all elective, inpatient surgical cases	68,213	159	111	67,550	178	127	95,492	178	124	77,125	143	120	60,574	161	120	48,485	164	140	122,625	124	105	65,957	166	119	
	Included are all elective, inpatient surgical cases																									
	Excluded are cases performed at separate OR locations																									

Notes: Data years 2005-3rd quarter 2013; UMC1 excluding data year 2010; UMC5 excluding data Thoracic Centre with separate OR location; UMC6 excluding data years 2010-2012

Interestingly, all ten respondents mentioned the purpose of networking. This network aspect in relation with benchmarking was not mentioned as one of the purposes found in literature and incorporated in the 4P-model. Respondents mentioned the national annual conference and the two-monthly focus group study meetings arranged by the project manager as opportunities to network with colleagues from other hospitals, working in the same professional field and dealing with the same professional issues. Afterwards, these networking events were found to make it easier for participants to contact an individual professional working in another hospital, to discuss today's challenges (the "mutual support function"), share more knowledge and organise site visits to each other's OR departments. This increased the understanding and the learning between members of the network.

Performance indicators. Benchmarking requires SMART indicators (specific, measurable, acceptable, relevant and time-framed), comparable indicator information and reliable data gathering and sharing (de Korne *et al.*, 2010, 2012). According to the document study, a considerable amount of time and effort was undertaken by the steering committee to develop a partnership agreement during the initiation phase of the collaborative. This agreement creates the foundation for trust and confidentiality between the eight participating hospitals. It describes goals and objectives, requirements, opportunities, organisational structure, finance and possible termination of the partnership. Confidentiality and ownership of benchmarking data are two delicate and important parts of the agreement.

During the first years the collaborative was directly and full time assisted by an independent academic department (University of Twente), in order to develop and harmonise data definitions of OR time periods, uniform methods of data registration and definitions of performance indicators among all participants.

Longitudinal data collection within the OR benchmarking collaboration started in 2005 and is still performed today. Every UMC registers details of each surgical case and time periods – e.g. "time patient enters the OR", "time surgery starts" – since multiple years. These time periods are prospectively and continuously measured, and registered electronically by the nursing staff in each Hospital Information System and validated by the responsible surgeon and anesthesiologist. Each UMC quarterly provides records for all performed surgical cases to a central OR benchmark database. This data focuses on the OR process, and not on outcomes (e.g. mortality, morbidity) or patient safety (e.g. surgical site infections).

An independent data management centre administers the central OR benchmark database. This centre provides professional expertise to facilitate the collection and processing of data records, as well as data reliability checks. The centre calculates all key performance indicators, based on the data provided by the UMCs: actual time periods are combined with the total amount of allocated OR session time.

The performance of one OR day, which is generally equal to eight hours of block time allocated to a specific surgical department, is commonly evaluated by the indicator "raw utilisation". The time when there is no patient present in the OR, so-called "non-operative time", can be evaluated by three performance indicators: first-case tardiness, turnover time and empty OR time at the end of the day, if cases finish earlier than scheduled. If cases run longer than the regularly scheduled hours of allocated block time, this is termed over-utilised time. All these performance indicators were calculated once per OR day (see Figure 2). Table II performance indicators measured in the Dutch OR benchmarking collaborative, including definitions.

These performance indicators, combined with (trend) analyses/benchmark reports, are shared between UMCs, which enables the identification of areas of improvement by comparing one's own performance to that of other similar organisations. The respondents indicated that all benchmarking participants can access the central database at any time using a highly secured web-based application/reporting tool.

All respondents ($n = 10$) indicated to be satisfied with the current set of performance indicators:

I believe every performance indicator which is now measured and benchmarked, is useful (UMC6 data-analyst).

From the start of this collaborative we carefully discussed our operating room processes and from there we developed these indicators, and I think they are still useful to apply in OR management today (UMC8 manager).

The current set of indicators is relevant and useful (UMC6 managing director).

Five respondents expressed their interest in the development and benchmarking of additional performance indicators regarding case cancellations on the day of surgery, productivity of OR personnel as well as OR cost-prices. The indicator "case cancellations on the day of surgery" was considered to be included in the original series of metrics. However, due to difficulties with respect to harmonisation of the definition and registration method of this indicator, it was not incorporated. Recently the steering committee has decided to expand the current series with two new performance indicators in the course of 2014: e.g. labour productivity (worked hours/OR minutes) and cost-prices (OR cost per minute).

Participants. Referring to the literature, there is no competition between participants, participation is voluntary and involved, and participating organisations have sufficient similarities in structure, process and outcomes (de Korne *et al.*, 2010, 2012). According to the document analysis, hospitals addressed to participate were all UMCs. Non-university major top-clinical hospitals as well as general hospitals were excluded. There were several reasons for this selection of participants, as supported by results of the document study: a small (eight centres) group is able to build trust between participants at short notice, which facilitates collaborative (inter-organisational) learning. The eight UMCs are comparable organisations regarding patient case mix – see Table III regarding Exceptional Medical Procedures (WBMV) – and their responsibility for tertiary care, clinical research, education and innovation, which

