Chapter 8

Benchmarking

Henrik Westh
KMA 445, Hvidovre Hospital, Kettergård Alle 30
DK-2650 Hvidovre, Denmark

Benchmark is a technical term used by surveyors to indicate a point of reference from which measurements may be made. For our purposes, benchmarking is a method for improving operations in our organization typically achieved through systematic comparison with other organization(s) recognized as best in the field. The object of benchmarking is not to compare key figures but to compare how tasks are performed. Learning can only be achieved by looking at those who are better than you. Properly applied benchmarking is an extremely effective way of improving your organization.

Before you try to benchmark your organization, there must be a perceived need for improvement and a willingness to improve. This is very important because otherwise results will only end up as a report. The benchmarking study and the results must be supported by those who later will have to implement changes and improvements.

A number of different definitions of benchmarking are found in Table 1. These definitions have in common keywords such as performance, comparison, measuring, outstanding, best, improvement, process, and practice.

1. DIFFERENT TYPES OF BENCHMARKING

Before you enter a benchmarking process, you will have to decide what to compare and whom to compare with, bearing in mind that the goal is to

Antibiotic Policies: Theory and Practice. Edited by Gould and van der Meer
Kluwer Academic / Plenum Publishers, New York, 2005

Table 1. Examples of definitions of benchmarking

• A process of finding, adapting, and implementing outstanding practices
• A process of identifying and importing best practice to improve performance
• Comparing the performance of your organization with that of others with outstanding performance to find fresh approaches and new ideas
• The process of comparing the performance of an individual organization against a benchmark, or ideal, level of performance. Benchmarks can be set on the basis of performance over time or across a sample of similar organizations, or against some externally set standard
• A continuous, systematic process for evaluating the products, services, and work processes of organizations that are recognized as representing the best practices for the purpose of organizational improvement

achieve improvements. When you have decided what to compare, three types of benchmarking are typically described.

- Process benchmarking
- Performance benchmarking
- Strategic benchmarking

When you have decided whom to compare with, three types of benchmarking are typically described.

- Internal benchmarking
- External benchmarking
- Generic benchmarking

Process benchmarking is learning from the best to improve one's own processes (comparison of methods and practices). Performance benchmarking is determining how good you are compared to others by comparing performance measures (either financial or operational). Strategic benchmarking is collecting information from other companies to improve one's own strategic planning and positioning.

Internal benchmarking is used to compare different units in the same organization. External benchmarking is the comparison with companies outside your organization that have similar or identical operations and processes. In external benchmarking, you will usually look for noncompeting organizations within your own field. In the private industry, one can run into problems of sensitive or confidential information, but this is rarely a problem in the public sector. Generic benchmarking involves comparison with unrelated industries that are worth learning from.

The benchmarking process is one of many tools for improving your institution or department. But it should be recognized that to be properly performed, it requires many resources. A good benchmarking process includes five major areas of activity.

1. Study and understand one's own process.
2. Find benchmarking partners.
3. Study the partner's process.
4. Analyse the differences between one's own process and that of the partner's.
5. Implement improvements based on what is learned.

It is important to spend enough time on the different elements of benchmarking. Typically 50% should be used in the planning phase, 30% on the study, and 20% in analysing the results. The timeframe for the implementation phase can only be estimated after the results are in.

2. PLANNING FOR BENCHMARKING

- Select a process to benchmark—think company strategy
- Identify needs
- Form a benchmarking team

Benchmarking is a method of improving performance. The goal is to identify and implement improvements by comparing your own operations with those of others who perform better. Almost anything that can be observed or measured can be benchmarked. This is often called gap. The planning and organizing of benchmark activities is extremely important. There must be a clear idea of goals and adequate resources allowing the benchmarking team time to fulfil its assignment. Tools must be found or developed for information and data gathering. Try to keep the goals specific (set limits), as too broad programmes will lead to an enormous amount of information and a lack of specific recommendations. This can often be achieved by identifying critical success factors (CSFs). A CSF could be patient satisfaction or adherence to hospital clinical guidelines or problems that have surfaced in audits. When you have identified CSFs, you have to evaluate your performance and also identify the processes that impact most on your CSFs. Look carefully at your organization and select indicators to benchmark against. Examples of potential indicators are listed in Table 2. It is also important to have a plan for how the results of the benchmarking will be used. The people in the benchmarking team structure must reflect your organization. Typically you will need a process owner, a process worker, a manager, and a user (customer).

