# Measuring quality of care for colorectal cancer care
# Comprehensive feedback information, driving quality improvement

Nicoline Elisabeth Kolfschoten

## Colofon

Measuring quality of care for colorectal cancer care, Comprehensive feedback information, driving quality improvement
Thesis, Leiden University, the Netherlands, 2014
Nicoline E. Kolfschoten, 2014 Then Netherlands
INSBN/EAN:

Cover: L. Bavinck, N. Kolfschoten
Layout: Chris D. Bor, Academic Medical Centre, University of Amsterdam
Printed by:

# Measuring quality of care for colorectal cancer care

Comprehensive feedback information, driving quality improvement

Proefschrift

Ter verkrijging van de graad van Doctor aan de Universiteit Leiden
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker
ingevolge het besluit van het College voor Promoties
in het openbaar te verdedigen op
Donderdag 29 januari 2015
Klokke 15.00 uur

Door

**Nicoline Elisabeth Kolfschoten**
Geboren op 12 april 1982 te Woerden

## Promotiecommissie

Promotores:

Prof. Dr. J. Kievit
Prof. dr. R.A.E.M. Tollenaar
Dr. M.W.J.M. Wouters
Dr. P.J. Marang-van de Mheen

Overige leden

Prof.  Dr. T. Wiggers
Prof. dr. E. Steyerberg
Prof. dr. C.J.H. van de Velde
Dr. M.E. Daniels

Dit proefschrift is tot stand gekomen met de financiële steun van:
KWF kankerbestrijding, Leids Universitair Medisch Centrum, Roche Nederland BV, Olympus Nederland B.V. Chipsoft, Covidien. Medisch centrum Haaglanden en het Dutch Institute for Clinical Auditing.

## Contents

# Chapter 1

## General introduction and outline of this thesis

## Quality of Care

### To err is human

Ever since the publication of the Harvard Medical Practice Study (1991)[1] and of the Institute of Medicine report "To Err is Human" (1999),[2,3] public attention has focused upon patient safety in health care, and the influence of the 'doctor' or 'hospital' on outcomes of care. A new field of research emerged, describing differences in clinical practice between hospitals and between care providers. Even so, as life expectancy increases, and new treatment options become available, healthcare costs are increasing exponentially, making quality assurance, while limiting budgets, one of the top priorities in health care politics all over the world. In the Netherlands, society's focus on hospital variations in quality of care resulted a strong call for transparency on hospital 'quality of care' for all stakeholders. To answer this call, various initiatives have been taken to measure and improve 'quality of care', and to increase transparency. Before we describe these different initiatives, it is helpful to first explain the concept 'quality of care'.

### Quality of care

Various definitions of quality of healthcare have been reported. The Organisation for Economic Co-operation and Development defined 6 aspects of quality of care: *effectiveness* (the degree of achieving desired outcomes), *safety* (the degree of prevention of adverse events), *patient centeredness* (the extent to which healthcare is organized around the patients' needs, rather than the doctor) *timeliness* (the accessibility of health services, in the Netherlands mostly defined by waiting lists), *equity* (the extent to which the system deals fairly with all concerned and guarantees the highest standards of care for all) and *efficiency* (is the system as productive as possible in terms of input and output).[4] The American Institute of Medicine defined quality of health care as "... the degree to which health services, for individuals and populations, increase the likelihood of desired health outcomes and are consistent with current professional knowledge."[5] This definition includes *safety* and *effectiveness* and *patient centeredness*. Although all aspects are equally relevant, this thesis will mainly focus on *safety* and *effectiveness*.

## Initiatives to measure and improve quality of care in the Netherlands

### Concentration of care

Due to the overwhelming evidence that a higher hospital procedural volume results in better outcome, the emphasis in quality assessment has shifted from outcome to volume, and procedural volume has become a surrogate measure for quality.[6-8] Accordingly, the political and professional focus in the Netherlands is now on concentrating care into high volume centres. Recently, the Association of Surgeons of the Netherlands has set a minimal annual volume standard for a large number of procedures. However, although for most of these

procedures an association between higher volume and better outcome has been proven, an evidence based minimal volume threshold has not been identified.[9] Moreover, these volume thresholds ignore the fact that low procedural volume does not rule out good quality of care, and high volume does not guarantee good quality. Therefore the Association of Surgeons of the Netherlands recognises that volume thresholds are only a pragmatic first step towards quality control, and that focus should be on quality rather than quantity.

## National evidence based guidelines

To reduce variations in treatment patterns and facilitate implementation of new scientific advances, national expert committees have developed evidence- and consensus-based guidelines for diagnosis, treatment and follow-up of different diseases. These guidelines are regularly updated when new evidence emerges. Among the most frequently used guidelines in the Netherlands are the oncological guidelines (accessible on www.oncoline.nl) developed by the Comprehensive Cancer Centres (CCC's). CCC's have a coordinating role in the regional multidisciplinary cancer networks, and are responsible for the implementation of national evidence based guidelines.

## Quality indicators

To monitor and improve quality of care and guideline adherence, the Dutch healthcare inspectorate introduced a set of quality indicators in 2003. Quality indicators, "measurable aspects of care, which reflect the quality of care", *(www.kiesbeter.nl)* have been developed to reveal either substandard or high quality of care.

## Types of quality indicators

Based on the Donabedian paradigm, quality indicators are commonly subdivided into three categories; structure, process or outcome indicators.[10] Structure indicators, reflecting 'what is there', or the availability of materials or means relevant for the treatment process, are easiest to measure, but often difficult to link directly to quality of care. Although outcome indicators are often seen as 'the bottom line' of what doctors do, valid outcome information, adjusted for a hospitals' *case-mix*, is not commonly available. Also, outcomes need a minimal event-rate to be *relevant* as an outcome indicator: as the event-rate of most specific adverse outcomes (specific complications or mortality) is rare, event-rates are largely influenced by random variation.[11,12] Process indicators on the other hand, usually selected from evidence based guidelines, give a fair reflection of what is done for a patient, and are more actionable than outcome indicators.[13] Moreover, they are less influenced by case-mix and random variation. However, the indicator must concern a process indicated for all selected patients, and have a proven relationship with relevant outcomes.

## Role of quality indicators in the Dutch Health Care System

After the introduction of quality indicators by the Dutch Health Care Inspectorate, many organisations followed, and by 2013 hospitals are overwhelmed by external requests for extensive lists of quality indicators, on various aspects of the care delivered. Nowadays, quality indicators are often used as hospital performance data and play a large role in the purchase policies of insurance companies. In addition, they are – sometimes carelessly - used by popular media to rate and rank hospitals in order to inform patients choosing their health care provider.

However, as quality indicators were designed only as 'signalling' measures to detect substandard care, their reliability and validity for these new purposes (for rating and ranking of hospitals, and as a basis for pay-for-performance) remains uncertain. [14,11]

## Reliability and validity of quality indicators

If quality indicators are used to assess and compare quality of care, it is important that they are both *reliable* and *valid*.

First, *reliability* of data is important. *Reliability* means that data are reproducible and that all patients concerned are included. Therefore, a uniform and valid data collection system for information on quality of care is vital for transparency of hospital quality. Currently, reliability of data used for quality indicators is insufficient, as for most quality indicators there is no clear, uniform definition, and data sources and methods for data collection differ between Dutch hospitals. Moreover, there is no data quality control system. Last, process and outcome indicators are often used regardless of relevance or case-mix.[15] It is therefore not surprising that hospital data on quality indicators often conflict with other sources, and result in unreliable and inconsistent hospital rankings.[16] These rankings and ratings result in unjustified negative publicity on quality of care, damaging societal and most importantly, patient's trust in the Dutch healthcare.

Second, a careful definition of quality indicators is important to assure *validity. Validity* means that the quality indicator measures what intents to measure. Presently, validity of the quality indicators used is uncertain: most indicators currently used are process indicators; ratio being that a good process will lead to good outcome. However, for most process indicators, a clear association with good hospital outcomes of care has not been established. In addition, in most cases a single indicator, only giving information on a small part of the care process, is used to assess the quality of the whole process of diagnostics, treatment and outcome of a specific disease. The use of such indicators bears the risk of a 'perverse incentive' leading to 'indicator driven practice': the focus being on optimization of results on a single indicator, rather than the quality of the whole care process. Also, good results on a single indicator can give a reassuring picture of the quality of care in a hospital, while reality may be different. The most striking example being the Staffordshire scandal, where the overall results on quality indicators were acceptable, while in fact quality of care was appalling.[17]

With the increasing societal call for data on hospital performances, there is a need for a new monitoring system and a new methodology for looking at hospital performances from a broader perspective, integrating the various aspects of quality of care. Therefore, a nation-wide database is needed, compiled by uniform data collection in all participating hospitals, with clear definitions, and with sufficient information to monitor structure, process as well as (case-mix adjusted) outcome at hospital level: a Clinical Audit.

## The introduction of Clinical Auditing

In 2009, the Dutch Cancer Society published a report on 'the quality of cancer care' in the Netherlands.[18] Main conclusions of this report were that quality of cancer care in the Netherlands was high, but large variations in treatment and outcome between Dutch hospitals existed. The reports' recommendations for improvement of quality of cancer care in the Netherlands were

- to develop quality standards describing the minimal requirements of infrastructure, volume and available medical specialties in a hospital to safely treat a disease,
- to centralize treatment into hospitals meeting these basic structural requirements, and
- to monitor the performance of these hospitals by the implementation of *clinical auditing:* the registration and feedback of detailed information on patients, processes and outcomes of care.

The information from these clinical audits can be used to reduce hospital differences in practice and outcome, to analyse national practice and to identify and implement best practices. Also, clinical audits can serve as a platform for the implementation of new techniques into clinical practice. Additionally, data from clinical audits can be used for hospital transparency in quality of care: to provide proof of adequate care to other stakeholders.

## The History of Clinical auditing

The idea of a clinical audit dates from the times of Florence Nightingale, who kept strict records of the mortality rates of injured or ill soldiers. After she implemented strict sanitary rules, and ensured that they were carried out, mortality dropped from 40 to 2%.[19] Another famous person who pioneered auditing in the medical profession was dr. Ernest Armory Codman (1869-1940). Codman, a surgeon in Massachusetts (USA), kept notes on all of his and his colleague's patients, treatments and their outcomes.[20] He suggested that *'every hospital should follow every patient it treats long enough to determine whether or not the treatment has been successful, and then to inquire, "If not, why not?" with a view to preventing similar failures in the future.'* However, confronting his peers with their failures and mistakes made him very unpopular in the medical society, and finally forced him to give up his medical practice.[21]

It was almost a century later that the value of clinical auditing was recognized in a broader medical audience. In 1975, the first national clinical audit for knee arthroplasty was implemented in Sweden. This example was quickly followed by many other audits in Sweden, but also internationally. Renowned examples are the Norwegian and Danish audits,

the National Quality Improvement Project (NSQIP) in the United States, and the various audits in the United Kingdom.

### Clinical auditing: process and requirements

Clinical auditing is a continuous process, often described as the '**audit cycle'**, which consists of

1. systematic, uniform registration of patient, treatment and outcome of all patients in a population, by those involved in the care process.
2. frequent comprehensive and meaningful feedback of information on performances.
3. improvement projects and identification of best practices to improve outcomes.
4. adjusting benchmarks and goals.

Basic requirements are

- a limited but meaningful dataset, including all relevant processes and outcomes, and all case-mix factors needed for risk-adjustments.
- full participation of all hospitals involved.
- a timely and frequent feedback system.
- involvement of professionals to realize meaningful feedback and case-mix adjustments.

## The Dutch Surgical Colorectal Audit

In 2009, a group of dedicated colorectal surgeons initiated the Dutch Surgical Colorectal Audit (DSCA): the first nationwide population based surgical outcome registration in the Netherlands. The aim of the DSCA is to monitor and improve outcomes of colorectal cancer resections in the Netherlands, by registering patient, diagnostic, treatment and outcome information and reporting this information back to the hospitals.

The dataset was developed by an expert committee and based on national evidence based guidelines. Data are entered by participating surgeons through a web-based interface in a highly secured database. In each participating hospital a single surgeon is appointed who is responsible for the data-entry. To secure data quality, hospitals receive data quality reports throughout the year, summing all patients with inconsistent or unusual data combinations, identified by a total of 70 queries. The responsible surgeon is asked to verify the data of these patients and to correct the data when indicated. Weekly feedback information on the number of registered patient files, and overall completeness of patient files combined with benchmarked performance indicators are placed online on a secured Internet page, accessible for the hospital only. Each year, outcomes of colorectal cancer care in the Netherlands are reported in an annual report, which is presented during an annual conference. Completeness and reliability of the data are cross-checked with the data from the Netherlands Cancer Registration.

## Colorectal Cancer treatment

Colorectal cancer is the third most common cancer in males and the second most common cancer in females. In the Netherlands, every year, 12,000 new patients are diagnosed with colorectal cancer, of which 8,600 with colon cancer, and 3,400 with rectal cancer.[22] Of these patients, 11,000 undergo a surgical resection of the tumour, or a Transanal Endoscopic Microsurgical (TEM) excision. Average 30-day mortality after colorectal cancer resections is 4.7%.[23] Complications occur in 34% and 11% of patients undergo a re-intervention. Long-term prognosis (5 year survival) for patients with colorectal cancer depends on tumour stage at diagnosis. Recent advances in treatment opportunities have increased survival, especially for rectal cancer patients. Overall 5-year survival is 59% for colon cancer and 61% for rectal cancer.[22] Evidence based guidelines describe the preferred treatment for each tumour stage, based on the latest scientific advances, in order to increase the likelihood of long-term disease-free survival.

The treatment process of colorectal cancer is described in the national evidence based guideline, available on 'Oncoline'. For each patient a tailored treatment plan should be made, using the information in the guidelines. However, little evidence is available on how to tailor treatment for specific patient groups such as elderly patients or patients with an emergent presentation.

## Hospital variation

Recent evidence has shown that for the treatment of colorectal cancer, there is a large variation between hospitals in adherence to these guidelines.[24,25] These variations may in part be explained by structural differences such as the availability of Magnetic Resonance Imaging or a radiotherapy department, but also by differences in hospital type, size, resources, organisation and logistics.[25] Last, the regional structure of multidisciplinary cancer networks may result in subtle differences in the implementation of guidelines. However, until recently these differences in guideline adherence were not observable for hospitals, doctors, or society, as guideline adherence was not systematically registered. Therefore, the DSCA aims to produce meaningful feedback information on structure, processes and outcomes of care in a hospital, which may be used to

- gain insight in and reduce hospital variations in practice, guideline adherence and outcome, to enable hospitals to **improve** their **outcomes** using the audit cycle, and to identify best practices
- gain more insight in **national practice** and performance, to set benchmarks, and identify aspects that need improvement
- monitor and assure the safe **implementation of new techniques**
- answer the need for **transparency** on quality of care to other stakeholders

## Challenges and general outline of this thesis.

The subject of this thesis is how data from clinical audits can be used to produce meaningful feedback information, supporting improvement of quality of care, using the DSCA as an example. It consists of 4 parts, elucidating the 4 different ways of using data from clinical audit to improve quality of care mentioned above: using the audit cycle as an instrument to **improve outcomes** and quality of care, evaluating **national practice** and performance, monitoring **implementation of new techniques**, and last, increasing **transparency** in quality of care.
In Part 1, **Chapter 2** we study the literature for evidence that clinical auditing, the registration and feedback of data on hospital quality of surgical care, indeed results in improvements. **Chapter 3** reports on the results of 3 years of clinical auditing on colorectal cancer treatment in the DSCA, and describes the key elements that led to the successful implementation of the DSCA.

In Part 2 we describe how data from clinical audits can be used to gain more insight in national practices and performances and identify aspects that need improvement, especially for high-risk patient groups. In **Chapter 4** we study the distribution of high-risk patients over the Dutch hospitals, using the 'expected mortality': an integrated measure for the effect of a patients risk factors on the likelihood of an unfavourable outcome. In **Chapter 5** we study the national results of non-elective colon cancer surgery compared to elective procedures in elderly patients.

In Part 3, **Chapter 6** shows how data from clinical audits can be used to monitor the implementation of new techniques, using the introduction of laparoscopic colorectal surgery as an example. We study the hospital variation of use and safety of laparoscopic surgery in the Netherlands.

In Part 4 we focus on how data from clinical audits can be used to evaluate quality of care and to increase transparency in quality of care. We evaluate how various aspects of quality of care cohere and on how these data can be combined into combined measures, which can be used by all stakeholders to evaluate quality of care as a whole. In **Chapter 7** we study the *validity* of available quality indicators to evaluate the quality of care for colorectal cancer, focussing on the construct validity and internal consistency of indicators. In **Chapter 8** process indicators are combined into composite measures reflecting guideline adherence. We study the hospital variation in guideline adherence and the association between guideline adherence and outcomes of care on patient and hospital level. In **Chapter 9** we describe how relevant outcome measures can be combined into a composite measure for a 'textbook outcome', which is valid and usable for all stakeholders, and prevents indicator driven practice. In **Chapter 10** we focus on the *relevance* of indicators. To judge quality of care, a minimal volume is needed. We therefore propose a combined measure for volume

and outcome, which selects hospitals with an acceptable outcome, but also a sufficient sample size to prove that their results are not just a 'lucky streak'.

In **Chapter 11**, the findings and implications of the studies included in this thesis are summarized, discussed and placed in a broader perspective. **Chapter 12** gives a summary of this thesis, translated in Dutch.

## References


1. Brennan TA, Leape LL, Laird NM, et al. Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. N Engl J Med. Feb 71991;324(6):370-376.
2. To Err is Human: Building a Safer Health System: Institute of medicine;1999.
3. Kohn L. To err is human: an interview with the Institute of Medicine's Linda Kohn. Jt Comm J Qual Improv. Apr 2000;26(4):227-234.
4. Hurst EKJ. Health Care Quality Indicators Project Conceptual Framework Paper: The Organisation for Economic Co-operation and Development; 2006.5. http://www.iom.edu.
6. Birkmeyer JD, Siewers AE, Finlayson EV, et al. Hospital volume and surgical mortality in the United States. N Engl J Med. Apr 11 2002;346(15):1128-1137.
7. Birkmeyer JD, Stukel TA, Siewers AE, Goodney PP, Wennberg DE, Lucas FL. Surgeon volume and operative mortality in the United States. N Engl J Med. Nov 27 2003;349(22):2117-2127.
8. Birkmeyer JD, Sun Y, Wong SL, Stukel TA. Hospital volume and late survival after cancer surgery. Ann Surg. May 2007;245(5):777-783.
9. Birkmeyer JD. Relation of surgical volume to outcome. Ann Surg. Nov 2000;232(5):724-725.
10. Donabedian A. Evaluating the quality of medical care. The Milbank Memorial Fund quarterly. Jul 1 1966;44(3):Suppl:166-206.
11. van Dishoeck AM, Lingsma HF, Mackenbach JP, Steyerberg EW. Random variation and rankability of hospitals using outcome indicators. BMJ quality & safety. Oct 2011;20(10):869-874.
12. van Dishoeck AM, Looman CW, van der Wilden-van Lier EC, Mackenbach JP, Steyerberg EW. Displaying random variation in comparing hospital performance. BMJ quality & safety. Aug 2011;20(8):651-657.
13. Birkmeyer JD, Dimick JB, Birkmeyer NJ. Measuring the quality of surgical care: structure, process, or outcomes? J Am Coll Surg. Apr 2004;198(4):626-632.
14. Brunelli A, Rocco G. Clinical and nonclinical indicators of performance in thoracic surgery. Thorac Surg Clin. Aug 2007;17(3):369-377.
15. Beperkt zicht; Onderzoek naar de betrouwbaarheid, validiteit en bruikbaarheid van prestatie indicatoren over de kwaliteit van de Nederlandse ziekenhuiszorg. Rapport NFU-consortium Kwaliteit van Zorg. 2012.
16. Maarsen HV, van, M. Nutteloze informatie, het ene beste ziekenhuis is het andere niet. Medisch Contact. 2004;59(44):1722-1724.
17. Francis R. Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry. (Report). House of Commons. ISBN 9780102981476. Retrieved 30 November 2013.
18. Jansen-Landheer MLEAW, M.W.J.M. Kwaliteit van kankerzorg in Nederland, door de signaleringscommissie Kanker van de KWF kankerbestrijding,. Oisterwijk: Vandenboogaard Print- en Mediamanagement; 2010.
19. Fee E, Garofalo ME. Florence Nightingale and the Crimean War. Am J Public Health. Sep 2010;100(9):1591.
20. Kauffmann RM, Landman MP, Shelton J, et al. The use of a multidisciplinary morbidity and mortality conference to incorporate ACGME general competencies. J Surg Educ. Jul-Aug 2011;68(4):303-308.
21. Brand RA. Ernest Amory Codman, MD, 1869-1940. Clinical orthopaedics and related research. Nov 2009;467(11):2763-2765.
22. www.cancerregistry.nl
23. Wouters MW, Eddes EH, Tollenaar RA, et al. Dutch Surgical Colorectal Audit, Jaarrapportage 2009. Uitkomst van zorg registratie; transparantie keuzes en kwaliteit van zorg. Arnhem: Dutch Surgical Colorectal Audit;2010.
24. Elferink MA, Wouters MW, Krijnen P, et al. Disparities in quality of care for colon cancer between hospitals in the Netherlands. Eur J Surg Oncol. Sep;36 Suppl 1:S64-73.
25. Elferink MA, Krijnen P, Wouters MW, et al. Variation in treatment and outcome of patients with rectal cancer by region, hospital type and volume in the Netherlands. Eur J Surg Oncol. Sep 2010;36 Suppl 1:S74-82.

# Part 1

The use of clinical auditing as an instrument for improvement of quality

Chapter 2

# ‘Clinical auditing’, a novel tool for quality assessment in surgical oncology

N.E. Kolfschoten[1*], N.J. van Leersum[1*], J.H.G. Klinkenbijl[2], R.A.E.M. Tollenaar[1] M.W.J.M. Wouters[3].

1 Leiden University Medical Center, dept Surgery, Leiden
2 Academic Medical Centre, dept Surgery, Amsterdam
3 Netherlands Cancer Institute, dept Surgical Oncology, Amsterdam
* The first two authors equally contributed to this paper.

Nederlands Tijdschrift voor Geneeskunde 2011;155(45):A4136

## Abstract

### Objective
To determine whether systematic audit and feedback of information about the process and outcomes improve the quality of surgical care.

### Design
Systematic review.

### Method
Embase, Pubmed, and Web of Science databases were searched for publications on 'quality assessment' and 'surgery'. The references of the publications found were examined as well. Publications were included in the review if the effect of auditing on the quality of surgical care had been investigated.

### Results
In the databases 2415 publications were found. After selection, 28 publications describing the effect of auditing, whether or not combined with a quality improvement project, on guideline adherence or indications of outcomes of care were included. In 21 studies, a statistically significant positive effect of auditing was reported. In 5 studies a positive effect was found, but this was either not significant or statistical significance was not determined. In 2 studies no effect was observed. 5 studies compared the combination of auditing with a quality improvement project with auditing alone; 4 of these reported an additional effect of the quality improvement project.

### Conclusion
Audit and feedback of quality information seem to have a positive effect on the quality of surgical care. The use of quality information from audits for the purpose of a quality improvement project can enhance the positive effect of the audit.

Conflict of interest: none declared. Financial support: none declared.

## Introduction

'Clinical Auditing' is a relatively new quality instrument in the Dutch healthcare system. Where regular evaluation of processes and end products is common in most branches, this is not the case for healthcare. In 1915, dr. Ernest Amory Codman, surgeon at Harvard University, advocated implementation of auditing, 'the systematic and critical analysis of quality of care delivered, including the process of diagnosis, treatment and outcomes of care, by those who deliver it', in medical practice. However, his visionary ideas were not appreciated by his colleagues. Only a century later, the use of auditing for quality improvement, transparency and accountability was internationally appreciated.
Clinical auditing is most commonly used in surgical oncology, as in this specialty, the relation between intervention and outcomes, or quality and costs is most obvious: a complication can result in repeated investigations, percutaneous interventions, reoperations, a long hospitalization and even treatment in an intensive care unit, all associated with substantial costs. Therefore, continuous improvement of quality of care is in the best interest of patients, but also of society.

In 2009 the 'Dutch Surgical Colorectal Audit' (DSCA,www.dsca.nl) was initiated, following previous international examples such as the National Surgical Quality Improvement Program (NSQIP; www.acsnsqip.org) in the United States and the 'National Bowel Cancer Project' (NBOCAP) in the United Kingdom (www.ic.nhs.uk/services/national-clinical-audit-support-programme-ncasp/cancer/bowel). The DSCA is a initiative of the Dutch Society for Surgical Oncology (NVCO), the Dutch Society for Gastro-intestinal Surgery (NVGIC) en de Dutch Colorectal Cancer Group (DCCG). By 2010, more than 20.000 patients are registered in this nationwide process and outcome registration for primary colorectal carcinoma. 98% of all Dutch hospitals participate, and from 2010 on, participation in the DSCA is a national performance indicator. Purpose of this registration system is to realize demonstrable quality improvement by means of systematic registration and feedback of reliable, case-mix adjusted information on the processes and outcomes of care delivered.

Recently, various medical professional associations have been facilitated by the Dutch Institute for Clinical Auditing (DICA; www.clinicalaudit.nl) to develop a clinical audit for breast, oesophagus, gastric and lung cancer, all according to the principles pioneered by the DSCA. These, and new developing audits now cover most of the surgical oncology field. However, clinical auditing also requires investments, not in the least of professionals, for whom the registration load is considerable. We therefore investigated the available evidence on whether measurement and feedback of information on process and outcome of surgical care result in improvement of process and outcomes of care by means of a systematic review of the available literature.

## Methods

### Search strategy

We searched for relevant articles in Pubmed, Web of Science and Embase, published before may 15th 2011. In this search, combinations of the 'medical subject headings' (MeSH-terms) 'surgery' (subdivided in 'surgical care' and 'operative procedure') and 'outcome- and process assessment' (subdivided in 'medical audit', 'outcome assessment', 'clinical audit', 'quality assurance' and 'benchmarking') were used. Outcome measures were process and/or outcomes of care, or guideline adherence. There were no restrictions on publication language. In addition, relevant websites and reference lists of included articles were screened for relevant articles.

### Article selection

Studies describing the effect of auditing on process and/or outcome indicators were selected. Auditing was defined as 'systematic measurement and feedback of structure, process and/or outcome information, in order to improve quality of care'; where needed, changes may be implemented at individual, team, hospital or national level and monitored by a new audit cycle.

Inclusion criteria were: a) at least one process or outcome indicator, or guideline adherence was measured, before and after the audit; b) the indicator or guideline was developed to evaluate quality of care, c) the indicator or guideline was focused on surgical care.

Relevant articles were selected by 2 independent researchers (NK en NvL) evaluating title and abstract of all retrieved publications. Discrepancies were discussed and when necessary, a third reviewer (MW) was consulted. Selected articles were included when all criteria were met. Included articles were subdivided in articles describing (a) the effect of auditing only, (b) the effect of auditing in combination with a quality improvement project and (c) comparing the effect of auditing with and without a quality improvement project. The level of evidence was assigned according to the CBO-guideline for 'Evidence-based Guideline development' (www.cbo.nl/thema/Richtlijnen/EBRO-handleiding/A-Levels-of-evidence/).

## Results

The search resulted in 2415 publications. After screening of titles and abstracts, 62 relevant articles were identified. After screening the reference lists of the selected articles, 9 more articles were selected. After reading the full text, 28 articles were included. (figure 1) Reasons for exclusion after reading the full text were: the audit did not fit our definition; the article did not describe original data, or the effect of the audit was not quantified.

Tables 1, 2 and 3 give an overview of the selected articles. Most articles were prospective cohort studies. Comparative studies (comparing two interventions) were summarized

in table 3. We found 2 randomized controlled trials (RCT) (table 3). Most studies were conducted in the United States in the last 5 years.

## Interventions and outcome measures

Nine studies described the effect of auditing only (table 1).[1-9] Twelve studies described the effect of auditing in combination with a quality improvement project (table 2),[10-21] such as the development of guidelines or checklists, in combination with educational meetings or newsletters. For example, one of these studies described the effect of a protocol for prevention of wound infections.[12]

Seven studies (2 RCT's and 5 prospective cohort studies, of which one longitudinal) described the effect of audits in combination with a quality improvement project compared with auditing only (table 3).[22-28] One of these studies compared results at three subsequent moments: before and after the start of the audit, and after the quality improvement project resulting from the audit.[28]

The manner and frequency of feedback varied. Information was presented in newsletters, websites or during specialist meetings, once or on weekly or annual basis. Three articles did not describe method nor the frequency of feedback.[20,22,25]

Most commonly described outcome measures were process indicators and guideline adherence (6 articles),[2,4,14,15,19,20] and the outcome indicators 'complications' and 'mortality' (13 articles),[1,5-12,18,22,23,28] or a combination of these (8 articles).[3,13,16,17,21,24,26,28] Outcomes were often compared with a baseline measurement.

## Effect of auditing

In 21 of 28 studies a statistically significant positive effect was described of auditing or of auditing in combination with a quality improvement project. In 5 studies, a positive effect was described, but no statistical tests were preformed.[5,8,10,13,15] In 1 study, the positive effect was not statistically significant ($p = 0.06$);[6] another study found no difference.[14] Six studies found a partial improvement, on some of the outcome measures investigated.[3,7,11,14,16,25]

## Effect of auditing in combination with quality improvement project

Three studies, as a part of the NSQIP compared the results of local improvement projects with other participants of the NSQIP (benchmarking).[24,26,27] Two of these studies described results of one hospital, which was a negative outlier in a previous report. In both studies, the improvement project resulted in the hospital returning to an average positing in the NSQIP. This was interpreted as a faster improvement than the total group of participating hospitals. One RCT investigated the effect of auditing with or without a quality improvement project consisting of implementation of a treatment guideline.[23] The study described an overall increase of guideline adherence, but no additive effect was found of the improvement project. In 3 of 4 comparative prospective cohort studies, a statistically significant

**Table 1.** Overview of prospective cohort studies investigating the effect of auditing in surgical interventions.

| Author, year | Type of surgery | Setting | Feedback | |
|---|---|---|---|---|
| | | | Type | Frequency |
| Antonacci, 2008[1] | All types of surgery | 3 hospitals | Meeting<br>Report | Weekly<br>Annual |
| Duxbury, 2003[2] | Colorectal cancer surgery | 1 hospital | Not specified | Once |
| Freeman, 2002[3] | Hip fractures | 10 hospitals | Not specified | Once |
| Galandiuk, 2004[4] | Colorectal surgery | 23 surgeons | Meeting<br>Report, newsletter | Every month<br>Annual<br>Not specified |
| Hall, 2009[5] | All types of surgery | NSQIP | Report | 2/year |
| Hammermeister, 1994[6] | Coronary bypass surgery | 45 hospitals | Report | 2/year |
| Henke, 2010[7] | All types of surgery | MSQC, NSQIP | 'Real time'-interface<br>Meeting | Continuous<br>4/ year |
| Khuri, 2002[8] | All types of surgery | NSQIP | Report | 2/year |
| Khuri, 2008[9] | All types of surgery | NSQIP | Report | 2/year |

NSQIP = National Surgical Quality Improvement Program (VS); MSQC = Michigan Surgical Quality Collaboration, a part of NSQIP; O/E = Observed/Expected (standardized for case-mix)
*Compared to baseline measurement before audit.
†Level B: prospective cohort study insufficiently controlled for confounders.
‡P < 0,05.
§Statistical significance not investigated

improvement was found in the group with an improvement project compared to the group with auditing only.

The second RCT investigated the effect of auditing in combination with a quality improvement project compared to no audit.[22] Auditing, combined with this improvement project resulted in a significant quality improvement. Another, observational study compared the effect of auditing or improvement projects with no intervention and found no differences.[25]

## Discussion

The results of our review suggest that the clinical auditing of process and outcomes of care, improves the quality of care. Clinical auditing can be combined with 'benchmarking',

| Effect* | Level of evidence† |
|---|---|
| Improvement:<br>Decrease of no of incidents in theatre† (wound infections, conversion, waste of implants and cancelled procedures) | B |
| Improvement:<br>Guideline adherence from 33 to 72%§ | B |
| Improvement:<br>Process improved‡<br>Morbidity decreased‡<br>Mortality unchanged | B |
| Improvement:<br>Guideline adherence improved‡ | B |
| Improvement:<br>In 66% of hospitals O/E mortality decreased ‡<br>In 82% of hospitals O/E morbidity decreased ‡ | B |
| Improvement:<br>Decrease of O/E mortality (p = 0,06) | B |
| Improvement:<br>Morbidity decreased from 15,8 to 13,8%‡<br>Mortality unchanged | B |
| Improvement:<br>Morbidity decreased 45%§<br>Mortality decreased 27% | B |
| Improvement:<br>Mortality decreased with 8,7%‡<br>Wound infections decreased with 9,1%‡<br>Renal complications decreased with 23,7%‡ | B |

comparing own results with those of other hospitals, or with improvement projects. The improvement of quality of care appears to be primarily accountable to the registration and feedback of information to professionals.

Previous reviews described similar results. A recent Cochrane review on the effect of auditing in on the quality of care in a broader perspective than surgical care only, reported a positive effect of auditing on the outcome measures.[29] However, the magnitude of improvement varied strongly between studies. A larger effect of auditing was found when the baseline situation was poor, and the feedback was more frequent and combined with educational sessions. The Cochrane review was limited to RCT's of which only two described surgical patients.

A second review in 1991, also found a positive effect of auditing on quality of care, especially when a target for improvement was set before the start of the audit.[30] When the auditing process, including feedback, was build into the process of care, the effect was found to be greater. The present study supports the previous findings of a positive effect of auditing of quality of surgical care. By expanding our search beyond RCT's we were able to include

**Table 2.** Overview of prospective cohort studies investigating the effect of auditing in combination with a quality improvement project in surgical interventions.

| Author, Year | Type of surgery | Setting | Feedback | |
|---|---|---|---|---|
| | | | Type | Frequency |
| Aitken, 1997[10] | All types of surgery | LSA | Meeting Report | Weekly Annual |
| Aletti, 2009[11] | Treatment of ovary cancer | 1 Hospital | Not specified | Not specified |
| Dellinger, 2005[12] | All types of surgery | 44 Hospitals | Report | 4/year |
| Doran, 1998[13] | All types of surgery | 2 Hospitals | Report | Every 2 weeks |
| Forbes, 2008[14] | All types of surgery | 1 Hospital | Report | Every month |
| Garnerin, 2007[15] | All types of surgery | 1 Hospital | Presentations | 4/year |
| Haynes, 2009[16] | All types of surgery | 3 Hospitals | Not specified | Once |
| Holman, 2004[17] | coronairy bypass surgery | 21 Hospitals | Not specified | Once |
| O'Connor, 1996[18] | coronairy bypass surgery | 5 Hospitals | Report | 3/year |
| Potenza, 2009[19] | All types of surgery | 1 Hospital | Meeting | Every month |
| Richardson, 1998[20] | All types of surgery | 1 Hospital | Not specified | Not specified |
| Tavris, 1999[21] | All types of surgery | 15 Hospitals | Not specified | Once |

LSA = Lothian Surgical Audit (Schotland)·
*compared to baseline measurement before audit·
Level B: prospective cohort study insufficiently controlled for confounders·
† Statistical significance not investigated·
‡P < 0,05·

more recent studies, reporting on various examples of clinical outcome registrations; apart from the RCT's we included 5 large prospective cohort studies with a level of evidence A2. However, most studies included had a longitudinal design, measuring the outcomes before and after implementation of the audit. A control group, in which no audit was conducted, was usually not available (level of evidence B). The observed improvements could therefore also be explained by autonomous evolvement of care instead of the clinical audit. Moreover, most studies only described short-term effects of clinical auditing. These effects could

| Improvement project | Effect* | Level of evidence |
|---|---|---|
| Specialized ward<br>Introduction of new methods | Improvement:<br>Decrease of mortality and complications† | B |
| seminars<br>cadaver training | Improvement:<br>Increase radical resections: 63 to 79%‡ | B |
| Development of guidelines for prevention of surgical site infections | Improvement:<br>Decrease in wound infections: 2.3 to 1.7%‡ | B |
| Development of guidelines<br>Adjustments to process of care | Improvement<br>Detubation within 6 hours: 5% to 70%<br>Decreased costs $18.200 to $14.700 per patient<br>Decreased median hospitalstay: 8.6 to 6.0 days† | B |
| Development of guidelines for prevention of surgical site infections | Improvement:<br>Guideline adherence improved‡<br>Surgical site infections: unchanged | B |
| Development of guidelines for prevention of 'wrong site/patient surgery' | Improvement:<br>Increased guideline adherence from 32 to 63% | B |
| surpass checklist | Decreased mortality: 1.5 to 0.8%‡<br>Decreased morbidity: 11 to 7%‡ | B |
| Defining performance-indicators<br>'site-visits'<br>Education | Improvement<br>Improved performance at most indicators‡<br>Outcomes unchanged | B |
| Annual meeting<br>Quality training<br>Site visits | Improvement:<br>Decreased mortality: 4.8 to 3.6%† | B |
| Development of guidelines for safe surgery | Improvement:<br>Increased guideline adherence: from 80 to 91% | B |
| Development of guidelines for ordering packed cells to reduce the crossmatch/transfusion ratio | Improvement: 'crossmatch/transfusion-ratio from 2.8 to 1.8† | B |
| Development of performance indicators for postoperative pain management | Improvement:<br>Improved performance on indicators 14 of 15 hospitals | B |

partly be explained by the Hawthorne-effect: the extra attention for the outcome measures brought on by the study, improves the medical practice for the duration of the study.

## The value of clinical auditing

Although clinical auditing cannot resolve all challenges of surgical oncology, it may improve treatment and survival of cancer patients. Previous studies such as the Dutch 'Total mesorectal excision' (TME)-trial, in which quality of rectal surgery was standardized and reviewed, showed how quality assurance of the surgical procedure can improve local control and survival in the study population.[31] However, patients included in studies often represent a specific, more favourable selection of the full population.

**Table 3.** Overview of studies comparing effect of auditing with auditing combined with an improvement project in surgical care.

| Author, year | Design* (Comparison) | Type of surgery | Setting | Feedback | |
|---|---|---|---|---|---|
| | | | | Type | Frequency |
| Berenguer, 20102[6] | Prospective cohort study (Audit + improvement project vs. audit) | Colorectal surgery | 1 hospital in NSQIP | Report | 2/year |
| Campbell, 20102[7] | Prospective cohort study (Audit + improvement project vs. audit) | All types of surgery | MSQC | Meeting<br>Report | 4/year<br>2/year |
| | | | NSQIP | Report | 2/year |
| Ferguson, 20032[2] | RCT (Audit + improvement project vs. control ‡) | Coronary bypass surgery | NCD | Not specified | Not specified |
| Guadagnoli, 20002[3] | RCT (Audit + improvement project vs. audit) | Breast cancer surgery | Not specified | Not specified | Once |
| Neumayer, 20002[4] | Prospective cohort study (Audit + improvement project vs. audit) | All types of surgery | NSQIP | Report | 2/year |
| Reilly, 20022[8] | Prospective cohort study (Audit, then improvement project) | All types of surgery | 1 hospital | Report | Every month |
| Sheikh, 20032[5] | Prospective cohort study (Audit + improvement project vs. control ‡) | Prostate cancer surgery | Not specified | Not specified | Not specified |

NSQIP = National Surgical Quality Improvement Program (VS); MSQC = Michigan Surgical Quality Collaboration, part of the NSQIP; NCD = National Cardiac Database SSI = Surgical Site Infection[5]
*Level of evidence: A2 (comparative clinical studies such as Randomized controlled trials or large cohort studies sufficiently corrected for confounders)·
†P < 0,05·
‡Control: no audit, no improvement project·

National clinical audits can be used to evaluate the effect of clinical practice on the full population, and to optimize practice when needed. Until recently, very little was known about the extent to which guidelines were followed, and the reasons for not adhering to guidelines. Clinical audits can be used as a platform for guideline evaluation, and implementation of new advances in technique or improvement projects. Based on information from these audits, best practices can be identified and implemented, and the effect of these best practices can be evaluated. In this way, professionals get more insight in the quality of care they deliver, but are also guided in how they can improve.

| Improvement project | Effect |
|---|---|
| Guideline for prevention of SSI | Audit + improvement project:<br>Guideline adherence improved from 38 to 92%†<br>Decrease of SSI from 13,3 to 8,3%†<br>Audit only (NSQIP):<br>Increase of SSI from 9,7 to 10,5% |
| MSQC: meetings and best practices in addition to audit and feedback<br>NSQIP: audit and feedback | MSQC: decreased morbidity rate from 10,7 to 9,7%†<br>NSQIP: no difference in morbidity rate (12,4%), no difference in mortality<br>Odds ratio for complications (MSQC vs |
| | NSQIP): 0,90† |
| Educational products, Presentations, Opinion leader, call to action letters | Larger improvement in preoperative bètablockade in intervention group than in control group†<br>Other process indicator not improved |
| Opinion leaders presentations and educational products | In both groups the possibility of a breast conserving treatment was more often discussed†<br>In both groups the frequency of breast conserving surgery increased†<br>no difference in effect between groups |
| Guideline for prevention of SSI | Decrease in SSI from 5,5 to 2,9%†<br>Hospital returned from negative outlier in NSQIP to average performing hospital |
| Guideline for prevention of SSI | SSI:<br>Before audit 14%<br>After audit 10%†<br>After improvement project 8%† |
| Presentations and information<br>Treatment guideline | No difference in radical prostatectomy rates between groups |

## Quality instrument

Clinical auditing is preferably used where a large effect can be established such as diseases involving large groups of patients or procedures that involve a considerable risk at adverse events. The data set should be based on an up-to-date evidence-based guideline, and an expert committee should be responsible for the definition of outcome measures and relevant case-mix factors (patient or disease related factors influencing the probability for the outcome measure). In this way, doctors are in the lead to define the essential processes which lead to the perfect hospitalization, and which will serve as their benchmarks. The success of clinical auditing therefore depends on the involvement and dedication of professionals.

For a frequent an timely feedback, short after the completion of the care process, data are collected from electronic patient files or by means of a 'web based' registration system.[7] With a complete national database, uniform definitions and the possibility to adjust for differences in case-mix and random variation, clinical auditing is a valid and reliable instrument for measuring and reporting on hospital quality of care. The results are of great value, not only for providers but also for policy makers, healthcare insurance companies, and patients. National clinical audits could also be used to support and control the imminent advances in oncological care such as centralization, regionalization and risk-based referral. Therefore, the implementation of a continuous clinical auditing cycle, consisting of guideline development and implementation, subsequent auditing, followed by education and visitation and finally auditing of the results, is strongly advised in any medical profession.

## Conclusion

Clinical auditing is a relatively new quality instrument in surgical oncology, which offers healthcare providers an insight in quality of care delivered. Clinical auditing may not only facilitate reviewing and benchmarking of providers' practices, but also offer insight in targets for quality improvement. Final goal is to assure that all Dutch patients receive optimal quality of surgical care.

### Take home message

- 'Clinical auditing' is defined as the systematic measurement and feedback of quality of care delivered, concerning patients, diagnostics, treatment and outcomes.
- The value of clinical auditing for practitioners should outweigh registration load
- Clinical auditing is increasingly used to monitor and improve quality of surgical oncological care.
- Clinical audits for the surgical treatment of bowel cancer, breast cancer, oesophagus and gastric cancer and lung cancer are now implemented in the Dutch healthcare system.
- Clinical auditing has a positive effect on the quality and outcomes of surgical care.
- Combining clinical auditing with a targeted quality improvement project, such as concentration of oncological care, or development of a treatment guideline, enlarges the effect.

## References


1 Antonacci AC, Lam S, Lavarias V, Homel P, Eavey RD. Benchmarking surgical incident reports using a database and a triage system to reduce adverse outcomes. Arch Surg. 2008;143:1192-7.

2 Duxbury MS, Brodribb AJ, Oppong FC, Hosie KB. Management of colorectal cancer: variations in practice in one hospital. Eur J Surg Oncol. 2003;29:400-2.

3 Freeman C, Todd C, Camilleri-Ferrante C, et al. Quality improvement for patients with hip fracture: experience from a multi-site audit. Qual Saf Health Care. 2002;11:239-45.

4 Galandiuk S, Rao MK, Heine MF, Scherm MJ, Polk HC. Mutual reporting of process and outcomes enhances quality outcomes for colon and rectal resections. Surgery. 2004;136:833-41.

5 Hall BL, Hamilton BH, Richards K, Bilimoria KY, Cohen ME, Ko CY. Does surgical quality improve in the American College of Surgeons National Surgical Quality Improvement Program: an evaluation of all participating hospitals. Ann Surg. 2009;250:363-76.

6 Hammermeister KE, Johnson R, Marshall G, Grover FL. Continuous assessment and improvement in quality of care. A model from the Department of Veterans Affairs Cardiac Surgery. Ann Surg. 1994;219:281-90.

7 Henke PK, Kubus J, Englesbe MJ, Harbaugh C, Campbell DA. A statewide consortium of surgical care: a longitudinal investigation of vascular operative procedures at 16 hospitals. Surgery. 2010;148:883-9.

8 Khuri SF, Daley J, Henderson WG. The comparative assessment and improvement of quality of surgical care in the Department of Veterans Affairs. Arch Surg. 2002;137:20-7.

9 Khuri SF, Henderson WG, Daley J, et al. Successful implementation of the department of Veterans Affairs' National Surgical Quality Improvement Program in the private sector: The patient safety in surgery study. Ann Surg. 2008;248:329-36.

10 Aitken RJ, Nixon SJ, Ruckley CV. Lothian surgical audit: a 15-year experience of improvement in surgical practice through regional computerised audit. Lancet. 1997;350:800-4.

11 Aletti GD, Dowdy SC, Gostout BS, et al. Quality improvement in the surgical approach to advanced ovarian cancer: the Mayo Clinic experience. J Am Coll Surg. 2009;208:614-20.

12 Dellinger EP, Hausmann SM, Bratzler DW, et al. Hospitals collaborate to decrease surgical site infections. Am J Surg. 2005;190:9-15.

13 Doran KA, Henry SA, Anderson BJ. Breakthrough change for adult cardiac surgery in a community-based cardiovascular program. Qual Manag Health Care. 1998;6:29-36.

14 Forbes SS, Stephen WJ, Harper WL, et al. Implementation of evidence-based practices for surgical site infection prophylaxis: results of a pre- and postintervention study. J Am Coll Surg. 2008;207:336-41.