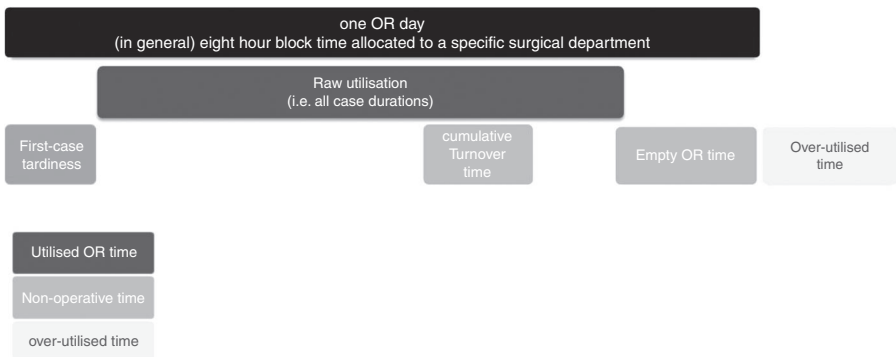


Figure 2.
Indicators to measure the performance of one OR day

Indicator and definition	UMC1		UMC2		UMC3		UMC4		UMC5		UMC6		UMC7		UMC8									
	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD						
1 Raw utilization (%): the total amount of time surgical patients are present in the OR, divided by the total amount of allocated block time per day >=100%, excluding the time spent in the OR by the staff, the time spent in the surgical department. The definition of raw utilization excluded turnover time and over-utilized OR time	16,550	84	14	36,756	83	15	50,726	85	16	33,892	80	17	30,260	84	15	25,011	76	20	41,885	84	15	33,477	82	15
2 Early start of the first surgical case of the day (in minutes): the difference between the scheduled starting time (generally 8:00 a.m.) and the actual room entry time of the first patient on that day (generally 8:00 a.m.) and the scheduled starting time (generally 8:00 a.m.)	16,550	6	4	36,756	5	4	50,726	7	5	33,892	8	5	30,260	4	4	25,011	8	8	41,885	6	5	33,477	4	4
3 Frequency early start (%): the percentage of the total number of operating rooms that started too early at a random workday	16,550	0.05		36,756	0.16		50,726	0.33		33,892	0.24		30,260	0.12		25,011	0.17		41,885	0.24		33,477	0.14	
4 First-case readiness (in minutes), or a "late start" of ready the first surgical case of the day. The difference between the scheduled starting time (generally 8:00 a.m.) and the actual room entry time of the first patient on that day (generally 8:00 a.m.) and the scheduled starting time (generally 8:00 a.m.) or exactly on the scheduled time. The common scheduled starting time was adjusted in case of an intentionally altered starting time. Every minute of lateness was calculated	16,550	30	35	36,756	21	37	50,726	26	40	33,892	36	50	30,260	16	36	25,011	32	46	41,885	19	36	33,477	20	31
5 Frequency of turnovers: the percentage of the total number of operating rooms that started too late at a random workday	16,550	0.71		36,756	0.64		50,726	0.54		33,892	0.59		30,260	0.56		25,011	0.64		41,885	0.63		33,477	0.74	
6 Turnover time: the absolute number of minutes per OR day. Turnover time was defined as the interval between two successive cases. The time between one patient leaving the OR and the next patient entering that OR1, also known as cleaning time	16,550	29	24	36,756	34	26	50,726	34	27	33,892	40	31	30,260	27	21	25,011	39	31	41,885	34	29	33,477	33	23
7 Frequency of turnovers: the absolute number of turnovers per operating room per workday	16,550	1.79		36,756	1.68		50,726	1.43		33,892	1.81		30,260	1.52		25,011	1.16		41,885	2.3		33,477	1.46	1
8 Scheduled time: the end of the day (also called "under-utilized time" at the end of the day). It was quantified by the difference between the actual and scheduled (generally 16:00h) room exit time of the last patient of the day, divided by the total number of operating rooms. The common scheduled allocated block time, more than the standard of eight hours	16,550	56	50	36,756	62	51	50,726	61	53	33,892	64	55	30,260	66	55	25,011	88	64	41,885	54	52	33,477	63	53
9 Frequency under-utilized time (%): the percentage of the total number of operating rooms that finished too early at a random workday	16,550	0.47		36,756	0.56		50,726	0.43		33,892	0.47		30,260	0.45		25,011	0.52		41,885	0.46		33,477	0.56	
10 Over-utilized time: was quantified by the difference between the actual and scheduled (generally 16:00h) room exit time of the last patient of the day, finishing after 16:00h. The common scheduled finishing time was adjusted in case of an intentionally extended finishing time (extended allocated block time, more than the standard of eight hours)	16,550	60	62	36,756	53	51	50,726	62	61	33,892	57	51	30,260	57	51	25,011	71	50	41,885	48	46	33,477	55	51
11 Frequency over-utilized time (%): the percentage of the total number of operating rooms that finished too late at a random workday	16,550	0.48		36,756	0.38		50,726	0.46		33,892	0.41		30,260	0.42		25,011	0.38		41,885	0.49		33,477	0.35	
12 Mean scheduling deviation (%): the percentage difference between the realized and expected/scheduled total case duration, divided by the expected total case duration		21		11			24			22			12			10								
13 Absolute scheduling deviation (%): the absolute difference between the realized and expected/scheduled total case duration, divided by the expected total case duration		35		29			36			36			28			29								
14 Number of surgical cases during the night: the absolute number of surgical cases operated on between midnight 00:00h and 07:00h AM	1,033			2,942			3,946			3,227			1,960			1,158								1,866

Notes: Included are all elective, inpatient surgical cases data years 2005-3rd quarter 2013; UMC1 excluding data year 2010; UMC5 excluding data Thoracic Centre with separate OR location; UMC6 excluding data years 2010-2012

Table II.
OR performance
indicators applied
for benchmarking

enables a fair comparison. Hence, all benchmark participants showed sufficient similarities in structure, process and outcomes. These participants also share the same interest in (current) issues regarding the OR environment.