3. FIND BENCHMARKING PARTNERS

- Look for long-lasting relationship
- Look for world champions or Best Practice guidelines
- Look out for differences in the scope of operation and in market conditions

Table 2.

Input indicators	Output indicators	Outcome indicators
Hospital type	DDD total antibiotics	Discharges
Antibiotic policies	DDD/1,000 beds	Operations
Drug formulary	Operation types	Surgical site infections, risk-adjusted infection rates
Budgets	Costs	Bed-days
Infection control policies	Surgical prophylaxis	Timing of surgical prophylaxis, length of operations
Staffing	Choice of prophylactic antibiotic regimes for	Survival rates for septicaemia, meningitis, community-
Best practice guidelines	orthopaedic surgery, etc.	acquired or hospital-acquired pneumonia
Microbiology service	Use of key antibiotics	Appropriateness of antibiotics usage
	3. gen cephalosporins	
	Betalactam-combinations, aminoglycosides,	
	Quinolones, monobactams, Vancomycin, others	
	Dosing	
	Time to first dose in meningitis, pneumonia, etc.	
	Administration route iv/oral	
	Resistance levels for key microorganisms	
	S. aureus	
	Pneumococci	

E. coli	
Klebsiella	
Coagulase negative staphylococci	
Acinetobacter	
Stenotrophomonas maltophilia	
Rate of Candida sepsis	
Nosocomial infection rates	
Device-associated infection rates	By type of device and type of ICU
Gloves bought	
Litres of hand disinfectant bought	
Hand disinfection bought	By hospital unit
Microbiology lab in hospital	
Infection control audits	
Blood cultures per 1,000 bed-days	
Positive rate of blood cultures	
CNS positive rate in blood cultures	Vancomycin consumption

It is very important to have an open exchange of information so it is important that you find a few partners who are better than you than go for a comparison of a large number of companies leading to superficial conclusions. You should also look outside your network. Partners can be found by searching scientific journals, books, newspapers, the Internet, etc.

4. STUDY PARTNERS

- Information gathering
- Questionnaires
- Definitions and explanations
- Document
- Get approval from managers

The data gathering process starts with obtaining a deep knowledge of your own organization. This leads to the development of a data gathering model. This model should be validated before using it on your study partner. What you are looking for in your partners is their performance level—how good they are and their practice—how do they do it.

5. ANALYSE

- Find differences, that is, performance gaps
- Quality control data
- Prepare report

Before you analyse, you must be satisfied that all information is correct. The goal is to find performance gaps that can lead to improvement of CSFs. It is therefore important to check and correct for comparability and quality control your data collection. You will often need to normalize data. For example, in antibiotic consumption, knowing the amount of antibiotics used in Defined Daily Doses (DDDs) is uninteresting if it is not corrected/normalized, typically by per 100 bed-days. The information accumulation leads to knowledge gathering, allowing you to understand why there are performance gaps. This allows you to prepare a report that can be used for the real reason for benchmarking which is to change practices. Several models for identifying causes for the gaps are available, for example, comparison of flow charts, relations diagrams, and root cause analysis (breaking a problem into smaller problems). After identifying gaps, ranking can be relevant, as the closing of gaps will have different costs, improvement potentials, and applicability to your organization. Remember that your conclusions should be adapted to your own organization.

6. IMPLEMENT IMPROVEMENTS

This is the final great challenge. The momentum of the project must be maintained. In effective implementation, you need to have a clear commitment from those involved and full management support. Factors that could block best practices must be identified (institutional culture, traditional cooperation, technology restraints, user groups, etc.)—all these elements can change best practice back to usual practice. The implementation process needs targets and always needs monitoring. The organization needs continuous orientation on progress and achievements. Remember that benchmarking results cannot be used as a carbon copy manual where best practices are copied directly into your own organization.