15 Garnerin P, Arès M, Huchet A, Clergue F. Verifying patient identity and site of surgery: Improving compliance with protocol by audit and feedback. Qual Saf Health Care. 2008;17:454-8.

16 Haynes AB, Weiser TG, Berry WR, Lipsitz SR, Breizat AH, Dellinger EP, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. N Engl J Med. 2009;360:491-9.

17 Holman WL, Sansom M, Kiefe CI, et al. Alabama coronary artery bypass grafting project: results from phase II of a statewide quality improvement initiative. Ann Surg. 2004;239:99-109.

18 O'Connor GT, Plume SK, Olmstead EM, Morton JR, Maloney CT, Nugent WC, et al. A regional intervention to improve the hospital mortality associated with coronary artery bypass graft surgery. The Northern New England Cardiovascular Disease Study Group. JAMA. 1996;275:841-6.

18 Potenza B, Deligencia M, Estigoy B, et al. Lessons learned from the institution of the Surgical Care Improvement Project at a teaching medical center. Am J Surg. 2009;198:881-8.

19 Richardson NG, Bradley WN, Donaldson DR, O'Shaughnessy DF. Maximum surgical blood ordering schedule in a district general hospital saves money and resources. Ann R Coll Surg Engl. 1998;80:262-5.

20 Tavris DR, Dahl J, Gordon D, et al. Evaluation of a local cooperative project to improve postoperative pain management in Wisconsin hospitals. Qual Manag Health Care. 1999;7:20-7.

21 Ferguson TB Jr, Peterson ED, Coombs LP, Eiken MC, Carey ML, Grover FL, et al. Use of Continuous Quality Improvement to Increase Use of Process Measures in Patients Undergoing Coronary Artery Bypass Graft Surgery. A Randomized Controlled Trial. JAMA. 2003;290:49-56.

22 Guadagnoli E, Soumerai SB, Gurwitz JH, Borbas C, Shapiro CL, Weeks JC, et al. Improving discussion of surgical treatment options for patients with breast cancer: Local medical opinion leaders versus audit and performance feedback. Breast Cancer Res Treat. 2000;61:171-5.

23 Neumayer L, Mastin M, Vanderhoof L, Hinson D. Using the veterans administration national surgical quality improvement program to improve patient outcomes. J Surg Res. 2000;88:58-61.

24 Sheikh K, Bullock C. Effectiveness of interventions for reducing the frequency of radical prostatectomy procedures in the elderly: an evaluation. Am J Med Qual. 2003;18:97-103.

25 Berenguer CM, Ochsner MG, Lord SA, Senkowski CK. Improving Surgical Site Infections: Using National Surgical Quality Improvement Program Data to Institute Surgical Care Improvement Project Protocols in Improving Surgical Outcomes. J Am Coll Surg. 2010;210:737-41.

26 Campbell DA Jr, Englesbe MJ, Kubus JJ, Phillips LR, Shanley CJ, Velanovich V, et al. Accelerating the pace of surgical quality improvement: The power of hospital collaboration. Arch Surg. 2010;145:985-91.

27 Reilly J, McIntosh J, Currie K. Changing surgical practice through feedback of performance data. J Adv Nurs. 2002;38:607-14.

28 Jamtvedt G, Young JM, Kristoffersen DT, O'Brien MA, Oxman AD. Does telling people what they have been doing change what they do? A systematic review of the effects of audit and feedback. Qual Saf Health Care. 2006;15:433-6.

29 Mugford M, Banfield P, O'Hanlon M. Effects of feedback of information on clinical practice: a review. BMJ. 1991;303:398-402.

30 Peeters KC, Marijnen CA, Nagtegaal ID, et al. The TME trial after a median follow-up of 6 years: increased local control but no survival benefit in irradiated patients with resectable rectal carcinoma. Ann Surg. 2007;246:693-701.

Chapter 3

# The Dutch Surgical Colorectal Audit

Van Leersum NJ*[1], Snijders HS*[1], Henneman D[1], Kolfschoten NE[1], Gooiker GA[1], ten Berge, M.G.[1], Eddes EH[3], Wouters MWJM[1, 2] and Tollenaar RAEM[1] on behalf of the Dutch Surgical Colorectal Audit Group**

1 Leiden University Medical Center, dept Surgery, Leiden
2 Academic Medical Centre, dept Surgery, Amsterdam
3 Netherlands Cancer Institute, dept Surgical Oncology, Amsterdam
* The first two authors equally contributed to this paper.
** The Dutch Surgical colorectal audit group: W.A. Bemelman, R.M. van Dam, M. Elferink, Th.M. Karsten, J.H.J.M. van Krieken, V.E.P.P. Lemmens, H.J. T. Rutten, E.R. Manusama, C.J.H. van de Velde, W.J.H.J Meijerink, Th. Wiggers, E. van der Harst, J.W.T. Dekker, J. Boerma

*Eur J Surg Oncol. 2013 Oct;39(10):1063-70.*

## Abstract

### Introduction

In 2009, the nationwide Dutch Surgical Colorectal Audit (DSCA) was initiated by the Association of Surgeons of the Netherlands (ASN) to monitor, evaluate and improve colorectal cancer care. The DSCA is currently widely used as a blueprint for the initiation of other audits, coordinated by the Dutch Institute for Clinical Auditing (DICA). This article illustrates key elements of the DSCA and results of three years of auditing.

### Methods

Key elements include: a leading role of the professional association with integration of the audit in the national quality assurance policy; web-based registration by medical specialists; weekly updated online feedback to participants; annual external data verification with other data sources; improvement projects.

### Results

In two years, all Dutch hospitals participated in the audit. Case-ascertainment was 92% in 2010 and 95% in 2011. External data verification by comparison with the Netherlands Cancer Registry (NCR) showed high concordance of data items. Within three years, guideline compliance for diagnostics, preoperative multidisciplinary meetings and standardised reporting increased; complication-, re-intervention and postoperative mortality rates decreased significantly.

### Discussion

The success of the DSCA is the result of effective surgical collaboration. The leading role of the ASN in conducting the audit resulted in full participation of all colorectal surgeons in the Netherlands. By integrating the audit into the ASNs' quality assurance policy, it could be used to set national quality standards. Future challenges include reduction of administrative burden; expansion to a multidisciplinary registration; and addition of financial information and patient reported outcomes to the audit data.

## Introduction

Several clinical audits have been initiated internationally, acknowledging the importance of reliable and valid quality information in health care. Clinical auditing has been recognised as an important tool for quality assessment and improvement, consequently leading to demonstrable improvements in patient outcome[1-4] Moreover, clinical audits are increasingly appreciated as a source of information for research on evidence based medicine as they provide 'real world' data on patients often not eligible for clinical trials.[5] However, the voluntary nature of existing audits may unintentionally lead to participation of mainly dedicated hospitals and underrepresentation of underperforming hospitals. Also, audit data are seldom transparent to other stakeholders involved in health care.

In 2009, the Dutch Surgical Colorectal Audit (DSCA) was initiated by the Association of Surgeons of the Netherlands (ASN) in collaboration with the Dutch Association for Surgical Oncology (NVCO), the Dutch Association for Gastrointestinal Surgery (NVGIC) and the Dutch Colorectal Cancer Group (DCCG). Their main goal was to evaluate and improve quality of care for primary colorectal cancer surgery in the Netherlands.

After one year of registration, participation in the audit had become a national performance indicator. Full participation of Dutch hospitals was realised within two years. Subsequent to this success, the Dutch Institute of Clinical Auditing (DICA) was founded in 2011 with the objective to facilitate and organise the start-up of new nation-wide audits. This article illustrates the introduction of the DSCA in the Netherlands by describing its main features and presenting the results of three years of auditing.

## Methods

### Main features of the DSCA

This section describes the organisational and structural key elements of the DSCA.

*1. The initiator: the professional organisation of surgeons*

All surgeons in the Netherlands are united in a professional organisation, the Association of Surgeons in the Netherlands (ASN). The ASN serves as a central protector of common interests of surgeons. Membership of the ASN is compulsory to all surgeons in the Netherlands. One of its main objectives is to assure that every surgical patient in the Netherlands receives high quality care. Furthermore, ASN continuously attempts to improve the quality of surgical care. The ASN uses different instruments to accomplish this, for example the development of evidence-based guidelines, surgical training programs and accreditation of surgeons in their surgical specialty. The initiation of clinical audits was necessary to facilitate the uniform measurement of quality of care and enhance the Association's quality improvement efforts.

*2. Dataset: involvement of all experts in the field*

The ASN formed a scientific committee of mandated clinical experts in colorectal cancer care (surgeons, oncologists, pathologists, epidemiologists) to initiate the first clinical audit. The scientific committee defined performance indicators and outcome measures, based on pre-existing evidence based guidelines, to highlight potential quality concerns, identify areas that need further investigation, and track changes over time. The committee defined a dataset using a Delphi method[6]. The dataset generally covers three aspects: case-mix variables (e.g. age, gender, co morbidity) necessary for hospital comparison; process variables (e.g. wait times and number of patients discussed in a multidisciplinary team); and outcomes of care (e.g. morbidity and mortality).

*3. Organizational structure*

In accordance with the format of the DSCA, the Dutch Institute of Clinical Auditing (DICA) was founded to enhance other clinical audit initiatives in the Netherlands. The main goal of the DICA was to support other clinical audits by facilitating on legal, technical, methodological and logistic issues. Three new audits have been initiated since the introduction of the DSCA: the breast cancer audit (NBCA), the upper GI cancer audit (DUCA) and the lung surgery audit (DLSA). The organization structure of the DICA is graphically presented in Figure 1.

*4. Funding*

The onset of the DSCA was funded by quality improvement grants donated by a health care insurance company. Since 2013, hospitals pay a subscription fee for participating in the DSCA. The subscription costs are returned to the hospitals as they are enclosed in the payments of treating patients with colorectal cancer. Costs of the data registration itself are not compensated and are borne by the hospitals.

*5. Online data is self-registered in a secured web form*

Each participating hospital appoints a surgeon responsible for (supervising) the data registration. The majority of the colorectal surgeons record the data themselves. The DSCA uses a generic internet based program to enable data entry in a secured web environment[7]. Depending on the complexity of the patient and perioperative course, a number of 56 to 179 variables have to be completed; registration time is approximately 20 to 30 minutes per patient. Data-entry can be entered either throughout patient's management or at the end of each admission. Data can be updated when necessary; for example when follow-up data is available. A third trusted party anonymises data regarding patient identification directly after data entry[8]. Definitions and helping texts are appointed to each variable in the dataset and are available during data entry. These guarantee that registration is performed uniformly. Also, frequently asked questions (FAQs) are available on the website and a front office can be contacted by data registrants for questions on both technical and content issues.

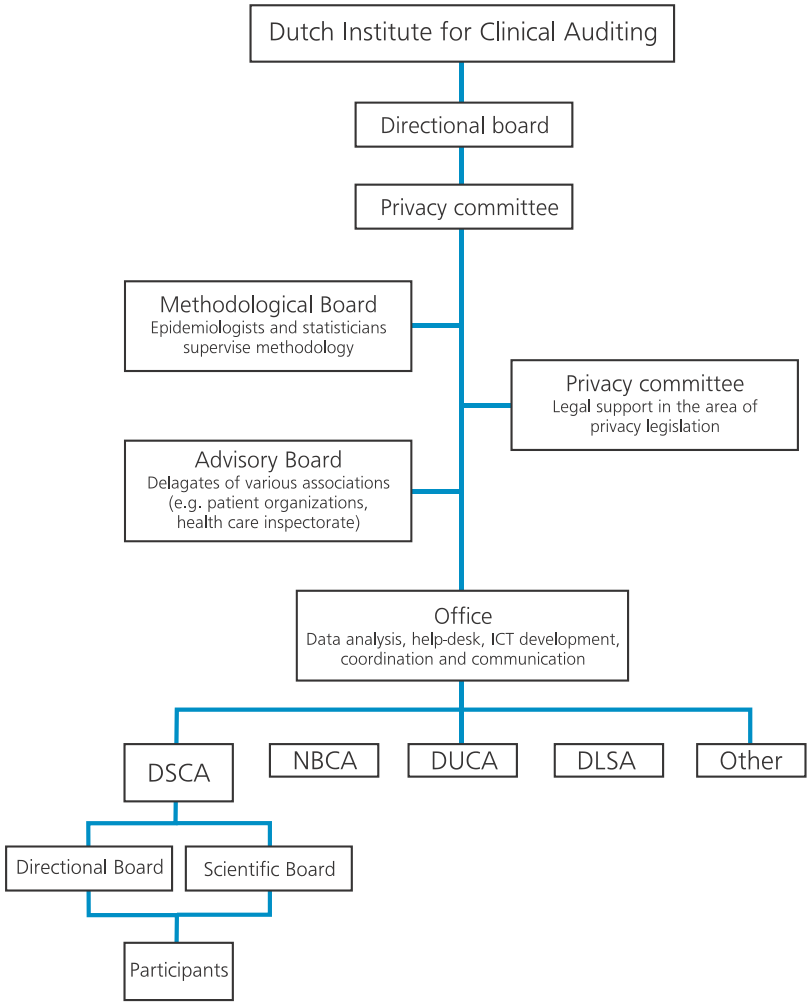



**Figure 1.** Organisational structure of the Dutch Institute for Clinical Auditing (DICA).
DSCA: Dutch Surgical Colorectal Audit; NBCA: Nabon Breast Cancer Audit; DUCA: Dutch Upper GI Audit; DLSA: Dutch Lung Surgery Audit.

### *6. Internal and external data verification*

Data validity is achieved and verified in various ways. The surgeon receives direct feedback on erroneous, missing or improbable data items during data entry through quality control tools that are built in the program. Hospitals receive feedback information on the number of patients and completeness of the data to encourage the participants to correct them when needed.

Data are annually compared with an external data registration, the National Cancer Registry (NCR), on completeness and accuracy.[1] The NCR registers all newly diagnosed malignancies in the Netherlands. Information on patient characteristics (e.g. age, gender)

tumour characteristics (TNM stage, localization, histology) treatment (surgical procedure, chemo and/or radiation therapy, laparoscopy, urgency of procedure) hospital of diagnosis, hospital of treatment and outcomes (30-day mortality, anastomotic leakage, CRM, lymph nodes), are collected from the medical records by specially trained registrars 9 months after diagnosis[9,10]. The NCR has an automatic linkage to many important and solid databases, among which the Municipal Administration (GBA), which allow the full enrolment of patients eligible for registration and notification for postoperative mortality. Quality of the NCR data is high; completeness is estimated to be at least 95%.[11] The registration of the NCR is linked to the Municipal Administration, which by law receives notification on all patients that decease in the Netherlands. The quality of the data in comparison to the NCR is described elsewhere[12].

*7. Online feedback is provided on a weekly basis*

Information regarding volume, performance indicators and outcomes of care are presented online to individual hospitals. Each participating hospital has access to its own secured website. Data are weekly updated. Results of the hospital are presented in relation to the national average and in relation to results of other anonymised hospitals.

*8. Outcomes are adjusted for differences in case-mix*

The methods to measure quality of care are described in detail elsewhere.[12,13] When comparing hospital outcomes differences in case-mix must be taken into account.[14] Therefore, a set of relevant case-mix variables specific for each outcome measure is embedded in the database. A standardised co morbidity module was developed using the Delphi method with incorporation of the Charlson Co morbidity Index.[15, 16] Case-mix adjusted hospital outcomes are presented in funnel plots using 95% confidence limits that vary in relation to the hospital volume.[17]

*9. Results and targets for quality improvement are presented in an annual report.*

An extensive national report presenting the results of the audit is published annually.[12] This report focuses on various themes for improvements in the scope of recent literature. The results are presented in a yearly conference accessible to clinicians, patients, patient advocates, health insurers and policy makers, politicians. The conference functions as a platform for all parties to address their (common) interests and to discuss diverse health care topics.

## Analysis of results of the DSCA

The completeness of the data on a national level is described by the percentage of participating hospitals and case ascertainment for each audit year. Patient, tumour and treatment characteristics are shown separately for patients with colon and rectal cancer. Then, the results of performance indicators on both process and outcomes of care were evaluated using a Chi square trend test was used to analyse changes over time. Last, hospital variation for preoperative multidisciplinary team discussions for rectal cancer surgery are presented in a scatter plot, illustrating changes in variation over time.

# Results

## Dataset

From 2009 to 2011, 26,511 patients undergoing surgical resection for colorectal carcinoma were registered by all 92 hospitals providing colorectal cancer care in the Netherlands (8 university, 47 teaching and 37 non-teaching hospitals). The national case ascertainment and completeness of the data per patient record was high. Compared with the data collected by the NCR, the DSCA included 80% of all eligible patients in 2009, 92% in 2010, and 95% in 2011. External data verification with the NCR showed nearly 100% completeness and high correspondence on almost all items of the dataset[12].

## Patients

Information on tumour localisation, date of surgery and mortality are minimal requirements for analysis of patient records. In total, 752 patients (2.8%) were excluded for this reason. Hospitals that failed to register more than 10 patients were excluded to minimise selection bias. In 2009, this concerned 5 hospitals registering a total of 37 patients. In 2010 and 2011, none were excluded. In the results presented in this article, patients with multiple synchronous tumours (n=894) were excluded as well. A total of 24,828 patients were included in the analysis. Patient, tumour and treatment characteristics are shown in Table 1, stratified by tumour location: colon (n=17,729) and rectal cancer (n=7,099). Patients in both groups differ in age, prevalence of preoperative complications, urgency of the resection and tumour stage. Treatment patterns differ as well. For example, the percentage of diverting stomas is 4% in colon cancer surgery compared to 33% in rectal resections. Preoperative radiation therapy is applied in 84% of rectal cancer patients, which is very high from an international perspective.[17]

**Table 1.** Patient, tumour and treatment characteristics of patients included in the DSCA, stratified by colon and rectum

| | | Colon | | Rectum | |
|---|---|---|---|---|---|
| | | N | % | N | % |
| Total | | 17729 | | 7099 | |
| Age | >70 | 10192 | 57.5% | 3155 | 44.4% |
| Gender | Male | 9212 | 52.0% | 4394 | 61.9% |
| ASA score | III | 4064 | 22.9% | 1133 | 16.0% |
| | IV-V | 410 | 2.3% | 65 | .9% |
| | Missing | 426 | 2.4% | 168 | 2.4% |
| Charlson score | 1 | 3965 | 22.4% | 1409 | 19.8% |
| | ≥ 2 | 4313 | 24.3% | 1327 | 18.7% |
| Body Mass Index | 25-30 kg/m² | 4701 | 26.5% | 1935 | 27.3% |
| | >30 kg/m² | 4752 | 26.8% | 2204 | 31.0% |
| | Missing | 5982 | 33.7% | 2073 | 29.2% |
| Abominal surgical history | Yes | 6009 | 34.6% | 2094 | 30.1% |

**Table 1.** *(Cont.)*

| | | Colon | | Rectum | |
|---|---|---|---|---|---|
| | | N | % | N | % |
| Tumour location | Right colon | 7917 | 44.7% | - | - |
| | Transversum/left colon | 2884 | 16.3% | - | - |
| | Sigmoid | 6928 | 39.1% | - | - |
| Distance of tumour from anal verge | < 5 cm | - | - | 2379 | 37.1% |
| | 5 - 10 cm | - | - | 2613 | 40.8% |
| | > 10 cm | - | - | 1417 | 22.1% |
| | Missing | - | - | 697 | 9.9% |
| Urgency of resection | Urgent | 3567 | 20.1% | 199 | 2.8% |
| Preoperative tumour complications | Tumour perforation | 354 | 2.0% | 41 | .6% |
| | Abces | 262 | 1.5% | 33 | .5% |
| | Ileus | 2290 | 12.9% | 176 | 2.5% |
| | Bleeding | 983 | 5.5% | 383 | 5.4% |
| Tumour stage (TNM) | I | 2974 | 16.8% | 2054 | 28.9% |
| | II | 6410 | 36.2% | 1804 | 25.4% |
| | III | 5500 | 31.0% | 2030 | 28.6% |
| | IV | 2319 | 13.1% | 566 | 8.0% |
| | X | 365 | 2.1% | 259 | 3.6% |
| Surgical preoperative treatment | Stoma | 182 | 9.6% | 560 | 9.8% |
| | Stent | 157 | 8.3% | 16 | .3% |
| | Metastasectomy/RFA | 35 | 1.8% | 96 | 1.7% |
| | Other | 24 | 1.3% | 34 | .6% |
| Preoperative radiotherapy | 5x5 Gy | - | - | 3312 | 46.7% |
| | Long course isolated radiotherapy | | | 595 | 7.9% |
| | Chemoradiation | - | - | 2033 | 28.6% |
| Surgical procedure | Ileocoecal resection | 258 | 1.5% | - | - |
| | Right hemicolectomy | 7785 | 43.9% | - | - |
| | Transversal resection | 553 | 3.1% | - | - |
| | Left hemicolectomy | 1762 | 9.9% | - | - |
| | Sigmoid/(low) nterior esection | 6489 | 36.6% | 4371 | 61.6% |
| | Abdominoperineal resection | - | - | 2168 | 30.5% |
| | Subtotal colectomy | 159 | 0.9% | 191 | 2.7% |
| | Panproctocolectomy | 148 | 0.8% | 43 | 0.6% |
| | Other | 289 | 1.7% | 126 | 1.8% |
| | Missing | 286 | 1.6% | 200 | 2.8% |
| Surgical approach | Laparoscopical | 6606 | 37.4% | 2690 | 38.1% |
| Anastomosis | Primary anastomosis | 15556 | 87.7% | 3252 | 45.8.% |
| | No anastomosis (end-colostomy)* | 2173 | 12.3% | 3847 | 54.2% |
| Diverting stoma** | Yes | 709 | 4.6% | 2123 | 65.3% |
| Extended resections | Minimal local extended resection | 1036 | 6.2% | 258 | 3.9% |
| | Maximal local extended resection | 810 | 4.8% | 280 | 4.2% |
| | Metastasectomy | 591 | 3.5% | 202 | 3.0% |

ASA: American Society of Anaesthesiologists risk score. RFA: radiofrequent ablation. *includes abdominoperineal resections; **percentage is related to the performed anastomoses.

## Performance indicators

A number of noticeable improvements on pre-defined performance indicators occurred since the introduction of the audit in 2009. These improvements concerned both processes as well as outcomes of care. Table 2 shows the results. Definitions of the various variables are provided in table 3.

### *Process*

From 2009 to 2011, the percentage of patients discussed in a preoperative multidisciplinary team increased significantly both in colon (46 to 68%, P<0.01) and rectal cancer surgery (80 to 96%, p<0.01). Moreover, the in-between hospital variation decreased during this time period (Figure 2). There was a significant increase in the implementation of guideline-recommended preoperative MR-imaging for rectal cancer surgery (80 to 83%, p<0.001), as well as an improved standard of pathological reporting of the circumferential resection margins (48% to 80%, p<0.01).

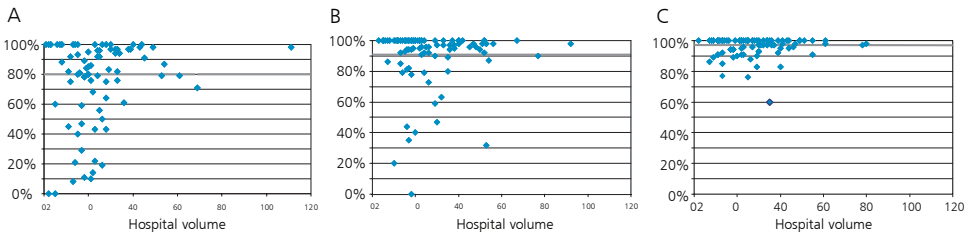



**Figure 2.** Variation between hospitals in the percentage of patients with rectal cancer that was preoperatively discussed in a multidisciplinary team; A) 2009; B) 2010; C) 2011. The gray line represents the average percentage of patients.

### *Outcomes*

Postoperative morbidity, length of hospital stay and postoperative mortality decreased significantly from 2009 to 2011 both for colon and rectal cancer surgery. The incidence of any postoperative complication decreased from 33 to 31% (p<0.01) after colon resections and from 40 to 38% (p<0.01) after rectal resections. The re-intervention rate decreased from 15 to 13% (p<0.001) after colon resections and from 17 to 14% (p<0.01) after rectal resections. Duration of hospital stay regressed with 2 days (both after colon and rectal resections). Postoperative mortality rates (both in-hospital and 30-day mortality) decreased from 5.8 to 4.0% (p=0.012) after colon resections and from 3.8 to 2.7% after rectal resections.

The percentage of patients with a positive circumferential resection margin (CRM) after rectal cancer surgery (≤1 mm distance tumour to CRM) decreased from 14% to 8.5% (p<0.001).

**Table 2.** Results of performance indicators for colorectal cancer care 2009 – 2011.

| | Colon | | | |
|---|---|---|---|---|
| | **2009** | **2010** | **2011** | **P-value** |
| Process | | | | |
| Cases discussed in preoperative MDT | 2286 46% | 3504 56% | 4255 68% | <0.01 |
| Total colonoscopy | 2931 61% | 3816 62% | 4149 67% | <0.01 |
| Preoperative MRI | | | | |
| CRM reported in pathology rapport | | | | |
| >10 lymph nodes in sample | 3623 73% | 4902 78% | 5423 84% | <0.01 |
| Outcomes | | | | |
| All complications | 1595 33% | 2062 33% | 1918 31% | <0.01 |
| Reintervention | 706 15% | 917 15% | 699 13% | <0.01 |
| Anastomotic leakage* | 328 7,5%% | 429 7,80% | 364 6,40% | <0.01 |
| Hospital stay (mean in days) | 13 | 12 | 11 | <0.01 |
| CRM positive margin | | | | |
| 30-day mortality | 223 4,50% | 255 4,10% | 210 3,40% | <0.01 |
| In-hospital mortality | 232 4,70% | 276 4,40% | 230 3,60% | 0.02 |
| In-hospital mortality/30 day mortality | 289 5,80% | 300 4,80% | 256 4,00% | <0.01 |
| Total | 4960 | 6293 | 6263 | |

MDT: Multidisciplinary Team; MRI: Magnetic Resonance Imaging; CRM: Circumferential Resection Margin * only for patients with a primary anastomosis.

**Table 3.** Definitions used in the DSCA.

| Term | Definition |
|---|---|
| Tumour perforation | Preoperative tumour perforation with clinical signs of faecal peritonitis. |
| Abscess | Preoperative abscess formation in the intraperitoneal or extraperitoneal spaces. |
| Ileus | Preoperatieve presence of (partial) mechanical bowel obstruction with symptoms of abdominal cramping, abdominal distention, nausea, vomiting or failure to pass gas or stool. |
| Bleeding | Preoperative tumour related blood loss that requires an intervention (transfusion, urgent operation) or leads to anemia (Hb <7 mmol/L in male patients and <6.5 mmol/L in female patients). |
| Total colonoscopy | Preoperative visualization of the entire colon including the ascending colon by colonoscopy or CT colonography. |
| (Low) anterior resection | Rectosigmoid or rectal resection according to the TME principle with anastomosis of the colon to the intra- or extraperitoneal rectum or anal canal. |
| Multidisciplinary team | A team that consists of all mentioned specialists: a surgeon, an oncologist, a radiologist, a radiotherapist, and a gastroenterologist. |
| Urgent procedure | Non-elective colorectal resection that was required and performed within 24 hours of admission. |
| Anastomotic leakage | Clinically relevant anastomotic leak requiring a radiological or surgical reintervention. |
| Reintervention | An invasive (surgical, radiological or endoscopical) measure to treat a complication (excluding superficial drainage abscess of a wound abscess on the patient ward; introduction of a nasogastric tube; a central venous catheter; or tracheostomy). |
| Positive CRM | A circumferential resection margin of 1 mm or less. |
| Negative outlier | A hospital with a significantly worse (adjusted) outcome than the population average of all hospitals in the registration. |

Hb=haemoglobin. CT=computed tomography. TME=total mesorectal excision. CRM=circumferential resection margin.

| Rectum | | | | | | |
|---|---|---|---|---|---|---|
| 2009 | | 2010 | | 2011 | | P-value |
| 1625 | 80% | 2249 | 91% | 2400 | 96% | <0.01 |
| 1467 | 76% | 1858 | 77% | 2016 | 83% | <0.01 |
| 1625 | 80% | 2016 | 81% | 2129 | 85% | <0.01 |
| 980 | 48% | 1472 | 59% | 2066 | 80% | <0.01 |
| 1182 | 58% | 1520 | 61% | 1700 | 68% | <0.01 |
| 793 | 40% | 1007 | 41% | 945 | 38% | <0.01 |
| 351 | 17% | 435 | 18% | 352 | 14% | <0.01 |
| 98 | 11,50% | 144 | 12,40% | 112 | 9,10% | <0.01 |
| 16 | | 14 | | 14 | | <0.01 |
| 138 | 14% | 175 | 12% | 168 | 8,50% | <0.01 |
| 48 | 2,40% | 48 | 1,90% | 54 | 2,20% | <0.01 |
| 55 | 2,70% | 55 | 2,20% | 64 | 2,50% | 0.663 |
| 77 | 3,80% | 58 | 2,30% | 69 | 2,70% | 0.035 |
| 2035 | | 2484 | | 2494 | | |

## Discussion

This paper reports the key elements of the Dutch Surgical Colorectal Audit that have been crucial for its success. Quality of care regarding guideline compliance and clinical outcomes for colorectal cancer patients in the Netherlands improved significantly.

Numerous international audit projects leading to substantial improvements in quality of care have preceded the DSCA. Many examples of successful clinical audits have been described in detail. [2,3,18-20] Often, the main goal of the audit is to generate valuable information for clinicians to receive feedback on the quality of care.

A unique feature of the DSCA is the use of the audit data to support the effectuation of the national quality assurance policy of the surgical professional association, the ASN. There is a common need for evidence based, professionally supported consensus on what high quality care means in order to set standards of care. Benchmarking hospital performances can support surgeons in determining the minimal requirements of the provided care. On a national level, outliers can be identified. The ASN initiated an independent audit committee to provide consultative advice to hospitals identified as negative outliers in the DSCA. Furthermore, the ASN can use the data for board certification of surgeons, accreditation of hospitals, national and local improvement projects and the provisioning of valid quality information for patients, health care insurers and policy makers.

The engagement of colorectal surgeons to participate was mainly achieved by a strong plea for auditing in national meetings and conferences. The ASN strongly believed that for a valid measurement of quality of care, quality measures should be designed, registered, and interpreted by surgeons themselves. From the onset, the initiative was supported by the majority of Dutch colorectal surgeons, despite the investment in time and costs. One year later, participation became a quality indicator for the health care inspectorate, which ensured an almost 100% participation rate.

The contents of the DSCA dataset as well as the pre-defined process and outcome measures are generally supported by colorectal surgeons in the Netherlands, since they are based on evidence based guidelines and developed by representatives of their own professional organization, who are experts in the field. The leading role of the professional association and its expert members in the design, development and conduct of the audit has important advantages. It produces meaningful and feasible quality information, valid in the face of participating surgeons. This may also have led to the high participation rate among colorectal surgeons and their tremendous efforts to enter high quality data in the registry.
In three years, a trend towards better performance indicator results was objectified. A significant reduction in postoperative morbidity and mortality was observed, as well as a reduced duration of hospital stay. Although promising, the continuation of these trends needs a longer period of registration to be confirmed.
Also, as was presented in Figure 2, the variation in guideline compliance between hospitals was reduced. Although, these improvements may have multifactorial causes, the active and integrated approach of the DSCA has at least resulted in increased awareness of surgeons for quality aspects of their practice and provided insight in areas of improvement. The potential of clinical registries to improve health care outcomes and lowering related costs was recently demonstrated in a study by Larsson et al. [21]
An important feature that supports the audit to function as a quality improvement tool, is the web based data collection system. This system facilitates timely registration of patients and automated feedback of benchmarked performance information on a weekly basis. These features may have contributed to the demonstrable improvements in quality of care presented here.
In recent years there has been an increasing demand for valuable and reliable information on the performance of health care providers from various perspectives. The ASN aimed at developing a system that responds to the exigencies of all major stakeholders in hospital care: patients, clinicians, managers, policy makers and insurance companies. Dutch surgeons have recently agreed to gradually publish publicly their hospital-specific audit results to provide transparency to all parties concerned. For the ASN, an important condition for external transparency is the validity and reliability of the data. This is assured by consistent quality checks on the registered data in the online system and the annual external validation with the National Cancer Registry.

A limitation of the DSCA concept is the administrative burden that is associated with data collection. The measurement of quality of care is complex, and requires the collection of multiple data points from different phases of the care process. The dataset is limited, but still entails detailed information to perform case-mix adjustment and in-depth analysis of observed variation in care processes. Structural data management support for the health care professionals is essential for a sustainable auditing process. Automated retrieval of data from electronic patient files is the logical next step. However, apart from the technical difficulties that have to be solved to extract data from the varying electronic systems in Dutch hospitals, it is essential that synoptic reporting is implemented in the administrative process of hospitals. Links between other databases like the Dutch Pathological Anatomical District Automatized Archives (PALGA) are being established to minimise the registration burden and to automate as much as possible.

In the future, to reach full potential of the audit, information on outcomes of care should be linked to patient reported outcomes and financial information. Feedback to clinicians on patients' satisfaction and quality of life enables them to improve their practice, attitude, facilities and outcomes. Cancer patient organizations in the Netherlands have already committed themselves to collaborate in providing the clinical audits with patient reported outcomes in the near future.

In conclusion, we demonstrated the feasibility of nationwide surgical audit programs, with national coverage and high case-ascertainment, accomplished in a relatively short period of time. The Dutch Surgical Colorectal Audit shows that substantial improvements can be realized within a time period of 3 years. Success factors include: a leading role for medical specialists, external data verification, weekly updated online feedback of benchmarked and meaningful quality information, and embedded in the quality assurance program of the professional association. In the Netherlands, this has been the recipe for the initiation of several other clinical audits, with a generic format consistent with the blueprint of the DSCA.

## References


1. Ozhathil DK, Li Y, Smith JK, et al. Colectomy performance improvement within NSQIP 2005-2008. The Journal of surgical research 2011;171:e9-13.
2. van Leersum NJ, Kolfschoten NE, Klinkenbijl JH, Tollenaar RA, Wouters MW. ['Clinical auditing', a novel tool for quality assessment in surgical oncology]. Ned Tijdschr Geneeskd 2011;155:A4136.
3. Khuri SF, Henderson WG, Daley J, et al. Successful implementation of the Department of Veterans Affairs' National Surgical Quality Improvement Program in the private sector: the Patient Safety in Surgery study. Ann Surg 2008;248:329-36.
4. Pahlman L, Bohe M, Cedermark B, et al. The Swedish rectal cancer registry. Br J Surg 2007;94:1285-92.
5. Dreyer NA, Garner S. Registries for robust evidence. Jama 2009;302:790-1.
6. Boulkedid R, Abdoul H, Loustau M, Sibony O, Alberti C. Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. PLoS One 2011;6:e20476.
7. http://www.msbi.nl/promise
8. http://www.zorgttp.nl
9. Elferink MA, Krijnen P, Wouters MW, et al. Variation in treatment and outcome of patients with rectal cancer by region, hospital type and volume in the Netherlands. Eur J Surg Oncol 2010;36 Suppl 1:S74-82.
10. Elferink MA, Wouters MW, Krijnen P, et al. Disparities in quality of care for colon cancer between hospitals in the Netherlands. Eur J Surg Oncol 2010;36 Suppl 1:S64-73.
11. Schouten LJ, Jager JJ, van den Brandt PA. Quality of cancer registry data: a comparison of data provided by clinicians with those of registration personnel. Br J Cancer 1993;68:974-7.
12. Dutch Institute for Clinical Auditing. Annual Reports 2011. http://www.clinicalaudit.nl.
13. Kolfschoten NE, van Leersum NJ, Gooiker GA, et al. Successful and Safe Introduction of Laparoscopic Colorectal Cancer Surgery in Dutch hospitals. Ann Surg 2012.
14. Kolfschoten NE, Marang van de Mheen PJ, Gooiker GA, et al. Variation in case-mix between hospitals treating colorectal cancer patients in the Netherlands. Eur J Surg Oncol 2011;37:956-63.
15. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40:373-83.
16. Charlson M, Szatrowski TP, Peterson J, Gold J. Validation of a combined comorbidity index. J Clin Epidemiol 1994;47:1245-51.
17. Spiegelhalter DJ. Funnel plots for comparing institutional performance. Statistics in medicine 2005;24:1185-202.
18. Cornish JA, Tekkis PP, Tan E, Tilney HS, Thompson MR, Smith JJ. The national bowel cancer audit project: the impact of organisational structure on outcome in operative bowel cancer within the United Kingdom. Surg Oncol 2011;20:e72-7.
19. Jung B, Pahlman L, Johansson R, Nilsson E. Rectal cancer treatment and outcome in the elderly: an audit based on the Swedish Rectal Cancer Registry 1995-2004. BMC Cancer 2009;9:68.
20. van Gijn W, van den Broek CB, Mroczkowski P, et al. The EURECCA project: Data items scored by European colorectal cancer audit registries. Eur J Surg Oncol 2012;38:467-71.
21. Larsson S, Lawyer P, Garellick G, Lindahl B, Lundstrom M. Use of 13 disease registries in 5 countries demonstrates the potential to use outcome data to improve health care's value. Health Aff (Millwood) 2012;31:220-7.

# Part 2

The use of data from clinical audits to evaluate and improve national practice

Chapter 4

# Variation in case-mix between hospitals treating colorectal cancer patients in the Netherlands

N. E. Kolfschoten[1], P.J. Marang-van de Mheen[2], G.A. Gooiker[1], E.H. Eddes[3], J. Kievit[1,2], R.A.E.M Tollenaar[1], M.W.J.M. Wouters[1,4], on behalf of the Dutch Surgical Colorectal Audit group*

[1] Leiden University Medical Centre, Dept of Surgery
[2] Leiden University Medical Center, Dept of Medical Decision Making
[3] Deventer Hospital, Dept of Surgery
[4] Netherlands Cancer Institute-Antoni van Leeuwenhoek Hospital, Dept of Surgical Oncology
* the Dutch Surgical Colorectal Audit group: W.A. Bemelman, O.R.C. Busch, R.M. van Dam, E. van der Harst, M.L.E.A. Jansen-Landheer, Th.M. Karsten, J.H.J.M. van Krieken, W.G.T. Kuijpers, V.E. Lemmens, E.R. Manusama, W.J.H.J Meijerink, H.J.T. Rutten, T. Wiggers, C.J.H. van de Velde.

Eur J Surg Oncol. 2011 Nov;37 (11):956-63

## Abstract

### Aims

The purpose of this study was to determine how expected mortality based on case-mix varies between colorectal cancer patients treated in non-teaching, teaching and university hospitals, or high, intermediate and low-volume hospitals in the Netherlands.

### Material and methods

We used the database of the Dutch Surgical Colorectal Audit 2010. Factors predicting mortality after colon and rectum carcinoma resections were identified using logistic regression models. Using these models, expected mortality was calculated for each patient.

### Results

8580 patients treated in 90 hospitals were included in the analysis. For colon carcinoma, hospitals' expected mortality ranged from 1.5 to 14%. Average expected mortality was lower in patients treated in high-volume hospitals than in low-volume hospitals (5.0 vs. 4.3%, $p<0.05$). For rectum carcinoma, hospitals expected mortality varied from 0.5 to 7.5%. Average expected mortality was higher in patients treated in non-teaching and teaching hospitals than in university hospitals (2.7 and 2.3vs. 1.3%, $p<0.01$). Furthermore, rectum carcinoma patients treated in high-volume hospitals had a higher expected mortality than patients treated in low-volume hospitals (2.6 vs. 2.2% $p<0.05$). We found no differences in risk-adjusted mortality.

### Conclusions

High-risk patients are not evenly distributed between hospitals. Using the expected mortality as an integrated measure for case-mix can help to gain insight in where high-risk patients go. The large variation in expected mortality between individual hospitals, hospital types and volume groups underlines the need for risk-adjustment when comparing hospital performances.

## Introduction

Colorectal carcinoma is the second most common cancer related cause of death in western countries, and incidence is rising. Surgical resection is still the cornerstone of treatment, though associated with a considerable peri-operative risk. It is also known that outcome after colorectal resections, especially postoperative mortality varies between hospitals[1]. Although differences in mortality rates are often studied with a focus on the inverse relationship with procedural volume[2, 3] similar variations have been reported between different types of hospitals, such as teaching and non-teaching hospitals[4-6].
Part of these variations can be explained by differences in case-mix (patient and tumour characteristics). Therefore case-mix correction is considered to be essential for reliable and valid outcome comparisons[7]. In the last decade, several countries have started nation-wide colorectal audit programs, which are coordinated on a European level by the EURECCA project of the European Society of Surgical Oncology. Considering this increasing interest in outcome measurements, information on differences in case-mix between (different types of) hospitals becomes more and more important.

A commonly used method for case-mix correction is the Observed/Expected mortality rate[8, 9], in which the observed mortality is the number of deaths in a hospital or group of hospitals, and the expected mortality is the sum of the patients' estimated probabilities for mortality, based on their case-mix factors. Previous studies have described variation in the distribution of specific case-mix factors between teaching and non-teaching hospitals, and high- and low-volume hospitals, with varying results[4-6]. However, to date, variations in expected mortality rates between different types of hospitals have not been studied in detail. More insight in the distribution of expected mortality can assist health care providers in recognizing high-risk patients treated in their hospital and focus quality improvement efforts on these patients.

Following the initiatives in Scandinavian countries and in the United Kingdom, the Dutch Surgical Colorectal Audit (DSCA) was initiated in 2009[10]. The DSCA is a nation wide, web based, interactive database in which detailed patient, tumour, diagnostic, procedural and outcome data are registered of patients who undergo a resection of a primary colorectal carcinoma in the Netherlands. This database gave us the opportunity to further investigate differences in expected mortality based on differences in case-mix between patients treated in different types of hospitals in the Netherlands.
The purpose of this study was to determine

1) which case-mix factors are predictors for postoperative mortality after the resection of colon and rectum carcinomas in the Netherlands
2) how expected mortality based on case-mix factors differs between individual as well as different types of hospitals: non-teaching, teaching and university hospitals and high-, intermediate and low-volume hospitals

3) which patient and tumour characteristics are responsible for these differences in expected mortality

## Patients and methods

### Patients

About 93% of all patients who underwent a resection of a primary colorectal carcinoma in the Netherlands from 1st of January 2010 until 31st of December 2010 were included in the DSCA [www.dsca.nl] on March 15th 2011. All Dutch hospitals participated. Data entry was web-based in a highly secured database. All participating hospitals appointed a surgeon who was responsible for the data-entry. Weekly feedback information on number of patients entered, number of registered patient files, and overall completeness of patient files combined with benchmarked performance indicators were placed online in a secured internet page, accessible for the hospital only. To secure data quality, hospitals received data quality reports throughout the year, summing all patients with inconsistent or unusual data combinations, identified by a total of 70 queries. The responsible surgeon was asked to verify the data of these patients and to correct the data when indicated.
Patients treated for a recurrence of a colorectal carcinoma or multiple synchronous colorectal tumours were excluded. Records that did not contain a tumour location, a date of surgery or did not specify if the patient had left the hospital alive were excluded.

### Definitions

Case-mix factors registered in the DSCA were age, gender, comorbidity (Charlson-comorbidity index), previous abdominal surgery, Body Mass Index, American Society of Anaesthesiologists (ASA) classification, TNM stage, neoadjuvant (chemo) radiation therapy, preoperative tumour complications, urgency of the resection, type of resection, additional resection for tumour invasion and/or metastasis.

Postoperative mortality was defined as mortality during the same hospital admission of the surgical resection or within 30 days after resection.
Hospitals were categorized as low-, medium- or high-volume, according to the criteria of the Dutch Cancer Society[1]. For colon carcinoma the categories are: less than 50, 50 to 100, and 100 or more patients treated surgically in the year 2010. For rectum carcinoma the categories are: less than 50 patients and 50 or more patients treated in 2010.

In the Dutch healthcare system, the 8 university hospitals function as tertiary referral centres for high-complex, low-volume care, (including locally advanced rectum and colon carcinoma resections and synchronous metastasectomies) and therefore treat a selected patient group. Of all university hospitals 5 were classified as low-volume hospital en 3 as medium-volume hospitals for colon carcinoma, all except 1 were classified as low- volume hospital for

rectum carcinoma. Each university hospital is affiliated with several teaching hospitals, and collaborates in training surgical residents. Most teaching hospitals function as referral centres for medium-complex high-volume surgical care such as locally advanced rectum carcinoma resections. Of all 45 teaching hospitals, 24 were classified as medium-volume hospital and 18 as high-volume hospitals for colon carcinoma, 11 were classified as high-volume hospital for rectum carcinoma. Of all 37 non-teaching hospitals, 14 were classified as medium-volume hospital for colon carcinoma, and all were classified as low-volume hospitals for rectum carcinoma.

## Analysis

As treatment and outcome of patients with a colon and rectum carcinoma are different, these groups were analysed separately. All case-mix factors that were predictive for postoperative mortality in univariate logistic regression ($p<0.10$) were included in the multivariate model. Using the coefficients from the multivariate model, an expected mortality was calculated for each patient. The average expected mortality in a group of patients formed the expected mortality rate of that group. The standardized mortality ratio (SMR) was calculated as the quotient of the observed mortality rate, and the expected mortality rate for a group of patients. Confidence intervals were calculated using the standard formulas[11].

For the analysis on hospital level, we calculated the average expected mortality, observed mortality and SMR of all patients treated in a hospital. Hospitals were evaluated individually and also categorized on their teaching status and procedural volume. Differences in expected mortality, observed mortality and standardized mortality rates between different types of hospitals and high and low-volume hospitals were investigated using the one-way ANOVA test. As the total number of hospitals in the Netherlands, and therewith the number of hospitals in each group may be too small to detect any differences between the groups investigated on a hospital level, analyses were repeated on the patient level.

On a patient level, we investigated the differences in expected mortality, observed mortality and standardized mortality rates between all patients treated in university, teaching and non-teaching hospitals, and in high and low-volume hospitals. When differences in expected mortality rates were detected, we further explored which case-mix factors were responsible for the differences in expected mortality rates by repeating the one-way ANOVA test, stratified for each case-mix factor. When the difference in expected mortality disappeared, the case-mix factor was considered to be (partly) responsible. Statistical significance was defined as $p<0.05$. All statistics were performed in PASW Statistics, Rel. 18.0.2009. Chicago: SPSS inc. and Microsoft Excel.

