Original Article

Selecting indicators for international benchmarking of radiotherapy centres

W.A.M. van Lent[1,4], R.D. de Beer[2], B. van Triest[3], W.H. van Harten[1,4]

[1]Division of Psychosocial Research and Epidemiology, Netherlands Cancer Institute—Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands, [2]Ministry of Health, Welfare and Sport, The Netherlands, [3]Radiotherapy department, Netherlands Cancer Institute—Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands, [4]Department of Health Technology Services Research, School of Management and Governance, University of Twente, The Netherlands

Abstract

Introduction: Benchmarking can be used to improve hospital performance. It is however not easy to develop a concise and meaningful set of indicators on aspects related to operations management. We developed an indicator set for managers and evaluated its use in an international benchmark of radiotherapy centres. The indicator set assessed the efficiency, patient-centeredness and timeliness of the services delivered.

Methods: We identified possible indicators from literature and professionals. Stakeholders' feedback helped to produce a shortlist of indicators. For this indicator set, data were obtained in a pilot that included four European radiotherapy centres. With these data, the indicators were evaluated on definition clarity, data availability, reliability and discriminative value.

Results: Literature produced a gross list of 81 indicators. Based on stakeholder feedback, 33 indicators were selected and evaluated in the benchmark. Six negatively evaluated indicators were adapted, together with eight positively evaluated indicators 14 indicators seemed feasible. Examples of indicators concerned utilisation, waiting times, patient satisfaction and risk analysis.

Conclusions: This study provides a pragmatic indicator development process for international benchmarks on operations management. The presented indicators showed to be feasible for use in international benchmarking of radiotherapy centres. The pilot identified attainable performance levels and provided leads for improvements.

Keywords

Correspondence to: W.A.M. van Lent, MSc, Division of Psychosocial Research and Epidemiology, Netherlands Cancer Institute − Antoni van Leeuwenhoek Hospital, PO Box 90203, 1006 BE Amsterdam, The Netherlands. Tel: ++31 (0)20 512 2855 Fax: ++31 (0)20 669 1449. E-mail: w.v.lent@nki.nl

INTRODUCTION

Improving the performance of quality of care was and is an important item on the agenda of

hospitals and radiotherapy departments. Improvement initiatives used to focus on clinical effectiveness and patient-centredness.[1,2] Gradually, a broader definition of quality was accepted that also included societal concerns over access to health care, effectiveness, efficiency and safety.[3]

Benchmarking, a technique that originated in operations management, is used to identify good and best practices.[4] It is a stepwise process whereby best practices are identified by comparing similar processes and are then transposed to other situations so as to achieve major process improvements.[5] To increase transparency on performance, an increasing number of medical and managerial performance indicators are presented in public reports[6] or offered through consultancy firms. At first glance, they seem to provide useful information, but they hardly explain how these results were achieved. Hospitals might benefit from more thorough national or international benchmarking methods that provide insight into the underlying organizational principles. As the topic seems to be covered mainly in popular management literature, there are few peer-reviewed publications on benchmarking and its use in health organisations.[7,8] Moreover, it is known that there can be considerable performance differences between countries (and regions),[9] so exploring *international* benchmarking on the operations management of hospitals or hospital departments can be relevant. Recent research on the process of international benchmarking on operations management showed that publications on this subject are scarce and that the selection of indicators is an important issue.[10]

This article shows the development of an indicator set for (international) benchmarks on operations management in hospitals. We selected and evaluated an indicator set assessing the efficiency, patient-centredness and timeliness of radiotherapy centres in an international setting. The set we obtained was evaluated in a benchmark exercise in four European radiotherapy centres that are also actively involved in research and training.

METHODS

Benchmarking process

Many benchmarks are based on the stepwise process described by Spendolini.[11] Van Hoorn et al.[12] adapted the process to compare hospitals using indicators that achieve consensus among stakeholders. The latter is important: those that receive the information may have different perspectives on performance and quality of care.[13] The indicators were primarily developed for managers; however, the researchers asked feedback from a broader range of radiotherapy stakeholders to increase support for this set. This resulted in a set that combined the perspectives of all stakeholders as performance on one aspect (for example, staffs used to treat the patients) affects other aspects (research outcomes). For the benchmark pilot, we further adjusted the Van Hoorn benchmarking process[12] for purposes of international comparison of radiotherapy centres (Figure 1). For more details into the process of benchmarking used in this case study, see van Lent et al.[10]

Indicator-selection process

Figure 2 summarizes the indicator-selection process. To develop a gross list of indicators relevant for our research purpose, we first performed a literature study. Initially, we started a search in PubMed, but as this produced very few relevant hits, we decided to add free text and to search databases that contain more management publications (Google scholar and PiCarta, i.e., the end-user web interface to the Dutch Union Catalogue). The following combinations of key words were used: *indicators*, *performance indicators*, *indicator development*, *quality*, *efficiency*, *radiotherapy*, *cancer*, *healthcare*, *hospital*. We also checked cross-references from the most relevant publications and checked who cited these publications. Also non-scientific publications released by agencies involved in benchmarking (such as the Dutch society for radiation oncologists (NVRO) and the Ministry of Health, Welfare and Sport) were included.