The respondents affirmed that the aim of this OR benchmarking initiative is learning and therefore the relationship between the UMCs is one of collaboration instead of competition. For six of the eight hospitals, the geographical distance is large enough since they are located in different provinces. Also the two UMCs situated in the same city in the Netherlands confirm that their relationship is not competitive, since there has always been collaboration between the two centres, which is intensifying in the near future because of concentration and task distribution of the most complex care, as well as a possible merger of the two centres.

The OR benchmarking collaborative was initiated by a surgeon and a manager working in one of the UMCs. The other seven centres were personally addressed to participate. Participation was not mandated by the government or other third parties but purely voluntary. The number of participants has remained unchanged since the start of the benchmarking collaboration. During OR benchmark meetings every UMC is represented; this was shown by document analysis of multidisciplinary focus group study meetings, conferences, steering and project committee meetings. The majority of the interview respondents pronounced to be satisfied with the content of OR benchmark meetings.

Performance management system. The 4P-model identified three conditions for the internal performance management systems of organisations participating in benchmarking (de Korne *et al.*, 2010, 2012):

- (1) managers must have knowledge about the performance indicators used and outcomes;
- (2) benchmarking findings have to be communicated to stakeholders in the organisation, to have any effect on performance; and
- (3) benchmarking needs to be incorporated in a continuous quality improvement model: the PDSA cycle.

Although all participants can access the central OR benchmark database at any time using a highly secured web-based reporting tool, the project manager also provides participants (solicited and unsolicited) with benchmark analyses and reports. These reports as well as actualities and urgent subjects concerning the OR, set the agenda for benchmarking meetings. The indicators used for benchmarking are indicators prevailing in general OR management (e.g. utilisation rates) and largely integrated in the local decision-making process of the participating UMCs. Four participants ($n = 4$) confirmed that benchmark results are habitually included into management reports for the board of directors. However, the incorporation of benchmark results into the local performance management and reporting system was not self-evident in every participating centre:

To a limited extent. Benchmark results are sometimes used or mentioned in presentations. We do not include the actual benchmark results into our standard management report. We should do so more often (UMC6 data-analyst and UMC5 OR scheduler).

Actually, at this moment I think it would be a good idea to include a summary of the benchmark report in a newsletter (UMC5 manager).

Document study revealed that since the start, two-monthly multidisciplinary focus group study meetings are organised to discuss the results of the data analysis and explore processes and practices “behind the data”. These focus group study meetings are usually visited by approximately 25 to 30 professionals per meeting from all eight UMCs; these professionals represent OR management, anesthesiologists, surgeons, OR nurses, anesthesia nurses and staff advisors. However, healthcare professionals that visit a focus group study meeting are not perpetual delegates since they are not obliged to visit the following meetings. As mentioned earlier, once per year a national invitational conference is organised to provide a broader learning and knowledge-sharing platform. The annual conferences are visited by approximately 200 professionals. Through these meetings, the collaborative tries to involve as many stakeholders and employees as possible in the eight participating hospitals. Through promoting dialogue between the participants a learning environment is created.

Recently, the Dutch OR benchmarking collaborative published a study in *The American Journal of Surgery*, showing that 43 per cent of all first operations start at least 5 min later than scheduled and 425,612 minutes are lost because of this annually, which has a respectable economic impact. This study also demonstrated that on an overall level of all UMCs, first-case tardiness has decreased since 2005. Moreover, it showed that four centres implemented successful interventions to reduce tardiness. These UMCs showed a stepwise reduction in variation of first-case tardiness, in other words a decrease in IQR during the years, which indicates an organisational learning effect (Sehwail and de Yong, 2003). The implemented interventions entailed, e.g. providing feedbacks directly when ORs started too late, new agreements between OR and intensive care unit departments concerning “intensive care unit bed release” policy, and a shift in responsibilities regarding transport of patients to the OR. One of the UMCs realised a reduction of 27,392 minutes of first-case tardiness in one year. Based on \$13.29 per regularly scheduled minute of OR time including labour costs, supply costs, indirect costs, anesthesiologist fees, and surgeon fees, this meant possible savings of \$364,040 that year.

First-case tardiness is merely one of the performance indicators measured and benchmarked in this collaborative. Even though the improvements in tardiness were driven by the Dutch OR benchmark, the “cyclic improvement”-approach needs more attention to guarantee similar achievements and secure the long-term existence of this collaborative.

Alternatively, the central OR benchmark database maintained by the collaborative is frequently used for multicentre research on scheduling and efficiency topics. This research proposes recommendations built on extensive data and statistical analyses, concerning the improvement of OR scheduling. Recent research results considering the influence of anesthesia-controlled time was published in the *Canadian Journal of Anesthesia* (van Veen-Berkx *et al.*, 2014b). This publication appears to be a start towards opening the discussion on this topic in several participating UMCs. Later on this might become a starting point for improvement.