Also remember that benchmarking is not a one-time event but a continuous process for improving your own processes. One of the many pitfalls in benchmarking is that it is not successfully integrated into the way the organization solves problems.

7. ETHICS IN BENCHMARKING

- Request only information that you would give out
- Respect confidentiality
- Full disclosure

As benchmarking is a continuous process, it is very important to have good ethics so you can sustain long-term relations with your benchmarking partners. You must be willing to provide the same information that you seek, to your benchmarking partners. Benchmarking must conform to legislation and moral codes. Otherwise, aspects such as industrial espionage, price fixing, customer allocation schemes, etc. can cloud issues.

8. BENCHMARKING AND PUBMED

Looking in PubMed on 1 October 2003, the keyword benchmarking gave 3,535 hits. The use of benchmarking and antibiotic as keywords gave 35 hits and benchmarking and microbiology gave 25 hits. Let us briefly look at two of the best of these papers and discuss their study design.

9. BENCHMARKING FOR REDUCING VANCOMYCIN USE AND VANCOMYCIN-RESISTANT ENTEROCOCCI IN US ICUs

This study was performed to improve compliance with process-of-care guidelines. The goal of the process was to reduce vancomycin consumption

Table 3. Prescribing practice changes implemented in response to benchmark data intervention, and mean rate of vancomycin use[a] before and after intervention, 50 Project ICARE ICUs, January 1996 to July 1999[b]

Vancomycin use prescribing practice change	No. of ICUs (%) (n=50)	Change absent		Change present		p value[c]
		Before	After	Before	After	
Hospitalwide[d]	22 (44)					
Drug use evaluation	19 (38)	74.2	80.5	105.3	94.1	0.62
Redistributed HICPAC guidelines on VRE	9 (18)	79.4	84.6	116.0	90.6	0.34
Prior approval of vancomycin required	3 (6)	87.2	84.7	67.2	99.4	0.25
Unit specific[d]	11 (22)					
ICU-specific education on appropriate vancomycin use	9 (18)	75.9	83.3	132.1	96.3	0.01
Removed vancomycin from surgical prophylaxis	3 (6)	82.0	85.9	149.1	82.2	0.01

[a]DDDs per 1,000 patient-days.
[b]ICARE, Intensive Care Antimicrobial Resistance Epidemiology; ICU, intensive care units; HICPAC, Healthcare Infection Control Practices Advisory Committee; VRE, vancomycin-resistant enterococci.
[c]Paired *t*-test.
[d]Components of each major category are not mutually exclusive, so one ICU may be represented in several components of each category.
Source: Data from Fridkin *et al*. (2002).

(Fridkin *et al*., 2002) (Table 3). The study is an example of benchmarking on the basis of performance across a sample of similar organizations. External benchmarking is the comparison with companies outside your organization that have similar or identical operations and processes.

In this Project ICARE study, preintervention data were collected during 1996–7. In part one of the project, the data collected were used to create a national benchmark defined as the aggregate summary data from 113 ICUs. The report included pooled means, medians, and key percentile distributions of the prevalence of VRE and MRSA and vancomycin use as DDDs/1,000 patient-days. This feedback report was presented in October 1997 to the participating hospitals (primarily to the infection-control committees).

In part two of this project, 50 ICUs in 20 hospitals participated in the postintervention period from April 1998 through July 1999 with at least 6 months of data collection. How the 1997 feedback report had been used by the hospitals was surveyed in September 1999. The feedback report survey looked for prescribing practice changes implemented in response to the

feedback report. These prescribing practice changes were assessed and compared with vancomycin use before and after intervention.

The ICU-specific use of vancomycin in the 50 ICUs at the 20 study hospitals after the intervention was 89.1 DDD/1,000 patient-days, a 2.8% increase over the preintervention rate of use. This increase in consumption could in part have been caused by an increasing median MRSA prevalence of 33.5% (preintervention) and 39% during the postintervention period.