The indicators identified were added to the gross list only when the following criteria were met: (a) they were relevant to the

international benchmarking of managerial aspects of performance and quality of care in radiotherapy, (b) the underlying characteristics could be influenced by decision makers,[12] (c) they were suitable for comparing organizations, and (d) they were discriminative.

Figure 1. Benchmarking process; visual representation of the research method.

For the selection process, we used triangulation, whereby indicators were selected on the basis of literature and of interviews with the main stakeholders within a single radiotherapy centre. After a stakeholder analysis,[14−16] one person of each stakeholder group − managers, radiotherapy department managers, radiation oncologists and clinical physicists − provided feedback on the relevance of the indicator for the research purpose, definition clarity, data availability and discriminative value of the indicators on the gross list. Thereafter, the researcher decided to refine the definitions of some indicators, to remove irrelevant indicators, and to add new and relevant indicators. As the goal of our indicator set was to use it in   an international benchmark on operations management to identify learning opportunities, a pragmatic approach seemed feasible.

The most relevant indicators were found in a paper on performance measurement in radiotherapy,[17] publications of the NVRO and in project descriptions on benchmarking within the Organization of European Cancer Institutes (OECI) that are not publicly accessible. The indicator-selection process resulted in a shortlist of 33 indicators (see Table 1) that were to be used in a pilot study.

Evaluation of indicators after the pilot

After collecting the data on the 33 indicators, we rated the face validity of the indicators using the responses of the contact persons at the radiotherapy centres involved on the basis of three criteria, based on de Korne et al.[8] and Cowper and Samuels[18]:

1. Definition clarity
2. Data availability (administrative burden?) and data reliability (comparable and reliable?)
3. Discriminative value of the indicator (useful to compare this indicator?)