Conclusion

This study investigated whether the collaborative, long-term approach of the Dutch OR benchmarking initiative leads to benefits in OR management and evaluated if the initiative met the requirements of the 4P-model (de Korne *et al.*, 2010, 2012). Based on

the findings we discuss the applicability of the 4P-model and present key findings useful for benchmarking in (university) hospital settings.

The findings of this investigation show that collaborative benchmarking appears to have benefits different from mainly performance improvement and identification of performance gaps. It is interesting to note that, since 2004, the OR benchmarking initiative still endures after already existing for ten years. A key benefit was pointed out in this recent study by all respondents as “the purpose of networking”, on top of the purposes recognised in the 4P-model. The networking events organised by the collaborative were found to make it easier for participants to contact and also visit one another. Apparently, such informal contacts were helpful in spreading knowledge, sharing policy documents and initiating improvement. One reason for this is that they could be used to discuss the tacit components of best practices, that are hard to share in more formal communication media. Respondents were satisfied with the content of these meetings and with the exchange of knowledge in an informal manner, the exchange of experiences including sharing best practices as well as discussing worries and today’s challenges in OR management. It enables understanding and learning from each other. These findings corroborate the idea of de Korne *et al.* (2010, 2012) that participating in benchmarking offers other advantages, such as generating discussions about how to deliver services and increasing the interaction between participants.

This case study showed that this benchmark largely met all key conditions of the 4P-model (de Korne *et al.*, 2010). However, the “cyclical plan-do-study-act improvement approach”, which is the third necessary condition with respect to the internal performance management system of organisations participating in benchmarking, was not applied in each activity arranged by the Dutch collaborative. Examples of successful application of this approach, e.g. first-case tardiness (van Veen-Berkx *et al.*, 2014a), do exist but a “continuing improvement cycle” was not completely incorporated.

Discussion

The OR benchmarking collaborative saves the eight participating UMCs from reinventing the wheel regarding several issues high on the agenda of OR departments. de Korne *et al.* (2010, 2012) has indicated that “taking part in an international benchmarking initiative is in itself seen as a powerful signal to stakeholders that the organisation is actively working on quality improvement”. Although the OR benchmark is a national initiative, the reputation it builds could be another possible explanation for the long-term commitment of the eight centres to the collaborative. At the end of every year there is a clear decision point whether every UMC wishes to continue its participation the upcoming year and is willing to pay the yearly participation fee charged on the OR budget.

During the initiation phase of the benchmark collaborative, a considerable amount of time (two years) and effort was undertaken by the steering committee to develop a collaboration agreement. As described in the findings, this agreement created the foundation for trust and confidentiality between the eight participating partners, because confidentiality and ownership of benchmarking data are two delicate and important parts of the agreement. These first years were also seized by the development and harmonisation of definitions of performance indicators. Common definitions are an essential base for external benchmarking (Fixler and Wright, 2013; Kazemier and van Veen-Berkx, 2013). The long-term commitment of the eight centres to

the OR benchmark collaborative is exceptional, yet might also be necessary to build and maintain trust between the centres, and also be necessary for uniform data registration and harmonisation of indicator definitions.

Benchmarking has often been approached as a competitive activity resulting in rankings and with a focus on creating competition between participants as driver for improvement. This study, however, clearly shows the advantages of a more collaborative approach. An important difference between public reporting and reporting arranged in this Dutch benchmarking collaborative is the fact that the performance as well as rankings are not publicly available elsewhere than to the eight participating UMCs. When information is publicly and freely available, it will be more difficult to build a relation of trust. This is not surprising, since attempts to increase transparency of professional work represent a potential threat to professional autonomy and therefore, professionals often react with suspicion and a certain amount of resistance (Levay and Waks, 2009). However, when professionals are actively involved in transparency technologies through translation and negotiation in expert networks, public quality reporting can actually become acceptable and advantageous. Advantageous with regard to retaining control over (external) evaluation criteria and drawing attention to professional activities and improvement efforts in order to gain legitimacy and support from external actors (Levay and Waks, 2009).

From the very first start, the initiators of the Dutch OR benchmarking collaborative as described in this study consistently and literally have avoided “naming and shaming” through publishing and vertical ranking of the eight UMCs, regarding the performance indicators measured. Lots of attention has been given to honest assessment and avoiding to compare apples and oranges. The physical, organisational characteristics and structure of all participating OR departments can be very different from one another. Contingency theory claims there is not “one best way for organising” because this is subject to the internal and external conditions of every organisation (Burns and Stalker, 1961; Lawrence and Lorsch, 1969; Mintzberg, 1979). Differences in organisational characteristics derive from differences in organisational conditions. Therefore, performance indicators used for benchmarking should take into account these differences, to avoid inaccurate interpretation of observed differences between organisations and to accomplish an honest comparison.

The character of benchmarks using DEAs is essentially different from the character of the Dutch OR benchmarking collaborative since it was initiated by the eight university hospitals themselves and not by a third external party. Moreover, data are derived from the local Hospital Information Systems, which are used for daily registration practices. The Dutch OR benchmarking collaborative is a “self-led” and voluntary collaboration with its own budget (paid for by the eight hospitals themselves). OR benchmark data are merely used by the participants and not by policy makers, the government or regulatory offices.