The only prescribing practice changes that led to a significant reduction in vancomycin use were ICU-specific education on appropriate vancomycin use and not surprisingly the removal of vancomycin from cardiac surgical prophylaxis. A risk-adjusted analysis was performed taking into consideration the ICU type and changes in MRSA prevalence (normalization). ICUs in which unit-specific practices were identified for improvement reported a 35–37% decrease in median vancomycin use (from median 132 to 96 DDD/1,000 patient-days for unit-specific education [9 units] and 149 to 82 DDD/1,000 patient-days for removal of prophylaxis [3 units]).

During the preintervention period, these ICUs reported a median VRE prevalence of 11.7% increasing to 14% during the postintervention period. However, when compared by type of practice change, the difference in VRE prevalence was significantly lower in ICUs in which unit-specific practice changes occurred, compared with other ICUs. Although many of the ICUs with decreases in vancomycin use reported increases in per cent VRE, all the ICUs noting a unit-specific practice change reported decreases in both per cent VRE and vancomycin use.

This study suggests that only focused efforts (i.e., ICU specific) were effective means of reducing excessive vancomycin use. The external benchmarks used were risk adjusted (i.e., stratified by ICU type) to account for the different rates of vancomycin used by different types of ICUs. This made comparison of local data more relevant (and more believable) to the ICU staff responsible for prescribing and other patient-care activities. The ICUs that used unit-specific changes had the highest prestudy rates of vancomycin use, and this excessive use may have made the ICU staff more amenable.

9.1. Comments

This study analysed performance and measured key figures such as DDDs and resistance levels to methicillin in *Staphylococcus aureus* and to vancomycin in enterococci. This focus on performance gives little information on how to improve or close the gap between the different departments. Most people would accept that there is an over-usage of antibiotics in hospitals. The usage of antibiotics is regulated by clinical guidelines and clinical practices.

The very wide range of vancomycin usage in this study suggests that the true benchmark for vancomycin usage lies somewhere within the range of usages found. If the aim of the study was to reduce vancomycin usage, benchmarking should have been used to analyse the ICUs with low vancomycin consumption and show that their outcome results were as good as best practice. By analysing the methods and practices of these ICUs, one might learn, from the best, to improve one's own processes. This could have lead to a best practice definition of the correct usage of vancomycin in the ICU. This information could be clinical guidelines, training, educational efforts, etc. An analysis of one's own performance would have led to the findings of gaps that then could be corrected. In this study the benchmark was defined as the current median usage of vancomycin in the studied ICUs. With this choice of a benchmark, current practices were accepted as the benchmark, or, in other words, a median/average was defined as best practice. The results of the study were that outliers changed practices and regressed to the mean.

10. THE HARVARD EMERGENCY DEPARTMENT QUALITY STUDY

This study was performed to improve compliance with process-of-care guidelines and patient-reported measures of quality. The study is an example of benchmarking on the basis of performance across a sample of similar organizations with use of some externally set standards (best practices). External benchmarking is the comparison with companies outside your organization that have similar or identical operations and processes.

Five Harvard teaching hospitals collaborated to improve quality in their emergency departments. The five areas chosen to improve were in patients presenting with abdominal pain, shortness of breath, chest pain, hand laceration, head trauma, or vaginal bleeding. A working group of experts reviewed the medical literature and existing guidelines and developed complaint-specific process-of-care data forms for medical record review. The goal was to improve compliance with process-of-care guidelines, patient satisfaction, and patient-reported problems with care.

In the preintervention phase, 4,876 medical records were evaluated, 2,327 patients completed onsite questionnaires, and 1,386 patients completed a 10-day follow-up questionnaire.

In the postintervention phase, 6,005 medical records were reviewed, 2,899 patients completed onsite questionnaires, and 2,326 patients completed a 10-day follow-up questionnaire.

Physician compliance with the process-of-care guidelines was the medical record based quality measure for the study and was evaluated by

physician-reviewers unaware of the purpose of the study. Patients were asked to report problems during their emergency department visit and patient satisfaction was evaluated through the follow-up telephone interview.

One year later, the results of the baseline investigation were provided and preintervention phase process-of-care criteria were distributed to all the emergency departments as clinical guidelines. Based on the preintervention data, the hospitals found 27 different quality improvement interventions. From this list, each hospital chose 8–10 quality improvement efforts for implementation in their hospital.