Selection of radiotherapy departments for pilot study

The structure, processes and outcomes of organizations involved in benchmarking should be sufficiently similar,[19] we used the following inclusion criteria: (a) the radiotherapy centres
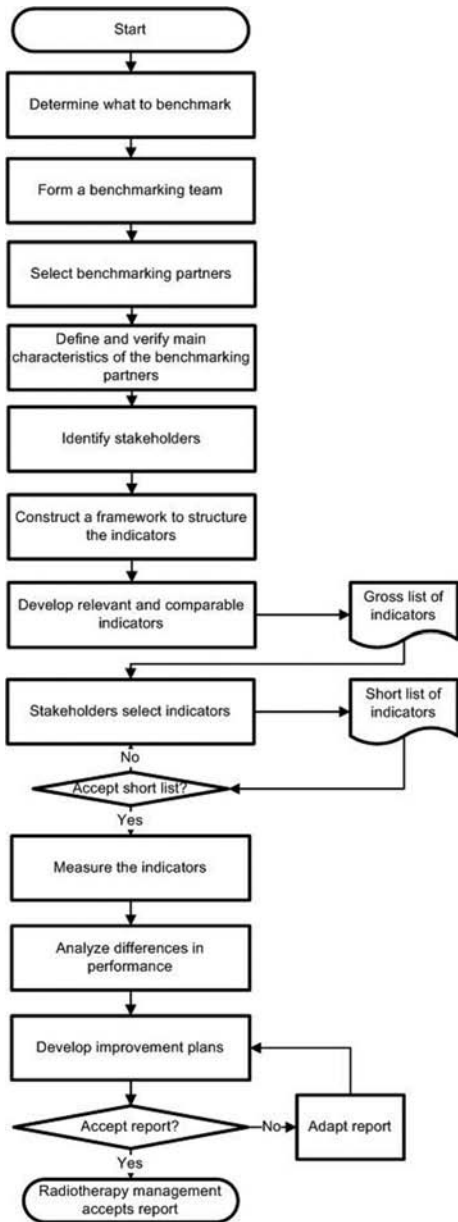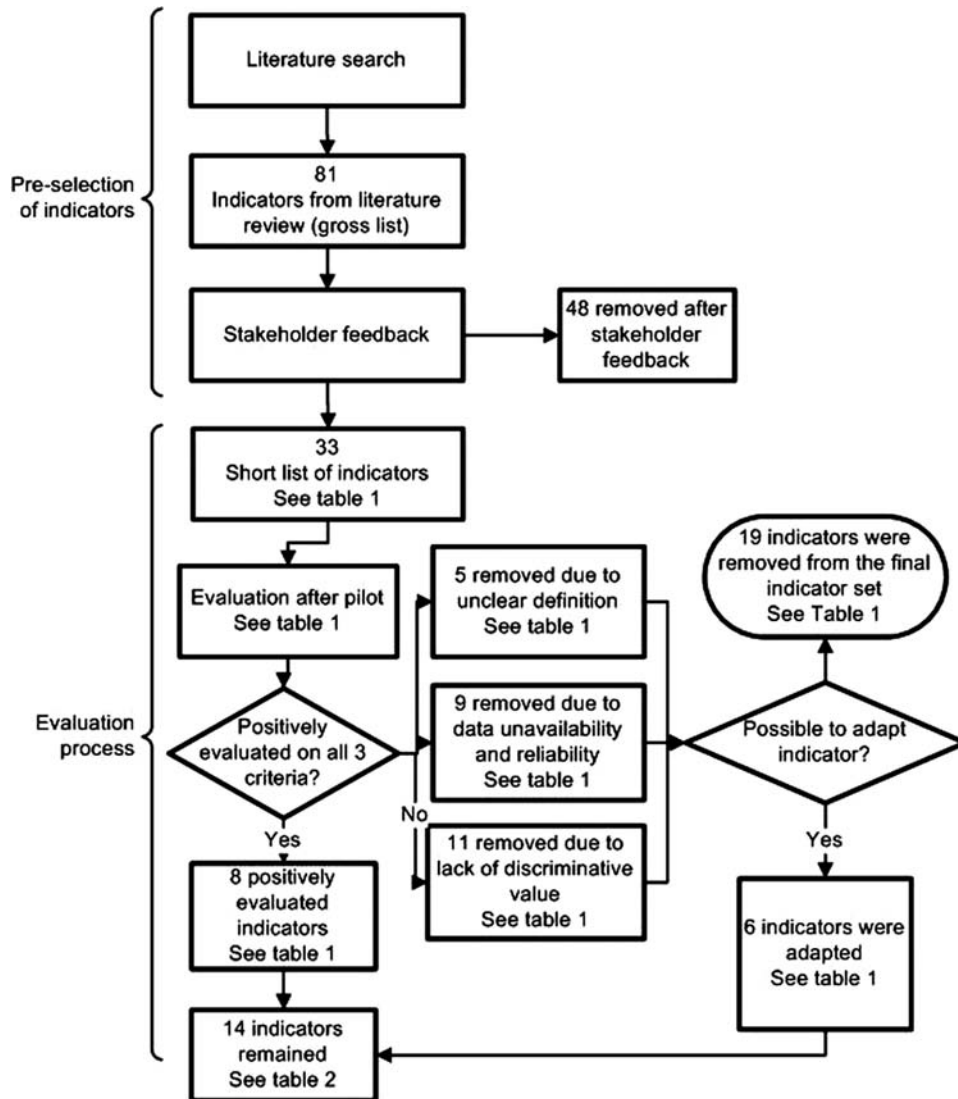
28

Figure 2. Results of the indicator selection and evaluation process.

should be situated within Europe, (b) each centre had to be part of a cancer centre that also delivered treatments other than radiotherapy, (c) should have a minimum of three linear accelerators and (d) had to be involved in research and training. This last aspect seems important as the time spent on research and training cannot be spent on patient treatment; organizations without research and training activities may probably see more patients per radiation oncologist. Data envelopment analysis on 213 hospitals has proven that teaching may attribute to up to 20% of the total inefficiency score of a hospital.[20]

Participants were approached through management contacts within the Organization of European Cancer Institutes (OECI). Four radiotherapy centres (in the Netherlands, Belgium, Germany and Sweden) fulfilled the criteria and agreed to participate in the benchmarking exercise. The centres are anonymously presented in the text as RT1, RT2, RT3 and RT4.