Another foundation of the collaborative benchmark described in this study, is the pursuit to learn from the organisational differences in structure, process designs, methods and performance. These differences can be a source of learning as they allow practitioners to compare relations between organisational characteristics and performance, especially in informal settings and networking. These differences also offer every participating OR department the opportunity to engage their own quality improvement pathway. Improvement starts with quantitative analyses and therefore performance indicators should be SMART. In this collaborative the inter-organisational

or “joint learning process” is more important than ranking participants or to identify “the best practice”. The OR departments of the eight UMCs are all providing the same healthcare product: perioperative care in a university hospital setting. It is important to gain insight into managing and controlling this process as well as insight into performance differences, to realise the “best fit” for each OR department.

The Dutch OR benchmark collaborative bears a resemblance to “quality improvement collaboratives” that became popular as the “breakthrough series”, an improvement method developed by the Institute for Healthcare Improvement (IHI) in Boston (IHI Innovation Series White Paper, 2003). Nembhard (2012) describes these collaboratives as “structured programmes in which multidisciplinary teams from different organizations work to improve care in one area of their operations (e.g. infection control). As part of a collaborative, teams attend a series of meetings where they learn about best practices in their target area, quality improvement techniques, and the experiences of others that have implemented new practices”. The OR benchmark focus group study meetings as described in our findings share the same goals as the collaboratives and have many similarities.

Nevertheless, the OR benchmark collaborative could learn from the IHI breakthrough series approach to develop a more structured PDSA-approach. Specifically with regard to the commitment of the participants during study meetings and the (learning) activities in between physical meetings. When a healthcare professional decides to participate in a breakthrough series, he commits to participate actively throughout the limited collaborative period. In general, this period is limited to 6-18 months, which is supposed to drive change. In between physical meetings, teams are expected to implement changes in their own organisation and it is mandatory to share implementation experiences with each other for collective learning through conference calls or digital (internet) platforms (Kilo, 1998, 1999; Nembhard, 2012). Clearly, this kind of “stable commitment” through continuous participation was not established in the OR benchmark collaborative in this recent study. Healthcare professionals that visit a focus group study meeting are not perpetual delegates since they are not obliged to visit the following meetings. The responsibility for improvement was kept an individual responsibility of each single UMC and not a collaborative responsibility. Future research should therefore concentrate on the investigation of the relation between benchmarking as instrument and the actual performance improvements realised through benchmarking in the local UMC’s.

The current study has the limitations accompanied with any qualitative research and particularly related to interviewing (Bowling, 2009; Silverman, 2000). First, qualitative research findings must be viewed within the context of the conducted case study (Bowling, 2009; Silverman, 2000). The perceptions and experiences of the respondents in this university hospital context in the Netherlands might not be transferable to other (general) hospital settings or other countries. Second, the number of conducted interviews is restricted; nevertheless, all other data sources are extensive. Third, while all members of the project committee and regular visitors of the OR benchmark meetings were invited to interview, it might be possible that “contribution bias” was present with the respondents who reacted the quickest, being hypothetically those who had more interest in the benchmark collaborative. Despite these limitations, this study provides valuable insights of experiences with benchmarking from a variety of participating centres representing different parts of the Netherlands. The context of the case study and conditions under which this specific benchmarking process took place, was comprehensively outlined, to allow for transferring of results

to other settings. In order to increase the validity of the study, data triangulation was applied and a variety of data sources were used. Moreover, construct validity was ensured by deploying several researchers to evaluate the analyses, operating separately from one another (Bowling, 2009; Silverman, 2000).

Benchmarking is defined as a “continuous process” (APQC, 2008) and encourages the use of a continuous quality improvement model (the PDSA cycle). Although this OR benchmark initiative, as many benchmark initiatives (Askim *et al.*, 2008), started with a stated aim to improve, actual (measurable) quality or performance improvements are not necessary for this initiative to endure. These findings further support the idea of de Korne *et al.* (2010, 2012) that benchmarking is relying on iterative and social processes in combination with structured and rational process of performance comparison. The relatively limited focus on OR utilisation in this benchmark seems to be a starting point for exchanging a variety of information and experiences considering the structure, process and performance of OR departments. More attention needs to be given to the relation between benchmarking as instrument and the actual performance improvements realised through benchmarking in the local UMC's. A collaborative approach in benchmarking can be effective because participants use its knowledge-sharing infrastructure which enables operational, tactical and strategic learning. Organisational learning is to the advantage of overall OR management. Benchmarking seems a useful instrument in enabling hospitals to learn from each other, to initiate performance improvements and catalyse knowledge-sharing.

Acknowledgements

The authors sincerely thank all members of the steering and project committee of the Dutch Operating Room Benchmarking Collaborative for their participation and contribution. Current Members of the Dutch Operating Room Benchmarking Collaborative (the Netherlands): Academic Medical Center Amsterdam: Ron Balm, MD, PhD, Department of Vascular Surgery; Diederich C.C. Cornelisse, MSc, Department of Operating Rooms; Maastricht University Medical Center: Wolfgang F. Buhre, MD, PhD, Division of Anesthesiology; Hub J. Ackermans, Department of Operating Rooms; Erasmus University Medical Center Rotterdam: Robert Jan Stolker, MD, PhD, Department of Anesthesiology and Division of Emergency, Perioperative and Intensive Care; Jeanne Bezstarosti, MD, Department of Anesthesiology and Department of Operating Rooms; Leiden University Medical Center: Rob C.M. Pelger, MD, PhD, Department of Urology; Roald R. Schaad, MD, Department of Anesthesiology; University Medical Center Groningen: Irmgard Krooneman-Smits, MBA, Department of Operating Rooms; Peter Meyer, MD, PhD, Department of Anesthesiology and Department of Operating Rooms; Radboud University Medical Center Nijmegen: Hein G. Gooszen, MD, PhD, Department of Operating Rooms; Mirjam van Dijk-Jager, Department of Operating Rooms; Simon A.W. Broecheler, MSc, Department of Operating Rooms and Department of Anesthesiology; University Medical Center Utrecht: A. Christiaan Kroese, MD, Department of Anesthesiology and Department of Operating Rooms; Jeffrey Kanters, Department of Operating Rooms; University of Twente: Johannes J. Krabbendam, PhD; Erwin W. Hans, PhD, Department of Operational Methods for Production and Logistics; VU University Medical Center: Marinus F. Caljouw, Division of Anesthesiology, Intensive Care and Emergency Medicine; Derk P. Veerman, MD, PhD, Division of Anesthesiology, Intensive Care and Emergency Medicine; Kjeld H. Aij, MBA, Division of Anesthesiology, Intensive Care and Emergency Medicine.

