Benchmark analyses in minimally invasive esophagectomy—impact on surgical quality improvement

Roxane D. Staiger, Christian A. Gutschow

Department of Surgery and Transplantation, University Hospital Zurich, Zurich, Switzerland

Contributions: (I) Conception and design: CA Gutschow; (II) Administrative support: None; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: None; (V) Data analysis and interpretation: None; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Prof. Dr. Christian A. Gutschow, MD. Department of Surgery, University Hospital Zurich, Rämistrasse 100, 8091 Zurich, Switzerland. Email: christian.gutschow@usz.ch.

Abstract: Over the last decades, benchmarking has become an established management tool to improve quality in commercial economics. It is a rather new concept in the healthcare industry, and a confusingly wide range of approaches referring to "benchmarking" have been employed in the field of minimally invasive esophageal cancer surgery. It is our conviction that benchmarking will be an essential element of surgical research in the future. Therefore, defining and implementing standards is not only a desirable, but a vital step. Recently, we have introduced a standardized method of establishing valid benchmarks for surgical quality improvement including ideal outcome thresholds for total minimally invasive transthoracic esophagectomy (ttMIE). The present article aims at discussing the actual literature on benchmarking in minimally invasive esophagectomy (MIE) and at fueling the debate on how to further improve the current practice of surgical outcome research.

Keywords: Benchmark; minimally invasive esophagectomy (MIE); outcome research; surgical quality

Introduction

Benchmarking is a well-established management tool, generally used to improve productivity, efficacy and product quality. In the field of esophageal cancer surgery, different approaches to adopt this concept for outcome research have been published in recent years. The key element of benchmarking is the comparison with the best possible outcome—the benchmark. Only the contrast to the best reveals the true potential for improvement. Just as a company has no interest in being average, it lays in every surgeon's nature to desire the best possible outcome for his patients. However, in the majority of studies with a focus on benchmarking minimally invasive esophageal cancer surgery, this principle is not entirely respected. The given thresholds are mostly based on large unselected data collections and simple averages drawn from their results. In contrast, the essence of a valid benchmark lays in the definition of the optimum, which is a most difficult endeavor—especially in the field of surgery.

Validity of benchmarks

Benchmarking is a cyclic process for quality improvement that consists of three important steps: First, defining the best; second, comparing to the best; and third, learning from the best (1,2). By comparison of own results to an optimal threshold, flaws in quality may be detected. These deficiencies may be corrected by adopting processes that were implemented by other centers that perform within the benchmark (1). There are numerous ways to create benchmarks in surgery, but all efforts pursue one common goal: to improve outcome (3). While simple comparison with other centers may detect quality deficits, only the

orientation towards a best possible result reflects the true potential for improvement (3).

In surgery, the patient and the surgeon are the two basic factors with a paramount impact on outcome: patients may present with various risk factors and a surgical team's experience in preparing and performing the intervention, but also in postoperative caretaking, may vary substantially (3). Optimizing a patient's nutritional and metabolic state prior to surgery has been shown to improve postoperative outcome (4-6). However, the best results for any operation or intervention are obtained when "ideal" patients are treated by experienced surgeons in international high-volume centers (3,7-9). In general, "ideal" patients are those with the fewest expected postoperative complications. Depending on the type of surgery, factors such as comorbidities (diabetes, chronic obstructive pulmonary disease, etc.), medication (steroid intake, anticoagulant therapy, etc.), lifestyle (smoking, obesity, etc.) or patient characteristics (age, sex) may have a great influence on outcome. For example, optimal results for hemihepatectomy may be assumed in living liver donors, since their mandatory prerequisite is to be young and healthy (7). Therefore, a benchmark calculation should exclusively be based on such "ideal" patients (3,7-9).

Over the last decade, centralization in healthcare has moved into the center of attention (10). It is commonly accepted that outcome after surgery does not only depend on the surgeon's performance, but rather on a finely tuned collaboration of all involved disciplines (11,12). As a general rule, postoperative complication rates and overall morbidity are lower, if the surgical intervention is frequently performed by a specialized team. Therefore, to grant their validity, only data from high-volume centers should be used to establish benchmark cut-off values (3,8).

Another restriction often imposed on published quality thresholds are national borders. Yet, good surgical quality is never restricted to one single country; therefore, best results are most realistically represented in international benchmarks including high-volume centers from all over the world (3).