Table 1. *Shortlist of indicators and the results of the evaluation*

	Indicator	Definition clear?	Data availability and reliability	Discriminative value	Adapted?
	Indicators that were evaluated positively				
1	Information technology use in multidisciplinary meetings	✓	✓	✓	No
2	Patient satisfaction	✓	✓	✓	No
3	Risk analysis	✓	✓	✓	No
	Percentage of treatment planning with a curative intent using a specific imaging technique	✓	✓	✓	No
4	Simulator				
5	CT				
6	MRI				
7	PET				
8	Utilization of the linear accelerators	✓	✓	✓	No
	Indicators that could be adjusted after negative evaluation				
9	Workload (per staff type)	—	TBD	TBD	Yes
10	Access times	✓	—	✓	Yes
11	Total number of publications and average impact points per publication	✓	✓	✓	Yes
12	Percentage of patients included in a clinical trial	✓	✓	✓	Yes
13	Percentage of patients treated with new technology, e.g. IMRT	—	✓	✓	Yes
14	Downtime for planned maintenance per linear accelerator	—	✓	✓	Yes
	Indicators that were removed after negative evaluation				
15	Sick leave	—	TBD	TBD	No
16	RT department overheads	—	TBD	TBD	No
17	Staff turnover rate	✓	—	TBD	No
18	Staff overtime	✓	—	TBD	No
19	No shows	✓	—	TBD	No
20	Equipment quality-control programs	✓	✓	—	No
21	Cone beams available	✓	✓	—	No
22	Segments per treatment	✓	—	TBD	No
23	Cross-hair position	✓	✓	—	No
24	Field symmetry of photon beams	✓	✓	—	No
25	Electron beam dosimetry	✓	✓	—	No
26	Gantry-angle dependence	✓	✓	—	No
27	Instrumentation for dosimetry and QC	✓	✓	—	No
28	Simulator efficiency	✓	✓	—	No
29	CT efficiency	✓	✓	—	No
30	MRI efficiency	✓	—	TBD	No
31	PET efficiency	✓	—	TBD	No
32	Treatments per radiation oncologist	✓	—	TBD	No
33	Idle time of linear accelerators	✓	—	TBD	No
	Total 33	5	9	9	6

Note: − = did not meet this criterion; ✓ = fulfilled criterion; TBD = To be determined, this information was not checked. The criteria were judged in the order provided in the table. When one criteria was not fulfilled it was impossible to check the other criteria.

Data collection for indicator evaluation

After the radiotherapy departments had agreed to participate, a site visit was made to each. Before these visits took place, the departments received an information letter and the complete indicator set. During the visit, one researcher collected the information needed to calculate the indicator outcomes. Most parameters were based on data from annual reports or calculated using information from the hospital information systems. During these visits, qualitative data needed for the indicators was collected by interviews with one of the stakeholders that had been earlier identified. The semi-structured interviews were also used to obtain more background information on the involved departments. Two indicators that were perceived as relevant — access time and percentage of

30

patients treated with new technologies − were measured with a convenience sample on site.

The contact persons at the radiotherapy centres verified the data and gave written permission for its use in this article.

RESULTS

We evaluated the indicator data against the set of criteria and the results of the pilot study. The latter shows how the indicators can be used to identify opportunities for improvement.

Indicator evaluation

In the benchmark pilot, the 33 indicators were evaluated, and Figure 2 summarizes the results. We identified 5 indicators whose definition was inadequate. Nine other indicators did not meet our reliability and data availability criteria, and 11 more had no discriminative value (Table 1). Thus, in total, 25 indicators were negatively evaluated. Based on suggestions of the stakeholders and the researchers, we were able to adapt 6 indicators in such a way that all criteria were met, the other 19 were not fit for use. Together with the 8 positively evaluated indicators, we have an indicator set of 14 indicators; their definitions are presented in Table 2.

Of the rejected 19 indicators, sick leave, staff turnover rate and overtime (Indicators 15−18 in Table 1) were removed because the length of the paid maternity leaves or because the tasks performed by radiation oncologists differed per country. In some countries, radiation oncologists also act as medical oncologist. This made the total number of staff members incomparable. Indicator 19, no shows, was excluded because the data were unreliable. We also excluded Indicators 20−27 (see Table 1) as they lacked discriminative value, or interpretation is related more to the safety of the treatments as such than to the management of a radiotherapy centre. The indicator on simulator utilisation (Indicator 28) was supposed to provide information on the efficiency of a radiotherapy department but proved to have no discriminative value. It seemed outdated as more advanced imaging techniques are cur-

rently being introduced, such as CT, PET and MRI. As a consequence, all departments have overcapacity on the simulator. Indicators 29−31 on the utilisation of CT, MRI and PET were excluded as some of the radiotherapy departments shared their equipment with the radiology department who used it for diagnostic purposes and local registries did not provide adequate insight in the exact division. The number of treatments per radiation oncologist (Indicator 32) was excluded as the activities of the radiation oncologists differed per country. The idle time of linear accelerators had to be excluded because the available production capacity excluding unexpected maintenance was not registered everywhere and uniform local definitions were lacking.

We also identified six indicators with an insufficient score on at least one of the evaluation criteria (Indicators 9−14 in Table 1) that could be redefined:

- We included workload per staff type (Table 1, Indicator 9). Comparison of the data was initially impossible because the tasks of the staff members differed per country. Therefore, in this exercise this indicator was adjusted to the number of patients treated per staff member of the radiotherapy department. Nevertheless, using the original indicator definition was thought to be preferable.
- Access time (Table 1, Indicator 10) was defined as the time between referral from the medical or surgical oncologist to the radiotherapy centre and the start of the first treatment. However, no department consistently measured access times according to this definition. Four points in time − Table 2 − were checked manually in patient records in a random sample of 15 breast-cancer patients and 15 prostate-cancer patients who had been treated in 2006. The interpretation of access times is complicated as these can be affected by different factors not related to the radiotherapy process, such as the start and end dates of chemotherapy and hormonal therapy. As all stakeholders saw the importance of this indicator, it remained on the list.