References

- APQC (2008), "Benchmarking code of conduct. Guidelines and ethics for benchmarkers", American Productivity and Quality Center (APQC), Houston, TX.
- Askim, J., Johnsen, A. and Christophersen, K.A. (2008), "Factors behind organizational learning from benchmarking: experiences from Norwegian municipal benchmarking networks", *Journal of Public Administration Research and Theory*, Vol. 18 No. 2, pp. 297-320.
- Berg, M., Meijerink, Y., Gras, M., Goossensen, A., Schellekens, W., Haeck, J., Kallewaard, M. and Kingma, H. (2005), "Feasibility first: developing public performance indicators on patient safety and clinical effectiveness for Dutch hospitals", *Health Policy*, Vol. 75 No. 1, pp. 59-73.
- Bevan, G. and Hood, C. (2006), "What's measured is what matters: targets and gaming in the English public health care system", *Public Administration*, Vol. 84 No. 3, pp. 517-538.
- Blank, J.L. (2008), "Innovations and productivity: an empirical investigation in Dutch hospital industry", *Advances in Health Economics and Health Services Research*, Vol. 18 No. 1, pp. 89-109.
- Blank, J.L. and Valdmanis, V.G. (2008), "Productivity in hospital industry", *Advances in Health Economics and Health Services Research*, Vol. 18 No. 1, pp. 3-12.
- Blank, J.L. and Valdmanis, V.G. (2010), "Environmental factors and productivity on Dutch hospitals: a semi-parametric approach", *Health Care Management Science*, Vol. 13 No. 1, pp. 27-34.
- Booth, M., James, P. and Stevanovic, V. (2005), "Benchmarking of hospital activity data: an international comparison", *Benchmarking: An International Journal*, Vol. 12 No. 6, pp. 515-522.
- Bowling, A. (2009), *Research Methods in Health. Investigating Health and Health Services*, Open University Press, Berkshire.
- Burns, T. and Stalker, G.M. (1961), *The Management of Innovation*, Tavistock, London.
- Camp, R. (1989), *Benchmarking: The Search for the Best Practices that Lead to Superior Performance*, QR, New York, NY.
- Camp, R.C. and Tweet, A.G. (1994), "Benchmarking applied to health care", *The Joint Commission Journal on Quality Improvement*, Vol. 20 No. 5, pp. 229-238.
- Creswell, J.W. (2003), *Research Design. Qualitative, Quantitative and Mixed Methods Approaches*, Sage Publications, London.
- Dattakumar, R. and Jagadeesh, R. (2003), "A review of literature on benchmarking", *Benchmarking: An International Journal*, Vol. 10 No. 3, pp. 176-209.
- de Korne, D.F., van Wijngaarden, J.D., Sol, K.J., Betz, R., Thomas, R.C., Schein, O.D. and Klazinga, N.S. (2012), "Hospital benchmarking: are US eye hospitals ready?", *Health Care Management Review*, Vol. 37 No. 2, pp. 187-198.
- de Korne, D.F., Sol, K.J., van Wijngaarden, J.D., van Vliet, E.J., Custers, T., Cubbon, M., Spileers, W., Ygge, J., Ang, C.L. and Klazinga, N.S. (2010), "Evaluation of an international benchmarking initiative in nine eye hospitals", *Health Care Management Review*, Vol. 35 No. 1, pp. 23-35.
- Devers, K.J., Pham, H.H. and Liu, G. (2004), "What is driving hospitals' patient safety efforts?", *Health Affairs*, Vol. 23 No. 2, pp. 103-115.
- Eijkemans, M.J., van Houdenhoven, M., Nguyen, T., Boersma, E., Steyerberg, E.W. and Kazemier, G. (2010), "Predicting the unpredictable: a new prediction model for operating room times using individual characteristics and the surgeon's estimate", *Anesthesiology*, Vol. 112 No. 1, pp. 41-49.
- Enthoven, A.C. and Tollen, L.A. (2005), "Competition in health care: it takes systems to pursue quality and efficiency", *Health Affairs*, Suppl. Web Exclusives: W5-420-33, 7 September.
- Field, A. (2013), *Discovering Statistics Using IBM SPSS Statistics*, 4th rev. ed., SAGE Publications Ltd, Thousand Oaks, CA.