In multivariate analyses, adjusting for site, age, urgency, and chief complaint, the mean compliance with guidelines for all complaints increased from 55.9% to 60.4% after interventions (see Table 4). For all sites combined, compliance with guidelines was significantly improved for abdominal pain, shortness of breath, and head trauma. There was no significant change in compliance with guidelines for chest pain, hand laceration, or vaginal bleeding. There were significant variations in intersite improvement rates in compliance with guidelines.

Changes in patient-reported problems were investigated by multivariate analyses adjusting for site, age, urgency, and chief complaint. The rate of

Table 4. Hospital specific (hospitals A–E) and total compliance with process-of-care guidelines

Complaint	Mean (95% confidence interval)[a]		*p* value
	Preintervention	Postintervention	
All complaints			
A (n = 3,291)	57.2 (55.2–59.2)	60.3 (58.5–62.1)	0.02
B (n = 2,903)	57.4 (55.4–59.4)	60.2 (58.4–61.9)	0.04
C (n = 1,881)	54.5 (52.1–56.9)	61.7 (59.5–63.9)	0.0001
D (n = 2,405)	52.7 (50.5–54.9)	58.6 (56.6–60.6)	0.0001
E (n = 456)	63.4 (58.7–68.1)	62.6 (58.3–66.9)	0.83
Total	55.9 (54.9–56.9)	60.4 (59.4–61.4)	0.0001
Abdominal pain			
A (n = 1,149)	57.2 (53.9–60.5)	60.0 (56.9–63.1)	0.23
B (n = 752)	58.4 (54.1–62.7)	60.6 (56.5–64.7)	0.45
C (n = 499)	53.8 (48.7–58.9)	62.5 (57.8–67.2)	0.02
D (n = 704)	55.4 (50.7–60.1)	57.9 (54.2–61.6)	0.42
E (n = 160)	71.3 (62.7–79.9)	65.7 (58.3–73.1)	0.34
Total	57.0 (55.0–59.0)	60.5 (58.7–62.3)	0.01
Shortness of breath			
A (n = 527)	72.0 (66.5–77.5)	70.0 (64.9–75.1)	0.61
B (n = 384)	31.6 (25.1–38.1)	52.1 (45.0–59.2)	0.0001
C (n = 332)	58.3 (49.1–67.5)	59.6 (50.4–68.8)	0.96
D (n = 417)	37.7 (30.8–44.6)	54.9 (48.6–61.2)	0.005
E (n = 100)	56.3 (35.1–77.5)	75.6 (54.4–96.8)	0.29
Total	52.1 (48.8–55.4)	60.9 (57.6–64.2)	0.0002

Table 4. *Continued*

Complaint	Mean (95% confidence interval)[a]		*p* value
	Preintervention	Postintervention	
Chest pain			
A ($n = 636$)	65.5 (62.4–68.6)	61.9 (59.0–64.8)	0.10
B ($n = 701$)	70.7 (67.6–73.9)	69.3 (66.4–72.2)	0.54
C ($n = 437$)	68.0 (64.5–71.5)	64.7 (61.2–68.2)	0.18
D ($n = 503$)	61.3 (57.8–64.8)	63.8 (60.7–66.9)	0.32
E ($n = 117$)	65.9 (59.8–72.0)	62.9 (56.6–69.2)	0.51
Total	66.7 (65.1–68.3)	65.0 (63.4–66.6)	0.13
Hand laceration			
A ($n = 176$)	58.5 (50.7–66.3)	57.6 (50.3–64.9)	0.86
B ($n = 293$)	55.0 (49.9–60.1)	56.9 (50.8–63.0)	0.65
C ($n = 196$)	55.7 (49.0–62.4)	68.1 (62.2–74.0)	0.008
D ($n = 178$)	66.2 (59.5–72.9)	67.8 (60.7–74.9)	0.76
E ($n = 31$)	68.9 (56.4–81.4)	68.2 (58.8–77.6)	0.94
Total	58.7 (55.6–61.8)	62.6 (59.5–65.7)	0.09
Head trauma			
A ($n = 589$)	40.5 (36.6–44.4)	52.7 (49.4–56.0)	0.0001
B ($n = 728$)	48.3 (44.8–51.8)	53.4 (50.0–56.7)	0.04
C ($n = 348$)	31.7 (26.8–36.6)	52.9 (48.6–57.2)	0.0001
D ($n = 441$)	32.0 (28.1–35.9)	46.2 (42.3–50.1)	0.0001
E ($n = 42$)	24.5 (10.6–38.4)	46.0 (29.1–62.9)	0.08
Total	40.0 (38.0–42.0)	51.4 (49.6–53.2)	0.0001
Vaginal bleeding			
A ($n = 206$)	64.0 (57.4–70.7)	70.2 (62.0–78.4)	0.47
B ($n = 34$)	81.4 (63.0–99.8)	72.9 (56.2–89.6)	0.52
C ($n = 52$)	66.6 (52.3 80.9)	70.7 (54.4–87.0)	0.66
D ($n = 152$)	73.3 (65.7–80.9)	73.3 (66.0–80.6)	0.8
E ($n = 0$)[b]			
Total	68.6 (64.1–73.1)	70.2 (65.3–75.1)	0.64