Table 2. *Final list of indicators for benchmarking radiotherapy centres*

Indicator	Objective	Numerator	Denumerator
Workload (also called patient-to-staff ratio)*	To measure staff productivity over a 1-year period	Number of patients treated in year X	Total number of staff members in year X (in full-time equivalents) Preferable: staff per function
Access times*	To measure the treatment delay between the day patients are referred to the radiotherapy department and the day they receive the first radiation fraction	Sum of the access times (in days) per treatment in year X. Total of four points: T1 = day of referral to the radiotherapy department T2 = day of the initial prescription T3 = day of the final prescription/treatment plan finished T4 = day the patient receives the first radiation fraction	Number of patients treated in year X
Patient satisfaction	To measure the quality aspect of patient-centeredness	Determine position in PDCA cycle: • Plan: develop method to collect data • Do: Collect, analyse and implement recommendations • Check: Did the actions result in an improvement? • Act: If necessary, change the method Start the cycle over again	
Risk analysis	To focus on the quality aspects of safety and patient-centredness	Determine position in PDCA cycle: • Plan: develop method to collect data • Do: Collect, analyse and implement recommendations • Check: Did the actions result in an improvement? • Act: If necessary, change the method Start the cycle over again	
Information technology use in multidisciplinary meetings (MDM)*	To measure differences in the organization of MDM. An MDM is already standard	Use of information technology in MDM: Is the information for the MDM digitally available in one computer system? Yes/no Are the conclusions of the meeting immediately digitally registered in this system? Yes/no	
Total number of publications and average impact points per publication*	To measure the quality of output of researchers in the RT research department on volume and quality	Total number of peer-reviewed published by RT research department over 3 years	Number of years included in the selected period (in this article, 1 year was used)

Table Continues

Table 2. *Final list of indicators for benchmarking radiotherapy centres*

Indicator	Objective	Numerator	Denumerator
Patients included in a clinical trial*	To describe the percentage of patients included in a trial and thus to indicate efficiency losses caused by research activities	Number of patients who participated in a clinical trial over the past 3 years	Number of patients treated over the past 3 years
Treatment planning with a specific imaging technique: Simulator CT MRI Pet	To measure the percentage of treatment planning with a curative intent using a specific imaging technique. This technique affects the quality of the treatment plan	Number of treatment plans based with a curative intent based on a specific imaging technique in year X	Total number of treatment plans with a curative intent made in year X
Percentage of patients treated with new technologies	To describe the centres' use of intensity-modulated radiotherapy (IMRT), image-guided radiotherapy (IGRT) and adapted radiotherapy (ART)	Number of patients treated with IGRT, IMRT or ART** in year X	Number of patients treated in year X
Use of the linear accelerators	To determine the overall equipment efficiency of the linear accelerators	Number of patients treated in year X	Number of linear accelerators in the centre in year X
Planned linear accelerator downtime*	To quantify unplanned maintenance of the linear accelerators during working hours	Number of hours of unplanned*** maintenance per linear accelerator during opening hours in year X	Total number of opening hours per linear accelerator in year X
		Total impacts points of all publications	Total number of peer-reviewed published by RT research department over 3 years

*This indicator has been adapted. A new definition for the indicator was suggested after evaluation but has not been tested in a case study. **Adjusted to what was state of the art at the time of measurement. ***Planned maintenance consists of planned maintenance, time needed for quality control and time reserved for research activities. All other maintenance activities are considered to be unplanned maintenance.

33

- Research output was measured on the basis of the number of published papers. Since this is interesting only in conjunction with their quality, we added the average impact factor per publication (Indicator 11).
- The percentage of patients included in a clinical trial (Indicator 12) is an indicator with a high variation per year. We adapted this indicator to measure for 3 years instead of 1.
- Percentage of patients treated with new technologies, e.g. intensity-modulated radiotherapy (IMRT; Indicator 13). Since radiotherapy is a rapidly advancing specialty that involves complex technologies, we examined the use of new technologies. Originally, the indicator asked about the use of a specific technology, but the verb ''use'' caused confusion. The adjusted indicator therefore examines the percentage of patients treated with IMRT, image-guide radiotherapy (IGRT) and adapted radiotherapy (ART) per tumour type.
- Downtime for unplanned maintenance per linear accelerator (Indicator 14). Linear accelerator downtime was redefined and specified to downtime for planned maintenance because some organizations did not register unplanned maintenance.