- Fixler, T. and Wright, J.G. (2013), "Identification and use of operating room efficiency indicators: the problem of definition", *Canadian Journal of Surgery*, Vol. 56 No. 4, pp. 224-226.
- Giard, R.W. (2006), "Top 100 hospitals: varying lists, varying reputations", *Nederlands Tijdschrift voor Geneeskunde*, Vol. 150 No. 43, pp. 2355-2358.
- Gift, R.G., Stoddart, T.D. and Wilson, K.B. (1994), "Collaborative benchmarking in a healthcare system", *Healthcare Financial Management*, Vol. 48 No. 9, pp. 80-82, 84-86, 88.
- Güven-Uslu, P. (2005), "Benchmarking in health services", *Benchmarking: An International Journal*, Vol. 12 No. 4, pp. 293-309.
- Hibbard, J.H., Stockard, J. and Tusler, M. (2005), "Hospital performance reports: impact on quality, market share, and reputation", *Health Affairs*, Vol. 24 No. 4, pp. 1150-1160.
- IHI Innovation Series White Paper (2003), *The Breakthrough Series: IHI's Collaborative Model for Achieving Breakthrough Improvement*, Original edition, Institute for Healthcare Improvement, Boston, MA, available at: www.IHI.org (accessed 27 January 2014).
- Joint Commission (2012), *Benchmarking in Health Care*, 2nd ed., Joint Commission Resources, Oakbrook Terrace, IL.
- Kacak, H., Özcan, Y.A., Lins, M.P.E., Silva, A.C.M. and Fiszman, R. (2014), "Determinants of efficiency for teaching hospitals in Brazil", *International Journal of Healthcare Management*, Vol. 7 No. 1, pp. 60-68.
- Kazemier, G. and van Veen-Berkx, E. (2013), "Comment on 'identification and use of operating room efficiency indicators: the problem of definition'", *Canadian Journal of Surgery*, Vol. 56 No. 5, pp. E103-E104.
- Kilo, C.M. (1998), "A framework for collaborative improvement: lessons from the Institute for Healthcare Improvement's breakthrough series", *Quality Management in Health Care*, Vol. 6 No. 4, pp. 1-13.
- Kilo, C.M. (1999), "Improving care through collaboration", *Pediatrics*, Vol. 103 No. 51E, pp. 384-393.
- Lawrence, P.R. and Lorsch, J.W. (1969), *Organization and Environment: Managing Differentiation and Integration*, Irwin, Homewood (IL).
- Levay, C. and Waks, C. (2009), "Professions and the pursuit of transparency in healthcare: two cases of soft autonomy", *Organization Studies*, Vol. 30 No. 5, pp. 509-527.
- Lingsma, H.F., Steyerberg, E.W., Eijkemans, M.J., Dippel, D.W., Scholte Op Reimer, W.J. and van Houwelingen, H.C., Investigators Netherlands Stroke Survey (2010), "Comparing and ranking hospitals based on outcome: results from the Netherlands Stroke Survey", *QJM*, Vol. 103 No. 2, pp. 99-108.
- Mant, J. (2001), "Process versus outcome indicators in the assessment of quality of health care", *International Journal of Quality in Health Care*, Vol. 13 No. 6, pp. 475-480.
- Marjamaa, R., Vakkuri, A. and Kirvela, O. (2008), "Operating room management: why, how and by whom?", *Acta Anaesthesiologica Scandinavica*, Vol. 52 No. 5, pp. 596-600.
- Mintzberg, H. (1979), *The Structuring of Organisations*, Prentice-Hall, New York, NY.
- Mosel, D. and Gift, B. (1994), "Collaborative benchmarking in health care", *Joint Commission Journal on Quality Improvement*, Vol. 20 No. 5, pp. 239-249.
- Nembhard, I.M. (2012), "All teach, all learn, all improve? The role of interorganizational learning in quality improvement collaboratives", *Health Care Management Review*, Vol. 37 No. 2, pp. 154-164.
- Porter, M.E. and Teisberg, E.O. (2006), *Redefining Healthcare: Creating Value-Based Competition on Results*, Harvard Business School Press, Cambridge.