Careful selection of the performance metrics of benchmarking is another important issue. The mortality rate alone does not reflect the true quality of surgical outcome, and therefore, postoperative morbidity and health related quality of life have lately moved into focus. Outcome measures for benchmark analysis should be easily available and routinely collected, such as overall postoperative morbidity, severe complications, and length of hospital or intensive care unit stay (3). Also, surgery-specific quality indicators are of interest (e.g., anastomotic fistula in esophageal surgery or graft failure in liver transplantation) (3). It is highly important that outcome parameters are clearly defined and that each center's results are documented separately; otherwise, center-specific differences cannot be included in the benchmark (3).

The determination of the benchmark cut-offs is another critical step. The idea of a benchmark is to represent a realistic best possible outcome. Therefore, the benchmark value cut-off has been set at the 75th percentile of the median proportion of each included center (3,8); as a result of which not only the top few, but 75% of the best outcome achieved in high volume centers on ideal patients represent the benchmark.

It is evident that such best possible outcome is not always achievable, in particular when operating on multimorbid patients or in high risk situations (3). Yet, the benchmark sets an anchor for surgeons to know what is possible. When applying benchmarks to own patients, it is important that the comparison group exclusively consists of low risk "benchmark patients"—just like the population that was deployed for calculation of the benchmark (3). Apples must be compared with apples of the same quality. Patients with higher comorbidity are likely to have inferior results. However, in benchmarking, the surgeon and his team are being put to the test (8). If they perform within the benchmark for "ideal" patients, it may be assumed that their results are also within the benchmark when operating on patients with higher comorbidity. Of note, the fundamental idea of benchmarking in surgery is not to judge, monitor, or supervise a center's or individual surgeon's performance, but to offer a pragmatic and individually accessible measure for quality and thereby fuel every surgeon's pursuit of perfection (3).

Benchmark analyses in minimally invasive esophagectomy (MIE)

MIE for surgical treatment of esophageal cancer was introduced around 25 years ago with the aim to reduce procedure-related morbidity (13,14). Lately, MIE has evolved to the procedure of choice in many centers around the world (8), although this trend is supported by a handful of randomized studies only (15-17).

It is important to keep in mind that MIE is an umbrella term for many different operations, including transhiatal and transthoracic esophagectomy, hybrid, total minimally

invasive or robotic-assisted procedures. All these techniques are referred to as "minimally invasive" in the current literature, which makes comparison often difficult and sometimes even impossible.

Outcome research on esophageal cancer surgery is typically performed on the basis of large national patient registries (18-21). However, these national data collections report the outcome of institutions with variable expertise and include patients with a large range of risk factors that underwent a motley variety of MIE procedures. Moreover, standardized or well-defined outcome parameters are not available in most studies. Consequently, these analyses, even if reporting "benchmarks", are usually biased and provide only a blurred snapshot of the situation in a specific region of the world.
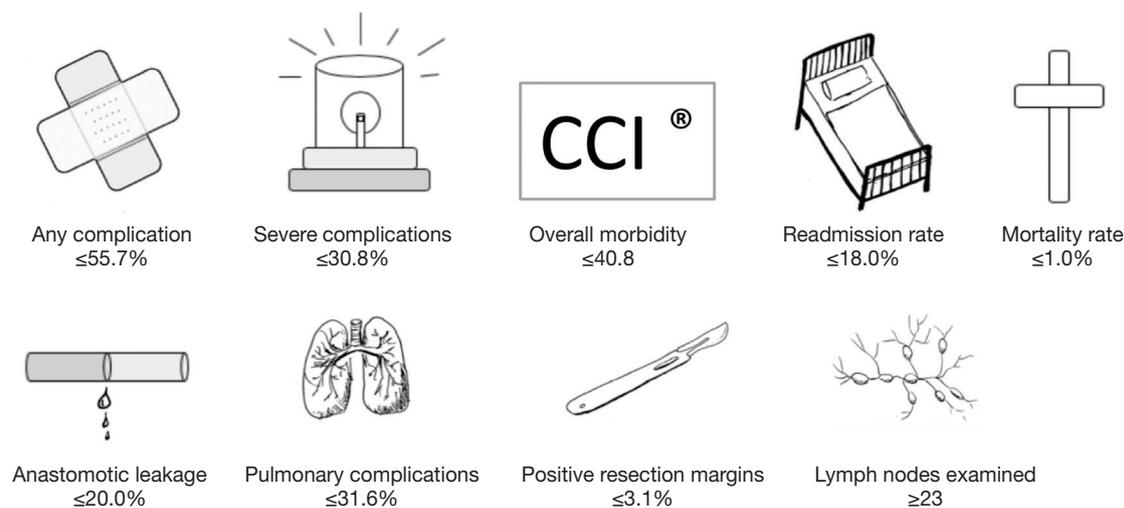
An alternative way to provide aggregated quality thresholds for esophagectomies is to gather data from renowned expert centers or high-volume institutions. Likewise, values on outcome after MIE—also referred to as benchmarks—were recently presented by the Esophagectomy Complications Consensus Group (ECCG) (22). Established from a high-quality international database including 24 expert centers from 14 different countries, this data collection holds a great amount of valuable information. Yet, for the determination of thresholds, the authors did not differentiate surgical approach or technique, nor did they provide different cut-off values for a specific patient's risk level. For instance, outcomes of patients with various comorbidity and different American Society of Anesthesiologists (ASA) scores that underwent all types of (open or minimally invasive) esophagectomy, or with tumors located in the proximal as well as distal esophagus etc. were mixed together. Further, points of reference are presented as a percentage of all patients included in the database; an approach that does not compensate for center specific differences. For one, hospitals contributing more patients to the dataset influence the threshold in a much greater extent; for the other, centers with better quality results may perish in the unity. For creating a valid benchmark that truly serves as point of reference of best possible outcome, the selection of the surgical approach as well as patients presumed to have the least amount of postoperative complications is of utter importance (3). Also, the results should be assessed for each center individually and each hospital's median result (or percentage of binominal outcomes) should be used for benchmark value calculation (3). Otherwise, these thresholds are rather a summary of the collected data than a true benchmark.