## Results

### Patients

At March 15th 2011, 90 hospitals (8 university hospitals, 45 teaching hospitals and 37 non-teaching hospitals) registered a total of 8,835 eligible patients with a date of surgery between January 1 and December 31 2010 in the DSCA. After exclusion of patients with multiple synchronous tumours (253 patients) a total of 8,580 patients were included in the analysis, 6,161 patients with colon cancer and 2,419 with rectum cancer. Average procedural volume was 95 patients per hospital (range 14 – 231). The distribution of case-mix factors in the population of patients with colon and rectum carcinoma is specified in table 1. Average observed mortality was 4.5% for colon carcinoma, and 2.3% for rectum carcinoma.

### Predictors for postoperative mortality

Age, gender, ASA-classification, Charlson comorbidity index, and urgency of the resection were individual significant predictors for mortality after colon carcinoma resections (Table 2). The C-statistic (95% confidence interval) of the model predicting postoperative mortality after a resection for colon carcinoma was 0.81 (0.78-0.83).
For RC, age, gender, and ASA-classification were significant predictors for postoperative mortality (Table 2). The C-statistic (95% confidence interval) of the model for rectum carcinoma was 0.87 (0.82-0.91).

### Differences between hospitals

*Colon carcinoma*
For colon carcinoma, hospitals' expected mortality ranged from 1.5 to almost 14% [Figure 1a]. The outlier hospital which had an expected mortality of 14% treated more elderly patients and patients with preoperative tumour complications than average. Hospitals' observed mortality varied from 0 to 15.9%, SMR's ranged from 0 to 4.04 (data not shown). We found no differences in expected mortality between non-teaching, teaching and university hospitals or low, medium or high-volume hospitals (table 3).

*Rectum carcinoma*
For rectum carcinoma, expected mortality ranged from 0.5 to 7.5% [Figure 1b]. Hospitals' observed mortality varied from 0 to 16.7%, SMR's ranged from 0 to 4.87 (data not shown). On the hospital level, non-teaching hospitals had a higher expected mortality than university hospitals (2.44 vs. 1.31%, $p<0.05$), there were no differences in expected mortality between high and low-volume hospitals. We found no differences in unadjusted or standardized mortality rates between the different types of hospitals and volume groups on the hospital level [Table 3].

**Table 1:** Case-mix factors for colon and rectum cancer patients in the DSCA.

| | | Colon | | Rectum | |
|---|---|---|---|---|---|
| | | N | % | N | % |
| N (patients) | | 6161 | 72 | 2419 | 28 |
| Age (years) | <70 | 2755 | 45 | 1443 | 59 |
| | 70-80 | 2062 | 34 | 715 | 30 |
| | 80+ | 1337 | 22 | 270 | 11 |
| Gender | Female | 2939 | 48 | 939 | 39 |
| Previous abdominal surgery | Yes | 2137 | 35 | 702 | 29 |
| ASA classification | I-II | 4511 | 73 | 2994 | 83 |
| | III | 1457 | 24 | 374 | 16 |
| | IV-V | 152 | 3 | 26 | 1 |
| | Missing | 41 | 1 | 15 | 1 |
| Charlson score | 2+ | 1231 | 20 | 334 | 14 |
| BMI category | <25 | 2212 | 36 | 945 | 39 |
| | 25-30 | 1870 | 30 | 801 | 33 |
| | >30 | 820 | 13 | 347 | 14 |
| | Missing | 1259 | 20 | 326 | 14 |
| TNMstage | I | 1077 | 18 | 712 | 30 |
| | II | 2205 | 36 | 646 | 27 |
| | III | 1830 | 30 | 763 | 32 |
| | IV | 921 | 15 | 240 | 10 |
| | X | 84 | 1 | 30 | 1 |
| Preop radiotherapie | 5x5 gy | - | - | 1132 | 47 |
| | >60 gy | - | - | 156 | 6 |
| | Chemoradiation | - | - | 674 | 28 |
| Preop tumourcomplication | Fistula | 139 | 2 | 15 | 1 |
| | Obstruction | 747 | 12 | 57 | 2 |
| | Else | 484 | 8 | 88 | 4 |
| | Missing | 209 | 3 | 88 | 4 |
| Urgency of resection | Urgent | 728 | 12 | 33 | 1 |
| | Acute | 444 | 7 | 20 | 1 |
| Surgical procedure | Right hemicolectomy | 2881 | 47 | - | - |
| | Left hemicolectomy (+transversum) | 796 | 13 | - | - |
| | Sigmoid | 1414 | 23 | - | - |
| | (L)AR | 834 | 14 | 1613 | 67 |
| | APR | - | - | 722 | 30 |
| | Else | 236 | 4 | 84 | 3 |
| Laparoscopic resection | Laparoscopic (+conversion) | 2296 | 37 | 888 | 37 |
| | Conversion to open | 366 | 6 | 117 | 5 |
| Extended resection | Locally advanced tumour | 598 | 10 | 172 | 7 |
| | Metastasectomia | 166 | 3 | 72 | 3 |

ASA = American Society of Anesthesiologists classification
BMI = Body Mass Index
TNM = Tumour Node Metastasis system
(L)AR = (Low) Anterior Resection
APR = abdomino perineale resectie

**Table 2:** casemix-factors included in the multivariate logistic regression model for mortality after colon and rectum carcinoma resections.

| | | Colon | | Rectum | |
|---|---|---|---|---|---|
| | | OR | 95% C.I | OR | 95% C.I |
| Age | <70 | 1.00 | | 1.00 | |
| | 70-79 | 2.06 | 1.43-2.99 | 3.15 | 1.51-6.58 |
| | 80+ | 5.10 | 3.56-7.29 | 3.55 | 1.54-8.22 |
| Gender | Female | 0.68 | 0.52-0.88 | 0.44 | 0.22-0.88 |
| Charlson score | 2+ | 1.38 | 1.04-1.83 | 1.81 | 0.95-3.46 |
| ASA score | I-II | 1.00 | | 1.00 | |
| | III | 2.44 | 1.83-3.25 | 4.79 | 2.50-9.16 |
| | IV-V | 7.47 | 4.72-11.82 | 6.24 | 1.72-22.61 |
| | Missing | 1.13 | 0.26-5.02 | 8.46 | 0.99-72.05 |
| TNM Stage | I | 1.00 | | 1.00 | |
| | II | 0.90 | 0.60-1.35 | 0.89 | 0.42-1.90 |
| | III | 1.00 | 0.65-1.52 | 0.71 | 0.33-1.53 |
| | IV | 1.53 | 0.97-2.41 | 1.39 | 0.53-3.68 |
| | X | 1.94 | 0.85-4.45 | 1.24 | 0.29-5.26 |
| Radiotherapy | None | | | 1.00 | |
| | 5x5 gy | | | 0.55 | 0.29-1.05 |
| | >60 Gy | | | 0.32 | 0.07-1.46 |
| | Chemoradiation | | | 0.32 | 0.11-1.00 |
| Tumourcomplication | None | 1.00 | | 1.00 | |
| | Perforation | 1.82 | 0.89-3.73 | - | |
| | Obstruction | 1.25 | 0.78-1.99 | 1.43 | 0.30-6.94 |
| | Else | 1.21 | 0.78-1.88 | 2.45 | 0.85-7.05 |
| | Missing | 0.96 | 0.45-2.05 | 2.56 | 0.88-7.49 |
| Urgency | Elective | 1.00 | | 1.00 | |
| | Urgent | 1.61 | 1.06-2.46 | 2.16 | 0.47-9.92 |
| | Acute | 2.13 | 1.26-3.59 | 2.74 | 0.38-19.69 |
| Procedure | Right hemicolectomy | 1.00 | | | |
| | Left hemicolectomy | 1.37 | 0.97-1.95 | | |
| | Sigmoid resection | 0.53 | 0.39-0.73 | | |
| | Else | 1.23 | 0.69-2.20 | | |
| Extended resection | Local | 1.45 | 0.98 | | |

CI = confidence interval
OR = Odds ratio
ASA = American Society of Anesthesiologists classificaion
TNM = Tumour Node Metastasis system
Bold printed numbers are significant odds ratios (p<0.05)

## Differences between patients treated in different types of hospitals and volume groups.

### *Colon Carcinoma*

Expected mortality varied between individual patients from 0.4 to 70.8%. Average expected mortality of patients treated in non-teaching, teaching and university hospitals and in patients treated in low-, medium- and high-volume hospitals are presented in table 4.

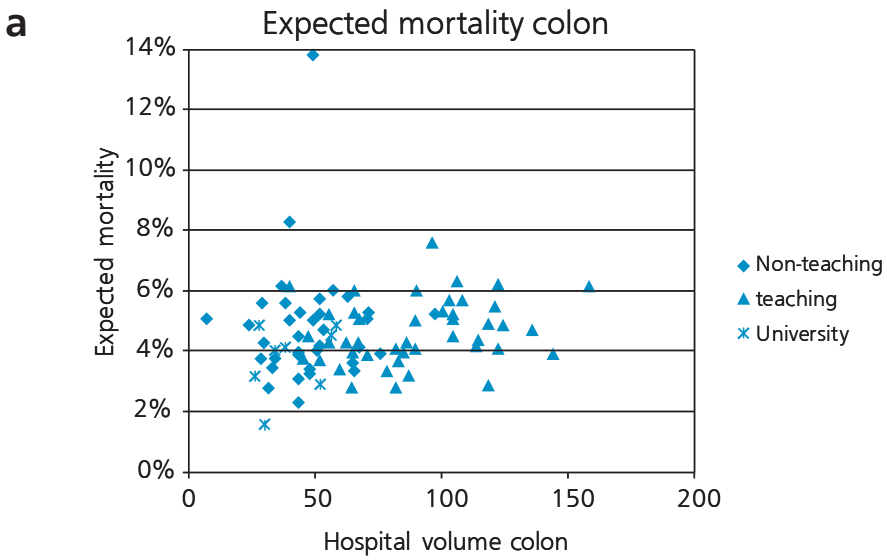
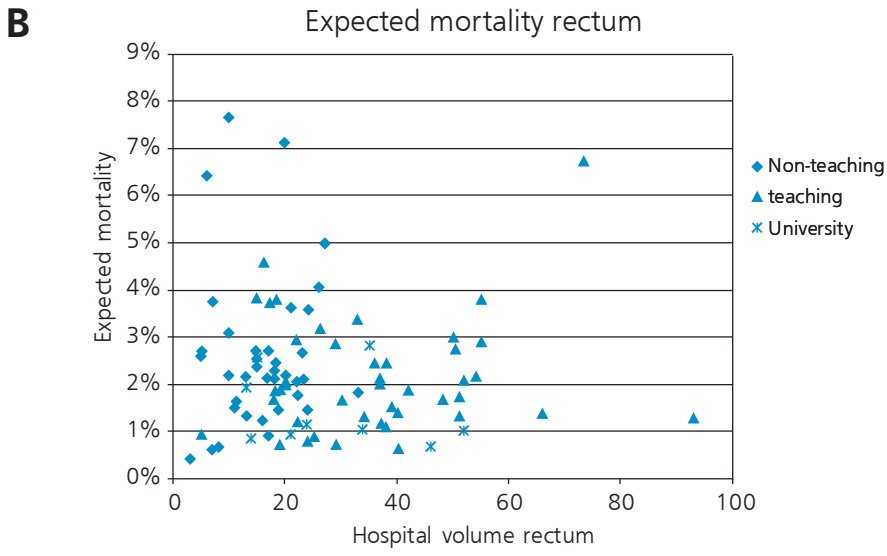



**Figure 1a**: Expected mortality colon carcinoma




**Figuur 1b:** Expected mortality rectum carcinoma

Average expected mortality did not differ between patients treated in university, teaching or non-teaching hospitals, though the observed mortality was higher in non-teaching hospitals. When outcome was adjusted for case-mix, we found no significant differences.
Expected mortality in patients treated in low-volume hospitals was significantly higher than in high-volume hospitals (5.0 vs. 4.3%, $p<0.05$). Despite the differences in expected

**Table 3:** Differences in Expected mortality on hospital level

| | N | Mean **expected** mortality (%) | CI 95% | Mean **observed** mortality (%) | CI 95% | SMR | CI95% |
|---|---|---|---|---|---|---|---|
| **Colon** | | | | | | | |
| Non-teaching | 37 | 4.76 | 4.41-5.11 | 5.07 | 3.52-6.61 | 1.22 | 0.88-1.57 |
| Teaching | 45 | 4.49 | 3.97-5.01 | 4.64 | 3.59-5.69 | 0.94 | 0.79-1.09 |
| University | 8 | 4.92 | 3.37-6.48 | 4.14 | 3.07-5.21 | 0.91 | 0.07-1.75 |
| <50 | 31 | 5.04 | 4.27-5.81 | 5.07 | 3.52-6.61 | 1.13 | 0.74-1.52 |
| 50-100 | 41 | 4.50 | 4.16-4.84 | 4.64 | 3.59-5.69 | 1.03 | 0.81-1.25 |
| 100+ | 18 | 4.27 | 3.86-4.68 | 4.14 | 3.07-5.21 | 0.97 | 0.71-1.23 |
| **Rectum** | | | | | | | |
| Non-teaching | 37 | 2.44* | 1.05-3.83 | 2.65 | 2.10-3.20 | 0.79 | 0.39-1.20 |
| Teaching | 45 | 2.17 | 1.44-2.90 | 2.22 | 1.86-2.58 | 0.97 | 0.62-1.32 |
| University | 8 | 1.32* | 0.28-2.92 | 1.33 | 0.74-1.92 | 1.06 | 0.43-2.55 |
| <50 | 78 | 2.18 | 1.43-2.92 | 2.28 | 1.97-2.59 | 0.89 | 0.60-1.17 |
| 50-100 | 12 | 2.39 | 0.74-4.03 | 2.53 | 1.53-3.53 | 1.03 | 0.38-1.69 |

*Non-Teaching vs University hospital: p<0.01
SMR: Standardized Mortality Ratio = Observed mortality/Expected mortality

**Table 4:** Differences in Expected mortality on patient level

| | N | Mean **expected** mortality (%) | 95% CI | Mean **observed** Mortality (%) | 95% CI | SMR | 95% CI |
|---|---|---|---|---|---|---|---|
| **Colon** | | | | | | | |
| Non teaching | 1760 | 4.8 | 4.4-5.1 | 5.5* | 4.5-6.7 | 1.18 | 0.97-1.43 |
| Teaching | 4080 | 4.4 | 4.2-4.6 | 4.1 | 3.5-4.7 | 0.93 | 0.80-1.08 |
| University | 321 | 4.8 | 4.0-5.6 | 3.7* | 1.7-5.8 | 0.78 | 0.42-1.32 |
| <50 | 1149 | 5.0@ | 4.6-5.4 | 5.3 | 4.0-6.6 | 1.06 | 0.82-1.36 |
| 51-100 | 2830 | 4.5 | 4.3-4.7 | 4.5 | 3.7-5.3 | 1.00 | 0.84-1.19 |
| 100+ | 2131 | 4.3@ | 4.0-4.5 | 4.1 | 3.3-4.9 | 0.96 | 0.78-1.18 |
| **Rectum** | | | | | | | |
| Non teaching | 589 | 2.7* | 2.3-3.1 | 2.6 | 1.3-3.8 | 0.95 | 0.57-1.57 |
| Teaching | 1591 | 2.3$ | 2.0-2.5 | 2.3 | 1.6-3.1 | 1.03 | 0.74-1.42 |
| University | 239 | 1.3*$ | 0.9-1.7 | 1.3 | 0.0-2.7 | 0.95 | 0.31-2.95 |
| <50 | 1717 | 2.2# | 1.9-2.4 | 2.2 | 1.5-2.8 | 1.00 | 0.72-1.38 |
| 50+ | 702 | 2.6# | 2.1-3.0 | 2.6 | 1.4-3.7 | 1.00 | 0.63-1.59 |

*Non-Teaching vs University hospital: p<0.01
$ Teaching vs University hospital: p<0.01
@ <50 vs 100+ p<0.01
# <50 vs 50+ p<0.05
SMR: Standardized Mortality Ratio = Observed mortality/Expected mortality

mortality, there were no differences in standardized, nor in unadjusted mortality rates between patients treated in different volume groups.

The differences in expected mortality between different groups of patients were further explored. Differences in expected mortality between high, medium and low-volume

hospitals disappeared when patients were stratified for Charlson score, ASA- classification and urgency of the resection, meaning that differences in expected mortality were mostly explained by differences in comorbidity score and urgency of the resection (data not shown).

*Rectum Carcinoma*

Expected mortality varied between individual patients from 0.0 to 52.5%. The expected mortality was higher in patients treated in non-teaching and teaching hospitals than in patients treated in university hospitals. (2.7 and 2.3 vs. 1.3% respectively, $p<0.01$). Also, patients treated in high-volume hospitals had a lower expected mortality than patients treated in low-volume hospitals (1.6 vs. 2.2% $p<0.05$). There were no differences in observed or SMR between different types of hospitals. Despite the differences in expected mortality, there were no differences in standardized, nor in unadjusted mortality rates between patients treated in different volume groups for rectum carcinoma. None of the individual case-mix factors could explain the differences in expected mortality between patients treated in university, teaching and non-teaching hospitals or high and low-volume hospitals (data not shown).

## Discussion

This study confirms the essential role of case-mix adjustments in clinical auditing for colorectal cancer. For individual hospitals, expected mortalities varied between 1.5 and 14% for colon cancer and from 0.5 and 7.5% for rectum cancer resections. Patients treated colon cancer in low-volume hospitals had a higher expected mortality than patients treated in high-volume hospitals, due to an unfavourable case-mix. These differences in expected mortality were mostly explained by low-volume hospitals treating patients with more comorbid diseases, with a higher ASA classification and more often in an urgent or acute setting. However, for rectal cancer resections expected mortality was higher in patients treated in high-volume hospitals and patients treated in non-teaching and teaching hospitals had a higher expected mortality than those treated in university hospitals. Despite these differences in case-mix, we found no differences in risk-adjusted outcome between patients treated in different types of hospitals.

One of the merits of the growing number of national audit programs is the United States and Europe is the increasing availability of prospectively collected case-mix information, which is needed to investigate the risk for postoperative mortality after colorectal cancer operations[12]. For this study we used the 2010 data from the Dutch Surgical Colorectal Audit, which contains an extensive set of case-mix variables. Based on this dataset we identified a risk profile for colon and rectum carcinoma resections separately, consistent with previous risk-adjustment models developed on different data-sets[13, 14].

Previous studies have described variations in case-mix distribution between teaching and non-teaching hospitals, and high- and low-volume hospitals with slightly different results. Zhang[5] reported that elderly patients with a low income, a high comorbidity index or tumour stage, were less likely to be treated in high-volume hospitals. Friese[4] investigated surgical oncology patients and found a younger population with a high comorbidity index and more patients with disseminated disease in the major teaching hospitals and high-volume hospitals. Khuri[6] showed that teaching hospitals serve a younger population, but have more emergency cases in general surgery. As these studies originate from different health care systems, it is difficult to assess their relevance for the situation in European countries. Even within Europe case-mix distribution between different types of hospitals can vary by country. To get more insight in the distribution of case-mix between different hospitals or countries, an integrated measure for case-mix is needed.

The present study introduces a commonly used measure for risk-adjustment: the expected mortality, as a measure for the integrated effect of case-mix factors, giving an insight in how high-risk patients are distributed in our healthcare system. We showed that there is considerable variation in risk-profile of patients treated in different hospitals in the Netherlands: in non-teaching, teaching and university hospitals, and in high and low-volume hospitals. Moreover, we showed marked differences in expected mortality between the individual hospitals, varying from 1.5 to 14% for colon cancer patients and from 0.5 and 7.5% for rectal cancer patients (Figure 1a and 1b).

As in many countries, the political focus in the Netherlands is increasingly on outcome measurement and concentration of care in specialized high-volume centres. However, we found no differences in observed or standardized mortality between patients treated in high, intermediate or low-volume hospitals in our study. We did find a higher unadjusted mortality rate among patients treated in non-teaching hospitals, but, after adjustment for case-mix variations, the SMR was not higher than for patients treated in teaching or university hospitals. Also, considerable variation in observed and standardized mortality between individual hospitals was found, irrespective of teaching status or procedural volume. Therefore, concentration of care should preferably be guided by outcomes of care of individual hospitals as measured in nation-wide clinical audits (outcome-based referral)[15, 16]. The variations in outcome between individual hospitals will be further investigated in a different study. In this context, case-mix adjustments based on hospitals' expected mortality are essential for reliable outcome comparisons.

## Limitations

The results of this study have to be interpreted in the light of several limitations. First, analyses were based on hospital specific data from one registration year, 2010. Differences in expected mortality between individual hospitals may change slightly due to random statistical variation when more registration years are evaluated, although we expect the observed

differences between types of hospitals to remain similar. Moreover, the number of patients treated in university hospitals, and the number of university hospitals in the Netherlands was relatively low. The expected mortality found for rectal resections was lower in university hospitals, than in teaching and non-teaching hospitals. Although it is a common perception that university hospitals perform more complex and therefore more high-risk procedures, procedural complexity had only a very small effect on the expected mortality. Age, gender and ASA-classification were the most important case-mix factors influencing the expected mortality for rectal resections. As university hospitals in the Netherlands are designated tertiary referral centres, preferably performing only high complex operative procedures, the majority of colorectal resections are performed in teaching and non-teaching hospitals. In an elective setting, young patients with more advanced cancer may more often choose to be treated in, or are more often referred to university hospitals, whereas elderly patients with a high comorbidity score will more often go to the nearest (non-)teaching hospital. Even though we found that none of the individual case-mix factors was responsible for the differences in expected mortality for rectum carcinoma, the combination of factors could explain the low expected mortality for university hospitals found in our study.

## Conclusion

The present study shows that high-risk colorectal cancer patients in the Netherlands are mostly characterised by an older age, comorbidity and an indication for non-elective resection. High-risk patients are not evenly distributed between individual and groups of hospitals. Using expected mortality as an integrated measure for their case-mix gives insight in the hospitals where high-risk patients go, which is important for the targeting of quality improvement programs. The large variation in expected mortality between individual hospitals, hospital types and volume groups underlines the need for risk-adjustment when comparing hospital performances, or even when comparing performances on a more international level as is intended in the EURECCA project of the European Society for Surgical Oncology[12]. A complete, population based dataset, and a uniform registration and methodology for case-mix adjustments are imperative for valid comparison of outcome of care between countries or healthcare systems[17].

### Acknowledgements

The authors would like to thank all surgeons, registrars, physician assistants and administrative nurses that registered all the patients in the DSCA, as well as the Dutch Surgical Colorectal Audit group and the methodological board for their advice.

## References


1. Elferink MA, Wouters MW, Krijnen P, et al. Disparities in quality of care for colon cancer between hospitals in the Netherlands. Eur J Surg Oncol. 2010;36 Suppl 1:S64-73.
2. van Gijn W, Gooiker GA, Wouters MW, Post PN, Tollenaar RA, van de Velde CJ. Volume and outcome in colorectal cancer surgery. Eur J Surg Oncol. 2010;36 Suppl 1:S55-63.
3. Borowski DW, Kelly SB, Bradburn DM, Wilson RG, Gunn A, Ratcliffe AA. Impact of surgeon volume and specialization on short-term outcomes in colorectal cancer surgery. Br J Surg. 2007;94:880-889.
4. Friese CR, Earle CC, Silber JH, Aiken LH. Hospital characteristics, clinical severity, and outcomes for surgical oncology patients. Surgery. 2010;147:602-609.
5. Zhang W, Ayanian JZ, Zaslavsky AM. Patient characteristics and hospital quality for colorectal cancer surgery. Int J Qual Health Care. 2007;19:11-20.
6. Khuri SF, Najjar SF, Daley J, et al. Comparison of surgical outcomes between teaching and nonteaching hospitals in the Department of Veterans Affairs. Ann Surg. 2001;234:370-382; discussion 382-373.
7. Wouters M. Outcome-based referral to improve quality of care in upper GI surgery. J Surg Oncol. 2008;98:307-308.
8. Daley J, Khuri SF, Henderson W, et al. Risk adjustment of the postoperative morbidity rate for the comparative assessment of the quality of surgical care: results of the National Veterans Affairs Surgical Risk Study. J Am Coll Surg. 1997;185:328-340.
9. Khuri SF, Daley J, Henderson W, et al. Risk adjustment of the postoperative mortality rate for the comparative assessment of the quality of surgical care: results of the National Veterans Affairs Surgical Risk Study. J Am Coll Surg. 1997;185:315-327.
10. Wouters MW, Jansen-Landheer ML, van de Velde CJ. The Quality of Cancer Care initiative in the Netherlands. Eur J Surg Oncol. 2010;36 Suppl 1:S3-S13.
11. Clayton DH, M. Statistical Models in Epidemiology. Oxford University Press, Oxford. 1993.
12. van Gijn W, Wouters MW, Peeters KC, van de Velde CJ. Nationwide outcome registrations to improve quality of care in rectal surgery. An initiative of the European Society of Surgical Oncology. J Surg Oncol. 2009;99:491-496.
13. Tekkis PP, Prytherch DR, Kocher HM, et al. Development of a dedicated risk-adjustment scoring system for colorectal surgery (colorectal POSSUM). Br J Surg. 2004;91:1174-1182.
14. Cohen ME, Bilimoria KY, Ko CY, Hall BL. Development of an American College of Surgeons National Surgery Quality Improvement Program: morbidity and mortality risk calculator for colorectal surgery. J Am Coll Surg. 2009;208:1009-1016.
15. Wouters MW, Karim-Kos HE, le Cessie S, et al. Centralization of esophageal cancer surgery: does it improve clinical outcome? Ann Surg Oncol. 2009;16:1789-1798.
16. Sonnenday CJ, Birkmeyer JD. A tale of two provinces: regionalization of pancreatic surgery in Ontario and Quebec. Ann Surg Oncol. 2010;17:2535-2536.
17. Marang-van de Mheen PJ, Hollander EJF, Kievit J, Effects of study methodology on adverse outcome occurrence and mortality. Int J Qual Health Care. 2007;19:399-406.

Chapter 5

# Non-elective colon cancer resections in elderly patients, results from the Dutch Surgical colorectal audit

N.E. Kolfschoten, MD[1], M.W.J.M. Wouters, Surgeon[2], E.H. Eddes PhD, Surgeon[3], J. Kievit, Professor in Medical Decision making, Surgeon[1,4], R.A.E.M Tollenaar, Professor in Surgical Oncology, Surgeon[1], P.J. Marang-van de Mheen, PhD[4], on behalf of the Dutch Surgical Colorectal Audit group*

[1] Leiden University Medical Centre, Dept of Surgery
[2] Dutch Cancer Institute, Amsterdam, Dept of Surgery
[3] Deventer Hospital, Dept of Surgery
[4] Leiden University Medical Center, Dept of Medical Decision Making
* Prof. dr. W.A. Bemelman, Dr. O.R.C. Busch, Drs. R.M. van Dam, Dr. E. van der Harst, Dr. M.L.E.A. Jansen-Landheer, Dr. Th.M. Karsten, Prof. dr. J.H.J.M. van Krieken, Drs. W.G.T. Kuijpers, Dr. V.E. Lemmens, Dr. E.R. Manusama, Dr. W.J.H.J Meijerink, Dr. H.J.T. Rutten, Prof. dr. C.J.H. van de Velde, Prof. dr. T. Wiggers

*Digestive Surgery. 2012;29(5):412-9.*
*Authorized reprint: Nederlands Tijdschrift voor Geneeskunde. 2013;157:A6426*

## Abstract

### Aims
The aim of the study was to assess which factors contribute to postoperative mortality, especially in elderly patients who undergo emergency colon cancer resections, using a nation-wide, population based database.

### Methods
6,161 patients (1172 non-elective) who underwent a colon cancer resection in 2010 in the Netherlands were included. Risk factors for postoperative mortality were investigated using a multivariate logistic regression model, for different age groups, elective and non-elective patients separately.

### Results
For both elective and non-elective patients, mortality risk increased with increasing age. For non-elective elderly patients (80+ years), each additional risk factor increased the mortality risk. For a non-elective patient of 80+ years with an ASA score of III+ and a left hemicolectomy or extended resection, postoperative mortality rate was 41% compared with 7% in patients without additional risk factors.

### Conclusions
For elderly patients with two or more additional risk factors, a non-elective resection should be considered a high-risk procedure with a mortality risk of up to 41%. The results of this study could be used to adequately inform patient and family and should have consequences for composing an operative team.

## Introduction

About 15 to 30% of all patients with a colon carcinoma present with symptoms of obstruction, perforation or bleeding. Most of these patients will undergo a resection in an emergency or acute setting. It is well known that non-elective colon resections result in poor outcomes, with a postoperative mortality rate of around 15%, which is much higher than the 2 to 5% postoperative mortality after elective resections[42-44]

As life expectancy in the western world increases and the incidence of cancer increases with age, a large proportion of patients with colorectal cancer are elderly. Elderly are known to be more likely to present with more advanced disease, and are therefore especially at risk for emergency surgery. Previous studies have shown that non-elective resections for elderly patients may have very poor outcome, with a postoperative mortality risk of 38%.[45] However, other studies have shown that the poor prognosis of elderly patients may also be explained by comorbidities and a higher tumour-stage at presentation, suggesting that, with adequate preoperative workup, a selected group of elderly patients may have similar outcome after major surgery as younger patients.[46] More insight in risk factors for mortality after elective and non-elective resections, especially for elderly patients, may support clinical decision making in colorectal cancer surgery.

To monitor and improve the quality of care of colorectal cancer patients, the Dutch Surgical Colorectal Audit (DSCA) was initiated in 2009. In the DSCA, detailed patient and tumour characteristics together with diagnostic, treatment, pathology and outcome data are registered. All Dutch hospitals participate in this registry and hospital-specific casemix-adjusted outcomes are reported back to the participating centers, on a weekly basis through a web-based module. In addition, data are analyzed and reported on a nation-wide level to set benchmarks, identify risk-groups and transfer best practices to all hospitals treating these patients.
Although overall outcomes after colon cancer resections were acceptable, results after non-elective resections were poor with a postoperative mortality of 14% in 2009 and 9% in 2010. (www.dsca.nl)
Aim of the present study was to assess which factors contribute to poor postoperative outcome, especially in elderly patients who undergo emergency colon cancer resections, using a nationwide, population based database.

## Patients and Methods

### Patients

The Dutch Surgical Colorectal Audit (DSCA) is a nationwide, web-based database in which case-mix, treatment and outcome data are registered of patients that undergo a resection

of a first presentation of a colorectal carcinoma in the Netherlands. For this study, no ethical approval was required. All patients with a date of surgery between the 1st of January 2010 and 31st of December 2010 who were included in the DSCA before March 15th 2011 were selected. All Dutch hospitals participated in the registration. The estimated completeness of the DSCA in 2010, if compared to the Netherlands Cancer Registry (NCR), was 93% (www.dsca.nl).[47,48] Details of this dataset are described elsewhere.[49]

In the Netherlands, the resection of a rectum carcinoma is associated with a lower postoperative mortality rate (www.dsca.nl), and non-elective resections for a rectum carcinoma are infrequently performed (<5%) because of the clear advantage of preoperative radiation therapy. Therefore, patients treated for a rectum carcinoma were excluded from the analysis. Also, patients treated for a recurrent disease or multiple synchronous colorectal tumours were excluded. Records that did not contain information about the tumour location, the date of surgery, urgency of the procedure, or hospital survival were excluded.

## Definitions

Patients were divided in three age categories: younger than 70 years, 70 to 79 years and 80 years or older. Non-elective resections were designated as such by the responsible surgeon. Unfortunately, as patients were only included in the DSCA when a definitive resection of the tumour was performed, patients who underwent a two-stage procedure (the creation of an ostomy or placement of a intraluminal stent in a first procedure, postponing the definitive resection of the tumour to a second, elective procedure) were only included when they underwent the second procedure. As this was a small group of patients, and no information was available on the clinical course of the first procedure, they were analysed in the elective group. Additional case-mix factors registered in the DSCA were gender, comorbidity (Charlson-score), previous abdominal surgery, American Society of Anaesthesiologists (ASA) classification, local tumour invasiveness, disseminated disease, type of resection (right hemicolectomy, left hemicolectomy (+transverse colon) and sigmoid resection and other, including subtotal or panproctocolectomies) additional resection for local tumour invasiveness and/or metastasis, and ileo/colostomy placement. For missing information on a case-mix factor, a separate category was created.
Outcome was assessed by postoperative mortality occurring either during the hospital admission or within 30 days after resection. As a more internationally recognized classification for morbidity as the Dindo classification was unfortunately not included in the registry, secondary endpoints were major morbidity with serious consequences: leading to a reintervention, a prolonged length of stay of 14 days or more or mortality; any morbidity, defined as any postoperative adverse outcome, reinterventions; prolonged length of stay of 14 days or more and irradical resection (microscopic or macroscopic).

## Analysis

As increasing age may have a different effect on outcome after elective and non-elective resections, elective and non-elective patients were analysed separately. We first assessed which case-mix factors were predictive for postoperative mortality for elective and non-elective patients separately. All case-mix factors predictive for mortality in the univariate logistic regression ($p<0.10$) were included in the multivariate model. Then we investigated the effect of each individual predictive risk factor in different age groups, for elective and non-elective patients with no other risk factors present. We also assessed the effect of the number of additional risk factors on mortality in different age groups for elective and non-elective patients, as well as for secondary endpoints. Differences in categorical variables were tested using the chi-square test. Statistical significance was defined as $p<0.05$. All statistics were performed in PASW Statistics for Mac, Rel 18.0.2009. Chicago: SPSS inc. and Microsoft Excel for Mac.

**Table 1:** Case-mix factors and outcome of elective and non-elective patients.

| | | Elective | | Non-elective | |
|---|---|---|---|---|---|
| | | N (4989) | % | N (1172) | % |
| **Case-mix factors** | | | | | |
| Age (years) | <70 | 2202* | 44 | 561* | 48 |
| | 70-80 | 1718 | 35 | 343 | 29 |
| | 80+ | 1069 | 22 | 268* | 23 |
| Gender | Female | 2371 | 48 | 565 | 48 |
| Previous abdominal surgery | Yes | 1788 | 36 | 347 | 30 |
| ASA-classification | III+ | 1195 | 24 | 414 | 35 |
| | Missing | 21 | 0 | 20 | 2 |
| Charlson score | 2 + | 1021 | 21 | 210 | 18 |
| T-stage (TNM) | T1-2 | 1235 | 25 | 85 | 7 |
| | T3 | 2971 | 60 | 710 | 60 |
| | T4 | 684 | 14 | 353 | 30 |
| | Tx | 96 | 2 | 24 | 2 |
| M-stage (TNM) | M1 | 625 | 13 | 295 | 25 |
| Surgical procedure | Right hemicolectomy | 2405 | 48 | 478 | 41 |
| | Left hemicolectomy (+transversum) | 597 | 12 | 198 | 17 |
| | Anterior/sigmoid resection | 1821 | 37 | 426 | 36 |
| | Other (subtotal colectomy etc.) | 166 | 3 | 70 | 6 |
| Extended resection | Locally advanced tumour | 412 | 8 | 186 | 16 |
| | Metastasectomia | 124 | 3 | 42 | 4 |
| Ileo/colostomy | | 441 | 9 | 373 | 32 |
| **Outcome** | | | | | |
| Postoperative Mortality | | 192 | 3.9 | 102 | 9.3 |
| Morbidity | | 1010 | 20 | 387 | 33 |
| Prolonged length of stay | 14 days or more | 887 | 18 | 380 | 33 |
| Reintervention | Surgical/minimal invasive | 687 | 14 | 209 | 18 |
| Irradical resection | Microscopic or macroscopic | 117 | 2 | 77 | 7 |

ASA= American Society of Anaesthesiologists classification
TNM = Tumour Node Metastasis classification

# Results

## Patients

A total of 8,835 patients that underwent a resection of a colorectal carcinoma in 92 hospitals between 1st of January and 31st of December 2010 were entered in the DSCA. After exclusion of hospitals that failed to register all patients (2 hospitals, 2 patients), patients with multiple synchronous tumours (253 patients) and patients with a rectum carcinoma (2,419 patients) 6,161 patients were included in the analysis, of whom 1172 underwent a non-elective resection (Table 1). Postoperative mortality in the total population was 4.9%. The postoperative mortality rate after elective and non-elective resections was 3.9% and 9.3% respectively. 95 patients underwent a two-stage procedure with a postoperative mortality of 4.2%

Predictors for postoperative mortality after elective resections were increasing age (70-79 and 80+), male gender, Charlson comorbidity score (2+), ASA-classification (III+), surgical procedure (left hemicolectomy) and disseminated disease (table 2). For non-elective patients, age (70-79 and 80+), ASA-classification (III+), and surgical procedure (left hemicolectomy or 'other') were independent significant predictors for postoperative mortality. Gender, Charlson-score and disseminated disease did not remain significant in the multivariate model for mortality after non-elective resections. The placement of an ileo- or colostomy after the resection was not protective for mortality after elective (Odds Ratio

**Table 2:** Predictive factors for postoperative mortality after elective and non-elective resections in univariate analysis, that were entered in the multivariate model, with their multivariate Odds ratios.

| | | Elective | | | Non-elective | | |
|---|---|---|---|---|---|---|---|
| | | OR | 95% CI | P | OR | 95% CI | P |
| Age | <70 | 1.00 | | | 1.00 | | |
| | 70-80 | 2.76 | 1.73-4.40 | <0.01 | 1.90 | 1.09-3.29 | 0.02 |
| | 80+ | 6.72 | 4.26-10.60 | <0.01 | 3.10 | 1.80-5.34 | <0.01 |
| Gender | Female | 1.00 | | | 1.00 | | |
| | Male | 1.77 | 1.30-2.42 | | 1.48 | 0.96-2.27 | 0.08 |
| ASA-class | I – II | 1.00 | | | 1.00 | | |
| | III+ | 2.46 | 1.79-3.37 | <0.01 | 4.01 | 2.53-6.36 | <0.01 |
| | Missing | 0 | 0-0 | 1.00 | 2.46 | 0.52-11.7 | 0.26 |
| Charlson score | 0-1 | 1.00 | | | | | |
| | 2+ | 1.48 | 1.07-2.04 | 0.02 | | | |
| M-stage (TNM) | M0 | 1.00 | | | | | |
| | M1 | 1.60 | 1.06-2.42 | 0.03 | | | |
| Procedure | Sigmoid resection | 1.00 | | 0.02 | 1.00 | | |
| | Right hemicolectomy | 1.60 | 1.11-2.30 | 0.01 | 1.45 | 0.86-2.46 | 0.17 |
| | Left hemicolectomy | 2.00 | 1.23-3.25 | 0.01 | 2.48 | 1.36-4.49 | <0.01 |
| | Other (Subtotal colectomy, etc) | 1.97 | 0.89-4.34 | 0.10 | 3.45 | 1.54-7.70 | <0.01 |
| Extended resection | Metastasectomy | | | | 2.41 | 0.94-6.18 | 0.07 |

ASA class = American Society of Anaesthesiologists classification
TNM = Tumour Node Metastasis classification

1.11, 95% Confidence Interval 0.73-1.69) or non-elective resections (Odds Ratio 1.42, 95% Confidence Interval 0.91-2.28) in the univariate analyses. The C-statistic for the prediction of postoperative mortality after elective and non-elective resections was 0.78 (0.76-0.81), and 0.76 (0.71-0.80) respectively.

Figure 1 shows the effect of each individual risk factor on postoperative mortality for different age groups in elective (a) and non-elective patients (b). This figure shows that for elective patients with no other additional risk factors than their age and male gender, a left hemicolectomy or Charlson comorbidity score 2+, mortality risk only slightly increased with increasing age (< 7%), and did not differ from the mortality risk of non-elective elderly patients with no additional risk factors (7%). However, elderly with an ASA score of III or higher or disseminated disease did have a higher mortality rate (> 10%). For non-elective patients, both risk factors also resulted in an increased mortality risk for elderly patients to over 20%.

a

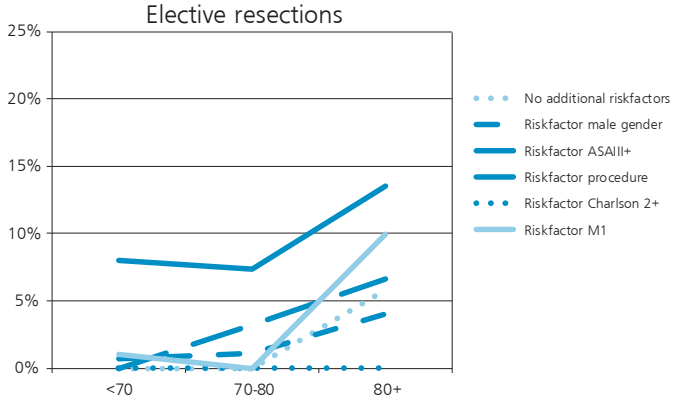



**Figure 1a:** mortality after elective resections, effect of the individual risk factors in combination with increasing age.

Risk factor male gender had no other risk factors than male gender, and age etc.
Risk factor Procedure: right or left hemicolectomy

b

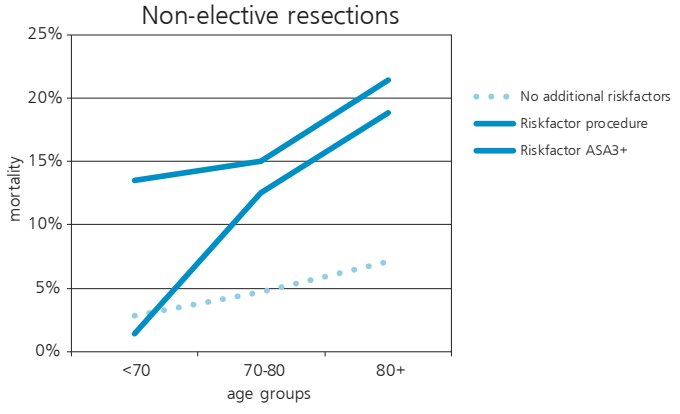



**Figure 1b:** mortality after elective resections, effect of the individual risk factors in combination with increasing age.

Risk factor ASA3+ had no other risk factors than ASA3+ and age, etc.
Risk factor procedure: left hemicolectomy or other (extended) resection

a

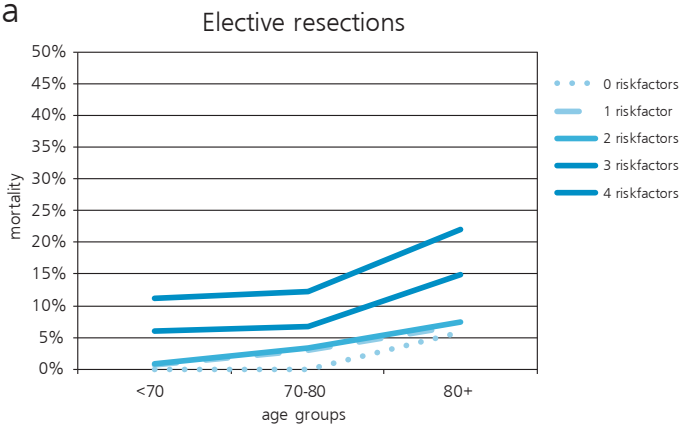



**Figure 2a:** Mortality after elective colon cancer resections, the effect of additional risk factors with increasing age.

Additional risk factors for elective patients: any combination of male gender, ASA III+, charlson score 2+, Metastatic disease or right or left hemicolectomy.

b

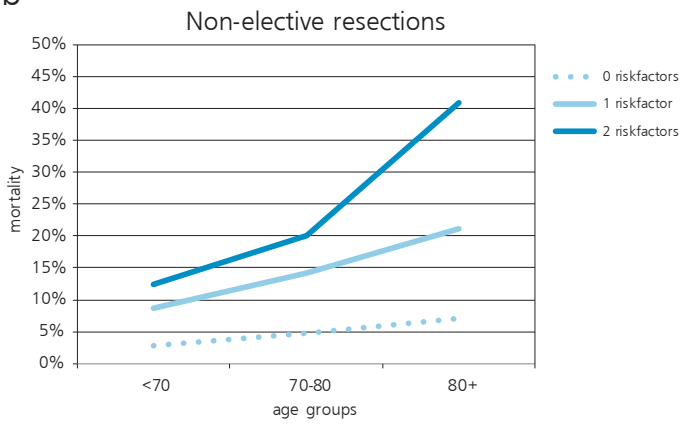



**Figure 2b:** Mortality after non-elective colon cancer resections

Additional risk factors for non-elective patients: any combination of ASA III+ or Left hemicolectomy or Other (extended) resection.

Figure 2 shows the effect of each additional risk factor on postoperative mortality for different age groups in elective (a) and non-elective patients (b). For elective patients, having one or two risk factors, additional to their age, did not result in an increased postoperative mortality. However, three and four risk factors resulted in a marked increased mortality risk, and even more so with increasing age (15 and 22% respectively). For non-elective patients, each additional risk factor increased the mortality risk, especially for elderly patients. For a non-elective patient of 80+ years with an ASA score of III+ and a left hemicolectomy or extended resection, postoperative mortality rate was 41% compared with 7% for non-elective 80+ patients without any additional risk factors.

## Elderly patients

Of all elderly patients, 1069 were treated electively, and 268 (20%) were treated non-electively (table 3). Of electively treated elderly 72 (6%) had 4 or more additional risk factors. Of all non-electively treated elderly, 98 (37%) had no additional risk factors and 23 (8%) had 2 additional risk factors. Secondary outcome parameters for elderly patients (80+) with 0 to 2, 3 and 4 additional risk factors for elective resections and 0, 1 and 2 additional

**Table 3:** Postoperative outcome for elderly (80+) elective and non-elective patients

| Elderly patients (80+ years) | | Elective | | | Non elective | | |
|---|---|---|---|---|---|---|---|
| No of risk factors | | 0-2 | 3 | 4 | 0 | 1 | 2 |
| No of elderly patients | | 783 | 214 | 72 | 98 | 147 | 23* |
| Outcome | | | | | | | |
| Postoperative Mortality | | 7% | 15% | 22% | 7% | 21% | 41% |
| Morbidity | | 25% | 38% | 38% | 27% | 50% | 77% |
| Prolonged length of stay | 14 days or more | 25% | 33% | 28% | 41% | 51% | 41% |
| Reintervention | Surgical/minimal invasive | 15% | 18% | 16% | 16% | 20% | 32% |
| Irradical resection | Microscopic or macroscopic | 2% | 2% | 4% | 5% | 5% | 9% |

* due to an error in the data processing, these numbers were published wrong in the first publication in Digestive surgery. (numbers as published: nonelective 80+, 2 riskfactors n= 22) However, this error did not change our results in any way.

risk factors for non-elective patients are shown in table 3. Non-elective elderly patients with one or two additional risk factors suffered morbidity in 50 and 77% of cases respectively. Morbidity rate for non-elective elderly with no additional risk factors was similar to elective elderly with up to two additional risk factors (27 versus 25%).