- Quartz, J., Wallenburg, I. and Bal, R. (2013), *The Performativity of Rankings. On the Organizational Effects of Hospital League Tables*, iBMG, Rotterdam.
- Scanlon, D.P., Swaminathan, S., Lee, W. and Chernew, M. (2008), "Does competition improve health care quality?", *Health Services Research*, Vol. 43 No. 6, pp. 1931-1951.
- Schut, F.T. and van de Ven, W.P.M.M. (2005), "Rationing and competition in the Dutch health care system", *Health Economics*, Vol. 14 No. S1, pp. S59-S74.
- Schut, F.T. and van de Ven, W.P.M.M. (2011), "Health care reform in the Netherlands: the fairest of all?", *Journal of Health Services Research and Policy*, Vol. 16 No. 1, pp. 3-4.
- Sehwal, L. and de Yong, C. (2003), "Six sigma in health care", *International Journal of Health Care Quality Assurance*, Vol. 16 No. 4, pp. 1-5.
- Silverman, D. (2000), *Doing Qualitative Research: A Practical Handbook*, Vol. 1, SAGE Publications, London.
- Spendingolini, M.J. (1992), *The Benchmark Book*, Amacom, New York, NY.
- Sower, V.E., Duffy, J.A. and Kohers, G. (2008), *Benchmarking for Hospitals. Achieving Best-in-Class Performance without Having to Reinvent the Wheel*, ASQ Quality Press, Milwaukee, WI.
- Van de Bovenkamp, H., Vollaard, H., Trappenburg, M. and Grit, K. (2013), "Voice and choice by delegation", *Journal of Health Politics, Policy and Law*, Vol. 38 No. 1, pp. 57-88.
- van Dishoeck, A.M., Lingsma, H.F., Mackenbach, J.P. and Steyerberg, E.W. (2011), "Random variation and rankability of hospitals using outcome indicators", *BMJ Quality and Safety*, Vol. 20 No. 10, pp. 869-874.
- van Lent, W.A., de Beer, R.D. and van Harten, W.H. (2010), "International benchmarking of specialty hospitals. A series of case studies on comprehensive cancer centres", *BMC Health Services Research*, Vol. 10 No. 253, pp. 1-11.
- van Veen-Berkx, E., Elkhuizen, S.G., Kalkman, C.J., Buhre, W.F. and Kazemier, G. (2014a), "Successful interventions to reduce first-case tardiness in Dutch university medical centers: results of a nationwide operating room benchmark study", *The American Journal of Surgery*, Vol. 207 No. 6, pp. 949-959.
- van Veen-Berkx, E., Bitter, J., Elkhuizen, S.G., Buhre, W.F., Kalkman, C.J., Gooszen, H.G. and Kazemier, G. (2014b), "The influence of anesthesia-controlled time on operating room scheduling in Dutch university medical centres", *Canadian Journal of Anaesthesia*, Vol. 61 No. 6, pp. 524-532.
- van Veen-Berkx, E., Elkhuizen, S.G., van Logten, S., Buhre, W.F., Kalkman, C.J., Gooszen, H.G., Kazemier, G. and Collaborative Dutch Operating Room Benchmarking (2014c), "Enhancement opportunities in operating room utilization; with a statistical appendix", *Journal of Surgical Research*, Vol. 194 No. 1, pp. 43-51.
- Watson, G.H. (1993), *Strategic Benchmarking: How to Rate Your Company's Performance Against the World's Best*, John Wiley & Sons, New York, NY.
- Welborn, C. and Bullington, K. (2013), "Benchmarking award winning health care organizations in the USA", *Benchmarking: An International Journal*, Vol. 20 No. 6, pp. 765-776.
- Wolfram Cox, J.R., Mann, L. and Samson, D. (1997), "Benchmarking as a mixed metaphor: disentangling assumptions of competition and collaboration", *Journal of Management Studies*, Vol. 34 No. 2, pp. 285-314.
- Yin, R.K. (2003), *Case Study Research. Design and Methods*, Sage Publications, London.

- Zairi, M. (1992), *Competitive Benchmarking: An Executive Guide*, TQM Practitioner Series, Technical Communications (Pub.), Letchworth and Hertfordshire.
- Zairi, M. and Ahmed, P. (1999), "Benchmarking maturity as we approach the next millennium", *Total Quality Management Journal*, Vol. 4 No. 5, pp. 810-816.
- Zairi, M. and Leonard, P. (1994), *Practical Benchmarking: The Complete Guide: A Complete Guide*, Springer Science and Business Media, New York, NY.

About the authors

Elizabeth van Veen-Berkx is a PhD Candidate at the Department of Operating Rooms at the Erasmus University Medical Centre Rotterdam, The Netherlands. Her research focuses on quality and efficiency in health care, in particular operating room management, organisation and performance benchmarking. Elizabeth holds an MSc in Health Policy and Management from the Erasmus University Rotterdam. Elizabeth van Veen-Berkx is the corresponding author and can be contacted at: e.berkx@erasmusmc.nl

Dirk F. de Korne is the Deputy Director of Health Innovation at Singapore National Eye Centre, SingHealth. He also holds positions as an Adjunct Assistant Professor at the Duke-National University of Singapore Graduate Medical School, Singapore, and Erasmus University Rotterdam, Netherlands. His research, teaching and implementation work focuses on understanding the diffusion of quality and patient safety innovations and include business process and systems redesign, human factors and ergonomics, and performance benchmarking. Dirk holds an MSc in Health Policy and Management from the Erasmus University Rotterdam and a PhD in Public Health from the University of Amsterdam.

Olivier S. Olivier holds a BSc in Health Policy and Management; he graduated in 2013 at the Erasmus University Rotterdam. His current position is Project Manager at Medworq. His focus areas are: business intelligence, data analysis and data management.

Roland A. Bal is the Chair of the Health Care Governance section at the Department of Health Policy and Management from the Erasmus University Rotterdam. Currently, he also functions as the Director of Studies, being responsible for all the undergraduate and graduate programs of the department. Roland is a Professor of Healthcare Governance and his main research interests are on the interaction of science, technology and society, as well as on science-policy relations, which are studied in several contexts and in several empirical domains.

Geert Kazemier is a Professor of Hepatobiliary Surgery and Transplantation at the VU University Medical Centre in Amsterdam. Currently, he holds the position of the Director of the VUmc Cancer Centre Amsterdam. Geert specializes in Liver, Pancreas and Biliary and Transplantation Surgery. Before he was Head of the Department of Operating Rooms at the Erasmus University Medical Centre in Rotterdam and Co-founder of "Erasmus Health Care Logistics", a joint venture between the Institute of Health Policy and Management of the Erasmus University and the Econometric Institute of the Erasmus School of Economics.