In an attempt to create a reliable benchmark for MIE, we recently performed a retrospective multicenter cohort study with an exclusive focus on total minimally invasive transthoracic esophagectomy (ttMIE) (Ivor Lewis and McKeown procedures only) (8). Some 1,057 patients from thirteen international expert centers for esophageal surgery (case load >20 esophagectomies per year) were included. We classified the outcome parameters according to the complications basic platform published by the ECCG (23) and postoperative morbidity was graded according to the Clavien-Dindo (CD) classification (24). "Ideal" patients meeting benchmark criteria were defined as having a low risk profile [Eastern Cooperative Oncology Group (ECOG) grade ≤1 and ASA score ≤2, age ≤65 years, and BMI 19–29 kg/m²]. Primary outcome measures for benchmark analysis were overall and major (CD ≥3a) morbidity, readmissions, anastomotic and pulmonary complications; all at 30 days after hospital discharge. In addition, positive resection margins, the number of examined lymph nodes, the 30- and 90-day Comprehensive Complication Index (CCI®) (25,26), and 30- and 90-day mortality rates were calculated (*Figure 1*). Benchmarks were defined as the 75th percentile of the median outcome parameters of the participating centers (3,8).

Our study was criticized because the reference points for benchmark values seemed high compared with other outcome research on MIE (27). However, this criticism reflects a common misinterpretation of the benchmark concept, because our cut-off values only represent upper limits of "best possible" results. This means that results from other centers should be within the thresholds of the benchmark (i.e., the best 75% of the median results of each center) to indicate acceptable outcome quality (28).

Discussion

Reducing the continuously increasing cost in health care is one of the greatest socioeconomic challenges of our time. In addition, there is increasing pressure on health care providers to introduce advanced surgical procedures or tools, such as minimally invasive or robotic-assisted surgery. However, from an economical point of view, the implementation of new technologies is often difficult to justify, because the benefits of those techniques are mostly short term with fewer postoperative complications and faster recovery after surgery. Furthermore, the scientific evidence supporting the upsides of these technologies is still weak and it remains debatable whether the benefits may counterbalance the greater financial expenditures.

Any complication
≤55.7%

Severe complications
≤30.8%

Overall morbidity
≤40.8

Readmission rate
≤18.0%

Mortality rate
≤1.0%

Anastomotic leakage
≤20.0%

Pulmonary complications
≤31.6%

Positive resection margins
≤3.1%

Lymph nodes examined
≥23

Figure 1 Benchmark values at 30 days after surgery. Results better than the given threshold are considered to be within the benchmark.

It has been shown that postoperative morbidity is a most critical factor in this debate. Thus, a strong correlation between overall postoperative morbidity and cost has just recently been demonstrated (29,30). Moreover, the reduction of postoperative complications not only complies with the strong economic requirements of today, but is of paramount importance for the patients' health, quality of life, and even long-term survival (30,31). However, it is a difficult task to work on outcome improvement, if the precise nature and extent of the quality issues are unknown. Benchmarks, which stand for the best possible outcome, may represent points of reference that help centers or individual surgeons on their path towards quality improvement.