Pilot study results: Usability of the final benchmark indicator set

All indicators in Table 1 were tested during our pilot, only Indicators 1−14 met the criteria for positive evaluation and the definitions are provided in Table 2. The outcomes as found in the pilot are presented in Table 3. Per indicator we discuss how the results provide opportunities for improvement:

- In the patient-to-staff ratio (Indicator 1 in Table 3), we included all staff that were paid from the radiotherapy budget and that were involved in the treatment of RT patients. Included staff members were radiation oncologists, radiation oncologists under training, radiation technicians, physicists, radiotherapy management, secretaries and researchers and other physicians working on radiotherapy treatments. The patient–to–staff ratio for RT3 is almost a third of RT1 and

RT4 and may provide leads for improving the efficiency of staff input.
- In access times (Indicator 2), we found large differences between the day of the actual prescription of radiotherapy (T2) and referral to the radiotherapy department (T1). These are due mainly to the differences in the preparation and treatment processes before the first radiotherapy fraction. Access times for breast cancer were short in RT1 and RT4. RT3 had the shortest prostate-cancer access time.
- The patient-satisfaction indicator (Indicator 3) measured whether the radiotherapy centre systematically collected and used patient satisfaction information to improve their results. This was measured using the Plan-Do-Check-Act cycle (PDCA):

  - Plan: construct a method to collect patient satisfaction information
  - Do: collect and analyze the data, determine improvement actions, and implement them
  - Check: determine whether the changes improved patient satisfaction
  - Act: if necessary, change the method so that it leads to improved patient satisfaction. Start the cycle over again.

None of the radiotherapy centres completed the cycle. Only RT1 and RT3 systematically provided all patients with a satisfaction questionnaire. RT3 did not analyze the results in a structured way. RT1 analyzed the results and formulated improvements which were reported to all radiotherapy employees every 2 months but did not complete the cycle.

- Indicator 4, the risk-analysis method, was also examined with the PDCA cycle. None of the centres completed the cycle. RT3 had no registration system for misses or near-misses, while RT4 registered only misses. RT2 registered misses and near-misses, which were published in monthly reports. However, we found no evidence that these reports led to improvement actions. RT1 discussed improvements on the basis of misses and near-misses in the department meetings but did not report on them structurally.

Table 3. Examples of indicators and their outcomes (collected for 2006)

Number	Indicators	RT1		RT2		RT3		RT4	
1	Work load: patients treated per staff member (in full-time equivalents)	21		Not available		32		22	
2	Access times (n = 15 for each patient group)*	B	P	B	P	B	P	B	P
	T1: day of referral to the radiotherapy department	−8	−78	−37	−27	−13	−21	−45	−10
	T2: day of the initial prescription	0	0	0	0	0	0	0	0
	T3: day of the final prescription/day on which the treatment plan finished	12	2	31	—	14	34	33	38
	T4: day the patient receives the first radiation fraction	17	10	41	—	18	43	36	42
3	Patient satisfaction, stage in PDCA cycle	Check-Act		Plan-Do		Do-Check		Plan-Do	
4	Risk analysis, stage in PDCA cycle	Do-Check		Do-Check		Plan-Do		Plan-Do	
5	Use of information technology in multi-disciplinary meetings (MDM): Information for MDM digitally available in one computer system?	No		Yes		Yes		Yes	
	MDM conclusions are immediately digitally registered in this system	No: registered in hardcopy patient file		Yes, imported		Yes, emailed to attendees		Yes	
6	Average impact points per publication and total publications	5.6 out of 297		2.3 out of 55		2.4 out of 33		Not available	
7	Percentage of patients in clinical trials	4.40%		0.70%		10.70%		3.45%	
	Percentage of treatment planning with:								
8	• Simulator	0		0		5		40	
9	• CT	91		98		1		56	
10	• MRI	8		2		74		0	
11	• PET	1		0		10		3	
12	Percentage of patients treated with	B:	P:	B:	P:	B:	P:	B:	P:
	• IMRT	31	39	0	1	0	6	0	0
	• IGRT	14	40	0	2	0	87	0	0
	• ART	0	17	0	1	0	78	0	0
13	Patients treated per linear accelerator per standard working hour	4.5		4.6		5.4		5.6	
14	Number of hours of downtime for planned maintenance per linear accelerator	156		173		47.5		84	

Note: B = breast cancer; P = prostate cancer. * = T1 is affected by other treatment prior to radiotherapy; therefore T2 was regarded as a starting point (value 0). Therefore T1 is negative. — = information was not available for publication.