## Discussion

Colorectal cancer surgery in elderly patients in an emergency setting, bares considerable risk for adverse outcomes. Nonetheless, this study shows that these risks are not only determined by age, but are also highly dependent on other risk factors, such as the condition of the patient at the time of surgery (ASA-score) and the surgical procedure performed (left sided colon resections or extended resections). Based on these risk-factors, there is a group of elderly patients in which non-elective surgery can bare a risk for postoperative mortality up to 41%. These results raise the question if increased awareness of the risks involved can help optimize the choice of surgical strategy and peri-operative care to improve the outcome in this patient group. On the other hand the present study also shows that elderly patients with no additional risk factors have a similar outcome to elective patients of similar age.

Increasingly, clinical practice is confronted with treatment of elderly patients with multiple comorbid diseases, for whom the risk for postoperative morbidity is high. Although various studies have identified risk factors for unfavourable outcome after colon resections,[38,39,50] with similar results and test characteristics of their models as in the present study,[38] the implications of these risk factors for elderly patients in 'real life' have not been elucidated. Our study, which is based on a nationwide population based registry, shows that for otherwise healthy elderly patients, operative risk is only slightly increased, and similar after elective and non-elective resections. For elective patients, mortality risk only starts to increase with 3 or more additional risk factors, while for patients undergoing a non-elective resection, one additional risk factor (e.g. a high ASA class or a left hemicolectomy) increases

the operative risk markedly. This indicates that a non-elective resection in elderly patients, without any additional risk factors can be performed in a safe way. Otherwise, for elderly patients, having additional risk factors, a non-elective resection should be considered very high-risk surgery.

One of the major strengths of this study is the large nationwide prospectively collected detailed dataset, which enabled to identify risk factors for mortality for both elective and non-elective resections. However, although Ghaferi has shown that voluntary reporting does not necessarily lead to selection bias[51], we cannot exclude that selection bias may have occurred caused by hospitals not registering all patients. However, comparisons with the NCR showed no evidence of selection bias and an overall estimated completeness of the dataset of 93% (www.dsca.nl), which is similar to the Scandinavian audits, and high compared to the National Bowel Cancer Audit Programme in the United Kingdom (UK) (completeness 68%) and the Belgium 'Project on Cancer of the Rectum' (40%)[37]. It is unlikely that the remaining 7% of all patients would change our results to a great extent.

Discriminating between high- and low-risk patients can be helpful in identifying 'best practices' in the treatment of elderly patients, for whom the clinical (oncological) decision making is complex and highly influenced by preoperative health, the risk of adverse events and overall life expectancy. The lack of information from randomized controlled trials about treatment strategies in elderly patients may be compensated by results from population-based clinical registries like the Dutch Surgical Colorectal Audit. As the mortality risk for non-elective patients seems to be mostly determined by their preoperative health status, the most promising way of improving outcome would be reducing the number of patients with an indication for a non-elective resection. This could be accomplished by increasing early detection of colon cancer through a national screening program as many nations are currently developing. Another strategy that could be suggested is the placement of an ileo- or colostomy to reduce the risk of anastomotic leakage, one of the major complications after colon cancer surgery. However, we found the placement of an ileo/colostomy was not protective for mortality in multivariate analysis. A possible explanation is that, due to an adequate selection strategy of Dutch surgeons, those patients who did not receive an ostomy, did not have a higher mortality risk. It is very likely that, when no ostomies would have been placed, the mortality rate would have been much higher. However, the placement of an ostomy only prevents anastomosis leakage, and is not protective for other possible complications. As an ostomy is often placed in patients with a poor preoperative condition, these patients may also be at risk for many other complications that may lead to unfavourable outcome, explaining the high mortality risk in non-elective patients. Previous observational studies have found similar results[42,52,53]
Interestingly, two-stage procedures had a similar postoperative mortality rate as elective procedures, indicating that postponing the definitive resection to a second procedure in the elective setting, allowing a careful preoperative workup may reduce the risk at postoperative

adverse events. However, results should be interpreted with great care as patients were only included in the DSCA when they underwent the second, definitive resection procedure. Patients, for whom the first-step, often non-elective procedure was not followed by the second step procedure, most likely because of an unfavourable outcome of the first step, were not included in the DSCA. When these patients would have been included in the database, and this group could have been analysed using the 'intention to treat' method, the mortality rate may have been higher. Further research, including patients who did not undergo a definitive resection, as well as long-term results, may help to further clarify these results.

Another opportunity for improvement of outcome after non-elective resections is, as Bilimoria et al. stated[54], risk based referral. He calculated a reduction in mortality rate for high-risk colon cancer patients from 7.9% to 6.6% when all high-risk patients would be referred to specialized centres. Unfortunately Bilimoria et al. stratified patients only for age and comorbidity index, and did not take into account the urgency of the resection. Ingraham[44] showed that outcome after non-elective resections also varies between hospitals, indicating that risk-based referral may also improve outcome for non-elective patients. Interestingly, the latter study found that hospitals with good outcome after elective resections did not necessarily have good results after non-elective resections. The present study shows that non-elective resections in elderly patients with additional risk factors, is extremely high-risk surgery, with outcomes comparable to acute abdominal aneurysm surgery[55]. Interestingly, until recently, non-elective colorectal surgery in the Netherlands was performed by almost all general surgeons on call, while other high risk procedures, such as acute aneurysm surgery, are performed by specialized surgeons only. Previous studies have demonstrated that the presence of a specialised surgeon during a non-elective operation improves outcome[56,57]. Unfortunately, our database did not contain information on the qualifications of the operating surgeon but it is likely that referral of high-risk non-elective patients to centres with a specialised surgeon available during on-call hours, improves outcome. However, referral of such patients may be a logistical challenge as for these patients, treatment delay due to referral may negatively influence outcome.

## Conclusions

The present study confirms well known risk factors for mortality after elective and non-elective colorectal cancer surgery in a large, national population based database. We found that only a small selection of patients is truly high-risk. These are elderly, non-elective patients, with a high ASA class and or undergoing a left hemicolectomy or extended resection. For these patients, mortality risk may rise up to 41%. Therefore, for these patients, a non-elective resection should be considered high-risk surgery. The different risk-profiles in colon cancer surgery presented in this study could be used to support clinical

decision-making and to adequately inform and advise patient and family, and should have consequences for composing an operative team.

## Acknowledgements

The authors would like to thank all surgeons, registrars, physician assistants and administrative nurses that registered all the patients in the DSCA, as well as the Dutch Surgical Colorectal Audit group and the methodological board for their advice.
During this research, N. Kolfschoten was partially funded by the Dutch Cancer Society.

## References


1. Tekkis PP, Kinsman R, Thompson MR, Stamatakis JD. The Association of Coloproctology of Great Britain and Ireland study of large bowel obstruction caused by colorectal cancer. Ann Surg. Jul 2004;240(1):76-81.
2. Engel AF, Oomen JL, Knol DL, Cuesta MA. Operative mortality after colorectal resection in the Netherlands. Br J Surg. Dec 2005;92(12):1526-1532.
3. Ingraham AM, Cohen ME, Bilimoria KY, et al. Comparison of hospital performance innonemergency versus emergency colorectal operations at 142 hospitals. J Am Coll Surg. Feb;210(2):155-165.
4. Waldron RP, Donovan IA, Drumm J, Mottram SN, Tedman S. Emergency presentation and mortality from colorectal cancer in the elderly. Br J Surg. Mar 1986;73(3):214-216.
5. Surgery for colorectal cancer in elderly patients: a systematic review. Colorectal CancerCollaborative Group. Lancet. Sep 16 2000;356(9234):968-974.
6. Schouten LJ, Hoppener P, van den Brandt PA, Knottnerus JA, Jager JJ. Completeness of cancer registration in Limburg, The Netherlands. Int J Epidemiol. Jun 1993;22(3):369-376.
7. Schouten LJ, Jager JJ, van den Brandt PA. Quality of cancer registry data: a comparison of data provided by clinicians with those of registration personnel. Br J Cancer. Nov 1993;68(5):974-977.
8. Kolfschoten NE, Marang van de Mheen PJ, Gooiker GA, et al. Variation in case-mix between hospitals treating colorectal cancer patients in the Netherlands. Eur J Surg Oncol. Nov 2011;37(11):956-963.
9. Cohen ME, Bilimoria KY, Ko CY, Hall BL. Development of an American College of Surgeons National Surgery Quality Improvement Program: morbidity and mortality risk calculator for colorectal surgery. J Am Coll Surg. Jun 2009;208(6):1009-1016.
10. Al-Homoud S, Purkayastha S, Aziz O, et al. Evaluating operative risk in colorectal cancer surgery: ASA and POSSUM-based predictive models. Surg Oncol. Aug-Nov 2004;13(2-3):83-92.
11. Tekkis PP, Prytherch DR, Kocher HM, et al. Development of a dedicated risk-adjustment scoring system for colorectal surgery (colorectal POSSUM). Br J Surg. Sep 2004;91(9):1174-1182.
12. Ghaferi AA, Osborne NH, Dimick JB. Does voluntary reporting bias hospital quality rankings? J Surg Res. Jun 15;161(2):190-194.
13. van Gijn W, Wouters MW, Peeters KC, van de Velde CJ. Nationwide outcome registrations to improve quality of care in rectal surgery. An initiative of the European Society of Surgical Oncology. J Surg Oncol. Jun 15 2009;99(8):491-496.
14. Faiz O, Warusavitarne J, Bottle A, et al. Nonelective excisional colorectal surgery in English National Health Service Trusts: a study of outcomes from Hospital Episode Statistics Data between 1996 and 2007. J Am Coll Surg. Apr;210(4):390-401.
15. Poon JT, Chan B, Law WL. Evaluation of P-POSSUM in surgery for obstructing colorectal cancer and correlation of the predicted mortality with different surgical options. Dis Colon Rectum. Mar 2005;48(3):493-498.
16. Bilimoria KY, Bentrem DJ, Talamonti MS, Stewart AK, Winchester DP, Ko CY. Risk-based selective referral for cancer surgery: a potential strategy to improve perioperative outcomes. Ann Surg. Apr;251(4):708-716.
17. Giles KA, Pomposelli FB, Hamdan AD, Wyers MC, Schermerhorn ML. Comparison of open and endovascular repair of ruptured abdominal aortic aneurysms from the ACS-NSQIP 2005-07. Journal of endovascular therapy : an official journal of the International Society of Endovascular Specialists. Jun 2009;16(3):365-372.
18. Elshove-Bolk J, Ellensen VS, Baatrup G. Logistics and outcome in urgent and emergency colorectal surgery. Colorectal Dis. Nov 10 2009.
19. Biondo S, Kreisler E, Millan M, et al. Impact of surgical specialization on emergency colorectal surgery outcomes. Arch Surg. Jan;145(1):79-86.colorectal surgery outcomes. Arch Surg. Jan;145(1):79-86.

# Part 3

The use of clinical auditing for the evaluation and monitoring of the implementation of new techniques

Chapter 6

# Successful and safe introduction of laparoscopic colorectal cancer surgery in Dutch hospitals

E. Kolfschoten[1], N.J. van Leersum[1], G.A.Gooiker[1], P.J. Marang van de Mheen[2], E.H. Eddes[3], J. Kievit[1,2] R. Brand[4], Pieter J. Tanis[5], W.A. Bemelman[5], R.A.E.M. Tollenaar[1], J. Meijerink[6], M.W.J.M. Wouters[7], on behalf of the Dutch Surgical Colorectal Audit group*

[1] Leiden University Medical Centre, Dept of Surgery
[2] Leiden University Medical Center, Dept of Medical Decision Making
[3] Deventer Hospital, Dept of Surgery
[4] Leiden University Medical Center, Dept of Biostatistics
[5] Academic Medical Centre, Amsterdam, Dept of Surgery
[6] VUMC, Amsterdam, Dept of Surgery
[7] The Netherlands Cancer Institute, Amsterdam, Dept of Surgery
* The Dutch Colorectal Audit Group: O.R.C. Busch, R.M. van Dam, E. van der Harst, M.L.E.A. Jansen-Landheer, Th.M. Karsten, J.H.J.M. van Krieken, W.G.T. Kuijpers, V.E. Lemmens, E.R. Manusama, H.J.T. Rutten, Prof. dr. C.J.H. van de Velde, T. Wiggers

Annals of Surgery 2013 May;257(5):916-21

## Abstract

### Objective

To investigate the safety of laparoscopic colorectal cancer resections in a nationwide population based study.

### Background

Although laparoscopic techniques are increasingly used in colorectal cancer surgery, little is known on results outside trials. With the fast introduction of laparoscopic resections, questions were raised about safety.

### Methods

Of all patients who underwent an elective colorectal cancer resection in 2010 in the Netherlands, 93% were included in the Dutch Surgical Colorectal Audit. Short-term outcome after laparoscopic, open and converted laparoscopic resection were compared in a generalized linear mixed model. We further explored hospital differences in laparoscopic resection and conversion rates.

### Results

7350 patients, treated in 90 hospitals were included. Laparoscopic resection rate was 41% with a conversion rate of 15%. After adjustment for differences in case-mix, laparoscopic resection was associated with a lower risk of mortality (OR 0.63, $p<0.01$), major morbidity (OddsRatio 0.72, $p<0.01$), any complications (OR 0.74, $p<0.01$), hospital stay > 14 days (OR 0.71, $p<0.01$), and irradical resections (OR 0.68, $p<0.01$), compared to open resection. Outcome after conversion was similar to open resections ($p>0.05$). A large variation in laparoscopic resection and conversion rates among hospitals was found, however, the difference in outcome associated with operative techniques was not influenced by hospital of treatment.

### Conclusions

Use of laparoscopic techniques in colorectal cancer surgery in the Netherlands is safe and results in better short-term outcome than open surgery, irrespective of the hospital of treatment. Outcome after conversion was similar to open resection.

## Introduction

After the introduction of the laparoscopic colectomy in 1991 by Jacobs,[58] the safety of laparoscopic resection (LR) for colorectal cancer became subject of debate. The main objections were the uncertainty about radicality of the resection and adequate mesenteric lymphadenectomy when using the laparoscopic approach, and whether enting of tumour cells would occur at the extraction site or other port sites.[59] Only after publication of several randomized clinical trials (Barcelona-, COST-, COLOR-I and CLASICC-trial), which convincingly demonstrated that there were no short or long term disadvantages, the use of LR for colorectal cancer markedly increased in the Netherlands, from less than 5% in 2003 to 36% of all elective resections in 2009 (www.dsca.nl).[60-65]

In clinical trials, LR has proven short term benefits over open resection (OR) for colorectal cancer: decreased perioperative blood loss and pain, faster postoperative recovery of bowel movements, decreased postoperative complication rate and a decreased length of stay.[60,65-68] These should be balanced against the possible disadvantages such as a longer operation time (although decreasing with experience), higher costs, and an extensive learning curve.[69] Moreover, there is a risk of conversion of LR to OR. Initial studies have shown that converted LR (CLR) is associated with longer operation time, longer hospital stay, and higher postoperative morbidity and mortality rates compared to completed LR.[70-74] However, comparing CLR with OR may be more appropriate. Although evidence is limited, several authors suggested that a CLR for colorectal cancer results in worse outcome than a primary OR.[75,76] Therefore, a careful patient selection is advised for laparoscopic surgery.

All these results are from randomized controlled trials (RCT), which are known to describe the work of dedicated surgeons, often on a select patient group, in high volume, specialized centres. Therefore, it is unclear whether the good results after LR can also be confirmed in a population based study. To monitor and improve the quality of care of colorectal cancer patients in the Netherlands, the Dutch Surgical Colorectal Audit (DSCA) was initiated in 2009. In the DSCA, detailed patient and tumour characteristics together with diagnostic, treatment, pathology and outcome data are registered. Compliance to this database is enforced by the Dutch Healthcare Inspectorate. Using this database, we were able to investigate the safety of LR in Dutch hospitals.

**The aim** of this study was to compare short-term outcome after LR (intention to treat), and OR, and after CLR and OR in a population based registry, including all colorectal cancer resections in the Netherlands, and to explore hospital differences in LR and CLR rates.

## Methods

### Patients

The dataset was retrieved from the DSCA, a nation wide, web-based database in which all patients undergoing a resection for primary colorectal cancer in the Netherlands are included. For this study, no ethical approval was required. All Dutch hospitals participated in the registration. The estimated completeness of the DSCA in 2010, if compared to the Netherlands Cancer Registry (NCR), was 93% (www.dsca.nl).[47,48] Details of this dataset are described elsewhere.[77]

### Inclusion and exclusion criteria

All elective patients with a date of surgery between the 1st of January 2010 and 31st of December 2010 and registered in the DSCA before March 15$^{th}$ 2011 were evaluated. Minimal data requirements for evaluation were information on tumour location, date of surgery and mortality. Resections in an urgent setting or for recurrent colorectal cancer or multiple synchronous colorectal tumours were excluded. To minimize the risk of selection bias, hospitals that failed to register more than 10 patients in 2010 were excluded.

### Definitions

LR was defined as any procedure started with the intention to resect the tumour using laparoscopic techniques, including CLR. CLR was defined as a procedure that was started with the intention to perform a LR, but was completed as an OR. Primary outcome measures were postoperative mortality, defined as in-hospital or 30-day mortality, and major morbidity, defined as an adverse outcome with serious consequences: leading to mortality, a reintervention (operative or percutaneous), or a postoperative hospital stay of at least 14 days. Secondary endpoints were adverse outcomes, occurring within 30 days after resection, reinterventions, prolonged hospital stay (>14 days), irradical resection (microscopic or macroscopic) and the number of lymph nodes investigated (10 or more).

Outcome was adjusted for case-mix. Available case-mix factors were age, gender, comorbidity (Charlson-score),[78] previous abdominal surgery, Body Mass Index (BMI), American Society of Anaesthesiologists (ASA) classification, operative procedure, tumour stage (TNM), preoperative radiation therapy and additional resections for tumour invasion and/or metastasis. All case-mix factors were categorized into discrete categories. Missing case-mix factors were analysed in a separate category.

### Analysis

First, we investigated the effect of operative technique on outcome at a patient level. Two separate analyses were performed. In the first analysis we compared LR (intention to treat) and OR. In the second analysis we compared CLR with OR. Differences in case-mix factors

between the patient groups were compared using the chi-square test. To investigate the combined effect of the case-mix, the predicted probability for mortality ('expected mortality') was calculated for each patient, based on an overall multivariable logistic regression model, constructed using a backwards-stepwise approach.[77] To compare differences in combined effect of case-mix between the three treatment groups, average expected mortality was compared using the student T-test. The independent effect of different operative techniques (LR vs. OR and CLR vs. OR) on primary and secondary outcomes was estimated using multivariate generalized linear mixed models in which all case-mix factors mentioned above were included. As the effect of operative techniques on outcome was similar for colon and rectum cancer, the analysis was performed on the full population, and the location of the tumor was included in the analysis as a covariate. Since there may be other unknown factors in a hospital, responsible for part of the variation in outcomes, we included a random effect for hospital of treatment in the model. To further investigate whether the influence of operative techniques (LR vs OR or CLR vs OR) on outcomes differed significantly between hospitals, we also added an interaction term between hospital treatment and operative techniques. When this interaction term significantly improved the model fit, this indicates that the influence of the operative technique on outcome differs between hospitals. This is roughly equivalent to a test on heterogeneity of the 'operative technique' among hospitals.

Secondly, we further investigated the differences in use of laparoscopic techniques at hospital level. The LR rate for each hospital was determined. As LR might be avoided in more complex cases, hospital variations in case-mix may affect the LR rate. Using a backwards-stepwise multivariate logistic regression analysis, hospitals' expected LR rate was calculated, based on the hospitals' case-mix. Subsequently, the quotient of the observed (O) and the expected (E) LR rate was determined: O/E LR ratio.[33,34] An O/E ratio of less than 1 means that less LR were performed than expected based on the hospitals case-mix and national average, while an O/E ratio of more than 1 indicates the opposite. A linear regression analysis was performed to explore the relation between hospital volume (e.g. the total number of elective colorectal resections) and the hospital's adoption of laparoscopic techniques (unadjusted LR rate, and O/E LR ratio).
In a similar way, factors predictive for conversion were identified and an O/E CLR ratio was calculated for each hospital. Hospitals with a significantly higher O/E ratio (lower limit of the 95% confidence interval above 1) were categorized as high outliers.[79] Low outliers were hospitals with a significantly lower O/E ratio (upper limit of the confidence interval below 1). We calculated the number of outlier hospitals for LR and CLR rates. The relation between the number of elective laparoscopic colorectal resections in a hospital (laparoscopic volume) and the unadjusted CLR rate and O/E CLR ratio was explored using a linear regression analysis. Statistical significance was defined as $p<0.05$. Statistics were performed using PASW Statistics, Rel 18.0.2009. Chicago: SPSS inc. Stata and R.

## Results

On March 15th 2011, 90 hospitals registered a total of 8835 eligible patients with a date of surgery between January 1 and December 31 2010 in the DSCA. After exclusion of patients with multiple synchronous tumours (N=253), and patients with an indication for urgent resection (N=1228), a total of 7350 patients (4986 colon cancer and 2364 rectal cancer) treated in 90 hospitals were included in the analyses.

Of all patients, 3113 (41%) underwent a LR. In 453 (15%) of these patients, LR was converted to OR. Of all patients with a colon cancer, 44% underwent a LR with a CLR rate of 15%; these percentages were 37% and 13% respectively for rectal cancer.
Table 1 shows differences in case-mix between LR and OR, and between CLR and OR. Case-mix for patients undergoing a LR was more favourable than case-mix of patients who underwent an OR; average expected mortality for LR was lower than for OR (2.9% vs 3.6% p<0.05). Interestingly, when we compared case-mix between CLR and OR, we found average expected mortality for CLR was similar to OR (table 1).

**Table 1.** Case-mix of patients who underwent a Laparoscopic resection (LR) for colorectal cancer compared to an open resection (OR), and for patients who underwent converted laparoscopic resection (CLR).

| | OR | LR | P* | CLR | P$ |
|---|---|---|---|---|---|
| No of patients | 4287 | 3063 | | 446 | |
| Age >70 years | 53% | 48% | <0.01 | 53% | 0.75 |
| Male | 54% | 57% | 0.01 | 62% | <0.01 |
| Previous abdominal surgery | 36% | 30% | <0.01 | 40% | 0.16 |
| Charlson 2+ | 20% | 16% | <0.01 | 21% | 0.77 |
| ASA 3+ | 23% | 19% | <0.01 | 24% | 0.49 |
| Stage III/ IV | 44% | 38% | <0.01 | 40% | 0.14 |
| Rectum | 35% | 29% | <0.01 | 26% | <0.01 |
| Preoperative radiotherapie | 30% | 24% | <0.01 | 20% | <0.01 |
| Additional resections | 11% | 4% | <0.01 | 11% | 0.84 |
| Average expected mortality | 3.6% | 2.9% | <0.01 | 3.7% | 0.76 |

ASA: American Society of Anesthesiologists classification
LR: laparoscopic resection (intention to treat)
OR: open resection
CLR: converted laparoscopic resection
* LR versus OR [chi2 or student T-test]
$ CLR versus OR [chi2 or student T-test]

### Differences in outcome between operative techniques

Outcome after LR was better than after OR with a postoperative mortality of 2.4% after LR vs 4.0% after OR (colon: 2.8% after LR and 4.9% after OR, rectum: 1.6% after LR and 2.6% after OR), and a postoperative major morbidity rate of 19% after LR vs 26% after OR (colon: 23 % after OR and 16% after LR, rectum: 29% after OR and 23% after LR). Also in high-risk patients (e.g. ASA III or Charlson score 2+), LR resulted in a lower postoperative mortality

(ASA III: 7.1% after LR vs 9.7% after OR, Charlson 2+: 4.5% after LR vs 8.2% after OR) and morbidity rate (ASA III: 27% after LR vs 36% after OR, Charlson 2+: 25% after LR vs 32% after OR).

After adjustment for case-mix and hospital of treatment, LR was associated with a lower risk for mortality (Odds Ratio 0.63, p<0.01), as well as major morbidity (Odds Ratio 0.72, p<0.01), any complications (Odds Ratio 0.74, p<0.01), prolonged hospital stay (Odds Ratio 0.71, p<0.01), and irradical resection (Odds Ratio 0.68, p<0.01). The percentage of patients in whom 10 or more lymph nodes were retrieved was similar for both groups (p=0.87). We did not find any differences in outcome between CLR and OR (p>0.05) [Table 2]. Adding an interaction term for hospital of treatment and operative technique did not improve the fit of any of the models (p>0.05), suggesting that differences in outcome associated with operative techniques did not differ between hospitals.

**Table 2.** Outcomes after laparoscopic resection (LR) of colorectal cancer and converted laparoscopic (CLR) resection compared to open resection (OR).

| | OR | LR | Odds-Ratio* (95%CI) | CLR | Odds-Ratio$ (95%CI) |
|---|---|---|---|---|---|
| No of patients | 4287 | 3063 | | 446 | |
| Mortality | 4.0% | 2.4% | 0.63 (0.47-0.86) | 4.7% | 1.09 (0.66-1.80) |
| Major Morbidity | 26% | 19% | 0.72 (0.63-0.83) | 28% | 1.07 (0.85-1.36) |
| Complication | 37% | 29% | 0.74 (0.66-0.84) | 39% | 1.05 (0.84-1.31) |
| Prolonged LOS | 24% | 17% | 0.71 (0.62-0.82) | 26% | 1.09 (0.86-1.39) |
| Reïntervention | 17% | 13% | 0.79 (0.68-0.92) | 15% | 0.78 (0.58-1.05) |
| Irradical resection | 3.8% | 2.1% | 0.68 (0.48-0.98) | 4.7% | 1.30 (0.76-2.23) |
| 10+ lymph nodes | 73% | 73% | 0.87 (0.76-1.00) | 74% | 0.90 (0.70-1.16) |

LR: laparoscopic resection
OR: open resection
CLR: converted laparoscopic resection
LOS: length of stay
CI: Confidence interval
* Odds Ratio of LR versus OR.
$ Odds Ratio of CLR versus OR.
Odds ratios were adjusted for differences in case-mix and hospital of treatment in a general linear mixed model.

## Differences in use of operative techniques between hospitals

Figure 1a shows the unadjusted LR rate for all hospitals in the Netherlands in a funnel plot. Hospital LR rates varied from 0% to 96%. Five hospitals did not perform any LR. Factors predictive for a laparoscopic resection were younger age, no previous abdominal surgery, a low ASA class, a Charlson score <2, a TNM stage I, a sigmoid or low-anterior resection and no additional resections. [Table 3] After adjustment for differences in case-mix, 30 (33%) hospitals had a significantly higher O/E LR ratio than average (high outlier); in 26 (29%) hospitals, the O/E LR ratio was significantly lower than average (low outlier). [figure 1b] No relation was found between hospital volume and hospital unadjusted LR rate (beta =-0.07, p=0.52) or O/E LR ratio (beta =-0.07, p=0.52)

Figure 2a shows the unadjusted CLR rate for all hospitals in the Netherlands in a funnelplot. Hospital CLR rates varied from 0% to 73%. Male gender, previous abdominal surgery,

high BMI, left hemicolectomy, and additional resections for locally advanced tumour or metastasectomy were independently predictive for conversion. An abdominal perineal resection (APR) was associated with a low risk of conversion [Table 3]. After adjustment for

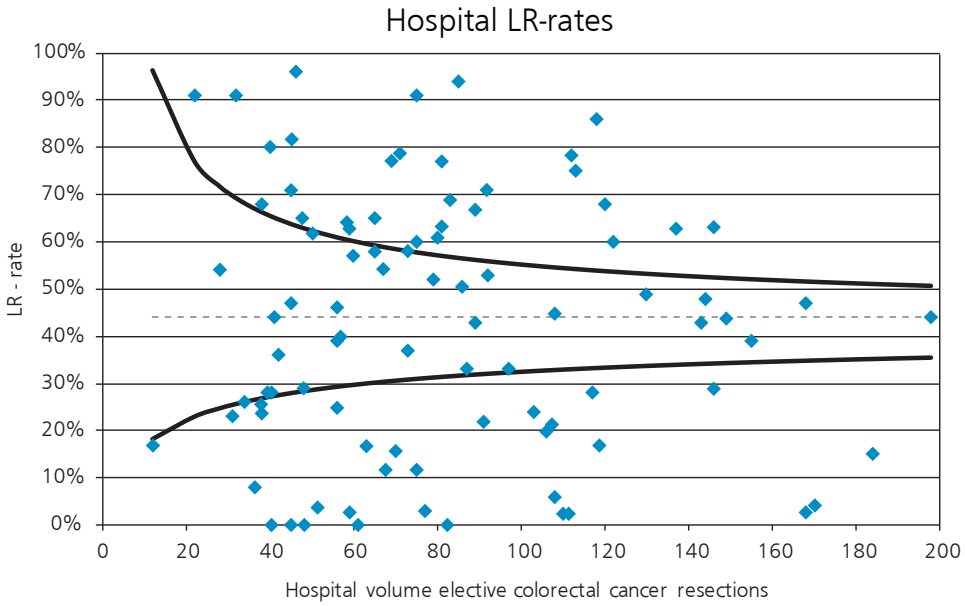



**Figure 1a.** Funnelplot of unadjusted percentage of laparoscopic resections (LR) for colorectal cancer in 2010 in hospitals included in the DSCA
LR= laparoscopic resection
Black lines represent the 95% confidence intervals of average (dotted line)

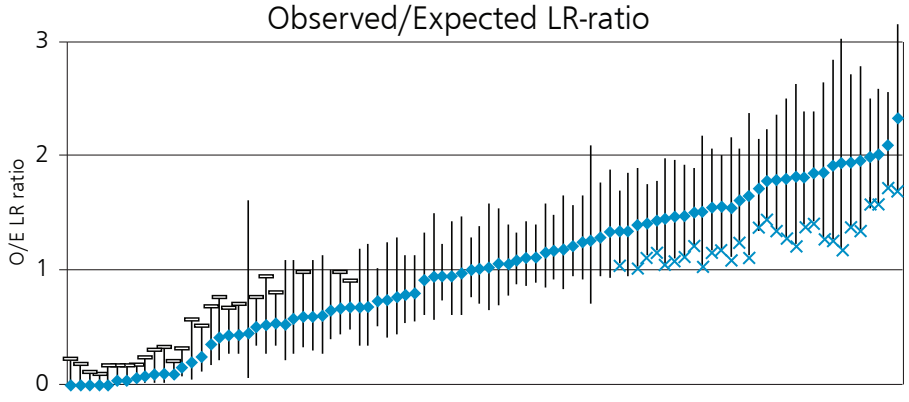



**Figure 1b**. Comparison of risk-adjusted percentage of laparoscopic colorectal cancer resections (LR) in 2010 in hospitals included in the DSCA
R= laparoscopic resection
marked hospitals are high outliers for laparoscopic resection rate
marked hospitals are low outliers for laparoscopic resection rates

differences in case-mix, only four (4.7%) hospitals had a significantly higher CLR rate than average; in five (5.9%) hospitals, the CLR rate was significantly lower than average. [figure 2b] We found no relation between hospital laparoscopic volume and hospital unadjusted CLR rate (beta =-0.065, p=0.56) or O/E CLR ratio (beta =-0.026, p=0.81).

**Table 3.** Risk-factors predictive for laparoscopic resection (LR) and converted laparoscopic resection (CLR) in the multivariate logistic regression model.

| Riskfactor | | LR<br>Odds Ratio (95% CI) | CLR<br>Odds Ratio (95% CI) |
|---|---|---|---|
| Age | < 60 | 1.00 (ref) | - |
| | 61 - 70 | 0.86 (0.75-1.00) | - |
| | 71 - 80 | 0.75 (0.65-0.87) | - |
| | 81+ | 0.67 (0.57-0.80) | - |
| Gender | Female | | 0.69 (0.55-0.86) |
| Abdominal history | | 0.82 (0.74-0.91) | 1.76 (1.41-2.20) |
| BMI | <25 | | 1.00 (ref) |
| | 25 - 30 | | 1.58 (1.21-2.06) |
| | >30 | | 2.67 (2.20-4.01) |
| | Missing | | 1.39 (1.00-1.93) |
| ASA | I - II | 1.00 (ref) | 1.00 (ref) |
| | III | 0.89 (0.78-1.01) | 1.23 (0.95-1.59) |
| | IV - V | 0.61 (0.40-0.94) | 3.05 (1.37-6.81) |
| Charlson score | 2+ | 0.83 (0.72-0.94) | - |
| Procedure | Right hemi | 1.00 (ref) | 1.00 (ref) |
| | Left hemi | 0.95 (0.79-1.15) | 1.73 (1.18-2.54) |
| | Sigmoid | 2.06 (1.78-2.38) | 0.84 (0.62-1.14) |
| | LAR | 1.47 (1.26-1.72) | 1.20 (0.92-1.57) |
| | APR | 1.23 (0.97-1.54) | 0.55 (0.33-0.90) |
| | Else | 0.68 (0.49-0.94) | 1.13 (0.55-2.33) |
| TNMstage | I | 1.00 (ref) | - |
| | II | 0.80 (0.70-0.91) | - |
| | III | 0.77 (0.68-0.88) | - |
| | IV | 0.54 (0.45-0.65) | - |
| | X | 0.49 (0.34-0.70) | - |
| Radiotherapy | None | 1.00 (ref) | - |
| | 5x5 Gy | 0.66 (0.55-0.79) | - |
| | >60 Gy | 0.89 (0.62-1.28) | - |
| | Chemoradiation | 0.40 (0.32-0.51) | - |
| Additional resection | Local | 0.32 (0.25-0.40) | 6.45 (4.24-9.81) |
| | Metastasis | 0.22 (0.14-0.36) | 2.91 (1.06-7.95) |
| C-statistic | | 0.65 (0.64-0.67) | 0.68 (0.65-0.70) |

LR = Laparoscopic resection (intention to treat)
CLR= Converted laparoscopic resection
95% CI = 95% confidence interval
BMI = Body Mass Index
ASA = American Society of Anesthesiologists
Hemi = hemicolectomy
LAR = Low anterior resection
APR= Abdominal perineal resection
TNM = Tumour Node Metastasis system

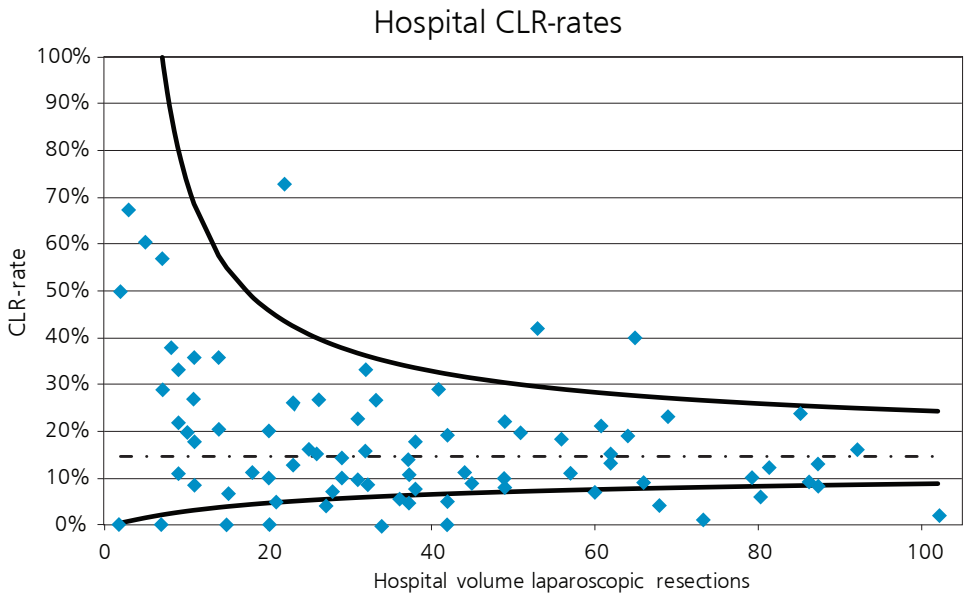



**Figure 2a**. Funnelplot of unadjusted converted laparoscopic resection rates after laparoscopic colorectal cancer resections in 2010 in hospitals included in the DSCA
CLR = Converted Laparoscopic Resection
Black lines represent the 95% confidence intervals of average (dotted line)

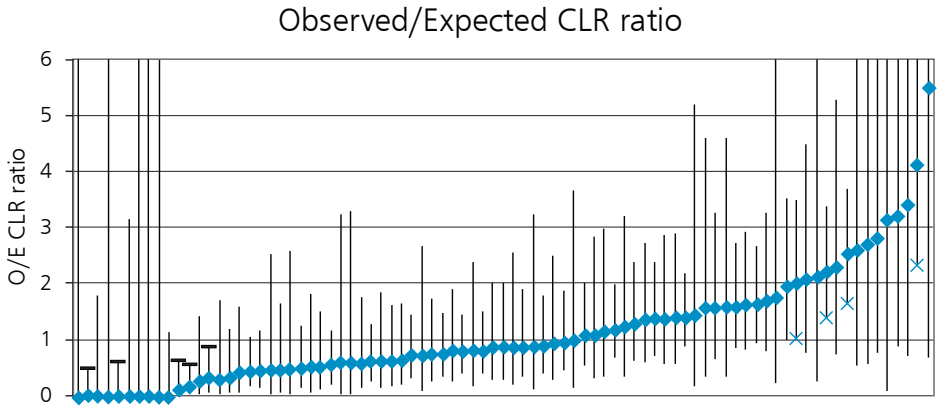



**Figure 2b**. Comparison of Observed/Expected (O/E) converted laparoscopic resection rate after laparoscopic colorectal cancer resections in 2010 in hospitals included in the DSCA
↓.R = Converted Laparoscopic Resection
↓ marked hospitals are high outliers for converted laparoscopic resection rate
T marked hospitals are low outliers for converted laparoscopic resection rate

## Discussion

This study shows that laparoscopic surgery is successfully used in colorectal cancer surgery in the Netherlands. More than 40% of all elective resections are performed using laparoscopic techniques, with an acceptable CLR rate of 15%. The group of patients selected for a LR has a more favourable case-mix than those selected for OR. However, the group of patients who underwent CLR had a similar case-mix as those selected for OR. After correction for differences in case-mix, outcome after LR was better than after OR, while outcome after CLR was not different from OR. A large variance in risk-adjusted LR rates between Dutch hospitals was observed. However, only four hospitals had a significantly higher, risk-adjusted CLR rate, nine had a significantly lower CLR rate. As the interaction term of hospital of treatment and operative technique did not improve our models, we may conclude that there is no evidence in our data that the effect of operative techniques on outcome differs between hospitals.

The percentage of LRs in the Netherlands is high compared to other countries, and conversion rates are low. In the United Kingdom, the LR rate is 30% with a conversion rate of 21%.[80] In the National Surgical Quality Improvement Program in the United States, the LR rate was 27-31%,[81,82] while CLR is not reported. Conversion rates in previous randomized controlled trials were 21% (COST), 19% (COLOR) and 30% (Classic).[61,62,83] A recent Cochrane review by Schwenk at al. described better short-term results after LR: a reduction of total morbidity, surgical site infections, postoperative pain and hospital stay.[65] However, no difference in postoperative general (non-surgical) morbidity and mortality was found. Outcome after CLR has been shown to be worse than after OR. Previous studies have found an increased complication rate,[70,71,84] increased wound infection rate,[70,71] an increased anastomotic leakage rate and a decreased disease free survival (in stage II disease) after CLR. [70,76]

In the present study, outcome, including mortality, was significantly better after LR than after OR, which is in contrast with the results of a Cochrane review by Schwenk at al.[65] Data from this review were exclusively derived from RCTs, in which patient selection may have led to inclusion of a group of patients with a lower risk for postoperative morbidity and mortality than might be found in the general population. The present study is an observational study using data from a prospective, nationwide registry including nearly all patients who underwent an elective colorectal cancer resection in the Netherlands. A previous study by McCloskey et al. reported that the positive effect of LR on outcome might be stronger in more high-risk patients than in the low-risk patients selected for RCTs.[66] This may explain the more favourable outcome, including a lower risk for mortality, after LR we found in the present study.
This study also reports no differences in outcome between CLR and OR. A possible explanation could be found in the reasons for conversion: as Yang stated, a pre-emptive conversion, due to a lack of progression or unclear anatomy, could be less unfavourable

than a reactive conversion due to an intra-operative complication[84]. Possibly, extensive training and increasing experience with laparoscopic techniques in the Netherlands has led to a tendency to early conversion of the procedure, before complications may occur. This could explain why in this study outcome after CLR was similar to outcome after primary OR. Unfortunately, detailed information on the reasons for and timing of conversion was not available in the DSCA.

We found a large variance in LR rates between hospitals in the Netherlands, which was not reduced after correction for differences in case-mix. Possibly, these differences could be explained by surgeon volume rather than hospital volume. In contrast, only four hospitals had a significantly higher CLR rate after adjustments for case-mix. The high CLR rates in these hospitals may indicate inadequate skills and experience with laparoscopic techniques. However, as these outlier hospitals all had a considerable procedural volume, it is unlikely that these hospitals were at the start of their learning curves. Possibly, the high CLR rates in these hospitals reflect a different selection strategy for LR. However, we found no evidence that the association between LR or CLR and outcome varied between hospitals. Therefore we may conclude that, given current LR practice, hospital LR and CLR rates were not associated with outcomes of care.

The results of this study should be interpreted in the light of several limitations. First, although the dataset of the DSCA 2010 was more complete than several other national registries,[37] the estimated completeness of the DSCA in 2010 was 93%. Data comparisons with the independently collected data of the NCR showed no overall differences in patient, tumour, procedural and outcome data. Therefore, it is unlikely that the results would have been influenced to a great extent by the missing 7% of patients. Another limitation is the absence of detailed information on the extent of the laparoscopic technique (laparoscopic assisted, hand assisted, intra/extracorporeal anastomosis etc.) and size of the incision. As no strict definitions for LR or CLR were applied, there may be differences in interpretation between hospitals. Also, reasons for conversion were not available. A word of caution is justified for the easily overlooked fact that these results have been obtained given the careful selection process that preceded the choice for LR. This selection is apparent from the highly significant differences for almost all case-mix factors between the 3 operative groups shown in table 1. It is an integral part of the professional expertise with which these excellent results have been obtained, and should not lead to the premature conclusion that “as LR is better than OR, it should be used in (almost) all patients, irrespective of case-mix”. Also, we did not have any information on the experience of the operating surgeon. When all LR would be performed by specialized surgeons, and most OR would be performed by general surgeons the observed differences in outcome may be explained by the operating surgeon rather than the operative techniques used. However, in most Dutch hospitals the majority of colorectal resections (laparoscopic and open) are performed by specialized surgeons, so that this is not likely to be the entire explanation. Last, no information on long

term results, such as long term survival or recurrence rates, is currently available. Although it would be most interesting to see whether the same long-term results can be achieved in routine practice as in clinical trials, these outcome measures were not yet available, as the DSCA was only initiated in 2009. Further studies will address long term results of different operative techniques, when data become available.

In 2007 the Dutch healthcare inspection presented a report on the safety of laparoscopic surgery, concluding that quality assurance for laparoscopic techniques was insufficient. As a result, the Dutch society for endoscopic surgery developed a quality assurance program for the introduction of new laparoscopic techniques. This system was based on a plan-do-check-act cycle involving the development of guidelines for the use and maintenance of instruments, a structured training and certification program and a registration and evaluation system. The present study shows how this quality assurance system has resulted in a safe and successful introduction of LR for colorectal cancer in the Netherlands. The LR rate in the Netherlands is high, with an acceptable conversion rate. Results from this large national database show that a LR for colorectal cancer is a safe approach. Short-term outcome after LR is better than after OR, even after correction for case-mix, while outcome after CLR was not different from outcome after OR. Although we found differences in LR and CLR rates between hospitals, we found no evidence that these differences affect the outcomes of care.

## Acknowledgements

We thank all surgeons, registrars, physician assistants and administrative nurses that registered all the patients in the DSCA, as well as the Dutch Surgical Colorectal Audit group (O.R.C. Busch, R.M. van Dam, E. van der Harst, M.L.E.A. Jansen-Landheer, Th.M. Karsten, J.H.J.M. van Krieken, W.G.T. Kuijpers, V.E. Lemmens, E.R. Manusama, H.J.T. Rutten, Prof. dr. C.J.H. van de Velde, T. Wiggers) and the methodological board for their advice.