Even if the crucial steps to create a valid benchmark are respected, a newly created benchmark metric may feature some subjectivity (3). It may be criticized for an individual selection of contributing centers or countries, patient inclusion criteria, or benchmark value cut-off points. Therefore, an intensive discussion about the appropriate performance metrics of benchmarking is paramount. The importance of this issue has recently been addressed in a study on risk adjustments for cancer esophagectomy (32). Parameters, such as the ASA score, the ECOG performance status, the number of comorbidities and anastomotic leakage rate were the strongest predictors of postoperative mortality (32). When establishing a new benchmark, such influential outcome measures must be considered when deciding on inclusion or exclusion criteria for "ideal" patients. However, when comparing own outcome data to the benchmark, additional risk adjustment is not required, as only "ideal"

patients meeting the identical inclusion criteria may be compared. Admittedly, centers that perform within the benchmark for "ideal" patients may not achieve the same level of quality in patients with higher comorbidity. This may be considered a drawback of the "ideal patient" benchmark approach. To answer this question, specific benchmarks for patients with higher perioperative risk would need to be established.

Another shortcoming of benchmarking relates to the effect of learning, which undoubtedly is a paramount factor in complex surgical procedures. Learning curves, particularly for novel technologies such as MIE (33) play a significant role in surgical benchmarking and it is evident that regular updates of benchmark cut-off parameters are necessary (28).

Center *vs.* individual surgeon volume is another aspect that has not been taken into account by previously published benchmark studies (7-9). Generally, due to the large experience of the surgical team, it is automatically assumed that high volume centers provide the highest level of surgical quality. However, the personal experience of an individual surgeon may play a more important role and a lower volume operated by a single surgeon may represent greater experience than a higher volume operated by many. Also, high volume centers often serve as training hospitals, which may bias results. Therefore, we recommend that future benchmark studies should also focus on individual surgeon volume.

In conclusion, compiling a valid benchmark that truly represents an anchor for best possible outcome is a tricky endeavor not only for minimally invasive procedures.

Careful evaluation of inclusion criteria to collect real benchmark data is essential.

Acknowledgements

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

References

1. Harrington HJ, Harrington JS. High Performance Benchmarking: 20 Steps to Success. New York: McGraw-Hill, 1996.

2. Zairi M, Leonard P. Practical benchmarking: the complete guide. London: Chapman & Hall, 1994.

3. Staiger RD, Schwandt H, Puhan MA, et al. Improving surgical outcomes through benchmarking. Br J Surg 2019;106:59-64.

4. Schiesser M, Kirchhoff P, Muller MK, et al. The correlation of nutrition risk index, nutrition risk score, and bioimpedance analysis with postoperative complications in patients undergoing gastrointestinal surgery. Surgery 2009;145:519-26.

5. Awad S, Lobo DN. Metabolic conditioning to attenuate the adverse effects of perioperative fasting and improve patient outcomes. Curr Opin Clin Nutr Metab Care 2012;15:194-200.

6. Evans DC, Martindale RG, Kiraly LN, et al. Nutrition optimization prior to surgery. Nutr Clin Pract 2014;29:10-21.

7. Rossler F, Sapisochin G, Song G, et al. Defining Benchmarks for Major Liver Surgery: A multicenter Analysis of 5202 Living Liver Donors. Ann Surg 2016;264:492-500.

8. Schmidt HM, Gisbertz SS, Moons J, et al. Defining Benchmarks for Transthoracic Esophagectomy: A Multicenter Analysis of Total Minimally Invasive Esophagectomy in Low Risk Patients. Ann Surg 2017;266:814-21.

9. Muller X, Marcon F, Sapisochin G, et al. Defining Benchmarks in Liver Transplantation: A Multicenter Outcome Analysis Determining Best Achievable Results. Ann Surg 2018;267:419-25.

10. Vonlanthen R, Lodge P, Barkun JS, et al. Toward a Consensus on Centralization in Surgery. Ann Surg 2018;268:712-24.

11. Özdemir-van Brunschot DM, Warlé MC, van der Jagt MF, et al. Surgical team composition has a major impact on effectiveness and costs in laparoscopic donor nephrectomy. World J Urol 2015;33:733-41.

12. Hartjes T, Gilliam J, Thompson A, et al. Improving Cardiac Surgery Outcomes by Using an Interdisciplinary Clinical Pathway. Aorn J 2018;108:265-73.