35

- Use of information technology in multidisciplinary meetings (Indicator 5). Since these meetings are standard in radiotherapy, this indicator examined digital information availability, and the immediate digital registration of the conclusions. At RT2 and RT4, the electronic patient record (EPR) was displayed, and the outcomes of the meeting were immediately imported online into the EPR for everyone present to see. RT3 developed a tool for presenting and registering the outcomes, which were e-mailed to the attending physicians. At RT1, the EPR was used only to read information. This was because the outcomes were written directly in the hardcopy patient record, with the radiation oncologist later importing the conclusions into the EPR.
- RT1 published most papers and presented the highest impact factor (Indicator 6). However, due to a lack of data concerning the total number of staff per function group, it remained unclear how this related to the staff numbers actually involved in research.
- Indicator 7 shows large differences in the percentage of patients included in clinical trials. Possible explanations are different recruitment procedures and the availability of specific technologies needed to stimulate participation.
- Percentage of treatment planning with a curative intent using a specific imaging technique, such as simulator, CT, MRI and PET (Indicators 8−11). Table 3 shows that RT4 is the only centre that still uses the simulator for 40% of its treatment plans. RT3 had the highest percentage of treatment planning involving PET and MRI.
- The percentage of patients treated with new technologies (Indicator 12) was examined for breast-cancer and prostate-cancer patients. RT1 treated most patients with IMRT, while RT3 was advanced in the use of IGRT and ART for prostate-cancer patients. RT2 used these technologies only for a small percentage of prostate-cancer patients as only one of its linear accelerators was equipped with a cone beam; plans were made to increase this to four in 2 years. This shows the dependency on investment policy of the functioning of these depart-

ments. RT4 did not use any of these technologies at the moment of benchmarking because new equipment was about to be installed.
- RT1 treated fewer patients per linear accelerator per standard working hour than RT 3 and RT4 (Indicator 13).
- Table 3 shows that RT1 had the highest planned linear accelerator downtime during working hours (Indicator 14), while RT3 had the lowest. Together with the previous indicator, this suggests that RT1 could increase its utilisation by reducing downtime by performing less planned maintenance during working hours.

DISCUSSION AND CONCLUSIONS

This study reported on the development of a set of 14 reliable, available and discriminative indicators which can be used as quantitative indicators in a comprehensive international benchmark. This study provided a pragmatic and feasible indicator development process for international benchmarks on operations management. The results of the pilot showed that the data produced for each relevant indicator can be used to identify attainable performance levels and that using them for benchmarking provide leads for improving the quality operations. The following sections subsequently describe the research implications and the practical implications of this study.

Research implications
Although we thoroughly searched the literature to select indicators for the gross list, some suitable indicators may have been missed due to the non-systematic search strategy. We also might have missed relevant indicators that are based on medical guidelines regarding radiotherapy. We did not check medical guidelines since they focus mainly on the medical aspects of the treatment.

We also used interviews with various stakeholders related to RT department management to reduce the possibility of missing relevant indicators. The stakeholders screened all indicators on the following criteria: relevance for this

benchmark, definition clarity, data availability and discriminative value. This resulted in the rejection of 48 indicators. Involving the stakeholders also generated support and resources for data collection.

Despite our indicator-selection process, defining good indicators remained difficult, especially in an international perspective. This could have been prevented by asking multiple stakeholders from different countries to grade the indicators. However, as a first step in international benchmarking on operations management, our pragmatic approach seemed feasible.

After the selection, five indicators still lacked a definition that covered every country's specific characteristics (see Table 1). Radiotherapy is part of a treatment chain, and when pre-radiation chemotherapy is given, the radiotherapy access time should reasonably start after that is finished. The radiotherapy centres found it difficult to distinguish the pre-radiation delay caused by chemotherapy. Distinguishing the pre-radiation delay, caused by chemotherapy, from other delays caused by the internal organization of the radiotherapy department is essential for benchmarking.

We found that the discriminative value of 11 indicators was insufficient. Radiotherapy is an evolving health care discipline that introduces new technologies in rapid succession. The indicators concerning the use of new technologies and the percentage of patients in clinical trials may be particularly affected by this evolution. We therefore recommend adjusting the indicator set to the latest developments.

Despite the thoroughness of the process whereby we developed this indicator set, 9 of the 33 selected indicators did not fulfil the criteria on 'availability and reliability of the data'. Due to the time constraints and the desire to keep administrative efforts low, the radiotherapy centres provided us primarily with information that was already being collected for administrative purposes. During the site visits, it became clear that specific radiotherapy information was usually collected on a department level. For some data, government regulations required a specific registration method that was incompatible with the purpose to obtain comparable data. For example, the length of the paid maternity leaves in the sick-leave statistics and differences between staff duties. As registration requirements differ per country, international comparisons are often more complex than national ones; a recent international benchmarking exercise in eye hospitals confirmed this.[8] Although differences in national health systems and social legislation inevitably lead to differences in the nature and availability of data, there is no reason to doubt the applicability of the approach used in this study in non-European countries such as the USA. As these differences often lead to different definitions and outcomes, consideration should be given to indicators that assess process characteristics and outcomes.[21]

All indicators were measured over a 1-year period (2006); however, for indicators with a considerable likelihood of strong variation per year, measuring over a prolonged period should be considered. Examples include the impact factor or the percentage of patients included in a clinical trial.