## References


1. Jacobs M, Verdeja JC, Goldstein HS. Minimally invasive colon resection (laparoscopic colectomy). Surg Laparosc Endosc 1991;1:144-50.
2. Berends FJ, Kazemier G, Bonjer HJ, et al. Subcutaneous metastases after laparoscopic colectomy. Lancet 1994;344:58.
3. Lacy AM, Garcia-Valdecasas JC, Delgado S, et al. Laparoscopy-assisted colectomy versus open colectomy for treatment of non-metastatic colon cancer: a randomised trial. Lancet 2002;359:2224-9.
4. Cost. study group A comparison of laparoscopically assisted and open colectomy for colon cancer. N Engl J Med 2004;350:2050-9.
5. Buunen M, Veldkamp R, Hop WC, et al. Survival after laparoscopic surgery versus open surgery for colon cancer: long-term outcome of a randomised clinical trial. Lancet Oncol 2009;10:44-52.
6. Jayne DG, Guillou PJ, Thorpe H, et al. Randomized trial of laparoscopic-assisted resection of colorectal carcinoma: 3-year results of the UK MRC CLASICC Trial Group. J Clin Oncol 2007;25:3061-8.
7. Kuhry E, Schwenk WF, Gaupset R, et al. Long-term results of laparoscopic colorectal cancer resection. Cochrane Database Syst Rev 2008:CD003432.
8. Schwenk W, Haase O, Neudecker J, et al. Short term benefits for laparoscopic colorectal resection. Cochrane Database Syst Rev 2005:CD003145.
9. McCloskey CA, Wilson MA, Hughes SJ, et al. Laparoscopic colorectal surgery is safe in the high-risk patient: a NSQIP risk-adjusted analysis. Surgery 2007;142:594-7; discussion 7 e1-2.
10. Schwenk W, Bohm B, Muller JM. Postoperative pain and fatigue after laparoscopic or conventional colorectal resections. A prospective randomized trial. Surg Endosc 1998;12:1131-6.
11. Stage JG, Schulze S, Moller P, et al. Prospective randomized study of laparoscopic versus open colonic resection for adenocarcinoma. Br J Surg 1997;84:391-6.
12. Wexner SD, Cohen SM, Ulrich A, et al. Laparoscopic colorectal surgery–are we being honest with our patients? Dis Colon Rectum 1995;38:723-7.
13. Lu KC, Cone MM, Diggs BS, et al. Laparoscopic converted to open colectomy: predictors and outcomes from the Nationwide Inpatient Sample. Am J Surg 2011;201:630-4.
14. Agha A, Furst A, Iesalnieks I, et al. Conversion rate in 300 laparoscopic rectal resections and its influence on morbidity and oncological outcome. Int J Colorectal Dis 2008;23:409-17.
15. Gervaz P, Pikarsky A, Utech M, et al. Converted laparoscopic colorectal surgery. Surg Endosc 2001;15:827-32.
16. Li JC, Lee JF, Ng SS, et al. Conversion in laparoscopic-assisted colectomy for right colon cancer: risk factors and clinical outcomes. Int J Colorectal Dis 2010;25:983-8.
17. Le Moine MC, Fabre JM, Vacher C, et al. Factors and consequences of conversion in laparoscopic sigmoidectomy for diverticular disease. Br J Surg 2003;90:232-6.
18. Belizon A, Sardinha CT, Sher ME. Converted laparoscopic colectomy: what are the consequences? Surg Endosc 2006;20:947-51.
19. Ptok H, Kube R, Schmidt U, et al. Conversion from laparoscopic to open colonic cancer resection - associated factors and their influence on long-term oncological outcome. Eur J Surg Oncol 2009;35:1273-9.
20. Schouten LJ, Hoppener P, van den Brandt PA, et al. Completeness of cancer registration in Limburg, The Netherlands. Int J Epidemiol 1993;22:369-76.
21. Schouten LJ, Jager JJ, van den Brandt PA. Quality of cancer registry data: a comparison of data provided by clinicians with those of registration personnel. Br J Cancer 1993;68:974-7.
22. Kolfschoten NE, Marang van de Mheen PJ, Gooiker GA, et al. Variation in case-mix between hospitals treating colorectal cancer patients in the Netherlands. Eur J Surg Oncol 2011.
23. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40:373-83.
24. Daley J, Khuri SF, Henderson W, et al. Risk adjustment of the postoperative morbidity rate for the comparative assessment of the quality of surgical care: results of the National Veterans Affairs Surgical Risk Study. J Am Coll Surg 1997;185:328-40.
25. Khuri SF, Daley J, Henderson W, et al. Risk adjustment of the postoperative mortality rate for the comparative assessment of the quality of surgical care: results of the National Veterans Affairs Surgical Risk Study. J Am Coll Surg 1997;185:315-27.
26. Daly L. Simple SAS macros for the calculation of exact binomial and Poisson confidence limits. Comput Biol Med 1992;22:351-61.
27. The National Bowel Cancer Audit annual report 2010.
28. Bilimoria KY, Bentrem DJ, Merkow RP, et al. Laparoscopic-assisted vs. open colectomy for cancer: comparison of short-term outcomes from 121 hospitals. J Gastrointest Surg 2008;12:2001-9.
29. Kiran RP, El-Gazzaz GH, Vogel JD, et al. Laparoscopic approach significantly reduces surgical site infections after colorectal surgery: data from national surgical quality improvement program. J Am Coll Surg 2010;211:232-8.
30. Jayne DG, Thorpe HC, Copeland J, et al. Five-year follow-up of the Medical Research Council CLASICC trial of laparoscopically assisted versus open surgery for colorectal cancer. Br J Surg 2010;97:1638-45.
31. Yang C, Wexner SD, Safar B, et al. Conversion in laparoscopic surgery: does intraoperative complication influence outcome? Surg Endosc 2009;23:2454-8.

32. van Gijn W, Wouters MW, Peeters KC, et al. Nationwide outcome registrations to improve quality of care in rectal surgery. An initiative of the European Society of Surgical Oncology. J Surg Oncol 2009;99:491-6.

# Part 4

Using clinical auditing to evaluate quality of care and give transparency to all stakeholders

Chapter 7

# Evaluating the validity of quality indicators for colorectal cancer care

G.A. Gooiker[1], N.E. Kolfschoten[1], E. Bastiaannet[1], C.J.H. van de Velde[1], E.H. Eddes[2], E. van der Harst[3], T. Wiggers[4], F.R. Rosendaal[5], R.A.E.M. Tollenaar[1] and M.W.J.M. Wouters[1]
On behalf of the Dutch Surgical Colorectal Audit group*

[1] Leiden University Medical Center, department of surgery1, and department of epidemiology5, Leiden
[2] Deventer hospital, department of surgery, Deventer
[3] Maasstad hospital, department of surgery, Rotterdam
[4] University Medical Center Groningen, department of surgery, Groningen
[*] The Dutch Surgical Colorectal Audit group: W.A. Bemelman, O.R.C. Busch, R.M. van Dam, M.L.E.A. Jansen-Landheer, Th.M. Karsten, J.H.J.M. van Krieken, W.G.T. Kuijpers, V.E. Lemmens, E.R. Manusama, W.J.H.J Meijerink, H.J.T. Rutten

J Surg Oncol. 2013 Dec;108(7):465-71.

## Abstract

### Background
Quality indicators (QI) have been developed to measure quality of colorectal cancer care in the Netherlands. The aim of this study is to evaluate if these QI correlate with each other (construct validity) and if these QI consistently assess the quality of colorectal cancer care in a hospital (internal consistency).

### Methods
The performance of 85 hospitals participating in the Dutch Surgical Colorectal Audit between the 1st of January 2010 and 31st of December 2010, were evaluated on nine QI: three process indicators for colon cancer, three process indicators for rectal cancer and three outcome indicators. Correlations between all QI were evaluated for colon and rectal cancer care separately and consistency between all QI was assessed.

### Results
Hospital performance on the nine QI ranged widely. Most evaluated process indicators for colorectal cancer care did not correlate with each other, but were associated with better hospital specific patient outcomes. There was little consistency between any of the combinations of process and outcome indicators in assessing hospital performance.

### Conclusion
QI on colorectal cancer care do provide complementary information, but individual QI are not suitable as a surrogate measure for the quality of colorectal cancer care.
More comprehensive measures are needed for true assessment of hospital performance.

## Introduction

In order to increase transparency and improve quality of healthcare, numerous organisations have developed quality indicators (QI) to measure the quality of colorectal cancer care. [2,4-6] QI are defined as "measurable aspects of care that reflect the quality of care" and serve as benchmarks by which healthcare providers, payers and policy makers can measure processes and outcomes of care.
Most often, QI reflect process measures. Process measures have the advantage that data are usually readily available from (administrative) databases and the influence of patient or tumour characteristics ('case-mix') is limited. Also, process measures usually are actionable. Outcome measures, on the contrary, reflect the results of care for the patient and therefore have a more intuitive relation with quality of care, but they are highly influenced by case-mix factors and more difficult to obtain.[7]

In the Netherlands, the Healthcare Transparency Program (HTP), a governmental project introduced to coordinate the development of quality indicators, has defined eight indicators for colorectal cancer care. [8-12] In addition to the HTP, the Dutch Healthcare Inspectorate (DHI) uses "the unplanned reoperation rate after colorectal surgery" as a QI for colorectal cancer care. [13,14] All these QI have been developed in expert consensus meetings and are derived from (inter-)national evidence-based guidelines. The QI were developed with apprehension of the criteria for a good quality indicator and are regularly revised and evaluated (see textbox).
These QI were originally developed as a screening tool for the quality of care, with the assumption that process indicators, reflecting the organization of care, correlate with each other and with patient outcomes at a hospital level: if a hospital performs well on one indicator, it will perform well in other areas. However, a consistent relation between favourable results on process indicators and patient outcome has not always been shown. [15,16] Moreover, adherence to process measures may even have led to unintended harm. [17]
With the increasing use and public reporting of hospital-specific QI results, a more robust scientific base is needed. Validity of QI can be evaluated by testing several aspects, including criterion, construct and content validity and internal consistency (see textbox). Although reports are emerging on the criterion validity of individual indicators, the other aspects remain underexposed.

To monitor and benchmark the quality of colorectal cancer care in the Netherlands, in 2009 the Dutch Surgical Colorectal Audit (DSCA) was initiated (www.clinicalaudit.nl). In this national registry, detailed data on both diagnostic and therapeutic processes and short-term outcomes are collected. Also, data on patient characteristics are registered, allowing adjustment for differences in case-mix between hospitals.
The detailed registration of the DSCA provides the opportunity to investigate the validity of QI for colorectal cancer care in the Netherlands on two aspects:

What makes a good Quality indicator (QI):[1,2]
**Important**: the QI must be relevant, involve a high-risk condition or represent an opportunity for improvement
**Scientific acceptable:** the measure must be reliable and valid.
- **Reliability** means that the indicator gives the same result on repeated measures; this requires the use of uniform definitions and complete data.
- **Validity**[3] means that the indicator measures what it is intended to measure: quality. This requires first that the methodological quality is good, i.e. that differences in case-mix and random variation are taken into account (criterion validity). Secondly, the number of indicators has to be a representative sample to give 'appropriate coverage' of the quality of care in a hospital (content validity) and the indicators have to correlate with each other (internal consistency). Third, an indicator has to be correlated with quality, thus with other performance measures and patient outcomes (construct validity).

**Feasible**: data for reporting QI should be feasible to obtain
**Usable:** the intended audience must understand results of the measure

**construct validity**: do process indicators correlate with each other and with short-term outcomes?
**Internal consistency**: do indicators consistently assess the quality of colorectal cancer care in a hospital?

## Methods

### Data

The dataset was retrieved from the DSCA, a nationwide, web-based database in which patient-, tumor-, diagnostic- and treatment characteristics as well as pathology and outcome data are registered for patients that undergo a resection of a primary colorectal carcinoma in the Netherlands. Details of this dataset regarding data collection and methodology have been published previously.[18] The dataset was based on Dutch evidence-based guidelines (www.oncoline.nl). After crosschecking with the Netherlands Cancer registry (NCR), estimated completeness in the year 2010 was 93%.[19,20] All Dutch hospitals participated in the registry. (www.clinicalaudit.nl)

### Hospitals and patients

All hospitals participating in the DSCA between the 1st of January 2010 and 31st of December 2010, were evaluated. Patients of hospitals that registered less than 30 patients with a date of surgery in 2010 were excluded. Furthermore, as case-mix correction is imperative for evaluating outcome of care, hospitals that failed to fill in the required case-mix factors for more than 15% of the registered patients were excluded.
Analyses of hospital performance were done on all patients with a date of surgery between the 1st of January 2010 and 31st of December 2010, and inclusion before March 15$^{th}$ 2011. Patients that underwent an urgent or acute resection, or were treated for a recurrence of a colorectal carcinoma or multiple synchronous colorectal tumours were excluded.

At March 15th 2011, 91 hospitals registered a total of 8835 evaluable patients with a date of surgery between January 1 and December 31 2010 in the DSCA. After exclusion of hospitals that registered less than 30 evaluable patients (3 hospitals, 16 patients), patients with multiple synchronous tumours (253 patients) and urgent and acute patients (1228 patients) and hospitals that had not filled in detailed case-mix factors for more than 15% of their patients (4 hospitals, 367 patients) a total of 85 hospitals treating 6971 patients, were included for analyses.

## Quality indicators

A team of medical experts developed QI used in the DSCA, using the Delphi method. [21] All QI are based on (inter-)national evidence based guidelines and reflect guideline adherence.[9,13] The definitions of five QI are equal to the QI used by two governmental agencies, the DHI and HTP. All QI are described in detail in the appendix.

## Process indicators

Hospital performance on process indicators was calculated for colon and rectum cancer care separately, as the process of care for patients with a colon carcinoma encompasses different aspects than for patients with a rectum carcinoma.
The following process indicators were selected in this study, each reflecting different stages of colorectal cancer care.

*Process indicators for the treatment of colon cancer:*

- The percentage of patients, who had a '*complete colonoscopy'* before the resection. Complete colonoscopy is defined as a complete visualisation of the colon until Bauhini's valve by colonoscopy or ct-colonography
- The percentage of patients who had *adequate staging* by visualisation of lungs and liver before resection, by either CT-thorax or X-thorax and CT-abdomen or ultrasound respectively.
- The percentage of patients for whom *more than 10 lymph nodes* retrieved are examined pathologically after the resection.

*Process indicators for the treatment of rectal cancer:*

- The percentage of patients, who are *discussed in a multidisciplinary meeting* before the resection. In a multidisciplinary meeting participation of at least a surgical oncologist, medical oncologist, pathologist, radiologist and radiotherapist are required.
- The percentage of patients who had *adequate tumour staging* before the resection by a MRI or CT of the pelvis.
- The percentage of patients with a *reported circumferential margin (CRM*) of the resection specimen in the pathology report.

## Outcome indicators

The following outcome indicators were selected for both colon and rectal cancer care:

- The percentage of patients who had an *unplanned reoperation*.
- The *risk-adjusted 30-day mortality*, defined as mortality within 30 days after resection.
- The *risk-adjusted morbidity* rate, defined as patients with postoperative complications requiring a re-intervention, patients who deceased or patients with a postoperative length of stay longer than 14 days.

The risk-adjusted mortality rate was calculated as the quotient of the observed mortality rate, and the expected mortality rate, multiplied by the average mortality rate in the population. An expected mortality was calculated for each patient using a multivariate model. The model included all case-mix factors registered in the DSCA: age, gender, comorbidity (Charlson Comorbidity Index), previous abdominal surgery, Body Mass Index, American Society of Anaesthesiologists classification (ASA), local tumour invasiveness (T-stage), disseminated disease (M-stage), neoadjuvant (chemo-)radiation therapy, preoperative tumour complications, type of resection, additional resection for tumour invasion and/or metastasis. The average expected mortality in a group of patients formed the expected mortality rate of that group.[7]
In a similar way, hospital risk-adjusted morbidity rates were calculated.

## Statistical analyses

QI intend to measure the quality of care in a hospital, therefore all analyses were performed *at a hospital level*. First, means and ranges of hospital performance were calculated for each of the 9 selected QIs. After this, different metrics were used to evaluate construct validity and internal consistency.

## Construct validity

The construct validity describes how process indicators correlate with each other and with short-term outcomes.[3] To evaluate criterion validity, the correlations between all combinations of process and outcome QI were calculated using a Pearson correlation test and results were presented in matrix scatterplots. Next, a Poisson regression, for colon and rectal cancer apart, was performed to test the correlation between process indicators and outcome indicators.

## Internal consistency

The content validity describes if indicators give appropriate coverage of the overall quality of care and do consistently assess the quality of colorectal cancer care in a hospital (internal consistency).[3] To evaluate this, internal consistency between all process and outcome indicators for colon and rectal cancer care, was measured using the Cronbach's alpha. Crohnbach's Alpha computes the inter-item correlations or covariances of all pairs

of variables and Crohnbach's Alpha statistic for the scale formed of them. The measure $\alpha$, indicates how different items test the same concept and is often used to validate psychometric tests, questionnaires and other scoring systems.[22] An $\alpha$ statistic $\geq 0.80$ is considered as good consistency, an $\alpha$ statistic 0.60- 0.80 as acceptable consistency. In the first step the inter-item correlation between all indicators is calculated. In the following steps, the indicator with the lowest value is removed from the model and the inter-item correlation is recalculated.
All statistics were performed using STATA version 10.0.

## Results

From January 1 to December 31 2010, 85 hospitals included a total of 6971 patients in the DSCA, 4732 patients with colon carcinoma and 2239 patients with rectal carcinoma. Hospital performance on the nine QI ranged widely (table 1).

**Table 1**: Hospital performance on process and outcome indicators for colorectal cancer care of 85 hospitals in the Netherlands treating patients that underwent a resection of a primary colorectal carcinoma between 1 January 2010 and 31 December 2010.

| Hospital level | Indicator | Mean (%) | Median (%) | IQR (%) |
|---|---|---|---|---|
| Colon | Imaging lung and liver | 91 | 95 | 89-98 |
| | Complete colonoscopy | 71 | 69 | 60-76 |
| | 10 or more lymph nodes | 79 | 79 | 73-86 |
| | Unplanned re-operation | 11 | 11 | jul-15 |
| | Risk-adjusted morbidity | 20.8 | 19.7 | 14.4-26.9 |
| | Risk-adjusted 30-day mortality | 3.6 | 2.9 | 1.4-5.2 |
| Rectal* | MRI pelvis | 88 | 94 | 84-100 |
| | MDC | 90 | 98 | 92-100 |
| | CRM | 62 | 64 | 50-81 |
| | Unplanned re-operation | 10 | 8 | 5-15 |
| | Risk-adjusted morbidity | 27.2 | 26.7 | 18.2-35.8 |
| | Risk-adjusted 30-day mortality | 2.1 | 0 | 0-3.5 |

MRI: MRI or CT of the pelvis for staging the tumor, before resection
MDC: discussion in multidisciplinary meeting, before resection
CRM: reported circumferential margin (CRM) of the resection specimen in the pathology report

### Construct validity

Figure 1 shows the correlation at a hospital level between all combinations of process and outcome indicators in matrix scatterplot for colon cancer (a) and rectal cancer (b). A significant correlation for colon cancer care was seen between 'complete imaging of liver and lungs' and 'more then 10 retrieved lymph nodes'. For rectal cancer care, there was a significant correlation between a high 'discussed in preoperative multidisciplinary meeting' and a lower hospital 'risk-adjusted mortality' rate.

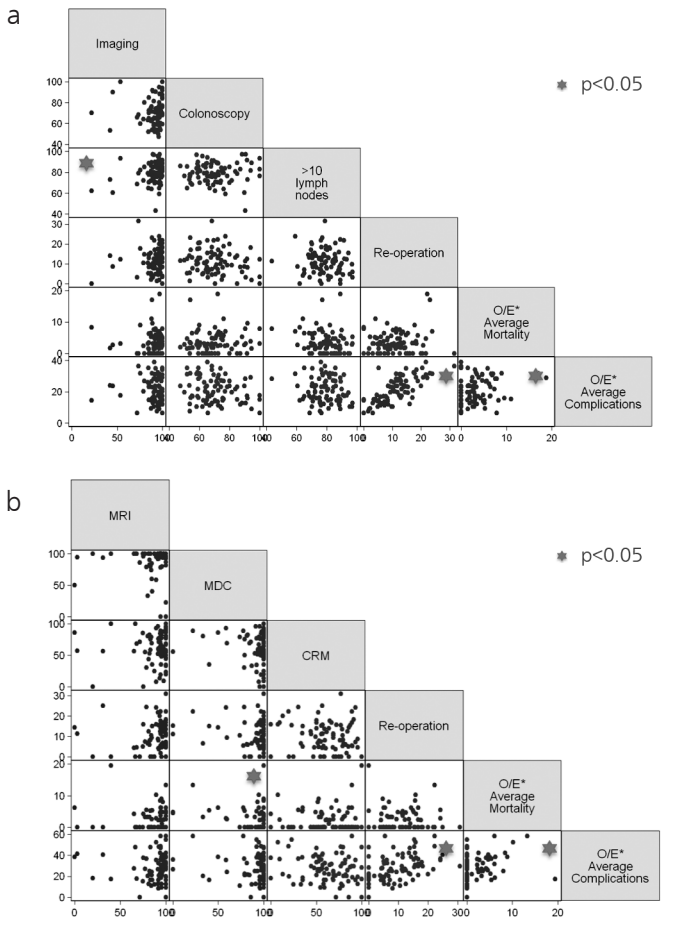



**Figure 1:** Matrix scatterplot for correlation at a hospital level between all combinations of process and outcome indicators for colon cancer (a) and rectal cancer (b)

**Table 2**: Average inter-item correlation and Cronbach's alpha for selected process indicators, measuring the performance of 85 hospitals treating patients that underwent a resection of a primary colorectal carcinoma between 1 January 2010 and 31 December 2010 in the Netherlands.

| Process indicators | | Cronbach's Alpha | Correlation |
|---|---|---|---|
| **Colon** | Imaging – colonoscopy – 10+nodes | 0.21 | 11.3 |
| **Rectum** | MRI – MDC – CRM | 0.10 | 15.8 |
| **Colorectal** | Imaging – colonoscopy – 10+nodes – MRI – MDC – CRM | 0.20 | 12.5 |

Table 2 shows the correlation between process indicators and outcome indicators at a hospital level using poison regression. In colon cancer care, a high rate of completed colonoscopies was correlated with a lower hospital specific 'unplanned re-operation rate' (RR 0.4 (0.3-0.8)), and lower risk-adjusted morbidity (RR 0.5 (0.4-0.8). The QI '10 or

more retrieved lymph nodes' was correlated with lower hospital specific morbidity (RR 0.5 (0.3-0.8) and lower '30-day mortality' (RR 0.1(0.05-0.4).
In rectal cancer care, a high rate of patients discussed in a multidisciplinary meeting was correlated with a lower hospital specific unplanned re-operation rate (RR 0.5 (0.4-0.8), lower risk-adjusted morbidity (RR 0.8 (0.6-0.9) and lower 30-day mortality (RR 0.2 (0.1-0.3). A high rate of reported circumferential margins was associated with lower re-operation rates (RR 0.6 (0.4-0.8) and risk-adjusted morbidity (RR 0.8 (0.7-0.9). A high rate of patients who had a preoperative MRI of the tumour was associated with lower 30-day mortality (RR 0.4 (0.2-0.7).

*Internal consistency*

The results of the consistency tests for all QI using average inter-item correlation and Crohnbach's alpha are shown in table 3 for colon and rectal cancer care separately. In both domains, there was little consistency between any of the tested combinations of process

**Table 3**: Evaluation of the correlation between process indicators and hospital specific patient outcomes, of 85 hospitals treating patients that underwent a resection of a primary colorectal carcinoma between 1 January 2010 and 31 December 2010 in the Netherlands. Correlation is tested by Poisson regression analysis. The results are shown for each selected outcome measure and are stratified for colon and rectal cancer care.

| Outcome | Process | Rate ratio | 95%CI | p-value |
|---|---|---|---|---|
| **Unplanned Re-operation *** | | | | |
| Colon | Imaging | 1.4 | 0.8-2.4 | 0.2 |
| | Colonoscopy | ***0.4*** | ***0.3-0.8*** | ***0.003*** |
| | 10 or more nodes | 0.6 | 0.3-1.3 | 0.2 |
| Rectum | MRI | 1.2 | 0.8-1.7 | 0.4 |
| | MDC | ***0.5*** | ***0.4-0.7*** | ***<0.001*** |
| | CRM | ***0.6*** | ***0.4-0.8*** | ***<0.001*** |
| **Risk-adjusted morbidity** | | | | |
| Colon | Imaging | 0.9 | 0.6-1.3 | 0.6 |
| | Colonoscopy | ***0.5*** | ***0.4-0.8*** | ***0.001*** |
| | 10 or more nodes | ***0.5*** | ***0.3-0.8*** | ***0.006*** |
| Rectum | MRI | 0.9 | 0.7-1.1 | 0.2 |
| | MDC | ***0.8*** | ***0.6-0.9*** | ***0.003*** |
| | CRM | ***0.8*** | ***0.7-0.9*** | ***0.002*** |
| **Risk-adjusted 30-day mortality** | | | | |
| Colon | Imaging | 0.7 | 0.3-1.4 | 0.3 |
| | Colonoscopy | 1.4 | 0.6-3.3 | 0.5 |
| | 10 or more nodes | ***0.1*** | ***0.05-0.4*** | ***<0.001*** |
| Rectum | MRI | ***0.4*** | ***0.2-0.7*** | ***0.001*** |
| | MDC | ***0.2*** | ***0.1-0.3*** | ***<0.001*** |
| | CRM | 1.0 | 0.5-1.8 | 0.9 |

Unplanned reoperation: not adjusted for casemix, % reresections/ hospital Morbidity and mortality: Adjusted for casemix: outcome Observed/Expected*average MRI: MRI or CT of the pelvis for staging the tumor, before resection. CRM: reported circumferential margin (CRM) of the resection specimen in the pathology report

and outcome indicators in assessing hospital performance. Only among the combined outcome indicators, acceptable consistency was observed, with Cronbachs $\alpha$ of 0.67 and 0.52 for colon and rectal cancer care respectively.

## Discussion

In this study, most evaluated process indicators for colorectal cancer care do not correlate with each other, but they are associated with better hospital specific patient outcomes. A high rate of completed colonoscopies, and 10 or more retrieved lymph nodes, were correlated with improved patient outcomes in colon cancer care. In rectal cancer care, better hospital outcomes were seen in hospitals with a high rate of patients discussed in a multidisciplinary meeting, a high rate of patients having a MRI before surgery and a high rate of reported circumferential margins. These results show that these process indicators have solid construct validity in the assessment of hospital performance in colorectal cancer care.
On the contrary, the evaluation of internal consistency showed minimal consistency in any of the tested combinations of both process and outcome indicators. This indicates that the current QI provide complementary information, reflecting all different stages of colorectal cancer care. Yet, QI are not suitable to act as a surrogate for the overall quality of care individually.

The present study builds on previous studies evaluating construct validity of QI. The construct validity of individual QI has been evaluated before. Especially the association between lymph node evaluation for colon cancer and survival, has been extensively studied.[8,16] At least, all evaluated QI have been tightly linked to patient outcomes in clinical trials or are included in clinical guidelines.[23,24] Yet, reports evaluating the construct validity or internal consistency of multiple QI are relatively scarce and show various results, depending on the domains covered by the QI.
For example, a study by Werner et al. evaluated the construct validity of the 'hospital compare measures': ten QI concerning the treatment of acute myocardial infection, heart failure and pneumonia in the USA. This study showed a modest correlation between the selected QI and risk-adjusted mortality rates. [25] In another study by Bradley et al. process measures designed to evaluate the quality of care for acute myocardial infarction alone, were evaluated, all largely reflecting the same domain: medication prescription practices. These QI did not correlate with hospitals' short-term mortality rates, but the inter-item correlation was acceptable. [26]
Similarly, a recent study evaluating QI of the US Surgical Care Improvement Project, showed that compliance to perioperative processes of care, covering infection and venous-thrombo-embolism prevention, was not correlated to risk-adjusted surgical outcomes. [15] Yet, these processes relate to secondary and less prevalent outcomes, which can explain the lack of correlation with risk-adjusted mortality rates that are likely influenced by many factors

independent of the selected QI. Comparable results were seen in a study examining the correlation between the Leapfrog Safe Practice Scores results and risk-adjusted mortality. [27] In the present study, a positive correlation was seen between most of the evaluated process indicators and hospital specific outcomes. All evaluated process indicators represent different critical components in the multi-disciplinary work-up and treatment of patients who undergo an elective resection of colorectal cancer. This implicates that hospitals where these organisational features are well implemented, have superior patient outcomes.
On the contrary, the present study showed weak consistency among the QI for colorectal cancer care. These process measures thus provide information that is complementary, each reflecting a unique dimension of quality. But none of these individual indicators can represent overall quality of care, since they lack content validity. The myriad components of high quality colorectal cancer care seem thus to be too complex to be captured in individual indicators.

Given this information, QI can be useful tool for screening for substandard care, though they are not suitable to assess the quality of colorectal cancer care as a whole. In the Netherlands, quality improvement programs have mainly focussed on measuring a small selection of process and outcome indicators. Meanwhile, the role of QI gets increasingly prominent in modern health care, with increasing use in pay for performance initiatives and public reporting. But, the provided information about the quality in a hospital can be misleading, as the assessment of the quality of care in a hospital relies on the choice of which QI is used.
Still, there is considerable variation in the quality of colorectal cancer care in the Netherlands. [28,29] The wide range in hospital performance on the QI in the present study, confirms these findings and show the need for quality measurement to improve the quality of care and reduce variation in hospital performances. [30] Yet, what aspects should be measured to truly measure hospital performance and provide meaningful information for both physicians and external parties is not elucidated. Ideally quality measurement should consist of measuring whether the right patient receives the right treatment at the right time and whether that treatment is effective. Further research is needed to evaluate the role of composite measures that capture all these aspects.
A better alternative is monitoring hospital performance by a valid outcome registration ('clinical audit'): constant monitoring of patient characteristics, the process of diagnostics and treatment and patient outcomes. In the last decade there have been various successful examples of auditing. including the US, UK and the Nordic countries, where auditing has led to impressive improvements in survival.[31-33] A beneficial 'side-effect' of auditing is its value for research, providing reliable population based observational data, when a nearly complete percentage of patients is entered. Also, auditing is an ideal platform for implementing quality improvement projects.

The present study is the first, testing the construct validity and internal consistency

of generally used QIs for colorectal cancer care in a population based, clinical database, covering almost all hospitals in the Netherlands and at least 85% of the patient population. Nonetheless, the results of this study should be interpreted in the light of several limitations. The exclusion of seven hospitals, which were underreporting, could have interfered with our findings. If these under*reporting* hospitals, were also under*performing* hospitals, a stronger correlation between the selected indicators and outcomes may have been observed.
In addition, although the completeness of the DSCA is high and estimated to be over 93% in 2010, patient selection could have influenced the results. The comparison with the NCR showed a similar distribution of baseline characteristics, procedures and postoperative outcomes. Although these results were not suggestive for patient selection, this cannot be excluded.
The sample size of 85 hospitals, (and the relative little variation on some QI) can limit the power of the observed correlations. Future projects in which European audits will be combined, will allow for more detailed evaluations of the validity of QI used in colorectal cancer care.[34]
Last, in this study a positive correlation was found between most of the evaluated process indicators and hospital specific outcomes. This implicates that hospitals that have good organisational features also have superior patient outcomes. The use of process indicators is mainly based on this assumption. Yet, the observed correlation can be confounded since it is possible that hospitals that perform well on the selected process indicators, also do other things that actually account for these superior outcomes.

In conclusion, although nearly all evaluated QI for colorectal cancer care were associated with improved hospital specific patient outcomes, our results show little consistency among the QI. Therefore QI do provide complementary information, but they cannot function as a surrogate for the quality of colorectal cancer care.
Current quality improvement projects focus mainly on a selection of process measures that are evidently insufficient to measure the overall quality of care.
Monitoring of both processes and outcomes, adjusted for differences in case-mix, in a 'clinical audit' is a better alternative and efforts should be done to implement process and outcome monitoring in daily routine. Further research is needed for the development of valid QI and composite measures of quality.

## References


1. Dimick JB. What makes a "good" quality indicator? Arch Surg 2010;145:295.
2. Patwardhan M, Fisher DA, Mantyh CR, et al. Assessing the quality of colorectal cancer care: do we have appropriate quality measures? (A systematic review of literature). Journal of evaluation in clinical practice 2007;13:831-45.
3. Bland JM, Altman DG. Statistics Notes: Validating scales and indexes. BMJ 2002;324:606-7.
4. McGory ML, Shekelle PG, Ko CY. Development of quality indicators for patients undergoing colorectal cancer surgery. J Natl Cancer Inst 2006;98:1623-33.
5. Gagliardi AR, Simunovic M, Langer B, Stern H, Brown AD. Development of quality indicators for colorectal cancer surgery, using a 3-step modified Delphi approach. Canadian journal of surgery Journal canadien de chirurgie 2005;48:441-52.
6. Vergara-Fernandez O, Swallow CJ, Victor JC, et al. Assessing outcomes following surgery for colorectal cancer using quality of care indicators. Canadian journal of surgery Journal canadien de chirurgie 2010;53:232-40.
7. Kolfschoten NE, Marang van de Mheen PJ, Gooiker GA, et al. Variation in case-mix between hospitals treating colorectal cancer patients in the Netherlands. Eur J Surg Oncol 2011;37:956-63.
8. Wong SL, Ji H, Hollenbeck BK, Morris AM, Baser O, Birkmeyer JD. Hospital lymph node examination rates and survival after resection for colon cancer. JAMA 2007;298:2149-54.
9. Kyllonen LE. Obstruction and perforation complicating colorectal carcinoma. An epidemiologic and clinical study with special reference to incidence and survival. Acta Chir Scand 1987;153:607-14.
10. Kapiteijn E, Marijnen CA, Nagtegaal ID, et al. Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer. N Engl J Med 2001;345:638-46.
11. Zegers M, de Bruijne MC, Spreeuwenberg P, Wagner C, Groenewegen PP, van der Wal G. Quality of patient record keeping: an indicator of the quality of care? Qual Saf Health Care 2011.
12. Wei JT, Miller EA, Woosley JT, Martin CF, Sandler RS. Quality of colon carcinoma pathology reporting: a process of care study. Cancer 2004;100:1262-7.
13. Watt AM, Faragher IG, Griffin TT, Rieger NA, Maddern GJ. Self-expanding metallic stents for relieving malignant colorectal obstruction: a systematic review. Ann Surg 2007;246:24-30.
14. Morris AM, Baldwin LM, Matthews B, et al. Reoperation as a quality indicator in colorectal surgery: a population-based analysis. Ann Surg 2007;245:73-9.
15. Nicholas LH, Osborne NH, Birkmeyer JD, Dimick JB. Hospital process compliance and surgical outcomes in medicare beneficiaries. Arch Surg 2010;145:999-1004.
16. Parsons HM, Tuttle TM, Kuntz KM, Begun JW, McGovern PM, Virnig BA. Association between lymph node evaluation for colon cancer and node positivity over the past 20 years. JAMA 2011;306:1089-97.
17. Hawn MT, Vick CC, Richman J, et al. Surgical site infection prevention: time to move beyond the surgical care improvement program. Ann Surg 2011;254:494-9; discussion 9-501.
18. Kolfschoten NE, Marang van de Mheen PJ, Gooiker GA, et al. Variation in case-mix between hospitals treating colorectal cancer patients in the Netherlands. Eur J Surg Oncol 2011.
19. Schouten LJ, Hoppener P, van den Brandt PA, Knottnerus JA, Jager JJ. Completeness of cancer registration in Limburg, The Netherlands. Int J Epidemiol 1993;22:369-76.
20. Schouten LJ, Jager JJ, van den Brandt PA. Quality of cancer registry data: a comparison of data provided by clinicians with those of registration personnel. Br J Cancer 1993;68:974-7.
21. Jones J, Hunter D. Consensus methods for medical and health services research. BMJ 1995;311:376-80.
22. Bland JM, Altman DG. Cronbach's alpha. BMJ 1997;314:572.
23. Comprehensive Cancer Center the Netherlands. National Guideline Colon carcinoma In; 2008.
24. Netherlands CCCt. National Guideline Rectal Carcinoma In; 2008.
25. Werner RM, Bradlow ET. Relationship between Medicare's hospital compare performance measures and mortality rates. JAMA 2006;296:2694-702.
26. Bradley EH, Herrin J, Elbel B, et al. Hospital quality for acute myocardial infarction: correlation among process measures and relationship with short-term mortality. JAMA 2006;296:72-8.
27. Kernisan LP, Lee SJ, Boscardin WJ, Landefeld CS, Dudley RA. Association between hospital-reported Leapfrog Safe Practices Scores and inpatient mortality. JAMA 2009;301:1341-8.
28. Elferink MA, Wouters MW, Krijnen P, et al. Disparities in quality of care for colon cancer between hospitals in the Netherlands. Eur J Surg Oncol 2010;36 Suppl 1:S64-73.
29. Elferink MA, Krijnen P, Wouters MW, et al. Variation in treatment and outcome of patients with rectal cancer by region, hospital type and volume in the Netherlands. Eur J Surg Oncol 2010;36 Suppl 1:S74-82.
30. van Leersum NJ, Kolfschoten NE, Klinkenbijl JH, Tollenaar RA, Wouters MW. ['Clinical auditing', a novel tool for quality assessment in surgical oncology]. Ned Tijdschr Geneeskd 2011;155:A4136.
31. Wibe A, Moller B, Norstein J, et al. A national strategic change in treatment policy for rectal cancer--implementation of total mesorectal excision as routine treatment in Norway. A national audit. Dis Colon Rectum 2002;45:857-66.
32. Khuri SF, Daley J, Henderson WG. The comparative assessment and improvement of quality of surgical care in the Department of Veterans Affairs. Arch Surg 2002;137:20-7.

33. Nedrebo BS, Soreide K, Eriksen MT, et al. Survival effect of implementing national treatment strategies for curatively resected colonic and rectal cancer. Br J Surg 2011;98:716-23.
34. van Gijn W, Wouters MW, Peeters KC, van de Velde CJ. Nationwide outcome registrations to improve quality of care in rectal surgery. An initiative of the European Society of Surgical Oncology. J Surg Oncol 2009;99:491-6.

Chapter 8

# Combining process indicators to evaluate quality of care for surgical colorectal cancer patients; are scores consistent with short-term outcome?

N.E. Kolfschoten[1], G.A. Gooiker[1], E. Bastiaannet[1], N.J. van Leersum[1], C.J.H. van de Velde[1], E.H. Eddes[3], P.J. Marang-van de Mheen[4], J. Kievit[1,4], E. van der Harst[5], T. Wiggers[6], M.W.J.M. Wouters[2], R.A.E.M Tollenaar[1], on behalf of the Dutch Surgical Colorectal Audit group*

[1] Leiden University Medical Centre, Dept of Surgery,
[2] Dutch Cancer Institute, Amsterdam, Dept of Surgery,
[3] Deventer Hospital, Dept of Surgery,
[4] Leiden University Medical Center, Dept of Medical Decision Making,
[5] Maasstadziekenhuis Rotterdam, Dept of Surgery,
[6] Groningen University Hospital, Dept of Surgery,
* W.A. Bemelman, O.R.C. Busch, R.M. van Dam, M.L.E.A. Jansen-Landheer, Th.M. Karsten, J.H.J.M. van Krieken, W.G.T. Kuijpers, V.E. Lemmens, E.R. Manusama, W.J.H.J Meijerink, H.J.T. Rutten

*BMJ Quality and Safety. 2012 Jun;21(6):481-9.*

## Abstract

### Objective
To determine if composite measures based on process indicators are consistent with short-term outcome indicators in surgical colorectal cancer care.

### Design
Longitudinal analysis of consistency between composite measures based on process indicators and outcome indicators for 85 Dutch hospitals.

### Setting
The Dutch Surgical Colorectal Audit database, the Netherlands

### Participants
4732 elective patients with colon carcinoma and 2239 with rectum carcinoma, treated in 85 hospitals were included in the analyses.

### Main outcome measures
All available process indicators were aggregated into five different composite measures. We investigated the association of the different composite measures with risk-adjusted postoperative mortality and morbidity, at patient and hospital level.

### Results
At a patient level, only one of the composite measures was negatively associated with morbidity for rectum carcinoma. At a hospital level, we found a strong negative association between composite measures and hospital mortality and morbidity rates for rectum carcinoma ($p<0.05$), and hospital morbidity rates for colon carcinoma.

### Conclusions
For individual patients, a high score on the composite measures based on process indicators is not associated with better short-term outcome. However, at a hospital level, a good score on the composite measures based on process indicators was consistent with more favourable risk-adjusted short-term outcome rates.

## Introduction

Society increasingly demands information on hospitals' quality of care. Quality of health care is defined as "... the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge[85]." Currently, quality indicators play a key role in monitoring quality of surgical care in many countries. Based on the Donabedian paradigm, quality indicators are often characterized as structure, process or outcome indicators.[10] Although outcome indicators are seen as 'the bottom line' of what surgeons do, valid outcome information, adjusted for a hospitals' *case-mix*, is commonly not available. Process indicators on the other hand, usually selected from evidence based guidelines, give a fair reflexion of what is done for a patient, and are more actionable than outcome indicators.[86] Therefore, most quality indicators are process indicators, based on the assumption that a 'good care process', e.g. being treated according to the evidence-based guideline, will lead to good outcome. At a hospital level this implies that a high percentage of patients treated according to the guideline will be related to favourable outcomes.

Most process indicators are selected on the basis of an assumed or proven relation with long-term or disease-free survival. However, long-term outcome reflects the quality of care delivered years before the outcome measurement and is therefore less indicative for the hospital's performance at that moment. Therefore, outcome of care is often measured by short-term outcome measures such as postoperative mortality, or major morbidity.[87] However, a clear relation between process indicators and short-term outcome indicators has not been established.[88,89] Moreover, though an individual process indicator may give useful information for targeted quality improvement programs, it is difficult to assess guideline adherence by means of a list of individual indicators: a hospital may have a high score on one indicator, but a low score on another indicator. Therefore, a summarizing measure is needed to give a comprehensive judgement of the care delivered in individual hospitals. For this purpose, indicators can be aggregated into a composite measure.[90] Reeves at all summed 5 commonly used methods of constructing composite outcome measures.[91]

To monitor and improve the quality of care colorectal cancer patients in the Netherlands, the Dutch Surgical Colorectal Audit (DSCA) was initiated in 2009. In the DSCA, patient and tumour characteristics together with diagnostic, treatment, pathology and outcome data are registered. Using this database, various composite process measures could be tested for their consistency with short-term outcome. The aim of the study was to determine whether a good score on composite process measures is associated with favourable short-term outcomes, on a patient and a hospital level.

## Methods

### Patients

The dataset was retrieved from the DSCA, a nation wide, web based database in which case-mix, diagnostic, treatment, pathology report and clinical outcome data are registered of patients that undergo a resection of a first presentation of a colorectal carcinoma in the Netherlands. The estimated completeness in the year 2010 compared to the National Cancer registry (NCR) was 93% (www.dsca.nl). Overall comparison with NCR was not suggestive for patient-selection (www.dsca.nl); the results of the comparison are communicated with the hospitals, encouraging completeness and validity of the data. For this study, no ethical approval was required.

#### *Inclusion and exclusion criteria*

All elective patients with a date of surgery between the 1st of January 2010 and 31st of December 2010 treated in one of the 92 Dutch Hospitals who were included in the DSCA before March 15$^{th}$ 2011 were evaluated. Patients who underwent an urgent or emergency resection, or were treated for a recurrence of a colorectal carcinoma or multiple synchronous colorectal tumours were excluded.

Hospitals that failed to register all patients in 2010 were excluded. Furthermore, as case-mix correction is imperative for evaluating outcome of care, hospitals that failed to fill in the required case-mix factors for more than 15% of the registered patients were excluded.

### Process indicators

As the process of care for colon cancer (CC) encompasses different aspects than for rectal cancer (RC), groups were analysed separately. The indicators used in this study were selected from the dataset of the DSCA. This dataset was constructed based on evidence based guidelines and extensive literature research, which resulted in a large dataset that was reduced to the present dataset by a team of medical experts using the Delphi method.[92,93] All variables available in the DSCA concerning (preoperative) guideline adherence were selected for evaluation. As postoperative guideline adherence, such as adjuvant chemotherapy for patients with stage 3 colon carcinoma is dependent on postoperative outcome,[94] postoperative processes of care (except for postoperative pathology reporting) were not included. For CC, four process indicators were selected, for RC, six process indicators were selected. For the indicator 'time to treatment < 7 weeks', records in which the date of first biopsy was not available were excluded for the indicator. As there are several reasons why a biopsy may not be obtained or the date of the biopsy would not be available (e.g. due to referral from another hospital), and no substitute for the date of diagnosis was available, the indicator was considered 'not indicated' for the patient. For all other indicators, the absence of information on the indicator was interpreted as 'indicator has not been met'; ratio being that when information on such important processes of care was unavailable to the caregiver, the process was not completed properly. This is in line with the current policy of the Dutch healthcare inspectorate.

## Outcome indicators

Outcome is registered at discharge or 30 days after the resection. Postoperative outcome was assessed by two separate outcome measures: postoperative mortality and major morbidity. Postoperative mortality was defined as mortality within 30 days after resection. Major morbidity was defined as any complication with serious consequences: leading to mortality, a reintervention, or to a length of stay longer than 14 days. Selected processes and outcome measures are shown in table 1.

**Table 1.** Process indicators in the Dutch Surgical Colorectal Audit

| Process indicator | Definition | Inclusion | Measures |
|---|---|---|---|
| Complete Colonoscopy | Was the colon completely visualized, caecal intubation performed, by colonoscopy or CT-colonograpy before the resection? | CC and RC | Staging |
| Diagnostic tests | Was preoperative visualisation of lungs and liver by either CT-thorax or X-thorax and CT-abdomen or ultrasound respectively, or Pet-CT performed? | CC* | Staging |
| MRI | Was the pelvis adequately staged with MR imaging before the resection? | RC | Staging |
| MDT | Was the patient discussed in a multidisciplinary team meeting? In a multidisciplinary team meeting at least a surgical oncologist, a medical oncologist, a pathologist a radiologist and a radiotherapist had to participate | RC | Staging |
| Radiotherapy | Did the patient receive adequate neoadjuvant radiation therapy? | Patients with RC with a clinical T2 (TNM classification) | Treatment |
| Time to treatment | Was the start of treatment within 7 weeks of diagnosis? [www.treeknorm.nl]. The date of diagnosis was defined as the date of the first tumour positive biopsy, the date of the start of treatment was the date of surgery, or, when neo-adjuvant treatment was given, the date of the first radiation therapy or chemotherapy | CC and RC | Waiting time |
| 10+ lymph nodes | Were 10 or more lymph nodes described in the pathology report? The required number of lymph nodes for adequate postoperative staging defined by the national guideline is 10 or more | CC | Treatment and pathology report |
| CRM reported | Was the involvement of the circumferential resection margin adequately reported pathology report? | RC | Pathology report |
| Outcome | | | |
| Mortality | Mortality within 30 days after resection, adjusted for case-mix variations | CC and RC | Outcome |
| Morbidity | Any complication resulting in a length of stay longer than 14 days, a reintervention, or mortality, adjusted for case-mix variations | CC and RC | Outcome |

CC = coloncarcinoma, RC = rectumcarcinoma, pTNM = pathological stage Tumour Node Metastasis system
CRM = circumferential resection margin
*due to an error in the database the imaging of lung and liver of patients with a rectumcarcinoma was not registered.