[a]Adjusted for age, urgency, and chief complaint. Total also adjusted for site.
[b]Hospital E had no patients with vaginal bleeding.

problems decreased overall from 24% to 20% and significant improvements were seen in four of the five sites. No improvements were seen in patient satisfaction.

10.1. Comments

This is a fine study where the use of benchmarking resulted in some improvement in emergency department quality of care. However, the focus on performance gives little information on how to improve or close the gap between the different departments.

This study used best practices as defined by the Harvard Emergency Department Quality Study team and internal benchmarking as the analyses were performed comparing preintervention and postintervention phases for each hospital. The benchmarking partners chosen for the study seem chosen for geographical reasons. A best practice Emergency Department was not part of the analysis; therefore, this is predominantly a study of equals. Looking at the Table 4, it can be seen that total compliance for "shortness of breath" was 72% in hospital A and 31.6% in hospital B. This difference is what benchmarking calls a gap. An analysis of the practices in hospital A that made it possible for them to reach this higher level of performance would have been perfect benchmarking. Although an analysis of differences between one's own and the partner's process was not described, the unblinding of the project results must have allowed for good opportunities to discuss the impacts of the different quality improvement interventions chosen by the different departments. However, the study team felt that the use of multiple interventions did not allow the team to evaluate which initiatives lead to improvement.

Lacking in this chapter is a discussion of the use of CSFs. An example of a CSF was that each unit in the study wanted to fulfil the American College of Emergency Physicians criteria for administration of thrombolytic therapy and achieved 100% compliance with the guideline (up from 65.3%).

11. CONCLUSION

In conclusion, benchmarking can become a valuable tool for improvements in healthcare. While key performance figures can be used to find gaps in performance, only an analysis of processes will allow one to understand the differences and plan for improvement.

SUGGESTED READING

Andersen, B. and Pettersen, P.-G., 1996, *The Benchmarking Handbook. Step-by Step Instructions*. Chapman & Hall, London, UK. ISBN 0 412 73520 2.

Benchmarking in the Public Sector. Some Methods and Experiences. March 2000. Ministry of Finance. ISBN 87-7856-331-3 www.fm.dk.

Burstin, H. R., Conn, A., Setnik, G., Rucker, D. W., Cleary, P. D., O'Neil, A. C. *et al*., 1999, Benchmarking and quality improvement: The Harvard Emergency Department Quality Study. *Am. J. Med*., **107**(5), 437–449.

Fridkin, S. K., Lawton, R., Edwards, J. R., Tenover, F. C, McGowan, J. E., Jr, Gaynes, R. P., 2002, Intensive Care Antimicrobial Resistance Epidemiology Project; National Nosocomial Infections Surveillance Systems Hospitals. Monitoring antimicrobial use and resistance: Comparison with a national benchmark on reducing vancomycin use and vancomycin-resistant enterococci. *Emerg. Infect. Dis*., **8**(7), 702–707.