Practical implications

The indicator set included indicators on efficiency, patient-centeredness and timeliness. For an appropriate and thorough identification of improvement opportunities the combination of quantitative (indicators) and qualitative information (site visits) is essential. The indicators standardize the comparison between the centres and the site visits enable a better understanding of the (underlying) processes.

We used the inclusion criteria to select radiotherapy centres that were rather comparable. For a proper comparison, case mix and complexity of treatments should be taken into account; the scope of our project did not allow us to expand on that.

The pilot results suggested RT1 to reduce planned downtime during regular working hours. RT2 was suggested to examine its inclusion rate for clinical trials and productivity of

research activities. RT4 had been working on a system to register misses and near-misses and used the data to determine the extent to which additional investments in manpower and equipment were needed to improve the safety and quality of treatments.

Out of the original long list of 81 indicators, 14 proved suitable for use in an international benchmark at radiotherapy centres. As the results are affected by the technologies available, obtaining information on access to technologies, investment policies, budgets and depreciation methods is essential. Future research should provide insight into variation of indicator scores over the years and to monitor improvement results.

ACKNOWLEDGEMENTS

References

1. Sanazaro PJ. Quality assessment and quality assurance in medical care. Annu Rev Public Health 1980; 1:37−68.

2. Laffel G, Blumenthal D. The case for using industrial quality management science in health care organizations. JAMA 1989; 262:2869−2873.

3. Committee on Quality of Health Care in America, Crossing the Quality Chasm: A New Health System for the 21st Century, Washington D.C., Institute of Medicine 2001.

4. Camp RC. A Bible for Benchmarking, by Xerox. Fin Exc 1993; 9:23.

5. Mosel D, Gift B. Collaborative benchmarking in health care. Jt Comm J Qual Improv 1994; 20:239−249.

6. Sheldon T. Promoting health care quality: what role performance indicators? Qual Health Care 1998; 7 Suppl: S45−S50.

7. Blank JL. Innovations and productivity: an empirical investigation in Dutch hospital industry. Adv Health Econ Health Serv Res 2008; 18:89−109.

8. de Korne DF, Sol KJ, van Wijngaarden JD, van Vliet EJ, Custers T, Cubbon M, Spileers W, Ygge J, Ang CL, Klazinga NS. Evaluation of an international benchmarking initiative in nine eye hospitals. Health Care Manage Rev 2010; 35:23−35.

9. Schoen C, Davis K, How SK, Schoenbaum SC. U.S. health system performance: a national scorecard. Health Aff (Millwood) 2006; 25:w457−w475.

10. van Lent WA, de Beer RD, van Harten WH. International benchmarking of specialty hospitals. A series of case studies on comprehensive cancer centres. BMC Health Serv Res 2010; 10:253.

11. Spendolini J. The Benchmarking Book. New York, Amacom 1992.

12. Van Hoorn A, Van Houdenhoven M, Wullink G. Een nieuw stappenplan voor benchmarking, Management Executive 2006: 1−15.

13. Campbell SM, Braspenning J, Hutchinson A, Marshall M. Research methods used in developing and applying quality indicators in primary care. Qual Saf Health Care 2002; 11:358−364.

14. Mitchell RK, Agle BR, Wood DJ. Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts. Acad Manage Rev 1997; 22:853−886.

15. Brugha R, Varvasovszky Z. Stakeholder analysis: a review. Health Policy Plan 2000; 15:239−246.

16. Carignani V. Management of change in health care organisations and human resource role. Eur J Radiol 2000; 33:8−13.

17. Cionini L, Gardani G, Gabriele P, Magri S, Morosini PL, Rosi A, Viti V; Italian Working Group General Indicators. Quality indicators in radiotherapy. Radiother Oncol 2007; 82:191−200.

18. Cowper J, Samuels M. Performance benchmarking in the public sector: the United Kingdom experience. Benchmarking, Evaluation and Strategic Management in the Public Sector, Paris, OECD 1997.

19. Gundmundsson H, Wyatt A, Gordon L. Benchmarking and Sustainable Transport Policy: Learning from the BEST Network. Transport Reviews 2005; 25:669−690.

20. Grosskopf S, Margatitis D, Valdmanis V. The effects of teaching on hospital productivity. Socio-economic Planning Sciences 2001; 35(3):189−204.

21. Camp RC, Tweet AG. Benchmarking applied to health care. Jt Comm J Qual Improv 1994; 20:229−238.