## Combining indicators

We evaluated all methods of combining indicators as described by Reeves and their relation to postoperative outcome, at a hospital level and a patient level [Table 2].[91] At a hospital level we tested five measures: the 'all or none', the '70% standard', the 'patient average', the 'indicator average' and the 'overall percentage' measure. Differences between the score on the latter three composite measures are defined by the proportion of indicators that are only relevant for a select group of patients, and the proportion of patients in a hospital for whom the indicator is relevant. At a patient level, we tested three measures: the 'all or none', the '70% standard', and the 'patient average'. The other two composite measures, 'Indicator Average', and 'Overall Percentage' were not applicable at the patient level, as they would give the same score as the 'patient average'. As pointed out by Reeves at al,

**Table 2.** Composite measures at hospital and patient level.

| Patient level | | |
|---|---|---|
| All-or-None | 'were all relevant indicators met for this patient?' | This is the strictest measure, as no points are given when one indicated process is missed. |
| 70% Standard | 'were 70% or more of all relevant indicators met for this patient?' | A less strict measure, the standard is met when, 70% or more of the indicated care is given. |
| Patient Average | 'the percentage of relevant indicators that were met for this patient.' | An analogue measure for the percentage of indicators that are met for a patient. |
| **Hospital level** | | |
| All-or-None | 'the percentage of patients for whom all relevant indicators were met'. | This is the strictest measure, as no points are given when one indicated process is missed. Nolan and Berwick found this score to be the most patient-centred score, as it represents the likelihood that 'all goes well' in a patients process[90]. |
| 70% Standard | 'the percentage of patients for whom 70% of more of all relevant indicators were met.' | A less strict measure, the standard is met when, 70% or more of the indicated care is given. |
| Patient Average | 'the mean rate at which relevant indicators were met for each patient.' | The percentage of relevant indicators met is averaged across all patients. This score represents the mean percentage of indicated care patients receive in the hospital[108]. |
| Indicator Average | 'the mean rate at which each indicator was met.' | For each indicator, the percentage of patients for whom the indicator was met is computed. The scores are averaged across all indicators[109]. This is the most commonly used method used by commercial parties in the Netherlands, as the individual indicator scores are usually the only publicly available information. |
| Overall Percentage | 'the percentage of relevant indicators that were met.' | The overall percentage of indicated care that was given in a hospital. Example: for 100 patients and 4 audited indicators, a total of 340 indicators were relevant, and 260 indicators were met, the overall percentage was 76%[90,110,111]. |

distributional properties of the composite measures are unknown, so that standard errors could not be calculated;[91] measures are presented as mean and range.

## Analyses

Outcomes were adjusted for case-mix factors including: age, gender, comorbidity (Charlson-score),[95] previous abdominal surgery, Body Mass Index (BMI), American Society of Anaesthesiologists (ASA) classification, procedure type (right hemicolectomy, left hemicolectomy including transversectomy, sigmoid resection, low anterior resection, abdominoperineal resection), local tumour invasiveness (T-stage), disseminated disease (M-stage) and additional resection for tumour invasion and/or metastasis.

We investigated the association of each of the three composite measures and postoperative outcome at a patient level and each of the five composite measures and postoperative outcome at a hospital level.

At the patient level, to test which of the composite measures was an independent predictor for mortality major morbidity, each composite measure was tested in a separate multivariable random effects logistic regression model, either for mortality or major morbidity. Odds ratios were adjusted for all case-mix factors mentioned above. Hospital of treatment was included as a random effect in the analyses. To take in account any possible complex relation between patients and the hospital of their treatment influencing the estimated effect of composite measures on outcome on a patients level, a sensitivity analysis was performed by means of a multivariable, multilevel hierarchical logistic regression model.

As the relationship between composite measures and outcome may be different on a hospital level, and possibly even have a different meaning, the data were also analyzed on a hospital level. At the hospital level, outcome was adjusted for case-mix by computing an expected outcome for each hospital. The hospitals' expected outcome was derived from logistic regression analysis in which all case-mix factors as listed above were included. The risk-adjusted mortality or major morbidity rate was computed by multiplying the observed/expected mortality and major morbidity ratio with the population average mortality and major morbidity rate.[33,34] We tested the relation of each of the five composite measures with hospitals' risk-adjusted mortality and major morbidity rate using Poisson regression analyses. As it is known that small volume hospitals are more prone to statistical chance fluctuations, and are therefore more likely to be outliers, the analysis was repeated for large volume hospitals (>50 CC or RC resections).

Statistical significance was defined as $p<0.05$, a trend towards significance was defined as $0.05>p<0.10$. All statistics were performed in PASW Statistics, Rel 18.0.2009. Chicago: SPSS inc. and Stata.

# Results

## Patients

At March 15th 2011, 92 hospitals (8 university hospitals, 46 teaching hospitals and 38 non-teaching hospitals) registered a total of 8835 eligible patients with a date of surgery between January 1 and December 31 2010 in the DSCA. After exclusion of hospitals that failed to register all patients (3 hospitals, 16 patients), patients with multiple synchronous tumours (253 patients) and urgent and acute patients (1228 patients) and hospitals that had not recorded case-mix details factors for more than 15% of their patients (367 patients, 4 hospitals) a total of 6971 patients (4732 CC and 2239 RC) treated in 85 hospitals were included in the analyses. Of these hospitals, 43 were high volume (>50 procedures) for CC and 9 were high volume for RC.

## Process indicators and outcome

Table 3 shows the indicators and outcome, at patient and hospital level, with their respective 95% confidence intervals. For all indicators missing data were less than 2%, except for time to treatment, in which date of first tumour positive biopsy was missing in 9.2%. For these patients, the indicator was considered irrelevant. Average scores on separate indicators were more than 80%, except for 'complete colonoscopy' (70-77%), 'time to treatment' for

**Table 3.** Average scores for the indicators and outcome for the patient population (patient level) and at a hospital level

| Level | Colon cancer Indicator | Population average | | Rectal cancer Indicator | Population average | |
|---|---|---|---|---|---|---|
| Patient | Complete Colonoscopy | 70% | | Complete Colonoscopy | 77% | |
| | Diagnostic tests | 90% | | MRI | 90% | |
| | 10+ lymph nodes | 79% | | MDT | 89% | |
| | Time to treatment | 86% | | CRM reported | 59% | |
| | | | | Time to treatment | 75% | |
| | | | | Radiotherapy | 88% | |
| | **Outcome** | | | **Outcome** | | |
| | Mortality | 3% | | Mortality | 2% | |
| | Morbidity | 21% | | Morbidity | 28% | |
| | | Hospital average | Range | | Hospital average | Range |
| Hospital | Complete Colonoscopy | 71% | 47-100% | Complete Colonoscopy | 77% | 33-100% |
| | Diagnostic tests | 91% | 22-100% | MRI | 88% | 0-100% |
| | Time to treatment | 85% | 41-100% | MDT | 90% | 0-100% |
| | 10+ lymph nodes | 79% | 43-97% | Time to treatment | 75% | 0-100% |
| | | | | Radiotherapy | 86% | 0-100% |
| | | | | CRM reported | 62% | 0-100% |
| | **Outcome** | | | **Outcome** | | |
| | Mortality$ | 3.6% | 0-19% | Mortality | 2.1% | 0-19% |
| | Morbidity$ | 21% | 6.4-39% | Morbidity | 27% | 0-58% |

$ risk adjusted outcome rate: observed outcome/expected outcome multiplied by the population average.

RC (75%), '10+ lymph nodes' (79%) and 'Circumferential Resection Margin (CRM) reported' (59%). Population mortality rate for CC was 3.6 and 2.1% for RC; Major morbidity rate was 20.8% for CC and 27.2% for RC.

## Composite measures

Although the 'patient average' was high (75% for CC and 80% for RC) for only 45% of CC patients and 29% of RC patients all indicators were met ('all or none' measure) [Table 4]. Similarly, average hospital scores on composite measures were high, but there was a wide range in performance. For CC, there were hospitals in which fewer than 10% of patients received all the indicated care ('all or none' measure), but also hospitals in which for 83% of patients all indicators were met. For RC, hospital score on the 'all or none' measure ranged from 0 to 100%.

**Table 4.** Average scores for the composite measures at patient and hospital level

| Level | Colon cancer Indicator | Population average | | Rectal cancer Indicator | Population average | |
|---|---|---|---|---|---|---|
| Patient level | All or none | 45% | | All or none | 29% | |
| | 70% standard | 78% | | 70% standard | 69% | |
| | Patient average | 75% | | Patient average | 80% | |
| | | Hospital average | Range | | Hospital average | Range |
| Hospital level | All or none | 45% | 8-83% | All or none | 29% | 0-100% |
| | 70% standard | 78% | 41-96% | 70% standard | 71% | 0-100% |
| | Patient average | 81% | 59-95% | Patient average | 80% | 50-100% |
| | Indicator average | 81% | 61-95% | Indicator average | 80% | 47-100% |
| | Overall percentage | 81% | 60-95% | Overall percentage | 80% | 50-100% |

## Composite measures and outcome, patient level

Table 5 shows the odds ratios of the three composite measures for mortality and major morbidity at patient level: the 'all or none', '70% standard' and 'patient average', after adjustment for case-mix and hospital effect in a multivariable analysis. This table shows that, for CC, none of the composite measures was associated with outcome. However, for RC the '70% standard' measure showed a negative association (OR =0.8, p=0.04) with major morbidity. A sensitivity analysis using a multilevel hierarchical model did not change the results substantially (table 5).

## Composite measures and outcome, hospital level

At hospital level, all composite measures were associated with lower risk-adjusted major morbidity rates for CC. However none of the composite measures for guideline adherence was associated with hospital risk-adjusted mortality, although the 'all or none' method showed a trend towards a negative association (p=0.07) [Table 6]. For RC, a high score on guideline adherence measured by any of the composite measures was significantly associated

**Table 5.** Odds ratios of composite measures for outcome, at a patients' level, as derived from a multivariable random-effects logistic regression model, and from multivariable multilevel hierarchical logistic regression model.

| Patient level | All or none | | 70% standard | | Patient average | |
|---|---|---|---|---|---|---|
| | **Log regression** | **Multilevel model** | **Log regression** | **Multilevel model** | **Log regression** | **Multilevel model** |
| | OR (95%CI) | OR (95%CI) | OR (95%CI) | OR (95%CI) | OR (95%CI) | OR (95%CI) |
| **Colon carcinoma** | | | | | | |
| Mortality | 1.0 (0.7-1.4) p=0.9 | 1.0 (0.1-7.4) p=0.9 | 0.8 (0.6-1.3) p=0.4 | 0.8 (0.1-8.0) p=0.9 | 0.9 (0.4-2.1) p=0.8 | 0.9 (0.1-107) p=0.9 |
| Morbidity | 0.9 (0.8-1.1) p=0.2 | 0.9 (0.8-1.1) p=0.3 | 1.1 (0.9-1.2) p=0.6 | 1.1 (0.9-1.3) p=0.6 | 0.8 (0.6-1.2) p=0.5 | 0.9 (0.6-1.3) p=0.5 |
| **Rectum carcinoma** | | | | | | |
| Mortality | 0.4 (0.1-1.2) p=0.1 | 0.4 (0.0-218) p=0.8 | 0.6 (0.3-1.3) p=0.2 | 0.6 (0.01-33.7) p=0.8 | 0.2 (0.04-1.6) p=0.1 | 0.2 (0.0-7952) p=0.8 |
| Morbidity | 0.9 (0.7-1.1) p=0.4 | 0.9 (0.7-1.1) p=0.4 | **0.8 (0.6-1.0) p=0.04** | 0.8 (0.6-1.1) p=0.06 | 0.7 (0.4-1.4) p=0.3 | 0.7 (0.4-1.4) p=0.4 |

**Table 6.** Rate ratios of composite measures for outcome, at hospital level, as derived from a Poisson regression model.

| | All or none | 70% standard | Patient average | Indicator average | Overall % |
|---|---|---|---|---|---|
| | RR (95% CI) | RR (95% CI) | RR (95% CI) | RR (95% CI) | RR (95% CI) |
| Colon Carcinoma | | | | | |
| Mortality$ | 0.5 (0.2-1.1) p=0.07 | 1.0 (0.4-2.8) p=0.9 | 0.4 (0.1-2.0) p=0.2 | 0.3 (0.1-1.8) p=0.2 | 0.3 (0.1-1.8) p=0.2 |
| Morbidity$ | **0.5 (0.4-0.7) p<0.001** | **0.5 (0.3-0.7) p=0.001** | **0.3 (0.1-0.6) p=0.001** | **0.3 (0.1-0.5) p<0.001** | **0.3 (0.1-0.8) p=0.001** |
| Rectum Carcinoma | | | | | |
| Mortality$ | **0.02 (0.01-0.1) p<0.001** | 0.6 (0.3-1.1) p=0.1 | **0.07 (0.02-0.3) p<0.001** | **0.1 (0.03-0.5) p=0.003** | **0.07 (0.02-0.3) p<0.001** |
| Morbidity$ | **0.4 (0.3-0.5) p<0.001** | **0.6 (0.5-0.8) p<0.001** | **0.4 (0.2-0.5) p<0.001** | **0.4 (0.3-0.7) p<0.001** | **0.4 (0.2-0.6) p<0.001** |

RR: rate ratio; CI: confidence interval; $ risk adjusted hospital mortality or major morbidity rate: Observed / Expected * population average

with a lower risk-adjusted mortality and major morbidity rate, except for the '70% standard' measure, which showed a trend towards a negative association with mortality (p=0.10). When the analysis was repeated for high volume hospitals only (50 or more procedures), all composite measures, accept for the '70% standard', were also significantly associated with risk adjusted mortality for CC. All other associations remained significant (data not shown).

## Discussion

The aim of this study was to measure consistency of composite measures based on process indicators, and outcome indicators. This study shows that a high hospital score on composite process measures is consistent with better risk-adjusted short-term outcome rates. However,

for the treatment of an individual patient, a high score on composite process measures was not clearly associated with better outcome. This indicates that being treated according to the guideline is not necessarily predictive for a better postoperative outcome for the individual patient, but a hospital's score on guideline adherence gives a good indication of the quality of care in a hospital.

The results of this study should be interpreted in the light of some limitations. First, although the dataset of the DSCA 2010 was more complete than several other national registries[96], the estimated completeness of the DSCA in 2011 was 93%. However, comparison with the NCR showed no overall differences in patient, tumour, procedural and outcome data. Therefore, it is unlikely that the results would have been influenced to a great extent by the missing 7%. Second, the indicators used in this study were similar to the indicators used in previous studies.[97-102] The validity of some of these indicators is still under debate,[103] as some of the selected processes, such as 'complete colonoscopy' and 'radiotherapy' are influenced by case-mix and patients preferences. However, as non-elective patients were excluded and outcome data were adjusted for differences in patient and tumour characteristics, it is unlikely that case-mix and patient preferences affected our results to a great extent. Another limitation is the absence of long-term follow-up results. As most quality indicators are based on evidence-based guidelines and a relation with long-term outcome, an association with long-term outcome would be expected. Unfortunately, long-term results were not yet available when this study was completed. However, although long-term outcome, combining both operative and oncologic outcome, may be a more valid outcome measure, it reflects the quality of care delivered one to five years before the outcome measurement and is therefore less indicative for the hospital's performance at that moment. Short-term outcome on the other hand, reflects the quality of recently delivered care, and is therefore commonly used for performance measurements.

Currently, hospital performances in the Netherlands are compared based on a list of process indicators. Previous studies have investigated the relation between process and outcome, with inconsistent results.[88,89] Possible explanations for the inconsistent results are sample sizes too small to reliably assess hospital performances on outcome indicators, and variation in hospital performances on different indicators.[89,104,105] Bradley at all. investigated the relation of individual indicators and outcome after acute myocardial infarction, and found varying results. However, when indicators were aggregated into a composite measure, there was a significant relation with outcome.[106] Similarly, Stullberg at al. found a significant association between a composite process measure for infection-prevention and a lower probability for postoperative infection, but could not reproduce this finding for any of the individual process measures.[107] Habib et al found no association between a composite measure based on process indicators and outcome for colorectal cancer resections in Australia.[102] However, outcome measures were not adjusted for hospital variations in case-mix.

In this dataset, none of the composite measures were associated with outcome on a patients level. However, we did find an association between composite measures and outcome on a hospitals' level. Possibly, on a patient level, other factors, such as patient related factors, are stronger predictors for outcome than the number of indicated process indicators met for the patient, while on a hospital level, composite measures for process indicators may reflect the underlying quality of care in a hospital. All composite measures showed a negative association with major postoperative morbidity. In the full analysis (including small volume hospitals), none of the composite measures was associated with hospital mortality rates for CC, but when analyses was repeated in only high volume hospitals, we found a strong relation with all but one of the composite measures. As outcome rates in low volume hospitals are more susceptible to chance fluctuations, results of small volume hospitals may have weakened the association between guideline adherence and mortality for CC at the hospital level. For RC the relation between guideline adherence and postoperative mortality was also clear in the full analysis. This may be the effect of the larger number of indicators used to assess guideline adherence for RC, but it may also reflect the more complex process of care for patients with an RC, in which only well organized and more dedicated hospitals succeed. As the association between guideline adherence and short-term outcome on a hospital level is unlikely to be explained by a causal relation, there may be other underlying processes and structural differences between hospitals, responsible for this association. Identification of these underlying processes and structures may help to improve quality of care in all hospitals.

As all composite measures showed a strong association with hospital major morbidity rates, we cannot recommend one specific composite measure based on our results. However, the '70% standard' was the only measure which was not consistent with mortality for RC and CC (high volume hospitals only). Possibly, the percentage of patients who received 70% of indicated care is an insufficiently strict measure to indentify good quality hospitals. These results support the ideas advocated by Nolan and Berwick, [90] that the relationship between the completeness of the process and outcome is not continuous, but determined by the completion of all indicated processes. These ideas are also adopted by the Institute for Healthcare Improvement (IHI, www.ihi.org), which developed 'bundles': a set of evidence based practices, that, when performed collectively and reliably, have been proven to improve patient outcomes. The IHI advocates measuring compliance as 'the percentage of patients who receive all elements of the bundle': the 'all or none' measure. Although our results do not support one of the remaining four measures, the choice for a composite measures may be dictated by other factors such as the availability of data at patient level, required for all measures but the 'indicator average', and patient's preferences. Future research should be directed to identifying the composite measure most comprehensible and informative for healthcare consumers.

## Conclusions

Guideline adherence in the process of care for colorectal cancer patients is not associated with better short-term outcome for the individual patient. However, hospitals with favourable scores on guideline adherence also have better postoperative outcome rates. When measuring quality of care for colorectal cancer patients by means of composite process measures, the 'all or none' the 'patient average', 'indicator average' or 'overall percentage' seem equally suitable measures as they were all consistent with hospital outcome indicators. However, the drive to have a good score on process indicators should never compromise good, personalized medicine: giving the right treatment at the right time to the right patient.

### Acknowledgements

We thank all surgeons, registrars, physician assistants and administrative nurses that registered all the patients in the DSCA, as well as the Dutch Surgical Colorectal Audit group and the methodological board for their advice.

## References


1. Institute of Medicine. Committee to Design a Strategy for Quality Review and Assurance in Medicine. The Institute of Medicine Report. In Medicare: A Strategy for Quality Assurance/ Committee to Design a Strategy for Quality Review and Assurance in Medicare. Division of Health Care Services, Institute of Medicine, Vol 1, report of a study, p 468 Washington, DC: National Academy Press 1990
2. Donabedian A. Evaluating the quality of medical care. The Milbank Memorial Fund quarterly 1966;44:Suppl:166-206.
3. Birkmeyer JD, Dimick JB, Birkmeyer NJ. Measuring the quality of surgical care: structure, process, or outcomes? J Am Coll Surg 2004;198:626-32.
4. Khuri SF, Daley J, Henderson WG. The comparative assessment and improvement of quality of surgical care in the Department of Veterans Affairs. Arch Surg 2002;137:20-7.
5. Thomas JW, Hofer TP. Research evidence on the validity of risk-adjusted mortality rate as a measure of hospital quality of care. Med Care Res Rev 1998;55:371-404.
6. Pitches DW, Mohammed MA, Lilford RJ. What is the empirical evidence that hospitals with higher-risk adjusted mortality rates provide poorer quality care? A systematic review of the literature. BMC Health Serv Res 2007;7:91.
7. Nolan T, Berwick DM. All-or-none measurement raises the bar on performance. JAMA 2006;295:1168-70.
8. Reeves D, Campbell SM, Adams J, et al.Combining multiple indicators of clinical quality: an evaluation of different analytic approaches. Med Care 2007;45:489-96.
9. Powell C. The Delphi technique: myths and realities. J Adv Nurs 2003;41:376-82.
10. Linstone HA TM. The Delphi method: techniques and applications [http://isnjitedu/pubs/ delphibook/], 2002- Accessed on the 31/05/2011.
11. Hendren S, Birkmeyer JD, Yin H, et al. Surgical complications are associated with omission of chemotherapy for stage III colorectal cancer. Dis Colon Rectum;53:1587-93.
12. Charlson ME, Sax FL, MacKenzie CR, et al. Assessing illness severity: does clinical judgment work? J Chronic Dis 1986;39:439-52.
13. Daley J, Khuri SF, Henderson W, et al. Risk adjustment of the postoperative morbidity rate for the comparative assessment of the quality of surgical care: results of the National Veterans Affairs Surgical Risk Study. J Am Coll Surg 1997;185:328-40.
14. Khuri SF, Daley J, Henderson W, et al. Risk adjustment of the postoperative mortality rate for the comparative assessment of the quality of surgical care: results of the National Veterans Affairs Surgical Risk Study. J Am Coll Surg 1997;185:315-27.
15. van Gijn W, van de Velde CJ, members of the Ec. Improving quality of cancer care through surgical audit. Eur J Surg Oncol;36 Suppl 1:S23-6.
16. Jackson GL, Melton LD, Abbott DH, et al. Quality of nonmetastatic colorectal cancer care in the Department of Veterans Affairs. J Clin Oncol;28:3176-81.
17. Chung KP, Chang YJ, Lai MS, et al. Is quality of colorectal cancer care good enough? Core measures development and its application for comparing hospitals in Taiwan. BMC Health Serv Res;10:27.
18. Gagliardi AR, Simunovic M, Langer B, et al. Development of quality indicators for colorectal cancer surgery, using a 3-step modified Delphi approach. Can J Surg 2005;48:441-52.
19. Schneider EC, Nadel MR, Zaslavsky AM, et al. Assessment of the scientific soundness of clinical performance measures: a field test of the National Committee for Quality Assurance's colorectal cancer screening measure. Arch Intern Med 2008;168:876-82.
20. Vergara-Fernandez O, Swallow CJ, Victor JC, et al. Assessing outcomes following surgery for colorectal cancer using quality of care indicators. Can J Surg;53:232-40.
21. Habib MR, Solomon MJ, Young JM, et al. Evidence-based and clinical outcome scores to facilitate audit and feedback for colorectal cancer care. Dis Colon Rectum 2009;52:616-22; discussion 22-3.
22. Patwardhan M, Fisher DA, Mantyh CR, et al. Assessing the quality of colorectal cancer care: do we have appropriate quality measures? (A systematic review of literature). J Eval Clin Pract 2007;13:831-45.
23. Dimick JB, Welch HG, Birkmeyer JD. Surgical mortality as an indicator of hospital quality: the problem with small sample size. JAMA 2004;292:847-51.
24. Lilford R, Mohammed MA, Spiegelhalter D, et al. Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. Lancet 2004;363:1147-54.
25. Bradley EH, Herrin J, Elbel B, et al. Hospital quality for acute myocardial infarction: correlation among process measures and relationship with short-term mortality. JAMA 2006;296:72-8.
26. Stulberg JJ, Delaney CP, Neuhauser DV, et al. Adherence to surgical care improvement project measures and the association with postoperative infections. JAMA 2010;303:2479-85.
27. Campbell SM, Roland MO, Middleton E, et al. Improvements in quality of clinical care in English general practice 1998-2003: longitudinal observational study. BMJ 2005;331:1121.
28. Jha AK, Li Z, Orav EJ, Epstein AM. Care in U.S. hospitals–the Hospital Quality Alliance program. N Engl J Med 2005;353:265-74.
29. McGlynn EA, Asch SM, Adams J, et al. The quality of health care delivered to adults in the United States. N Engl J Med 2003;348:2635-45.
30. Premier. Inc. Summary of the composite quality scoring methodology. Available at: http:// www.premierinc.com/quality-safety/tools-services/ p4p/hqi/resources/top-performer-summary.pdf. . Accessed April 17, 2007.

Chapter 9

# Focusing on desired outcomes of care after colon cancer resections; hospital variations in 'textbook outcome'

N. E. Kolfschoten[1], J. Kievit[1,2], G.A. Gooiker[1], N.J. van Leersum[1], H.S. Snijders[1], EHEddes[3], R.A.E.M Tollenaar[1], M.W.J.M. Wouters[4], P.J. Marang-van de Mheen[2], on behalf of the Dutch Surgical Colorectal Audit group*

[1] Leiden University Medical Centre, Dept of Surgery
[2] Leiden University Medical Center, Dept of Medical Decision Making
[3] Deventer Hospital, Dept of Surgery
[4] Netherlands Cancer Institute-Antoni van Leeuwenhoek Hospital, Dept of Surgical Oncology
* W.A. Bemelman, O.R.C. Busch, R.M. van Dam, E. van der Harst, M.L.E.A. Jansen-Landheer, Th.M. Karsten, J.H.J.M. van Krieken, W.G.T. Kuijpers, V.E. Lemmens, E.R. Manusama, W.J.H.J Meijerink, H.J.T. Rutten, T. Wiggers, C.J.H. van de Velde.

*European Journal of Surgical Oncology 2013 Feb;39(2):156-63*

## Abstract

### Aims

We propose a summarizing measure for outcome indicators, representing the proportion of patients for whom all desired short-term outcomes of care (a 'textbook outcome') is realized. The aim of this study was to investigate hospital variation in the proportion of patients with a 'textbook outcome' after colon cancer resections in the Netherlands.

### Methods

Patients who underwent a colon cancer resection in 2010 in the Netherlands were included in the Dutch Surgical Colorectal Audit. A textbook outcome was defined as hospital survival, radical resection, no reintervention, no ostomy, no adverse outcome and a hospital stay < 14 days. We calculated the number of hospitals with a significantly higher (positive outlier) or lower (negative outlier) Observed/Expected (O/E) textbook outcome than average. As quality measures may be more discriminative in a low-risk population, analyses were repeated for low-risk patients only.

### Results

A total of 5582 patients, treated in 82 hospitals were included. Average textbook outcome was 49% (range 26 to 71%). Eight hospitals were identified as negative outliers. In these hospitals a 'textbook outcome' was realized in 35% vs. 52% in average hospitals ($p<0.01$). In a sub-analysis for low-risk patients, only one additional negative outlier was identified.

### Conclusions

The textbook outcome, representing the proportion of patients with a perfect hospitalization, gives a simple comprehensive summary of hospital performance, while preventing indicator driven practice. Therewith the 'textbook outcome' is meaningful for patients, providers, insurance companies and healthcare inspectorate.

### Introduction

Society increasingly demands information on hospitals' quality of care, and wants to know which hospitals deliver 'Exemplary Care and Outcome'[1]. Quality of health care is defined as "the degree to which health services ... increase the likelihood of desired health outcomes."[112] This implicates that quality of care would be best represented by the proportion of patients for whom all desired health outcomes are realized, in other words: the proportion of patients with a 'textbook outcome'. Although such measure has been suggested,[3,4] it was never presented for outcome after colon cancer resections.

Previous studies have reported on hospital variations in outcome after colon cancer resections.[4-6] Although information on individual outcome indicators are useful for targeted quality improvement programs, it is difficult to compare hospital performance by means of individual indicators: a hospital may have a high score on one indicator, but a low score on another. This is especially true for outcome indicators as they can be interrelated: a high ostomy rate may result in a low reintervention rate. Therefore, a summarizing measure could give an additional, comprehensible, overview of hospital quality of care for patients, payers and providers.

Information on hospital performance, including outcome indicators, is increasingly available within the public domain. However, Marshall et al showed that patients rarely searched for this information and did not understand or trust it.[7] New evidence suggests that patients would be more likely to use information on differences in quality of care when presented as a summary measure, for example as the percentage of patients with a 'textbook outcome'.[3,8] Hospital rates of 'textbook outcome' may however be largely affected by preoperative patient and tumor characteristics. Therefore, hospital performances on outcome indicators should be adjusted for case-mix. Moreover, Coory at al. showed that the systematic variation in hospital mortality after acute myocardial infarction was largest in a low-risk patient group,[9] while for high-risk patients, outcome was more likely determined by other factors than hospital quality of care. They concluded that therefore, hospital performances should be assessed on results in low-risk patients.

**The aim** of this study was to investigate hospital variation in the number and proportion of patients with a 'textbook outcome' after colon cancer resections in the Netherlands. In addition, we investigated if hospital variation in 'textbook outcome' differed when only low-risk patients were included.

## Patients and Methods

### Patients

The dataset was retrieved from the Dutch Surgical Colorectal Audit (DSCA), a prospective national database, in which patient-, tumor-, diagnostic- and treatment characteristics as well as pathology and outcome data are registered for patients that undergo a resection

of a primary colorectal carcinoma in the Netherlands. Details of this dataset regarding data collection and methodology have been published previously.[10] Aim of the DSCA is to improve quality of care for colorectal cancer patients in the Netherlands by providing reliable feedback and benchmark information for all Dutch hospitals. The dataset and outcome indicators were based on Dutch evidence-based guidelines (www.oncoline.nl). Estimated completeness in the year 2010 compared to the Netherlands Cancer registry (NCR) was 93%,[11,12] all Dutch hospitals participated. (www.dsca.nl)

## Inclusion and exclusion criteria

All colon cancer patients with a date of surgery between the January 1st 2010 and December 31st 2010 treated in one of the 92 Dutch Hospitals who were included in the DSCA before March 15th 2011 were evaluated. For this study, no ethical approval was required. As rectal cancer treatment involves different outcome indicators, this study was limited to colon cancer. Patients who were treated for recurrent colon cancer or multiple synchronous colorectal tumors were excluded.

Hospitals that failed to register all patients in 2010 were excluded (2 hospitals, 2 patients). Hospitals that failed to register all outcome parameters for more than 15% of patients were excluded (7 hospitals, 551 patients). Furthermore, as case-mix correction is imperative for evaluating outcome of care, and clustering of absent case-mix information makes risk-adjustment unreliable, hospitals that failed to fill in the required case-mix factors for more than 15% of patients were excluded (1 additional hospital, 28 patients). A total of 5582 patients, treated in 82 hospitals were included in the analyses.

## Definitions

Outcome was registered at discharge or 30 days after the resection. 'Textbook outcome' was assessed by 6 separate 'desired outcome' measures: hospital survival, radical resection, no reintervention, no ostomy placement, no adverse outcomes and a hospital stay of 14 days or less. Postoperative mortality was defined as mortality within the hospitalization or 30 days after resection. Radical resection was defined as a microscopic radical resection. A reintervention was defined as an adverse outcome requiring a reoperation or percutaneous reintervention. Adverse outcomes were defined as any adverse outcome occurring within 30 days after resection. A 'textbook' hospital stay was set at the 75th percentile of the population: a hospital stay of 14 days or less. When all 6 desired health outcomes were realized, a 'textbook outcome', was achieved.

Available case-mix factors were age, gender, Body Mass Index (BMI), comorbidity (Charlson-score)[13], previous abdominal surgery, American Society of Anaesthesiologists (ASA) classifications, tumor stage, urgency, tumor complications, operative procedures and additional resections for tumor invasion and/or metastasis. All case-mix factors were categorized into discrete categories. Missing case-mix factors were analyzed in a separate category.

## Analysis

First, we calculated the number and proportion of patients for whom each outcome indicator was realized. We then calculated the number and proportion of patients for whom each consecutive outcome indicators was realized; conditional to that the previous had succeeded. Outcome indicators were ranked in decreasing order of importance, with hospital survival as most important, and hospital stay as least important outcome indicator. This resulted in the number and proportion of patients for whom all desired health outcomes were realized and thereby a 'textbook outcome' was achieved. For each hospital, the proportion of patients with a 'textbook outcome' was calculated. Hospitals' unadjusted 'textbook outcome' rates were compared in a funnel plot.[14]

Secondly, as differences in case-mix may explain a large part of the hospital variation in 'textbook outcome' rates,[4,13] hospital performances were adjusted for case-mix by calculating Observed/Expected (O/E) outcome ratios.[15,16] For this calculation, the observed outcome was the number of patients with a 'textbook outcome' in a hospital and the expected outcome is the sum of all patients' estimated probabilities for a textbook outcome. Patients' probability estimates were derived from a backwards-stepwise multivariate logistic regression model, fitted on the data of all included hospitals, and using all case-mix factors mentioned above. For an average performing hospital, the observed outcome will be equal to the expected outcome, resulting in an O/E outcome ratio of 1. Hospitals that perform better than average have an O/E outcome ratio higher than 1, hospitals that perform worse than average have an O/E ratio lower than 1. For each hospital the exact Poisson 95% confidence interval for its O/E outcome ratio was calculated.[17] When the lower limit of the 95% confidence interval is above 1, the hospitals' O/E outcome is 'significantly better than average', and the hospital was considered a positive outlier. When the upper limit of the 95% confidence interval is below 1, O/E outcome is significantly worse than average and the hospitals was a negative outlier.

Lastly, as previous studies suggested that, as for high risk patients outcomes are largely determined by other factors than hospital quality of care, differences in hospital performances are best assessed in a low-risk population,[9] hospital variation in 'textbook outcome' was also assessed in a selection of 'low-risk' patients. Risk stratification was based on the patients estimated probability for 'textbook outcome' as derived from logistic regression analysis (i.e. low risk patients were patients with a high probability for a 'textbook outcome'). Patients with a predicted probability of more than the 75$^{th}$ percentile of 'textbook outcome' were classified as low-risk patients. We investigated hospital variation in O/E 'textbook outcome' rates for low-risk patients and identified outlier hospitals.

All statistics were performed in PASW Statistics for Mac, Rel 18.0.2009. Chicago: SPSS inc. and Microsoft Excel.

## Results

Average number of colon cancer resections per hospital was 67 patients (range 7 to 144). Table 1 shows the number and proportion of patients for whom each desired health outcome was realized. Next, performances on each outcome indicator were calculated conditionally to that the previous outcome indicator was met, in decreasing order of importance of outcomes. For each outcome indicator added, the percentage of patients for whom that specific outcome and all previous outcome indicators were realized, was reduced

**Table 1.** Outcome after colon cancer resections, population average on individual outcome indicators, and on outcome indicators, each conditional to the previous indicator being fulfilled.

| | Population | | Conditional | |
|---|---|---|---|---|
| | N | % | N | % |
| Patients | 5582 | | | |
| Hospital survival | 5312 | 95% | 5312 | 95% |
| Radical resection | 4965 | 89% | 4735 | 85% |
| No reintervention | 4716 | 85% | 4076 | 73% |
| No ostomy | 4747 | 85% | 3544 | 64% |
| No adverse outcome | 3731 | 67% | 2918 | 52% |
| Hospital stay <14 days | 4319 | 77% | 2721 | 49% |
| **Textbook outcome** | | | 2721 | 49% |

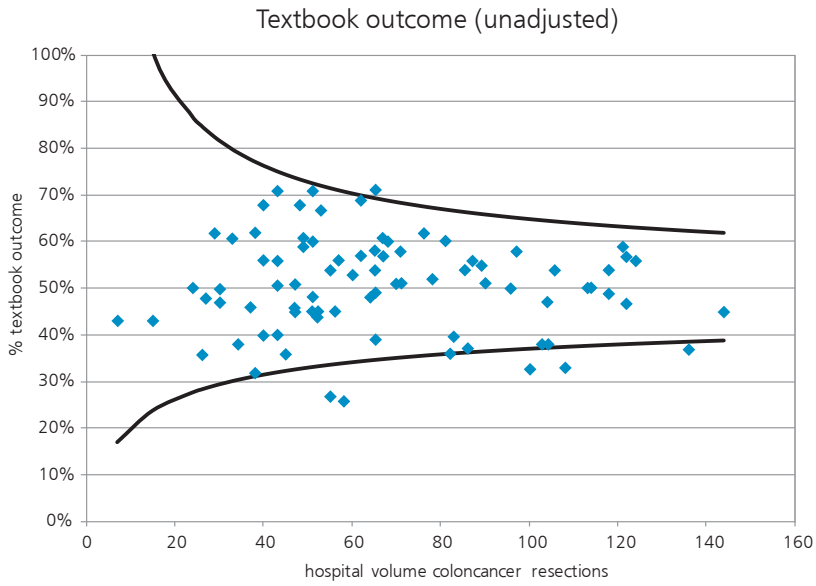



**Figure 1.** Hospital variation in percentage of patients with 'textbook outcome', unadjusted. Black lines represent the 95% confidence interval of population average. Gray thin line represents population average 'textbook outcome'.

with approximately 10%, except for the last outcome indicator, a hospital stay of 14 days or less. For only 49% of all patients, a 'textbook outcome' was realized.

## Hospital differences in 'textbook outcome'

Figure 1 shows the hospital variation in percentage of patients with 'textbook outcome'. Unadjusted hospital 'textbook outcome' rates varied from 26 to 71%. Five hospitals performed significantly worse than average, based on unadjusted outcome. One hospital was significantly better than average.

Factors which reduce chances at a 'textbook outcome' are: older age, male gender, previous abdominal surgery, ASA score of III or more, urgent resection, tumor complications, left hemicolectomy, low anterior resection or 'other resection' (e.g. subtotal colectomy or panproctocolectomy), a higher tumor stage and additional resections for local tumor invasiveness. Charlson co-morbidity score and Body Mass Index (BMI did not remain statistically significant in the logistic regression model. The C-statistic was 0.75 (0.74-0.76), indicating a moderately good prediction model.[18] (Table 2)

**Table 2.** Case-mix factors predictive for 'textbook outcome'.

| Case-mix Factor | | N | % | OR | (95% CI) |
|---|---|---|---|---|---|
| Age (years) | >70 | 2403 | 45% | 1.00 (ref) | |
| | 70-79 | 1769 | 33% | 0.66 | (0.57-0.75) |
| | 80+ | 1212 | 22% | 0.49 | (0.41-0.57) |
| Gender | Female | 2562 | 52% | 1.51 | (1.34-1.71) |
| Abdominal history | | 1856 | 34% | 0.82 | (0.72-0.93) |
| ASA score | III+ | 1388 | 26% | 0.52 | (0.46-0.60) |
| Urgency | Urgent/acute | 1014 | 19% | 0.44 | (0.35-0.56) |
| Tumor | None | 4092 | 76% | 1.00 (ref) | |
| complication | Perforation | 125 | 2% | 0.40 | (0.23-0.69) |
| | Obstruction | 671 | 13% | 0.66 | (0.51-0.86) |
| | Else | 454 | 8% | 0.94 | (0.75-1.17) |
| Procedure | Right hemicolectomy | 2527 | 47% | 1.00 (ref) | |
| | Left hemicolectomy | 715 | 13% | 0.77 | (0.64-0.92) |
| | Sigmoid resection | 1263 | 23% | 1.08 | (0.92-1.25) |
| | (L)AR | 715 | 13% | 0.46 | (0.38-0.55) |
| | Else | 173 | 4% | 0.29 | (0.20-0.43) |
| Stage | I | 949 | 18% | 1.00 (ref) | |
| | II | 1932 | 36% | 0.74 | (0.62-0.88) |
| | III | 1626 | 30% | 0.74 | (0.62-0.88) |
| | IV | 819 | 15% | 0.21 | (0.16-0.25) |
| | X | 65 | 1% | 0.59 | (0.34-2.01) |
| Additional resection | Local tumor invasion | 507 | 9% | 0.44 | (0.36-0.55) |
| C-statistic | | | | 0.75 | (0.74-0.76) |

ASA: American society of Anaesthesiologists
LAR: Low anterior resection
OR: Odds Ration
CI: Confidence interval

**Table 3.** Case-mix and Outcome parameters for Low-risk and High-risk patients

| Case-mix | | Low-risk (4041) | High –risk (1541) | p |
|---|---|---|---|---|
| Age (years) | <70 | 48% | 35% | |
| | 70-79 | 33% | 34% | |
| | 80+ | 19% | 31% | |
| Gender | Female | 51% | 37% | <0.01 |
| Abdominal history | | 34% | 37% | <0.05 |
| ASA score | III+ | 19% | 45% | <0.01 |
| Urgency | Urgent/acute | 7% | 55% | <0.01 |
| Tumor complication | None | 87% | 43% | <0.01 |
| | Perforation | 1% | 9% | |
| | Obstruction | 5% | 36% | |
| | Else | 7% | 12% | |
| Procedure | Right hemicolectomy | 47% | 35% | <0.01 |
| | Left hemicolectomy | 12% | 15% | |
| | Sigmoid resection | 26% | 16% | |
| | (L)AR | 9% | 25% | |
| | Else | 2% | 9% | |
| Stage | I | 22% | 3% | <0.01 |
| | II | 39% | 26% | |
| | III | 33% | 23% | |
| | IV | 5% | 47% | |
| | X | 1% | 1% | |
| Additional resection | Local tumor invasion | 5% | 23% | <0.01 |
| Outcome | | | | |
| Hospital survival | | 97% | 91% | <0.01 |
| Radical resection | | 96% | 73% | <0.01 |
| No reintervention | | 86% | 81% | <0.01 |
| No ostomy | | 93% | 67% | <0.01 |
| No adverse outcome | | 71% | 56% | <0.01 |
| Hospital stay <14 days | | 83% | 66% | <0.01 |
| **Textbook outcome** | | **60%** | **21%** | <0.01 |

ASA: American society of Anaesthesiologists
LAR: Low anterior resection
Low-risk patients: Predicted probability for 'textbook outcome'> 35%.

Figure 2 shows hospital adjusted O/E 'textbook outcome' ratios and the concomitant 95% confidence intervals. O/E 'textbook outcome' ranged from 0.52 to 1.40. Eight hospitals were identified as negative outliers, with a significantly worse O/E textbook outcome than average. No positive outliers were identified. Negative outlier hospitals realized a 'textbook outcome' of in 35 vs. 52% of patients in hospitals with average performance (p<0.01). Patients treated in a negative outlier hospital had a lower chance at the desired health outcomes of 'no reintervention' (78 vs 86%, p<0.01), 'no adverse outcome' (53 vs. 69%, p<0.01) and a hospital stay <14 days (75 vs. 79%, p<0.01).

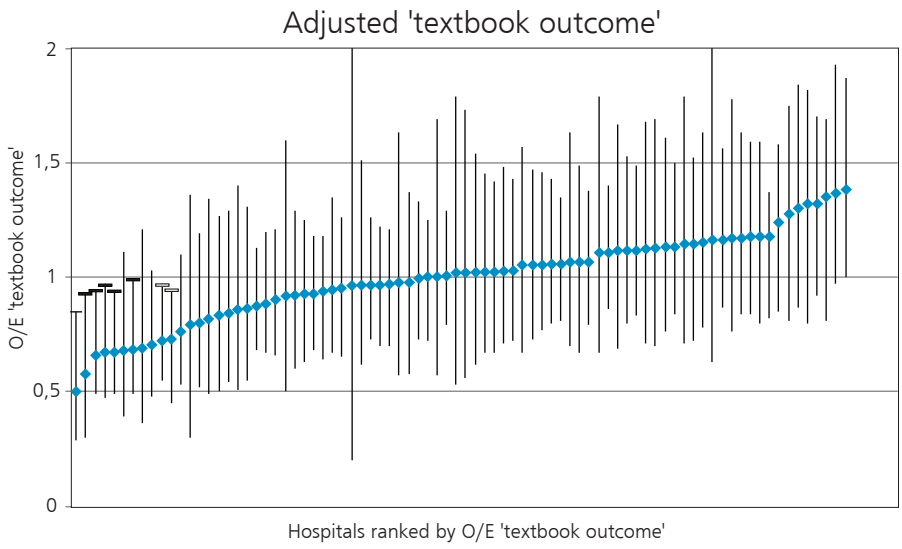



**Figure 2.** Hospitals ranked by risk-adjusted, i.e. observed/expected (O/E) 'Textbook outcome'.
Hospitals with an O/E 'textbook outcome' of less than 1 have less patients with a textbook outcome than expected based on their case-mix and thus perform worse than average. Hospitals with an O/E 'textbook outcome' of more than 1 have more patients with a textbook outcome than expected based on their case-mix and thus perform better than average.
Hospitals marked with a – are negative outliers, and had significantly less patients with a 'textbook outcome' than average.

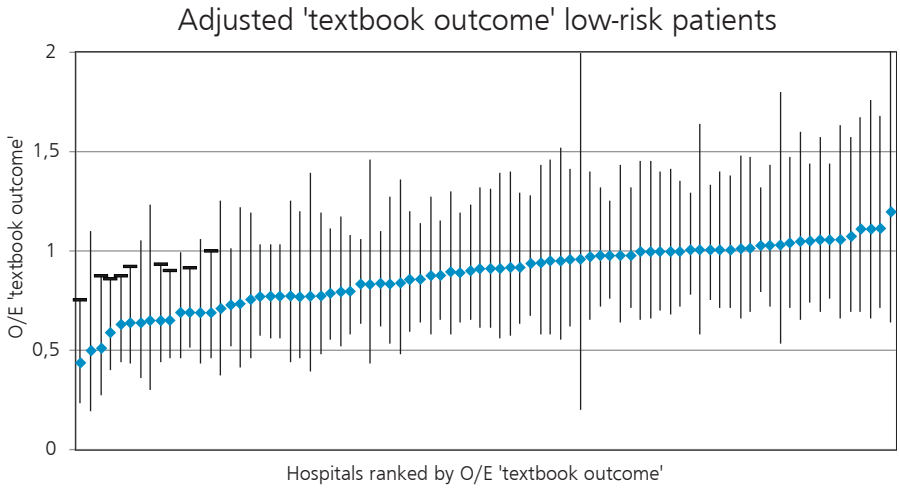



**Figure 3.** Risk-adjusted, i.e. observed/expected (O/E) 'textbook outcome' for low-risk patients.
Low risk patients: patients with a high probability for 'textbook outcome' (predicted probability > 35%).
Hospitals with an O/E 'textbook outcome' of less than 1 have less patients with a textbook outcome than expected based on their case-mix and thus perform worse than average. Hospitals with an O/E 'textbook outcome' of more than 1 have more patients with a textbook outcome than expected based on their case-mix and thus perform better than average.
Hospitals marked with a – are negative outliers, and had significantly less patients with a 'textbook outcome' than average.

### Low-risk patients

Of all patients, 4041 (75%) were classified as 'low-risk' patients. Low risk patients were younger, more often female, had a lower ASA-score, mostly scheduled in an elective setting, more often had a right sided or sigmoid tumor and had a lower tumor stage (Table 3). Average 'textbook outcome' for low-risk patients was 60% vs. 21% for high-risk patients ($p<0.05$). Hospital volume for low-risk patients ranged from 4 to 102 patients. Hospitals' 'textbook outcome' rate for low-risk patients ranged from 31% to 82%, O/E 'textbook outcome' ranged from 0.44 to 1.20. Nine hospitals were negative outliers for 'textbook outcome' in low-risk patients. Amongst these nine hospitals were all eight hospitals, which were negative outliers in the whole population. (Figure 3) The newly identified outlier hospital also had a worse than average O/E textbook outcome in the full analysis, however this was not statistically significant.