13. Giugliano DN, Berger AC, Rosato EL, et al. Total minimally invasive esophagectomy for esophageal cancer: approaches and outcomes. Langenbecks Arch Surg 2016;401:747-56.

14. Cuschieri A, Shimi S, Banting S. Endoscopic oesophagectomy through a right thoracoscopic approach. J R Coll Surg Edinb 1992;37:7-11.

15. Straatman J, van der Wielen N, Cuesta MA, et al. Minimally Invasive Versus Open Esophageal Resection: Three-year Follow-up of the Previously Reported Randomized Controlled Trial: the TIME Trial. Ann Surg 2017;266:232-6.

16. Briez N, Piessen G, Bonnetain F, et al. Open versus laparoscopically-assisted oesophagectomy for cancer: a multicentre randomised controlled phase III trial - the MIRO trial. BMC Cancer 2011;11:310.

17. van der Sluis PC, van der Horst S, May AM, et al. Robot-assisted Minimally Invasive Thoracolaparoscopic Esophagectomy Versus Open Transthoracic Esophagectomy for Resectable Esophageal Cancer: A Randomized Controlled Trial. Ann Surg 2019;269:621-30.

18. Messager M, Pasquer A, Duhamel A, et al. Laparoscopic Gastric Mobilization Reduces Postoperative Mortality After Esophageal Cancer Surgery: A French Nationwide Study. Ann Surg 2015;262:817-22; discussion 822-3.

19. Seesing MFJ, Gisbertz SS, Goense L, et al. A Propensity Score Matched Analysis of Open Versus Minimally Invasive Transthoracic Esophagectomy in the Netherlands. Ann Surg 2017;266:839-46.

20. Takeuchi H, Miyata H, Ozawa S, et al. Comparison of Short-Term Outcomes Between Open and Minimally Invasive Esophagectomy for Esophageal Cancer Using a Nationwide Database in Japan. Ann Surg Oncol 2017;24:1821-7.

21. Thirunavukarasu P, Gabriel E, Attwood K, et al. Nationwide analysis of short-term surgical outcomes of minimally invasive esophagectomy for malignancy. Int J Surg 2016;25:69-75.

22. Low DE, Kuppusamy MK, Alderson D, et al. Benchmarking Complications Associated with Esophagectomy. Ann Surg 2019;269:291-8.

23. Low DE, Alderson D, Cecconello I, et al. International Consensus on Standardization of Data Collection for Complications Associated With Esophagectomy: Esophagectomy Complications Consensus Group (ECCG). Ann Surg 2015;262:286-94.

24. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. Ann Surg 2004;240:205-13.

25. Slankamenac K, Graf R, Barkun J, et al. The comprehensive complication index: a novel continuous scale to measure surgical morbidity. Ann Surg 2013;258:1-7.

26. Clavien PA, Vetter D, Staiger RD, et al. The Comprehensive Complication Index (CCI®): Added Value and Clinical Perspectives 3 Years "Down the Line". Ann Surg 2017;265:1045-50.

27. Helminen O, Mrena J, Sihvo E. Defining Benchmarks for Transthoracic Esophagectomy: A Multicenter Analysis of Total Minimally Invasive Esophagectomy in Low-risk Patients. Ann Surg 2018. [Epub ahead of print].

28. Gutschow CA. Response to: "Letter to editor: 'Defining Benchmarks for Transthoracic Esophagectomy: A Multicenter Analysis of Total Minimally Invasive Esophagectomy in Low-risk Patients'". Ann Surg 2018. doi: 10.1097/SLA.0000000000003008

29. Vonlanthen R, Slankamenac K, Breitenstein S, et al. The impact of complications on costs of major surgical procedures: a cost analysis of 1200 patients. Ann Surg 2011;254:907-13.

30. Staiger RD, Cimino M, Javed A, et al. The Comprehensive Complication Index (CCI) is a Novel Cost Assessment Tool for Surgical Procedures. Ann Surg 2018;268:784-91.

31. Fransen L, Berkelmans G, Asti E, et al. FA01.02: the effect of postoperative complications after MIE on long-term survival: a retrospective, multi-center cohort study. Dis Esophagus 2018;31:1.

32. Fischer C, Lingsma H, Hardwick R, et al. Risk adjustment models for short-term outcomes after surgical resection for oesophagogastric cancer. Br J Surg 2016;103:105-16.

33. van Workum F, Stenstra MHBC, Berkelmans GHK, et al. Learning Curve and Associated Morbidity of Minimally Invasive Esophagectomy: A Retrospective Multicenter Study. Ann Surg 2019;269:88-94.