## Discussion

### Major findings

The present study uses a composite outcome measure representing the percentage of patients with a 'textbook outcome', for the assessment of quality of care in colon cancer surgery. This composite measure illustrates how clinical audits can be used to provide case-mix adjusted outcome information, which is simple and meaningful for all involved parties. We showed that for only 49% of all colon cancer patients, all desired health outcomes were realized. Using this composite measure we were able to identify eight outlier hospitals, with a 'textbook outcome' rate significantly lower than expected based on their case-mix and the population average. When hospital performances were assessed based on low-risk patients only, one more outlier hospital was identified. This implies that hospital outlier status was hardly determined by an excess of high-risk patients, and moreover, that outlier hospitals also performed worse than average for their low-risk patients.

Previous studies have investigated different methods of constructing composite measures. Reeves at al. described five methods for combining indicators and they found different methods to be useful in different situations.[19] They found that the 'all or none' method (the percentage of patients in whom all indicators are met) is especially suitable when success depended on meeting all indicators. Habib et al. defined a composite outcome measure for colorectal cancer surgery in Australia, based on hospitals' performance relative to the national average, but did not account for hospital variations in case-mix or random variations.[20] O'Brien, tested four methods for combining indicators in adult cardiac surgery: an opportunity-based approach, (weighted) averaging of item-specific estimates, 'all or none' scoring, and latent trait analysis, and found the 'all or none' approach the most suitable for constructing a composite measure.[21] The 'all or none' method is simple and comprehensive, and can be adjusted for differences in case-mix, it sets a high benchmark for the ideal hospitalization and shows the largest inter-provider variability.

## Strengths of the 'textbook outcome' measure

The composite 'all or none' measure 'textbook outcome', as presented in this study is a comprehensive score, which combines all relevant short-term outcome indicators for colon cancer resections. The 'textbook outcome' is simple and seems meaningful to all stakeholders (high face-validity). For patients, it represents their chances for the most favourable ('textbook') outcome in a specific hospital. For care-givers it gives information on how often treatment is successful and therewith may drive quality improvement. For insurance companies and hospitals, it may be useful in selective contracting as it summarizes indicators on patient safety, effectiveness and efficiency. For the healthcare inspectorate this score may guide surveillance-programs. By combining important outcomes of care in one comprehensive measure, the 'textbook outcome' prevents defensive, indicator driven practice. For example: the placement of an ostomy to prevent anastomotic leakage can be a good clinical decision, but a hospital policy to accept an excess ostomy rate to get a better score at the indicator 'no reintervention' may not be in the patients best interest. Good clinical decision-making will therefore result in an optimal score on 'textbook outcome'. Moreover, this measure seems discriminative, as we identified eight outlier hospitals, in which quality of care was worse than average, on several aspects.

Coory at al. showed that the systematic variation in hospital mortality after acute myocardial infarction was largest in a low-risk patient group.[16] However, a large sample size (i.e 10 years) was required to prove an increased systematic variation in outcome in low-risk patients. The present study showed that outlier hospitals for low-risk patients were consistent with outlier hospitals in the whole population, and identified only one additional outlier in the low-risk sub-analysis. Possibly, when data of multiple years are combined, we may find an increased systematic variation in low-risk patients. Also, we may be able to identify positive outliers on 'textbook outcome', and the best practices leading to these outcomes.

## Limitations and possible refinements to this measure

The major limitation of this study is the completeness of the database: although the dataset of the DSCA 2010 was more complete than several other national registries, the estimated completeness of the DSCA in 2011 was 93%. Data comparisons with the independently collected data of the NCR showed no overall differences in patient, tumour, procedural and outcome data. (www.clinicalaudit.nl) Therefore, it is unlikely that the results would have been influenced to a great extent by the missing 7% of patients.

The present study found that the percentage of patients for whom all desired health outcomes were realized is less than 50%. This percentage is higher for low-risk patients, who are more likely to be in the position to select their hospital of treatment as they are often treated in an elective setting and fit to seek and travel for better quality of care. However, the percentage of low-risk patients with 'textbook outcome' is still only 60%, indicating that the measure may be too strict. The outcome indicators with the lowest population average were 'no adverse outcomes' and 'hospital stay <14 days'. Adverse outcomes which do not result in a reintervention or permanent morbidity, may be considered 'minor morbidity', and

it could be argued that these should not be included in the definition of 'textbook outcome'. However, when we exclude 'no hospital stay > 14 days' and 'no adverse outcomes', still only 64% of all patients, had a textbook outcome. Moreover, we found that outlier hospitals had a 'worse than average' performance on several outcome indicators, indicating that outlier status was not explained by a high 'minor morbidity' rate only. Another limitation of this study is the absence of long-term follow-up results, which were unfortunately not available when this study was completed. Moreover, although long-term outcome, combining operative and oncologic outcome, may be a more valid outcome measure, it reflects the quality of care delivered one to five years earlier. Short-term outcome on the other hand, reflects the quality of recently delivered care, and is therefore commonly used for performance measurements.

Possibly the 'textbook outcome' measure could be refined by adding weights to different outcome measures, making one outcome more important than another. However, no clear data or rationale from which to derive these weights is available. Therefore, any weights added to the textbook outcome measure would be subjective and therefore would diminish its simplicity and usefulness.

### Implications and directions for future research

Previous studies have shown that patients are willing to travel further for better quality of care,[22] and that they would like to use outcome information, summarized as 'textbook outcome' for their choice for a hospital of treatment.[3] However, in order to produce such outcome information, a uniform, complete and reliable clinical registry, in which at least case-mix, process and outcomes of care are registered, is required. In the Netherlands such a registry exists for a few disease entities, but many more registries are needed. Furthermore, the measure 'textbook outcome' is composed of clinical outcomes considered relevant by providers. Extensive literature has shown that doctors' preferences are inconsistent with patients' preferences, even when doctors are asked to imagine that they are patients.[23,24] A patients' survey, investigating outcomes of care considered relevant by patients would be of additional value.

## Conclusions

The present study introduces a composite measure, which reflects the percentage of patients for whom all desired health outcomes are realized. Not only may this measure be helpful for patients to base their choice for a hospital of treatment on, it may also provide useful feedback information for care providers. For low-risk patients, overall chance at a 'textbook outcome' is 60%. This number can be used to inform patients during the preoperative clinical consult. However, improvement of quality of colon cancer care in the Netherlands is required. Using the 'textbook outcome' measure, in combination with the individual

outcome indicators, may facilitate such improvement of outcomes of colon cancer care in the Netherlands.

## Acknowledgements

We thank all surgeons, registrars, physician assistants and administrative nurses that registered all the patients in the DSCA, as well as the Dutch Surgical Colorectal Audit group and the methodological board for their advice.

## References


1. Kievit, J., VA National Quality Scholars (VAQS) Fellowship Program. 10th Annual Summer Institute. Fairlee, Vermont 2008. 2008.
2. Institute of Medicine. Committee to Design a Strategy for Quality Review and Assurance in Medicine. The Institute of Medicine Report. In Medicare: A Strategy for Quality
3. Dijs-Elsinga, J., Otten, W., Versluis, M.M., et al., Choosing a hospital for surgery: the importance of information on quality of care. Med Decis Making, 2010. **30**(5): p. 544-55.
4. Dijs-Elsinga, J., Marang-van de Mheen, P.J., Smeets, H.J., et al, [Quality of care and operative complications: comparison of the results of 3 medical departments]. Ned Tijdschr Geneeskd, 2011. **155**: p. A1703.
5. Elferink, M.A., Wouters, M.W., Krijnen, P., et al., Disparities in quality of care for colon cancer between hospitals in the Netherlands. Eur J Surg Oncol. **36 Suppl 1**: p. S64-73.
6. van Dishoeck, A.M., Lingsma, H.F., Mackenbach, J.P., Steyerberg, E.W., Random variation and rankability of hospitals using outcome indicators. BMJ Qual Saf, 2011. **20**(10): p. 869-74.
7. Marshall, M.N., Shekelle, P.G., Leatherman, S., Brook, R.H., The public release of performance data: what do we expect to gain? A review of the evidence. JAMA, 2000. **283**(14): p. 1866-74.
8. Marang-van de Mheen, P.J., Dijs-Elsinga, J., Otten, W., et al., The Relative Importance of Quality of Care Information When Choosing a Hospital for Surgical Treatment: A Hospital Choice Experiment. Med Decis Making, 2010.
9. Coory, M. and I. Scott, Analysing low-risk patient populations allows better discrimination between high-performing and low-performing hospitals: a case study using inhospital mortality from acute myocardial infarction. Qual Saf Health Care, 2007. **16**(5): p. 324-8.
10. Kolfschoten, N.E., Marang-van de Mheen, P.J., Gooiker, G.A. et al., Variation in case-mix between hospitals treating colorectal cancer patients in the Netherlands. Eur J Surg Oncol, 2011.
11. Schouten, L.J., Hoppener, P., van den Brandt, P.A., Knottnerus, J.A., Jager, J.J., Completeness of cancer registration in Limburg, The Netherlands. Int J Epidemiol, 1993. **22**(3): p. 369-76.
12. Schouten, L.J., J.J. Jager, and P.A. van den Brandt, Quality of cancer registry data: a comparison of data provided by clinicians with those of registration personnel. Br J Cancer, 1993. **68**(5): p. 974-7.
13. Charlson, M.E., et al., A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis, 1987. **40**(5): p. 373-83.
14. Mayer, E.K., et al., Funnel plots and their emerging application in surgery. Ann Surg, 2009. **249**(3): p. 376-83.
15. Daley, J., Khuri, S.F., Henderson, W., et al., Risk adjustment of the postoperative morbidity rate for the comparative assessment of the quality of surgical care: results of the National Veterans Affairs Surgical Risk Study. J Am Coll Surg, 1997. **185**(4): p. 328-40.
16. Khuri, S.F., Daley, J., Henderson, W., et al., Risk adjustment of the postoperative mortality rate for the comparative assessment of the quality of surgical care: results of the National Veterans Affairs Surgical Risk Study. J Am Coll Surg, 1997. **185**(4): p. 315-27.
17. Daly, L., Simple SAS macros for the calculation of exact binomial and Poisson confidence limits. Comput Biol Med, 1992. **22**: p. 351-361.
18. Hosmer, D.W.L., S, Applied logistic regression. 2000, Hoboken: John Wiley and Sons.
19. Reeves, D., Campbell, S.M., Adams, J., et al, Combining multiple indicators of clinical quality: an evaluation of different analytic approaches. Med Care, 2007. **45**(6): p. 489-96.
20. Habib, M.R., Solomon, M.J., Young, J.M., et al, Evidence-based and clinical outcome scores to facilitate audit and feedback for colorectal cancer care. Dis Colon Rectum, 2009. **52**(4): p. 616-22; discussion 622-3.
21. O'Brien, S.M., Shahian, D.M., DeLong, E.R., et al., Quality measurement in adult cardiac surgery: part 2–Statistical considerations in composite measure scoring and provider rating. Ann Thorac Surg, 2007. **83**(4 Suppl): p. S13-26.
22. de Groot, I.B., Otten, W., Smeets, H.J., Marang-van de Mheen, P.J., CHOICE-2 study group., Is the impact of hospital performance data greater in patients who have compared hospitals? BMC Health Serv Res, 2011. **11**(1): p. 214.
23. Covinsky, K.E., Fuller, J.D., Yaffe, K., et al., Communication and decision-making in seriously ill patients: findings of the SUPPORT project. The Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments. J Am Geriatr Soc, 2000. **48**(5 Suppl): p. S187-93.
24. Fagerlin, A., Ditto, P.H., Danks, J.H., Houts, R.M., Smucker, W.D., Projection in surrogate decisions about life-sustaining medical treatments. Health Psychol, 2001. **20**(3): p. 166-75.

Chapter 10

# A combined measure of procedural volume and outcome to assess hospital quality of colorectal cancer surgery, a secondary analysis of clinical audit data

N.E. Kolfschoten[1,2], P.J. Marang-van de Mheen[2], M.W.J.M Wouters[3], E.H.Eddes[4], R.A.E.M. Tollenaar[1], T. Stijnen[5], J. Kievit[1,2], on behalf of the Dutch Surgical Colorectal Audit group*

1 Leiden University Medical Centre, Dept. of Surgery
2 Leiden University Medical Centre, Dept. of Medical Decision Making
3 Netherlands Cancer Institute, Dept. of Surgical Oncology
4 Deventer Hospital, Dept. of Surgery
5 Leiden University Medical Centre, Dept. of Biostatistics
* The Dutch Surgical Colorectal Audit group: W.A. Bemelman, O.R.C. Busch, R.M. van Dam, E. van der Harst, M.L.E.A. Jansen-Landheer, Th.M. Karsten, J.H.J.M. van Krieken, W.G.T. Kuijpers, V.E. Lemmens, E.R. Manusama, W.J.H.J Meijerink, H.J.T. Rutten, C.J.H. van de Velde, T. Wiggers

*Plos-one, Accepted on 27th January 2014*

## Abstract

### Objective
To identify, on the basis of past performance, those hospitals that demonstrate good outcomes in sufficient numbers to make it likely that they will provide adequate quality of care in the future, using a combined measure of volume and outcome (CM-V&O). To compare this CM-V&O with measures using outcome-only (O-O) or volume-only (V-O), and verify 2010-quality of care assessment on 2011 data.

### Design
Secondary analysis of clinical audit data

### Setting
The Dutch Surgical Colorectal Audit database of 2010 and 2011, the Netherlands

### Participants
8911 patients (test population, treated in 2010) and 9212 patients (verification population, treated in 2011) who underwent a resection of primary colorectal cancer in 89 Dutch hospitals.

### Main outcome measures
Outcome was measured by Observed/Expected (O/E) postoperative mortality and morbidity. CM-V&O states 2 criteria; 1) outcome is not significantly worse than average, and 2) outcome is significantly better than substandard, with 'substandard care' being defined as an unacceptably high O/E threshold for mortality and/or morbidity (which we set at 2 and 1.5 respectively).

### Results
Average mortality and morbidity in 2010 were 4.1 and 24.3% respectively. 84 (94%) hospitals performed 'not worse than average' for mortality, but only 21 (24%) of those were able to proof they were also 'better than substandard' (O/E<2). For morbidity, 42 hospitals (47%) met the CM-V&O. Morbidity in 2011 was significantly lower in these hospitals (19.8 vs. 22.8% p<0.01). No relationship was found between hospitals' 2010 performance on O-O en V-O, and the quality of their care in 2011.

### Conclusion
CM-V&O for morbidity can be used to identify hospitals that provide adequate quality and is associated with better outcomes in the subsequent year.

## Introduction

Increasingly, society demands that health care providers demonstrate that the quality of the care they provide is adequate. However, it is not clear how quality should be measured and judged. Quality of health care has been defined as "the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge"[1]. Patient safety (the prevention of unintended harm) is an essential element of quality of care. Assessing quality, in particular patient safety, on the basis of outcome, i.e. the occurrence of adverse outcomes, has proven unreliable. The incidence of adverse outcomes is usually low, so the absence of adverse outcomes in a small series of patients is likely even if care were substandard. In the same way, when procedural volume is low, even an adverse event rate of 3 or 5 times average may still be 'not significantly worse than average'[2].

For this reason, and because higher procedural volume is associated with better outcome, the emphasis in quality assessment has shifted from outcome to volume. Procedural volume has thus become a surrogate measure for quality[3,4]. Accordingly, political focus in the Netherlands now aims at concentrating care into high volume centres. Recently, the Association of Surgeons of the Netherlands (ASN) has set the minimal annual procedural volume for colorectal resections at 50 procedures per hospital per year. However, any volume criterion is arbitrary and ignores the fact that lower volumes (e.g. 45 per year) do not exclude high quality, just as high volumes do not rule out substandard care. Therefore assessing quality on the basis of volume only, ignoring outcome, is as inadequate as assessing it by outcome only, ignoring volume. We therefore propose a quality measure that combines volume and outcome, and corrects for case-mix variation to provide statistical evidence that care is both 'not significantly worse than average' as well as 'significantly better than substandard'. A hospital that meets both criteria deserves public confidence that its quality of care is adequate.

**The aim** of the present study is to elucidate and test the proposed method and compare hospitals by three measures to define adequate quality: 1) outcome only (O-O), 2) volume only (V-O), and 3) a combined measure of volume and outcome (CM-V&O). We aim to demonstrate that the CM-V&O not only has a better theoretical basis, but that it also identifies hospitals with better outcome in the subsequent year.

## Patients and Methods

### Patients

We used the database of the Dutch Surgical Colorectal Audit (DSCA, www.clinicalaudit.nl), which has been created in accordance with principles pioneered earlier in the UK and the Scandinavian countries[5]. Details on procedures, data registration and data validity have been described in a recent paper[6]. For the present study (for which no ethical approval was required) we used data from patients treated in 2010 as study database, and validated our

measures on patients treated in 2011. The study population consisted of 8,911 patients who underwent a resection for a primary colorectal cancer during 2010 in one of 89 participating Dutch hospitals. The verification population consisted of 9,212 patients, treated in the same 89 hospitals in 2011. These databases include 93% of all patients treated, and 96% of Dutch hospitals. An observation period of 1 year was chosen, as this is the time-span commonly used for benchmarking in quality assurance. Data included 15 case-mix factors (age, gender, Body Mass Index (BMI), preoperative ASA-score, Charlson comorbidity-score[7], history of previous abdominal surgery, Tumour Node Metastasis (TNM) stage, preoperative radiation therapy, preoperative tumour complications (perforation, obstruction or other), multiple synchronous tumors, urgency and type of procedure (right, left/transverse, sigmoid, low anterior or abdomino-perineal resection, and/or extended resection for locally advanced tumour or metastases), as well as outcome. Outcome was assessed by postoperative mortality, occurring either during hospital admission or within 30 days after resection, and/or by serious morbidity, i.e. leading to an intervention (operative or percutaneous) or to prolonged hospital stay (14 days or more).

## Analyses

Risk-adjusted Observed/Expected (O/E) outcome ratio was used as the basic measure of hospital specific quality of care.[8,9] Observed outcome is the number of adverse outcomes (mortality or morbidity) that occur in a particular hospital in one year, while expected outcome is the sum of all patients' estimated probabilities for these outcomes in that same hospital that year. These probability estimates for patients' mortality and morbidity were derived from a backwards-stepwise multivariate logistic regression model, fitted on the data of that year, of all hospitals. For each of the 89 hospitals studied, O/E outcome ratios (including the exact Poisson 95% confidence intervals) were calculated both for mortality and morbidity, and for 2010 and 2011 separately[10]. For a hospital with average performance, the observed outcome will equal the expected outcome, resulting in an O/E outcome ratio of 1. Hospitals that perform better than average have an O/E outcome ratio lower than 1, while this ratio is higher than 1 in hospitals with poorer than average performance.

We compared the 89 hospitals according to 3 different quality measures:

**Outcome-only**, this (historical) measure assesses whether patient outcomes are not worse than a predefined 'threshold for substandard care'. For the present study, the base-case threshold for substandard care was defined as an O/E outcome ratio of 2 for mortality, and of 1.5 for morbidity (and was varied in sensitivity-analyses to 1.5 and 3). Note that this approach ignores the existence and size of a confidence interval (which depends partly on volume) around the point estimate of the O/E ratio.

**Volume-only**, a more recently proposed measure, assesses whether the volume of procedures (irrespective of outcome) is at least as high as the (arbitrary) threshold of 50 colorectal resections per year set by the Association of Surgeons of the Netherlands (including those for benign diseases). As we did not have any information on benign

procedures, for the present study the volume threshold was defined as at least 50 colorectal cancer resections in the year 2010.

A **combined measure of volume and outcome,** assesses not only whether outcome is adequate, but in addition whether patient volume is sufficiently high to narrow the uncertainty around the observed outcome to an acceptable range. In the CM-V&O, the minimal volume is therefore not a normative threshold, but a statistical condition for reliable assessment of hospital outcome. To pass this measure, a hospital should meet two criteria:

- its O/E outcome ratio must be 'not significantly worse than average', i.e. the lower limit of the 95% confidence interval (CI95min for short) of its O/E outcome ratio should be no higher than 1, and
- its O/E outcome ratio must be 'significantly better than substandard', i.e. the upper limit of the 95% confidence interval (CI95max for short) of its O/E outcome ratio should be lower than the predefined threshold for substandard care (see A).

Conceptually, these two criteria mean 'no proof of care being bad' and 'sufficient proof of care being OK' respectively, with the burden of proof lying in particular with the hospital for the second criterion.

### Verification

A good measure is not only *discriminative*, meaning that it will identify the adequately performing hospitals and detect the hospitals with insufficient quality, but also *reliable*, meaning that it will not only do so for the year measured, but also the following year. To validate the reliability of the CM-V&O, we compared 2011 outcomes between hospitals meeting or failing different measures in 2010, to see if hospitals' quality of care in 2011 was predicted by their performance on the measures O-O, V-O and CM-V&O in 2011. To translate the O/E outcomes to clinical outcomes, we also calculated the risk-adjusted outcomes (O/E outcome multiplied by the national average outcome). Also, we also performed the reversed analysis: we investigated if the different measures also detected hospitals with insufficient quality the following year by calculating how many of the hospitals that were 'significantly worse than average' in 2011, were detected by the different quality measures as delivering substandard care in 2010 (i.e. NOT meeting the different measures).

All statistics were performed in PASW Statistics for Mac, Rel 18.0.2009. Chicago: SPSS inc. and Microsoft Excel.

## Results

### Patients

The study population (2010) consisted of 8,911 patients treated in 89 hospitals, with an average procedural volume of 100 patients per year (range 14 – 241). The population average mortality and morbidity rates were 4.1% and 24.3% respectively. Seven out of

15 case-mix factors (age, gender, ASA, Carlson score, type of procedure, preoperative radiation therapy and urgency) predicted mortality with good accuracy. Morbidity was predicted by 12 factors (all except BMI, TNM-stage and synchronous metastasectomy). The model predicting mortality, and that predicting morbidity, had similar test-characteristics (C-statistics 0.80 CI95 0.78-0,82 and 0.75 CI95 0.72-0.78 respectively). The two models were used to calculate expected mortality and expected morbidity, and O/E mortality and morbidity ratios for each hospital. The average hospital O/E mortality ratio in 2010 was 1.04 (CI95 0.90-1.19), the average O/E morbidity ratio 0.97 (CI95 0.90-1.04). (Figure 1a)

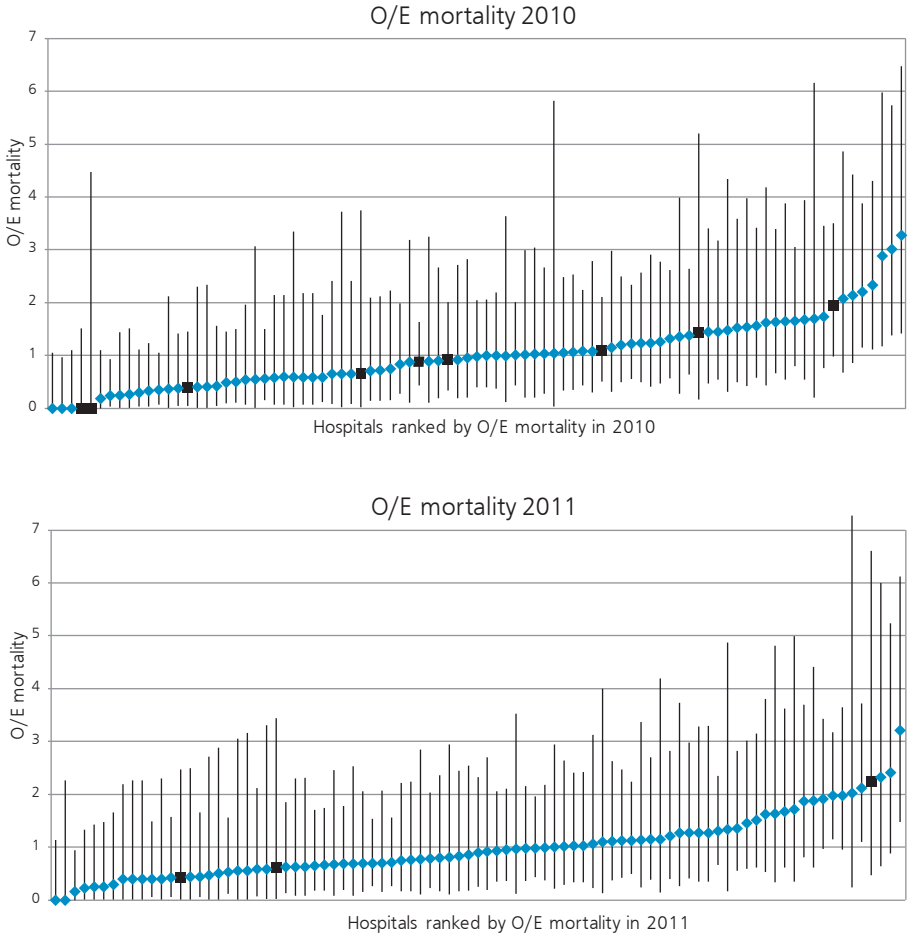



**Figure 1a.** Observed/Expected (O/E) mortality ratio of all hospitals in the Dutch Surgical Colorectal Audit in 2010 and 2011
Measure of 'Outcome only': requires that a hospital's O/E ratio, regardless of its confidence interval, is below or equal to 2 (red fat printed line), met the 'outcome only' criterion.
Measure of 'Volume only': requires that a hospital meets the 'volume only' criterion of > 50 procedures per year (red squares)
Combined measure of 'volume and outcome (CM-V&O)': the lower limit of the confidence interval around the hospitals O/E ratio is below or equal to 1, i.e. that the hospital is not significantly worse than average. In addition CM-V&O requires that the higher limit of the confidence interval is below 2 (fat printed red line), i.e. that the hospital is significantly better than substandard.

The verification population (2011) consisted of 9,212 patients, treated in the same 89 hospitals. Average morbidity and mortality rates decreased significantly in 2011, as compared to 2010: 3.7% ($p<0.01$) and 21.5% ($p<0.01$) respectively, as described previously.[11] Table 1 shows patient, tumour and treatment characteristics and outcome in the DSCA in the study population of 2010 and the verification population of 2011. Figure 1b shows the O/E mortality and morbidity ratios in 2011.

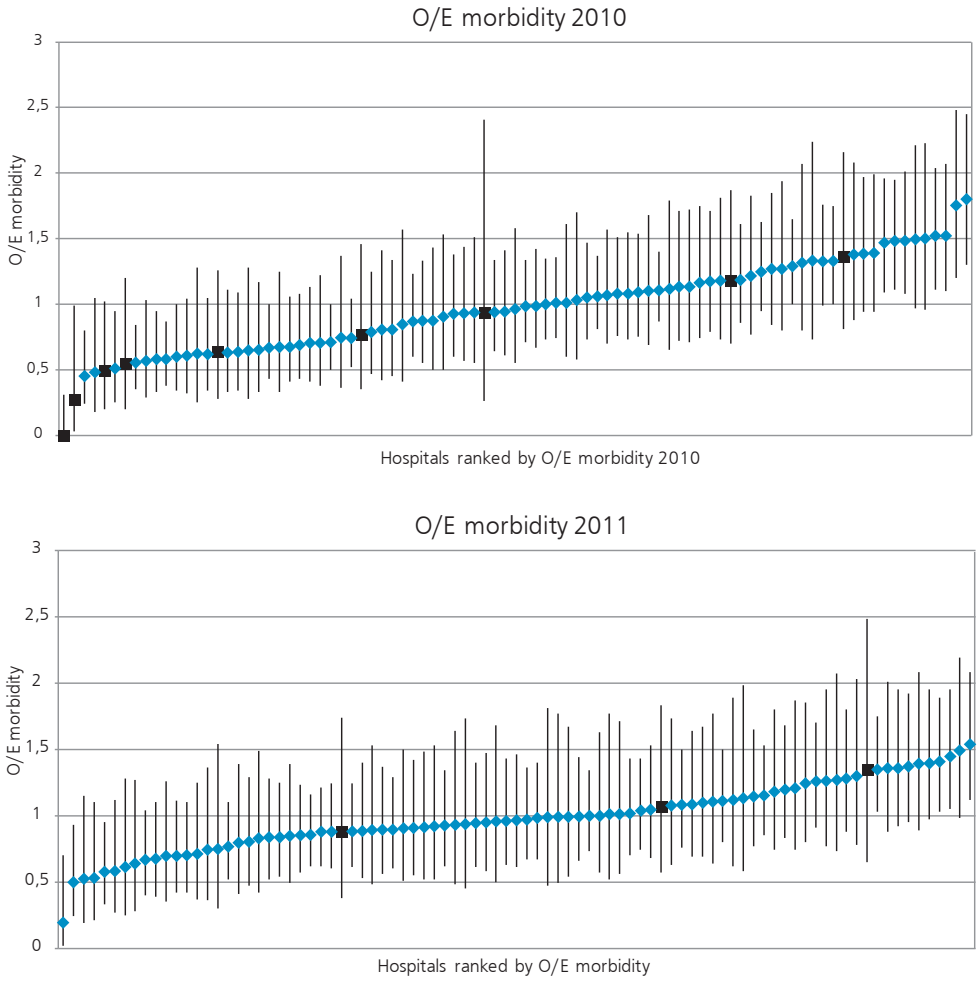



**Figure 1b.** Observed/Expected (O/E) morbidity ratio of all hospitals in the Dutch Surgical Colorectal Audit in 2010 and 2011
Measure of 'Outcome only': requires that a hospital's O/E ratio, regardless of its confidence interval, is below or equal to 1.5 (red fat printed line), met the 'outcome only' criterion.
Measure of 'Volume only': requires that a hospital meets the 'volume only' criterion of > 50 procedures per year (red squares)
Combined measure of 'volume and outcome (CM-V&O)': requires that the lower limit of the confidence interval around the hospitals O/E ratio is below or equal to 1, i.e. that the hospital is not significantly worse than average. In addition CM-V&O requires that the higher limit of the confidence interval is below 1.5 (fat printed red line), i.e. that the hospital is significantly better than substandard.

**Table 1.** Patient, tumour and treatment characteristics and outcome in the DSCA 2010.

| Case-mix factors | | 2010 N (%) | 2011 N (%) |
|---|---|---|---|
| Age (years) | <70 | 4006 (45) | 4222 (46) |
| | 70-79 | 2924 (33) | 3038 (33) |
| | 80+ | 1975 (22) | 1935 (21) |
| Gender | Male | 4914 (55) | 5065 (55) |
| Previous abdominal surgery | Yes | 2925 (33) | 3087 (34) |
| | Missing | 132 (1.5) | 117 (1.3) |
| ASA | III+ | 2094 (24) | 2156 (23) |
| | Missing | 82 (1) | 60 (1) |
| Charlson | 2+ | 969 (11) | 1054 (11) |
| BMI | <25 | 3248 (36) | 3678 (40) |
| | 25-30 | 2769 (31) | 3468 (38) |
| | >30 | 1169 (13) | 1362 (15) |
| | Missing | 1725 (19) | 703 (8) |
| TNM stage | I and II | 4884 (55) | 4900 (53) |
| | III | 2769 (31) | 2896 (31) |
| | IV | 1053 (12) | 1053 (11) |
| | X | 205 (2) | 363 (4) |
| Synchronous tumours | Yes | 246 (3) | 311 (3) |
| Neoadjuvant therapy | Radiotherapy 5x5 gy | 1257 (14) | 1268 (14) |
| | Radiotherapy >60 gy | 173 (2) | 256 (3) |
| | Chemo-radiation | 727 (8) | 789 (9) |
| Preoperative complication | Feacal peritonitis | 153 (2) | 117 (1) |
| | Obstruction | 798 (9) | 103 (11) |
| | Other | | |
| Urgency | Urgent | 794 (9) | 679 (7) |
| | Acute | 487 (6) | 648 (7) |
| Procedure | Right hemi-colectomy | 3142 (35) | 3013 (33) |
| | Left hemi-colectomy* | 830 (9) | 892 (10) |
| | Sigmoid / LAR | 3783 (43) | 4107 (45) |
| | Abdomino-perineal resection | 732 (8) | 812 (9) |
| | Other ^ | 442 (5) | 388 (4) |
| Additional resection | Locally advanced tumour | 840 (9) | 875 (10) |
| | Metastasis | 250 (3) | 351 (4) |
| Mortality | | 365 (4.1) | 343 (3.7) |
| Morbidity | | 2167 (24.3) | 1982 (21.5) |

ASA= American Society of Anaesthesiologists classification BMI= Body Mass Index, TNM= Tumour Node Metastasis classification system, gy= gray LAR= Low anterior resection, * = including transverse colectomies, ^ = including subtotal or proctocolectectomy

*Three measures for adequate quality in 2010*

Using O-O as the measure for adequate quality, 82 (92%) of the hospitals met the O/E ratio ≤2 for mortality, and 84 (94%; see table 2) met the O/E ratio ≤ 1.5 for morbidity.

With respect to V-O, 9 hospitals (10%) performed less than 50 resections for primary colorectal cancer, while the remaining 80 hospitals (90%) met the volume threshold of 50 or more (table 2).

**Table 2.** Number of hospitals meeting the measure of 'outcome-only' (better than substandard), of 'volume-only' (50 procedures or more), or the combined measure of volume and outcome' in 2010, using different thresholds for 'substandard care'.

| Quality measure | | Outcome only | Volume only | Combined measure of volume and outcome | | |
|---|---|---|---|---|---|---|
| Outcome measure | Threshold for substandard care | O/E <= substandard (%) | N >= 50 (%) | Not worse than average: CI95min < =1 (%) | Better than substandard: CI95max < threshold (%) | Both (%) |
| Mortality | O/E < 1.5 | 71 (80) | 80 (90) | 84 (94) | 13 (15) | 13 (15) |
| | O/E < 2 | 82 (92) | | | 21 (24) | 21 (24) |
| | O/E < 3 | 87 (98) | | | 59 (66) | 56 (63) |
| Morbidity | O/E < 1.5 | 84 (94) | 80 (90) | 81 (90) | 47 (53) | 42 (47) |
| | O/E < 2 | 89 (100) | | | 77 (87) | 74 (83) |
| | O/E < 3 | 89 (100) | | | 89 (100) | 81 (90) |

N = number of patients
O/E = Observed/Expected outcome ratio
CI95min = Lower limit of 95% confidence interval around the O/E ratio
CI95max = Upper limit of 95% confidence interval around the O/E ratio

**CM-V&O** states, as explained before, 2 criteria:

- *Criterion 1): being 'not significantly worse than average'*

For mortality, 84 hospitals (94%) had an O/E mortality ratio which was 'not significantly worse than average', CI95min <= 1 (table 2). Three of these hospitals even performed significantly better than average (i.e. even CI95max < 1). The remaining 5 hospitals had an O/E mortality ratio worse than average (i.e. CI95min > 1). For morbidity, 81 (90%) hospitals had an O/E morbidity ratio that was 'not significantly worse than average' (table 2), of which 11 hospitals were even significantly better than average. In the remaining 8 hospitals, O/E morbidity ratio was significantly worse than average. One hospital performed significantly worse than average on both mortality and morbidity, and one hospital was significantly better than average on both outcome measures. None of the hospitals that were significantly better than average on one outcome measure were 'significantly worse than average' on the other.

- *Criterion 2): being 'significantly better than substandard'*

At an O/E outcome ratio of 2 as the threshold for substandard care, 21 hospitals (24%) were 'significantly better than substandard' (i.e. CI95max < 2, see table 2). All of these hospitals were also 'not significantly worse than average' and thus met both criteria of CM-V&O. For morbidity, 77 hospitals were 'significantly better than substandard' (CI95max <2), of which 74 met both criteria. All hospitals that met both criteria for mortality also did so for morbidity.

*Using different thresholds for 'substandard care'*

If we had used a stricter threshold for substandard care, such as an O/E outcome ratio of 1.5, the number of hospitals meeting CM-V&O would have dropped to just 13 for mortality,

and to 42 for morbidity (table 2). Using a more lenient threshold for substandard care of 3, 56 hospitals would have met CM-V&O for mortality, and 81 for morbidity (table 2).

*Comparing the three quality measures*

Table 3 shows that 61 hospitals met the O-O measure for mortality, but had insufficient procedural volume to assess hospital postoperative outcome reliably. As a result of this their CI95max ranged up to 7 times the expected mortality. For morbidity, 37 hospitals met the O-O measure, but not the CM-V&O. Similarly, 61 hospitals met the V-O measure for mortality but outcomes were inadequate to meet the CM-V&O (for morbidity: 39 hospitals). Among these 61 hospitals there were also the 5 hospitals with an O/E mortality ratio significantly worse than average (for morbidity: 8 hospitals). On the other hand, there were also 2 hospitals that did meet the CM-V&O, but not the V-O measure (for morbidity: 6 hospitals).

**Table 3.** Comparing the number of hospitals meeting (or failing) the 3 different measures

| Quality measures met in 2010 | | | Combined measure of volume and outcome 2010*) | |
|---|---|---|---|---|
| | | | yes | No |
| Mortality | outcome only* | O/E =< 2 | 21 | 61 |
| | | O/E >2 | 0 | 7 |
| | volume only | N => 50 | 19 | 61 |
| | | N < 50 | 2 | 7 |
| Morbidity | outcome only* | O/E =< 1.5 | 47 | 37 |
| | | O/E >1.5 | 0 | 5 |
| | volume only | N => 50 | 41 | 39 |
| | | N < 50 | 6 | 3 |

O/E = Observed/Expected outcome ratio; N = Hospital procedural volume; *) using O/E ratio of > 2 as the threshold for substandard care for mortality, and O/E ratio of > 1.5 as the threshold for substandard care for morbidity.

## Verification

Hospitals meeting the CM-V&O for mortality in 2010 had a lower risk-adjusted mortality, than hospitals that did not meet the CM-V&O in 2010, but the difference was not statistically significant (3.3 vs 3.9%, n.s.) [Table 4]. Hospitals that met the V-O measure in 2010 had a higher, rather than lower, risk-adjusted mortality in 2011. The CM-V&O detected 2 of 3 hospitals which performed significantly worse than average in 2011 (these hospitals did not meet the CM-V&O in 2010), while none of the other measures detected any of the significantly worse than average hospitals in 2011 (data not shown). For morbidity, hospitals meeting the CM-V&O in 2010 had a significantly lower risk-adjusted morbidity in 2011 (19.8 vs 22.8%, $p<0.05$), [Table 4] while this effect was not found for the O-O or V-O measure. The CM-V&O detected 3 out of 4 hospitals that were significantly worse than average in 2011, while the O-O measure detected only one, and the V-O measure detected none of the significantly worse than average hospitals in 2011 (data not shown).

**Table 4.** Verification: Outcomes in 2011 of hospitals meeting the different measures in 2010

| Mortality 2010 | | O/E Mortality 2011 (CI95%) | Mortality 2011 (risk-adjusted) |
|---|---|---|---|
| Outcome <2 2010 | Yes (82) | 0.98 (0.88-1.09) | 3.6% |
| | No (7) | 1.36 (0.93-1.98) | 5.0% |
| Volume >50 in 2010 | Yes (80) | 1.01 (0.90-1.12) | 3.7% |
| | No (9) | 0.89 (0.59-1.35) | 3.3% |
| CM-V&O 2 2010 | Yes (21) | 0.89 (0.72-1.10) | 3.3% |
| | No (68) | 1.04 (0.92-1.18) | 3.9% |
| CM-V&O 3 2010 | Yes (56) | 0.98 (0.80-1.19) | 3.6% |
| | No (33) | 1.01 (0.89-1.15) | 3.7% |
| Morbidity 2010 | | O/E Morbidity 2011 (CI95%) | |
| Outcome <1.5 | Yes (84) | 0.99 (0.95-1.04) | 21.3% |
| | No (5) | 1.18 (0.99-1.41) | 25.5% |
| Volume >50 | Yes (80) | 1.00 (0.96-1.05) | 21.6% |
| | No (9) | 0.94 (0.78-1.14) | 20.3% |
| **CM-V&O 1.5*** | **Yes (47)** | **0.92 (0.89-1.00)** | **19.8%** |
| | **No (42)** | **1.06 (1.01-1.14)** | **22.8%** |
| **CM-V&O 2*** | **Yes (74)** | **0.97 (0.92-1.02)** | **20.8%** |
| | **No (15)** | **1.16 (1.05-1.28)** | **25.0%** |

CM-V&O = combined measure for volume and outcome; O/E = Observed/Expected outcome ratio; CI95% = 95% confidence interval around the O/E ratio; * $P < 0.05$

## Discussion

In the present study we propose and test a (risk-adjusted) combined measure of volume and outcome to assess the quality of care provided by hospitals. Hospitals that meet this quality measure have not only demonstrably good health care outcomes, but had sufficient annual numbers to demonstrate that their results are not just a 'lucky streak', but a manifestation of consistently good underlying quality of care. We demonstrated in the verification population that hospitals meeting the CM-V&O for morbidity had a significantly lower morbidity in the following year. A similar trend was found for mortality; however this did not reach statistical significance. Both the 'volume only' and 'outcome only' measure did not identify the hospitals with better outcomes the following year. Also, the CM-V&O performed better in detecting hospitals that performed significantly worse than average the next year.

The advantage of our study is that all analyses were performed on large, reliable databases, containing almost nine thousand patients per year from 89 hospitals. However, this also illustrates that such calculations can only be performed if outcome registration is excellent and relevant case-mix factors are included. Fortunately, increasing awareness of the need for quality assurance has led to an increasing dedication to reliable outcome registration by means of National Clinical Audits. In the Netherlands, the first Clinical Audit, the DSCA was initiated in 2009. Although participation, completeness and validity were overwhelmingly good after only one year of registration, the availability of weekly online feedback data on hospital performances was relatively new in the two years described in this study. In three

years after the introduction of clinical auditing, a significant improvement in various process and outcome measures was observed, while variation in hospital performances decreased.[11] These improvements may have interfered with the results of our study as hospitals with substandard outcomes may have felt a greater incentive for improvement of these outcomes, reducing the possible differences in outcome in 2011 between hospitals that did and did not meet the different measures. Possibly, when this selective quality improvement would not have taken place, hospitals meeting the CM-V&O in 2010 would also have had a lower risk adjusted mortality rate in 2011.

The ultimate measure for hospital performances on outcomes of care would be *discriminative* and *reliable,* but also understandable for all stakeholders. Although previous studies have found that a high procedural volume for colorectal cancer is associated with better outcome,[12] we found that the recently implemented minimal volume chosen by the ASN (the 'volume only' measure) was not *discriminative*, as it excluded hospitals with reliable and good outcomes, and included hospitals with significantly worse than average outcomes. We found that the measure was not *reliable*, as hospitals with insufficient volume according to the ASN, although few, did not perform worse than the hospitals with sufficient numbers in 2011, but that their outcomes, although not significant, were even better. Possibly, a cut-off point of 50 procedures is too low to identify high-quality hospitals. However, our results are in line with recent evidence that suggests that centralization only results in an improvement of outcome if patients are referred to hospitals with a better outcome: outcome-based referral[13-16]. On the other hand, we have also shown that selecting hospitals based on 'outcome only', was also not *reliable*, as it did not detect hospitals performing significantly worse than average in 2011. Therefore, we propose to judge hospitals using the CM-V&O. This measure identified the hospitals with better outcomes the following year, and detected most of the hospitals performing significantly worse than average the following year, and was therewith most discriminative and reliable. Moreover, this simple measure is also understandable for all stakeholders. It may be argued that the CM-V&O, is too strict for the Dutch hospitals, as only 24% of all hospitals met the measure. By varying the 'level of substandard care' to for example an O/E mortality of 3 (e.g. 3 times higher mortality rate than expected based on the hospitals' case-mix) the CM-V&O may be adjusted so that more hospitals meet the measure. However, as the number of hospitals in relation to the size of the Dutch population is exceptionally high, and therefore the average procedural volume low, it may be argued that the CM-V&O is not 'too strict', but that it exposes the limited accountability of Dutch hospitals in their current number. Possibly the CM-V&O is even more discriminative and reliable, and selects more hospitals when tested in a different, larger healthcare system.

Other studies have proposed similar composite measures using volume and outcome to assess hospital performances,[17] with similar results. However, they used the Empirical Bayes method to adjust outcomes for the procedural volume. The main difference between our

approach and the Empirical Bayes method is that the Empirical Bayes method uses hospital rankings (from best to worst performing) instead of rating (better or worse than average), and therefore also takes into account the uncertainty of the position of the hospital, based on its performances, in relation to the position of the other hospitals. To identify outlier hospitals using the Empirical Bayes method, large procedural volumes are needed. Therefore, in the Dutch population, using the Empirical Bayes method, results in a 'flat line' with all hospitals performing 'average', and fails to identify better performing or underperforming hospitals. Moreover, a recent study describes that after the introduction of the DSCA not only national average outcomes improved, but also differences in hospital performance were reduced,[11] meaning that the position of one hospital, relative to another becomes more uncertain. This study proposes a more simple method, looking only at the position of the hospital in relation to the national average. Therewith, the CM-V&O does not classify the one hospital as performing better than the other, but simply classifies a hospital good enough or not good enough.

Although the CM-V&O performed better than the other two measures in identifying the better performing hospitals and detecting the underperforming hospitals, we found no significant difference in mortality in 2011 between hospitals that did and did not meet the CM-V&O in 2010. As Dishoeck et al. showed in their study, the 'rankability' of hospitals (the part of the heterogeneity between hospitals that is due to unexplained, hospital dependent differences) is highly dependent of the number of events in the different hospitals.[18] For mortality, the rankability is rather low meaning that most of the differences between hospitals is due to random variation and may thus be less suitable to rank hospitals. For morbidity on the other hand, rankability is much better as there is more systematic variation. This may also explain why the CM-V&O performed better for morbidity than for mortality. As mortality is just one aspect of quality, the CM-V&O should preferably be applied for both mortality and morbidity, or even for composite quality measures that combine both short and long term outcome, adverse as well as desirable.[19]

Our study resembles clinical audits or quality registration programs in other countries, such as the National Surgical Quality Improvement Program in the United States of America, or the various nation-wide registries in European countries. Some of these registrations also use the combination of volume and outcome to produce annual hospital ratings. However, they identify positive and negative outliers, but leave the majority of hospitals unclassified[20,21], arguing that it cannot be proven that quality of care in these hospitals is insufficient. This line of reasoning differs from the relationship between providers and clients across many other areas of society, where the burden of proof for a good product lies with the provider, instead of the burden of proof for substandard quality lying with the client. The analogue in healthcare is that nowadays society will not settle for the lack of statistical proof that care is substandard, but demands evidence that the quality of care is adequate, in particular for high-risk procedures. The CM-V&O that we propose does exactly that.

Policy makers in many countries increasingly respond to societal concerns about health care safety and quality. In the Netherlands, societal demand for transparency has been formulated by the Dutch Health Care Inspectorate as the need for "justified trust". Hospital volume has been chosen as a proxy for quality, backed up by enforced volume-based referral in an attempt to improve outcomes.[3,22]. The present study suggests that CM-V&O is on both theoretical and practical grounds, better suited than the volume-measure to provide the "justified trust" in quality of care that society demands.

## Acknowledgements

We thank all surgeons, registrars, physician assistants and administrative nurses that registered all the patients in the DSCA, as well as the Dutch Surgical Colorectal Audit group and the methodological board for their advice.

## References


1. Medicine Io. Committee to Design a Strategy for Quality Review and Assurance in Medicine. The Institute of Medicine Report. In Medicare: A Strategy for Quality Assurance/Committee to Design a Strategy for Quality Review and Assurance in Medicare. Division of Health Care Services, Institute of Medicine, Vol 1, report of a study, p 468 Washington, DC: National Academy Press 1990
2. Dimick JB, Welch HG, Birkmeyer JD. Surgical mortality as an indicator of hospital quality: the problem with small sample size. JAMA 2004;292:847-51.
3. Birkmeyer JD, Siewers AE, Finlayson EV, et al. Hospital volume and surgical mortality in the United States. N Engl J Med 2002;346:1128-37.
4. van Gijn W, Gooiker GA, Wouters MW, Post PN, Tollenaar RA, van de Velde CJ. Volume and outcome in colorectal cancer surgery. Eur J Surg Oncol;36 Suppl 1:S55-63.
5. van Gijn W, Wouters MW, Peeters KC, van de Velde CJ. Nationwide outcome registrations to improve quality of care in rectal surgery. An initiative of the European Society of Surgical Oncology. J Surg Oncol 2009;99:491-6.
6. Verhaart IE, Heemskerk H, Karnaoukh TG, et al. Prednisolone treatment does not interfere with 2'-O-methyl phosphorothioate antisense-mediated exon skipping in Duchenne muscular dystrophy. Human gene therapy 2012;23:262-73.
7. Elferink MA, Krijnen P, Wouters MW, et al. Variation in treatment and outcome of patients with rectal cancer by region, hospital type and volume in the Netherlands. Eur J Surg Oncol 2010;36 Suppl 1:S74-82.
8. Daley J, Khuri SF, Henderson W, et al. Risk adjustment of the postoperative morbidity rate for the comparative assessment of the quality of surgical care: results of the National Veterans Affairs Surgical Risk Study. J Am Coll Surg 1997;185:328-40.
9. Khuri SF, Daley J, Henderson W, et al. Risk adjustment of the postoperative mortality rate for the comparative assessment of the quality of surgical care: results of the National Veterans Affairs Surgical Risk Study. J Am Coll Surg 1997;185:315-27.
10. Daly L. Simple SAS macros for the calculation of exact binomial and Poisson confidence limits. Comput Biol Med 1992.; 22:351-61.
11. Van Leersum NJ, Snijders HS, Henneman D, et al. The Dutch Surgical Colorectal Audit. Eur J Surg Oncol 2013.
12. van Gijn W, Gooiker GA, Wouters MW, Post PN, Tollenaar RA, van de Velde CJ. Volume and outcome in colorectal cancer surgery. Eur J Surg Oncol 2010;36 Suppl 1:S55-63.
13. Wouters M. Outcome-based referral to improve quality of care in upper GI surgery. J Surg Oncol 2008;98:307-8.
14. Wouters MW, Krijnen P, Le Cessie S, et al. Volume- or outcome-based referral to improve quality of care for esophageal cancer surgery in The Netherlands. J Surg Oncol 2009;99:481-7.
15. Simunovic M, Urbach D, Major D, et al. Assessing the volume-outcome hypothesis and region-level quality improvement interventions: pancreas cancer surgery in two Canadian Provinces. Ann Surg Oncol;17:2537-44.
16. Sonnenday CJ, Birkmeyer JD. A tale of two provinces: regionalization of pancreatic surgery in Ontario and Quebec. Ann Surg Oncol;17:2535-6.
17. Dimick JB, Staiger DO, Baser O, Birkmeyer JD. Composite measures for predicting surgical mortality in the hospital. Health Aff (Millwood) 2009;28:1189-98.
18. van Dishoeck AM, Lingsma HF, Mackenbach JP, Steyerberg EW. Random variation and rankability of hospitals using outcome indicators. BMJ quality & safety 2011;20:869-74.
19. Kolfschoten NE, Kievit J, Gooiker GA, et al. Focusing on desired outcomes of care after colon cancer resections; hospital variations in 'textbook outcome'. Eur J Surg Oncol 2013;39:156-63.
20. Khuri SF, Daley J, Henderson WG. The comparative assessment and improvement of quality of surgical care in the Department of Veterans Affairs. Arch Surg 2002;137:20-7.
21. Westaby S, Archer N, Manning N, et al. Comparison of hospital episode statistics and central cardiac audit database in public reporting of congenital heart surgery mortality. BMJ 2007;335:759.
22. Blamey RW, Catalotti, L. . EUSOMA: the requirements of a specialist breast unit; first revision in European guidelines for quality assurance in breast screening and diagnostic European Communities, 4th edition 2006:343-54.

# Chapter 11

## Discussion and Summary of this thesis

Recent reports on variation in process and outcomes of care between hospitals, in combination with several healthcare scandals, have led to a societal call for more information on hospital quality of care. These calls have been further amplified by the rapidly increasing costs in healthcare. Care providers are more and more being held accountable for the quality of care delivered and for the spending of health care means. In order to maintain societies trust in quality of care in the Dutch hospitals, and to control healthcare costs, valid information on hospital performance is needed. To answer these needs, a new system for systematic registration and feedback of information on treatment processes and outcomes of care was introduced: Clinical auditing.

In 2009 the Dutch surgeons introduced clinical auditing in the Netherlands, with the start of the Dutch Surgical Colorectal Audit (DSCA, www.clinicalaudit.nl). The goals of the se surgeons were

- to give hospitals an instrument for *improvement* of their practices and *outcomes*, and to gain insight in and reduce hospital variations in practice, guideline adherence and outcome,
- to gain more insight in *national* practice and *performance*, to set benchmarks, and identify aspects that need improvement and to find and implement best practices
- to evaluate and monitor the *implementation of new techniques*
- to answer the need for *transparency* to all stakeholders

However, a uniform methodology to realize these goals was not available. One of the most urgent research questions was the definition and content of 'valid and meaningful feedback information'. Therefore, the aim of this thesis was to investigate how data from clinical audits can be used to produce valid and meaningful feedback information, which may support improvement of quality of care, using the DSCA as an example. **Part one** aims to evaluate the *use of clinical auditing as an instrument for the improvement of quality of care in surgery.* **Part two** investigates *how data from clinical audits can be used to monitor and improve national practice and performance*, especially for high-risk patient groups. **Part three** investigates the use of clinical auditing for the *evaluation and monitoring of the implementation of new techniques on a national level.* **Part four** investigates how data can be used to *evaluate quality of care and give transparency to all stakeholders.* We investigate how various aspects of quality of care are related and how these data can be used in composite-measures, which can be used by all stakeholders to evaluate quality of care as a whole.

The aim was to develop and improve methodologies that may be used by many other clinical audits.

# Part 1. The use of clinical auditing as an instrument for improvement of quality of care

Many studies report on the positive effects of clinical auditing on quality of care. In clinical audits various aspects of the care process and its outcomes are registered and used to provide feedback of benchmarked performance information to individual hospitals. For example the national surgical quality improvement program (NSQIP) was initiated 20 years ago in the Veterans Affairs hospitals in the United States, after reports that the quality of care in these hospitals was far below the national average. After the introduction of this clinical audit a 45% reduction in postoperative morbidity and a 27% reduction in mortality after surgery was observed in 7 years[1]. Similar results were reported after introduction of this auditing system in the private sector. Other examples of the positive effect of clinical auditing are a 25% reduction in postoperative mortality after Coronary Artery Bypass Graft surgery, only one year after feedback of outcome data to the surgeons[2], or the reduction of local recurrence after rectal cancer resections in Norway after the introduction of clinical auditing from 28% to 7%.[3] We showed in a systematic review that clinical auditing has a positive effect on the outcomes of surgical care. [**Chapter 2**] A similar result was found for quality of care in non-surgical care in a recent Cochrane review.[4]
However, it remains difficult to prove that the quality improvements shown can be attributed to clinical auditing alone, as there may be more mechanisms that contribute to the quality improvement.

- The first mechanism is *regression to the mean*: a hospital with a high mortality rate in one year is very likely to have a lower mortality rate the next year as a result of statistical random variation. Therefore, when a hospital participating in a clinical audit, takes action to improve substandard results of the previous year, it is highly likely that an improvement is found. The effect of random variation is larger when hospital volumes are smaller, or the event-rate of the outcome measure is lower.[5] However, the continuous overall improvement as described after implementation of national clinical audits cannot be discounted by this effect.[1,6]

- The second mechanism is that of *autonomous quality improvement*. The simplest example for autonomous quality improvement is the implementation of a new treatment during the clinical audit, which may contribute to the improvement outcomes. Also, autonomous quality improvement is often seen during participation in a study as a result of the increased focus on outcomes: the Hawthorne effect. However, this effect may also be seen as the basis of clinical auditing: the constantly renewing focus on outcomes of care creates a continuous 'Hawthorne effect'. Although the Hawthorne effect is commonly used in industrial psychology as a strategy to enhance human performance,[7,8] the idea of using this Hawthorne effect in medicine, to improve quality of care is relatively new. Clinical auditing adds a control system to this 'Hawthorne effect', by providing timely

feedback on the magnitude of the improvement needed, and the results of improvement interventions. Where the Hawthorne effect, in its traditional form (e.g. without feedback) is thought to be 'fading over time',[9] clinical auditing provides periodic change of outcome measures, and constant motivation and improvement impulses , which may help to constantly renew the Hawthorne effect.[9,10]

- The third mechanism is the *perverse incentive*. Although clinical auditing is designed to monitor the treatment process as a whole, there are always 'holes in the maze'. For example, in-hospital mortality may be improved by rapidly discharging patients to a hospice. After a few years of auditing, some hospitals may discover these 'holes in the maze', and learn to use them to improve their outcomes in the registration, without improving their quality of care.[11] Thorough data validation with other data sources may prevent this *perverse incentive*.

## The missing link

Although in most industries it is common practice to frequently 'audit' process and outcomes, in order to improve efficiency and results, this idea is relatively new in medicine. Most industries work according to a PDCA circle (Plan-Do-Check-Act), which enables a constant quality improvement. The audit results in an improvement plan, which is carried out, the effect of the improvement intervention is evaluated and additional improvement interventions are initiated where needed. In medicine however, too often, innovation is limited to a ' PD' cycle: an innovation is designed and implemented, but no 'check' is performed to see if these new innovations are effective. Also, audits to identify flaws in the care process were uncommon. Clinical auditing is the missing link in medicine, which enables caregivers to monitor and improve, and be transparent on the quality of their care, on a hospital level, but also on a national level. However for an (inter)national clinical audit to be successful, good data quality and timely availability is essential.

## A successful audit

The first key to a successful (inter)national clinical audit is an up-to-date complete and valid database, which is described on five levels.

- *Participation*: all hospitals in the target area participate. For example, the Project on Cancer of the Rectum (PROCARE) Belgium has a voluntary participation policy resulting in only 40% of all patients treated in Belgium being registered in the audit. All other hospitals do not benefit from the audit. Also, results from this audit may not give a fair reflexion of national performance, as participating hospitals most likely to be the more dedicated hospitals with better results.
- *Case-ascertainment:* all cases in each hospital are registered. In the United Kingdom, participation in the National clinical Bowel Cancer Audit (NBOCAP) is also voluntary. Although participation has increased from 44.3% in 2006 to 90.1% in 2011, case ascertainment is still poor, and 23.3% of all hospitals still fail to submit more than 50%

of their colorectal workload.[12] Recent studies have shown that hospitals submitting all data on all patients had a significantly lower 30-day mortality compared to hospitals submitting less than 10% of all cases.[12] A possible explanation is that the hospitals submitting all data were the more dedicated hospitals with good outcomes, resulting in a selection bias.[13,14] Using such data as a national reference may result in an unrealistic representation of national performances, depicturing average performing hospitals as negative outliers, and discouraging average and underperforming hospitals to submit their data.

- *Completeness*: all relevant clinical information of each case is registered. For example, when case-mix information is incomplete, hospitals risk-adjusted outcomes may be insufficiently adjusted for their case-mix, and quality of care may thus appear worse (or better) than it actually is.
- *Timeliness*: registration is preferably synchronized with clinical reports, or at least done shortly after information becomes available. Previous studies have shown that frequent and timely feedback is an important driver for quality improvement.[15] In the Netherlands, the NKR registered outcomes of cancer care long before the DSCA was initiated, however, data only became available a year after the date of surgery. Therefore, reports based on these data were considered 'out-dated' and were seldom used to improve quality of care.
- *Validity*: registration is truthful, and can be validated against other data sources.

The Dutch Surgical Colorectal Audit (DSCA), initiated in 2009, was designed after these many international examples. Within two years, all Dutch hospitals participated in the audit. Case-ascertainment was 92% in 2010 and 95% in 2011, and completeness was 100% on almost all items. [**Chapter 3**] We identified the 'driving' key elements that lead to this successful implementation of this audit, and describe recent results after introduction of this audit. These 'driving' key elements were

- a leading role of the professional association in the development of the dataset and outcome measures
- integration of the audit in the national quality assurance policy of the professional association
- a web-based registration system and registration by medical specialists
- weekly updated online valid and meaningful feedback to participants
- annual external data verification with other data sources
- quality standards set by the professional association and introduction of improvement projects to meet these standards, the first standard being participation and full case-ascertainment in the audit. [**Chapter 3**]

These key elements have recently been confirmed in a systematic review.[16] The DSCA was designed as a blueprint for auditing in the Netherlands, and many clinical audits have followed since.

### Results from the DSCA

Hospitals participating in the DSCA receive weekly updated feedback data on their performance, compared to those of all other hospitals in the Netherlands. These feedback data include data on patient and tumour characteristics, and treatment and outcome. Feedback data on treatment concerning all crucial steps required in the treatment of colorectal cancer, as described by the national evidence based guidelines. From 2009 to 2011, a significant increase in overall guideline adherence was found on several processes, such as the percentage of patients discussed in a preoperative multidisciplinary team, the use of preoperative Magnetic Resonance Imaging (MRI) for rectal cancer surgery, and the standard of pathological reporting of the Circumferential Resection Margin (CRM). Moreover, we also found that variation in hospital performance decreased. [**Chapter 3**] After 2011 guideline adherence has even further improved.[17]
Feedback data on outcomes of colorectal cancer care also included postoperative morbidity, length of hospital stay and mortality. Postoperative morbidity, length of hospital stay and postoperative mortality and complications decreased significantly from 2009 to 2011 both for colon and rectal cancer surgery. Also, duration of hospital stay diminished by 2 days (both after colon and rectal resections). [**Chapter 3**] After 2011 outcomes improved further on a national level.[17]

As stated before, other mechanisms than clinical auditing only may have contributed to these improvements. However, the continuous improvement on various outcomes, the reduction of hospital variations, and the thorough validation with the Netherlands Cancer Registry support the idea that clinical auditing can be used as an instrument for the improvement of surgical quality of care.

## Part 2. The use of data from clinical audits to evaluate and improve national practice

Audit data can be analysed at a national level, to investigate current national practice and results, to identify high-risk patients and evaluate current referral patterns for these high-risk patients. Extensive research has described risk factors for unfavourable outcome after colorectal surgery, and hospital differences in case-mix. However, neither the distribution of 'high-risk patients' between different hospitals in the Netherlands nor the impact of known risk-factors on outcomes in the Dutch population had been studied before. As high-risk patients are often not included in clinical trials, information on outcomes of treatment for these patients, and scientific evidence on their best treatment are limited. Data from national clinical audit may help to gain more insight in treatment and outcomes for these patient groups.

We have shown that the 'expected mortality': the patient's predicted mortality risk during or after colorectal cancer surgery in the Netherlands, based on the patients' case-mix factors, can help to gain insight on how high-risk patients are distributed between Dutch hospitals. We found considerable differences in 'expected mortality' between individual hospitals and different types of hospitals. We also found that patients with colon cancer with a high risk for postoperative mortality in the Netherlands were more likely to be treated in low-volume hospitals or non-teaching or teaching hospitals, rather than in university hospitals. The higher expected mortality in these hospitals was mostly explained by these hospitals treating patients with a higher ASA-classification, with more comorbid diseases, and more often in a non-elective setting. We found no differences in Standardized Mortality Rates between different types of hospitals. [**Chapter 4**] In the Netherlands, referral of colorectal cancer patients to university hospitals, or high volume specialized centres is based on the stage of disease, and the existence of complex co-morbidities. However, this referral system does not result in a higher expected mortality in these hospitals. Although patients with complex co-morbidities are often referred to larger or academic centres, the low-volume, non-teaching or teaching hospitals treat more elderly patients with multiple comorbidities. Especially when these patients are treated in a non-elective setting, mortality risk is extremely high and can go up to 41% [**Chapter 5**].
A possible explanation is that patients in need of an urgent resection may not have the time to choose or be referred for treatment in a high-volume, specialized hospital, and therefore are treated in the local, low volume hospital. Although we failed to find a significant difference in case-mix adjusted outcome between different types of hospitals, [**Chapter 4**] previous research has shown that a higher annual volume of colorectal cancer resections results in better postoperative outcome. However, a minimum or maximum number needed to treat has never been identified.[18] Also, previous studies have demonstrated that the presence of a specialised surgeon during a non-elective operation improves outcome.[19,20] Although referral of high-risk patients may be a logistical challenge, it is likely that referral of such patients to specialised centres, where proper facilities for peri-operative care including a specialised surgeon during on-call hours are available, improves outcome. The Association of Surgeons in the Netherlands is currently working on a statement on minimal requirements of hospitals treating patients with colorectal cancer, including the availability of a specialized surgical oncologist or gastro-intestinal surgeon during on-call hours.

By analysing data from clinical audits on a national scale, high-risk patient groups can be identified, and their distribution among Dutch hospitals can be studied. More insight is gained in the current state of treatment and outcome for high-risk patients in the Netherlands, and de magnitude of their operative risk. These data can be used to inform patients and guide clinical decision-making. Also, the relationship between treatment processes and outcomes for these high-risk patients, who are generally excluded from clinical trials, can be analysed. This can lead to the identification of best practices to improve outcomes for these patients.

# Part 3. The use of clinical auditing for the evaluation and monitoring of the implementation of new techniques

Clinical audits also form the perfect platform for quality assurance, and the evaluation and monitoring of the implementation of new advances in (surgical) treatment. Although new techniques are extensively investigated in randomized controlled trials before implementation in clinical practice, results after implementation are not often studied. However, the study population included in clinical trials is known to be a specific, often younger and healthier subgroup, whose outcomes may not reflect the outcomes of the full population. Therefore studying the results of a new technique after implementation in the full population may be seen as the essential 'check' step of the PDCA cycle. Also, differences in use and experience may be evaluated using data from a national clinical audit.

A leading example of quality assurance is the Dutch Total Mesorectal Excision (TME) trial, in which the new TME technique was compared to the standard operative technique at the time. To reduce variation in skill and interpretation of the technique between surgeons, participating surgeons were properly trained using workshops videotapes and supervision during the first five procedures. Also, surgeons received immediate feedback on their performance by the pathologist. The new TME technique resulted in a 50% reduction of local recurrence.[21] However, after the trial was completed, focus on the TME resection diminished. As a result for only 48% of all patients with a rectum carcinoma, a CRM was reported in the DSCA in 2009. 3 years after the introduction of the DSCA, the reporting rate of the CRM had increased to 80%, [**Chapter 3**] illustrating the importance of a continuous feedback system.

Another example of quality assurance using clinical audits is the introduction of laparoscopic surgery in the Netherlands. In 2007 the Dutch society for endoscopic surgery developed a quality assurance program for the introduction of new laparoscopic techniques. This system was based on a plan-do-check-act cycle involving the development of guidelines for the use and maintenance of instruments, a structured training and certification program and a registration and evaluation system. We showed how this quality assurance system has resulted in a safe and successful introduction of laparoscopic resections for colorectal cancer in the Netherlands. The laparoscopic resection rate in the Netherlands is high compared to international standards with an acceptable conversion rate. [**Chapter 6**] We also showed that short-term outcome after laparoscopic resections in the Netherlands was better than after open resections, even after correction for case-mix, while outcome after conversion was not different from outcome after open resection. This is in contrast with previous studies, which did not find better outcomes after laparoscopic surgery.[22] A possible explanation is that previous randomized controlled trials (RCT's) included a selected, low-risk population, while the benefit of laparoscopic surgery may be larger for more high-risk patients,[23] who are included in national clinical audits such as the DSCA. Although we found differences in laparoscopic resection and conversion rates between hospitals, we found no evidence that these differences affect hospital outcomes of care. [**Chapter 6**]

# Part 4. Using clinical auditing to evaluate quality of care and give transparency to all stakeholders

Overwhelming media attention for variation in hospital quality of care, has lead to a strong societal call for transparency in quality of care. Data on hospital performances on treatment and outcomes of care may be used as 'quality indicators', to evaluate quality of care, and give transparency on quality of care to different stakeholders. These data are thought to improve outcomes of care via two pathways:[24]

*Selection*: data may be used by patients to choose a hospital for treatment, or by insurance companies to selectively contract hospitals, ratio being that this 'free market' policy will drive hospitals to improve quality of care and reduce costs. Also, healthcare inspectorate may use these data to identify underperforming hospitals.

*Improvement*: transparency of quality of care data to other stakeholders may further stimulate hospitals to use data from clinical audits improve their practice and outcomes. [**Chapter 2** and **Chapter 3**]

Quality of care is often evaluated using *'quality indicators'*. Although quality indicators are widely used, the ideal quality indicator, which is appreciated by all stakeholders, has not yet been described.

## Quality indicators

Quality indicators are defined as measurable aspects of care that reflect quality of care as a whole. A good quality indicator has several requirements, concerning *importance*, *scientific acceptability* and *usability.*[25]

- *Importance:* Is the indicator relevant to a large population at considerable risk, and is there an opportunity for improvement? For example, it could be argued that the opportunity for quality improvement is greater in a low risk population, as in a high risk population the influence of the patient and disease related risk-factors is much stronger than the possible influence of quality of care provided. Also, the event-rate of the indicator must be adequate to allow hospital comparisons.[5,26] However, the definition of an 'adequate' event-rate remains unclear.
- *Scientific acceptability:* data must be reliable and valid.
  - Reliability means that data are reproducible and that all patients concerned are included. Uniform data collection and clear definitions are of major importance for reliable quality indicators. Also, to reliably compare hospital performances, results must be adjusted for hospital differences in case-mix and random variation.
  - Validity means that the quality indicator measures what intents to measure. This means that there is a clear and consistent relation between a good result on the quality indicator, and high quality of care. Currently, hospital quality of care is most often assessed on their performances on process indicators, based on the assumption that a hospital that treats patients according to the latest guidelines will also have good

outcomes. However, for most process indicators currently used in colorectal cancer surgery, an association with outcomes of care has not been proven or studied.[27] Moreover, as hospital scores on different indicators are known to be inconsistent,[28,29] and the use of a single indicator to evaluate the quality of care as a whole bears the risk of a perverse incentive, indicators may also be combined into composite measures.

- *Usability*: the results must be understandable and usable for the intended audience. The usability of an indicator may vary for different stakeholders. For example, the Healthcare Inspectorate may be most interested in safety and effectiveness, while insurance companies may be more interested in efficiency. For patients, information on safety, effectiveness, timeliness, and patient centeredness may be more important. Therefore, when choosing a quality indicator for transparency, *usability* for the intended audience must be kept in mind.

## Clinical audits and quality indicators

In the Netherlands, societal call for transparency in quality of care has led to an overwhelming list of quality indicators, which are used to rank and rate hospitals in media and play a large role in the purchase policies of insurance companies. However, for most of these quality indicators, importance, reliability, validity and usability have not been studied, or have proven to be questionable.[30,31] This is not surprising, as before the implementation of clinical auditing in the Dutch healthcare system, no registration system for quality of care information was available to study quality indicators. Clinical Audits contain detailed, uniformly registered information on processes and outcomes of care, which enable the research for 'good' quality indicators for the Dutch healthcare system.

## Process and outcome

Most commonly used quality indicators for colorectal carcinoma are process indicators, as they are less influenced by case-mix and random variation than outcome indicators. However, valid process indicators must concern a process relevant to all selected patients, and have a proven association with relevant outcomes. We studied the association between process and outcome indicators in the DSCA and found that, for some of the process indicators, a good hospital score was associated with favourable scores on outcome indicators; however, this was not true for all process indicators. [**Chapter 7**] This varying association may be explained by a lack of consistency between different process indicators: hospitals with a good score on one indicator did not necessarily have a good score on another process indicator. Therefore, combining process indicators into one composite measure may be preferable. We have shown how a good score on composite measures for guideline adherence is associated with good postoperative outcome. [**Chapter 8**] As most processes described in guidelines are thought to have an effect on long term (disease free) survival, but do not influence patients chances for good postoperative outcome, it is not surprising that on a patient level, there was no association between guideline adherence and a patients' chance for good postoperative outcome. [**Chapter 8**] However, when data were

analysed on a hospitals' level, there was a strong association between guideline adherence and good postoperative outcome. These associations were even stronger when only high volume hospitals were included, indicating that the association was only ameliorated by less reliable results from small volume hospitals. [**Chapter 8**] Although in some cases there may be good reasons for multidisciplinary team to defer from the guidelines, this will concern the minority of patients. To realize a full guideline adherence for all other patients is an organizational challenge. Therefore, rather than measuring a causal effect, it is more likely that the association between guideline-adherence and outcome represents the 'hospital effect'; the influence of the efforts of dedicated caregivers and a well-organized logistical process, which results in favourable outcome.

## Composite measures

Although some outcome measures may be of special importance to a specific patient group, a list of individual outcome indicators representing the quality of care in a hospital may be difficult to interpret by patients: how does one choose between a higher risk for complications, and a lower risk for postoperative mortality? However, previous research has shown that patients would be interested in using a composite measure for outcome indicators: the 'textbook outcome' measure. 'Textbook outcome' represents the percentage of patients for whom all desired (postoperative) health outcomes were accomplished, to choose a hospital of treatment.[32,33] This measure is simple and *usable* to all stakeholders. For patients, it represents their chances for the most favourable outcome in a specific hospital. For caregivers it gives feedback information on how often treatment is successful. For insurance companies and hospitals, it summarizes patient safety, effectiveness and efficiency, and may therefore be useful in selective contracting. For the healthcare inspectorate this score may guide surveillance-programs. By combining important outcomes of care in one comprehensive measure, the 'textbook outcome' prevents defensive, indicator driven practice. Good clinical decision-making will therefore result in an optimal score on 'textbook outcome'. Moreover, this measure seems discriminative, as only 49% of all patients had a textbook outcome, but hospital 'textbook outcome' rates ranged between 27 and 71%. Even after case-adjustments, there were hospitals with half as many textbook outcomes than expected, but also hospitals with 1.4 times as many textbook outcomes than expected based on their case-mix. We identified eight outlier hospitals, in which quality of care was worse than average, on several aspects. [**Chapter 9**]

## Case-mix adjustments

As hospital processes and outcomes of care are largely influenced by the characteristics of the patients and diseases treated, appropriate case-mix adjustment is thought to be imperative for comparison of hospital performances. However, as registration of case-mix factors is time consuming, questions are raised if case-mix adjustments are necessary for *all* outcome measures. Dimick showed that for coronary artery bypass surgery, unadjusted and adjusted outcomes are highly related, and equally well predicted hospital performances in

the next years. From this research it was concluded that for coronary artery bypass surgery, case-mix did not differ between hospitals. A possible explanation is that the patient group undergoing coronary artery bypass surgery is relatively homogenous.[34] Similarly, Snijders at al showed that the large hospital variation in anastomotic leakage rates after colorectal cancer resections in the Netherlands could not be explained by differences in case-mix. This suggests that there are other factors that explain the hospital variation in anastomotic leakage rates.[35] We showed, by calculating the expected mortality for all patients treated for colon cancer in Dutch Hospitals, based on only their patient and disease characteristics, hospital average expected mortality ranged from 1.5 to 14%, clearly illustrating the need for case-mix adjustments when comparing hospitals based on mortality after colorectal cancer resections. [**Chapter 4**] These results are supported by the recent findings of Snijders et al, showing that hospital variation in postoperative mortality was significantly reduced after adjustments for case-mix.[35]

Although hospital variations in outcome indicators may be partly explained by differences in case-mix, this may not be equally important for all outcomes. However, case-mix adjustments still remain imperative for the face-validity of hospital comparisons based on outcome indicators.[34,35] Without these adjustments, care providers may lose their trust in the validity of hospital comparisons, reducing the incentive for quality improvement, and possibly even resulting in a perverse incentive to, for example, refer the sickest patients in order to improve outcomes. Therefore, when comparing hospitals based on outcome indicators, results should always be adjusted for case-mix.

### Relevance, the volume needed to assess quality

Another major difficulty in comparing hospital performances is the 'problem of small sample size': when procedural volume is low or adverse events are rare, it is difficult to assess hospital performance, as the confidence intervals for the adverse event rate will be very wide. As a result of this statistical phenomenon, outcomes of low volume hospitals may be five times higher than expected, but there will still be no proof that they are out of the normal range. [**Chapter 10**] We therefore propose a combined measure for volume and outcome to identify hospitals that deliver reliable proof of good quality of care within an acceptable observation period. For these hospitals the combination of their outcomes and the number of patients treated are such that they provide sufficient proof that their outcomes next year, (with consistent performance) will not be worse than a predefined level of substandard care. We found that this combined measure for volume and outcome performed better than measures for volume-only or outcome-only in predicting hospital performance the following year. However, this combined measure for volume and outcome performed better for morbidity than for mortality. [**Chapter 10**] Previous studies have shown that the influence of random variation on hospital performances is larger when outcome measures have a low event-rate, while the proportion of variation between hospital performances explained by systematic variation such as hospital practices is larger when the outcome measure occurs more frequently.[26] This may explain why the combined measure for volume and outcome

performed better for morbidity than for mortality, which has a much lower event-rate. Therefore, when comparing hospitals based on outcome measures, the event-rate of the outcome measure must be taken into account.

## 'Good' quality indicators

When comparing hospital performances, process measures may be used to evaluate implementation of specific processes in hospitals. However, results on a single indicator will not always give a fair reflexion of the quality of the whole process of care. When combined together into a composite measure for guideline adherence, a good hospital performance, reflecting a well-organized care-process and a dedicated team, is highly consistent with good outcome. Therefore a composite measure for guideline adherence is more *valid* and *usable* as an indicator for quality of care. When comparing outcomes of care, one should keep in mind that outcomes of care are often interrelated, and that a good outcome not only involves safety but also effectiveness and efficiency. Furthermore, outcome indicators should be adjusted for case-mix to assure face-validity, and the event-rate of the outcome indicator should be adequate to assure that outcomes are reliable.

Other studies have suggested composite measures combining both process and outcome,[28,36,37] and even adding structure indicators to this measures, such as hospital volume, teaching status, or nurse to patient ratio's.[38,39] These studies found that hospital rankings based on these composite measures differ very much from hospital rankings based on simple outcome measures such as mortality or morbidity alone, showing again that quality of care encompasses more that mortality or morbidity alone. Dimick found that his composite measure combining volume, morbidity and mortality and other structural aspects of care, performed well in predicting future hospital performances.[38] However, indicators included in these scores were selected and weighted completely based on their ability to predict hospital performances, and not based on (clinical) rationale. Although these scores have a good statistical performance, they may be difficult to interpret for different stakeholders, therewith impairing their *usability.* A hospital failing the measure, will not be able to explain on what basis it failed the measure, or improve performances accordingly. For example, although hospital procedural volume is included in the measure, a hospital failing the measure will not be able to extrapolate the number of procedures needed to improve outcomes the following year.

Therefore we propose a more simple measure combining volume and outcome, which selects hospitals with adequate performances and a sufficient volume to assure that outcomes are likely to be adequate the following year. This combined measure may also be used to evaluate hospital scores on 'textbook outcome', to select those hospitals with a high percentage of patients for who all desired (postoperative) health outcomes were accomplished, in a sufficient volume to assure future performances will be adequate.

## The way forward

The successful and fast implementation of the DSCA, and the impressive improvement of quality of care after its implementation have led to the initiation of many new audits. Using the DSCA as a blueprint, new audits can be developed and running within one-year. Also, auditing is not limited to the surgical or oncological field: the audit for melanoma has included systemic treatment after surgery, the Dutch vascular surgeons have initiated two audits evaluating treatment of abdominal aortic aneurysms and carotid surgery, for paediatric surgery a clinical audit is being developed, and there is a clinical audit for the treatment of cerebrovascular accidents and Parkinson's disease. We expect and hope that in a few years time, all major disease entities are evaluated in a clinical audit. A recent report of the Dutch Health Organization also underlines that clinical auditing may be the best way forward to transparency and improvement of quality of care.[40]

However, increasing the number of clinical audits also means increasing the amount of administrative work. Although most Dutch hospitals are nowadays equipped with an electronic patient file (EPF) system, only very few hospitals have managed to integrate data collection within their daily practice, automatically submitting required data when writing up their patients notes. Although these techniques are widely available, and are likely to result in a more reliable and complete registration, implementation is often complex and laborious. However, the success of clinical auditing has pushed this issue to the top of the priority list of many stakeholders. Recently the Dutch federation of university hospitals has published a report advocating implementing a new, uniform way of documentation, which enables data collection straight from the electronic patient file system.[41]

## Best practices and regional 'quality of care' conferences

Not only can feedback data be used by hospitals to improve their practice and outcomes, but data can also be used to discuss regional differences in 'quality of care' conferences. By sharing feedback data with other hospitals, lessons may be learned on how to change and improve performances. Those hospitals with better results may share the practices that lead to their results: best practices, which may be implemented in other hospitals in the region. Also these data may be used for organizational arrangements or referral of specific patient groups. For example the Dutch province Friesland has hired a consultant to develop an optimal referral system in the region based on data from clinical audits that improves outcomes and maybe even reduces costs.

## Transparency of care

Until recently, data from the DSCA on individual hospital performances have not been publicly reported. Insufficient trust in validity of the data and fear of registration fraud, perverse incentives or naming and blaming or even repercussions, has prevented early disclosure of these data. However, recently the Boston Consulting Group published a report on quality

improvement using clinical auditing. Using the Swedish audits as an example they showed that disclosure of valid hospital-specific performance data does not result in a naming and shaming scandal, but in fast, nationwide improvement of quality of care, and reduction of healthcare costs.[42] Therefore, in 2012, the ASN decided on a staged transparency program disclosing more data every year: the first year, only hospital case-volume is disclosed, the following year process indicators are published, and the third year, outcome indicators are disclosed. Such a staged transparency program anticipates time for data quality control, but also fulfils societal expectations and needs for insight in hospital quality of care. Twice a year medical specialists from the various audits and representatives of their professional organizations (e.g. the ASN), together with insurers and representatives of the patient organizations, define that years' quality indicators. Together they agree on the quality indicators that will be used that year to guide selective contracting and to inform patients.

The most important stakeholder in this case may be the patient. However, previous studies have shown that patients do not often use hospital quality of care information for their choice for a hospital of treatment, because they do not understand or trust it.[43,44] Using the data from clinical audits, important, reliable, valid and usable information on hospital performances can be produced for patients to base their choice for a hospital of treatment on. Previous research has shown that patients would like to use a measure like the 'textbook outcome' to base their choices on.[32] Further research should further explore how patients' needs for hospital quality of care information might best be fulfilled; e.g. the type of information patients are most interested in, and the best way of presenting this information so that is usable for patients. Furthermore, further research should evaluate if public availability of quality of care information, presented in the way patients find it most usable, also results in an increased understanding, trust and use of this information.

However, data from clinical audits only reflect one aspect of quality of care. Recently, the Dutch Federation of Cancer Patient Organizations (NFK) has defined four aspects of quality of care, which they consider relevant to patients. These are:
Structure information, describing what is available in a hospital, and whether a hospital meets the standards set by the professional associations concerned
Process and outcome information from clinical audits, describing if the hospital meets the standards for process and outcome indicators set by the professional associations concerned.
Patient Reported Outcome Measures (PROM's), describing results of treatment and postoperative quality of life reported by patients.
Consumer Quality index, describing the subjective experience of patients treated in that hospital.
The Dutch Institute for Clinical Auditing, together with the NFK and the Dutch Healthcare Insurance companies have recently started a transparency project integrating these four aspects of quality of care, using colorectal cancer as a showcase. Different aspects of quality of care are reported on a transparent and publicly available website. This website, of which

the design and contents are dictated by patients, may help regain societal trust in quality of healthcare in the Netherlands.

## References


1. Khuri SF, Daley J, Henderson WG. The comparative assessment and improvement of quality of surgical care in the Department of Veterans Affairs. Arch Surg. Jan 2002;137(1):20-27.
2. O'Connor GT, Plume SK, Olmstead EM, et al. A regional intervention to improve the hospital mortality associated with coronary artery bypass graft surgery. The Northern New England Cardiovascular Disease Study Group. JAMA. Mar 20 1996;275(11):841-846.
3. Wibe A, Moller B, Norstein J, et al. A national strategic change in treatment policy for rectalcancer--implementation of total mesorectal excision as routine treatment in Norway. A national audit. Dis Colon Rectum. Jul 2002;45(7):857-866.
4. Ivers N, Jamtvedt G, Flottorp S, et al. Audit and feedback: effects on professional practice and healthcare outcomes. Cochrane Database Syst Rev. 2012;6:CD000259.
5. Dimick JB, Welch HG, Birkmeyer JD. Surgical mortality as an indicator of hospital quality: the problem with small sample size. JAMA. Aug 18 2004;292(7):847-851.
6. van Gijn W, Wouters MW, Peeters KC, van de Velde CJ. Nationwide outcome registrations to improve quality of care in rectal surgery. An initiative of the European Society of Surgical Oncology. J Surg Oncol. Jun 15 2009;99(8):491-496.
7. Hsueh Y. The Hawthorne experiments and the introduction of Jean Piaget in American industrial psychology, 1929-1932. History of psychology. May 2002;5(2):163-189.
8. Gale EA. The Hawthorne studies-a fable for our times? QJM : monthly journal of the Association of Physicians. Jul 2004;97(7):439-449.
9. Lehmann C, Nowak A. The Hawthorne effect: can it be measured and utilized? British journal of anaesthesia. Apr 2013;110(4):658-659.
10. Usichenko TJ, A; Wendt, M. 'Yes, we can' utilize the Hawthorne effect to improve postoperative analgesia. British journal of anaesthesia. 2013;110(4):659.
11. Pouw ME, Peelen LM, Lingsma HF, et al. Hospital standardized mortality ratio: consequences of adjusting hospital mortality with indirect standardization. PLoS One. 2013;8(4):e59160.
12. Almoudaris AM, Burns EM, Bottle A, Aylin P, Darzi A, Faiz O. A colorectal perspective on voluntary submission of outcome data to clinical registries. Br J Surg. Jan 2011;98(1):132-139.
13. Engel A. This paper is about a rather small colorectal unit trying to come to grips with the frustration produced by the discrepancy between their perceived performance and NBOCAP performance. Colorectal Dis. Aug 2012;14(8):1019.
14. Lewis RS, Graham DG, Watson JD, Lunniss PJ. Learning the hard way: the importance of accurate data. Colorectal Dis. Aug 2012;14(8):1015-1018.
15. Jamtvedt G, Young JM, Kristoffersen DT, O'Brien MA, Oxman AD. Does telling people what they have been doing change what they do? A systematic review of the effects of audit and feedback. Qual Saf Health Care. Dec 2006;15(6):433-436.
16. Dixon N. Proposed standards for the design and conduct of a national clinical audit or quality improvement study. Int J Qual Health Care. May 21 2013.
17. Wouters MWJM. DICA rapportages 2012: Dutch Institute for Clinical Auditing;2013.
18. van Gijn W, Gooiker GA, Wouters MW, Post PN, Tollenaar RA, van de Velde CJ. Volume and outcome in colorectal cancer surgery. Eur J Surg Oncol. Jul 6.
19. Elshove-Bolk J, Ellensen VS, Baatrup G. Logistics and outcome in urgent and emergency colorectal surgery. Colorectal Dis. Nov 10 2009.
20. Biondo S, Kreisler E, Millan M, et al. Impact of surgical specialization on emergency colorectal surgery outcomes. Arch Surg. Jan;145(1):79-86.
21. Peeters KC, van de Velde CJ. Surgical quality assurance in rectal cancer treatment: the key to improved outcome. Eur J Surg Oncol. Aug 2005;31(6):630-635.
22. Schwenk W, Haase O, Neudecker J, Muller JM. Short term benefits for laparoscopic colorectal resection. Cochrane Database Syst Rev. 2005(3):CD003145.
23. McCloskey CA, Wilson MA, Hughes SJ, Eid GM. Laparoscopic colorectal surgery is safe in the high-risk patient: a NSQIP risk-adjusted analysis. Surgery. Oct 2007;142(4):594-597; discussion 597 e591-592.
24 Berwick DM, James B, Coye MJ. Connections between quality measurement and improvement. Med Care. 2003 Jan;41(1 Suppl):I30-8.
25. Dimick JB. What makes a "good" quality indicator? Arch Surg. Mar 2010;145(3):295.
26. van Dishoeck AM, Lingsma HF, Mackenbach JP, Steyerberg EW. Random variation and rankability of hospitals using outcome indicators. BMJ quality & safety. Oct 2011;20(10):869-874.
27. Patwardhan M, Fisher DA, Mantyh CR, et al. Assessing the quality of colorectal cancer care: do we have appropriate quality measures? (A systematic review of literature). J Eval Clin Pract. Dec 2007;13(6):831-845.
28. Habib MR, Solomon MJ, Young JM, Armstrong BK, O'Connell D, Armstrong K. Evidence-based and clinical outcome scores to facilitate audit and feedback for colorectal cancer care. Dis Colon Rectum. Apr 2009;52(4):616-622; discussion 622-613.
29. Lilford R, Mohammed MA, Spiegelhalter D, Thomson R. Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. Lancet. Apr 3 2004;363(9415):1147-1154.

30 Gooiker GA, Verbeek L, van der Geest LG, et al. The quality indicator 'tumour positive surgical margin following breast-conserving surgery' does not provide transparent insight into care. Ned Tijdschr Geneeskd. 2010;154:A1142.

31 Beperkt zicht; Onderzoek naar de betrouwbaarheid, validiteit en bruikbaarheid van prestatie-indicatoren over de kwaliteit van de Nederlandse ziekenhuiszorg. Rapport NFU-consortium Kwaliteit van Zorg. 2012.

32. Marang-van de Mheen PJ, Dijs-Elsinga J, Otten W, et al. The relative importance of quality of care information when choosing a hospital for surgical treatment: a hospital choice experiment. Medical decision making : an international journal of the Society for Medical Decision Making. Nov-Dec 2011;31(6):816-827.

33. Dijs-Elsinga J, Otten W, Versluijs MM, et al. Choosing a hospital for surgery: the importance of information on quality of care. Medical decision making : an international journal of the Society for Medical Decision Making. Sep-Oct 2010;30(5):544-555.

34. Dimick JB, Birkmeyer JD. Ranking hospitals on surgical quality: does risk-adjustment always matter? J Am Coll Surg. Sep 2008;207(3):347-351.

35. Snijders HS, Henneman D, van Leersum NL, et al. Anastomotic leakage as an outcome measure for quality of colorectal cancer surgery. BMJ quality & safety. Sep 2013;22(9):759-767.

36. Brunelli A, Berrisford RG, Rocco G, Varela G. The European Thoracic Database project: composite performance score to measure quality of care after major lung resection. Eur J Cardiothorac Surg. May 2009;35(5):769-774.

37. Brunelli A, Rocco G, Van Raemdonck D, Varela G, Dahan M. Lessons learned from the European thoracic surgery database: the Composite Performance Score. Eur J Surg Oncol. Sep 2010;36 Suppl 1:S93-99.

38. Dimick JB, Staiger DO, Osborne NH, Nicholas LH, Birkmeyer JD. Composite measures for rating hospital quality with major surgery. Health services research. Oct 2012;47(5):1861-1879.

39. Dimick JB, Staiger DO, Hall BL, Ko CY, Birkmeyer JD. Composite measures for profiling hospitals on surgical morbidity. Ann Surg. Jan 2013;257(1):67-72.

40. Gezondheidsraad. Rapport: Publieke indicatoren voor kwaliteit van curatieve zorg. De stand van de discussie. 2013

41. NFU. Rapport: Registratie aan de bron. 2013.

42. Kuenen JM, R; Larsson, S; Leeuwen, W van. Zorg voor waarde. Boston Consultancy Group. 2011.

43. Marshall MN, Shekelle PG, Leatherman S, Brook RH. The public release of performance data: what do we expect to gain? A review of the evidence. JAMA. Apr 12 2000;283(14):1866-1874.

44. Faber M, Bosch M, Wollersheim H, Leatherman S, Grol R. Public reporting in health care: how do consumers use quality-of-care information? A systematic review. Med Care. Jan 2009;47(1):1-8.

# Publication list

## This Thesis


1. N.E. Kolfschoten en N.J. Van Leersum, Klinkenbijl JH, Tollenaar RA, Wouters MW.. Clinical auditing', a novel tool for quality assessment in surgical oncology Ned Tijdschr Geneeskd. 2011;155(45):A4136.

2. Van Leersum NJ, Snijders HS, Henneman D, Kolfschoten NE, Gooiker GA, ten Berge MG, Eddes EH, Wouters MW, Tollenaar RA; Dutch Surgical Colorectal Cancer Audit Group, Bemelman WA, van Dam RM, Elferink MA, Karsten TM, van Krieken JH, Lemmens VE, Rutten HJ, Manusama ER, van de Velde CJ, Meijerink WJ, Wiggers T, van der Harst E, Dekker JW, Boerma D. The Dutch surgical colorectal audit. Eur J Surg Oncol. 2013 Oct;39(10):1063-70.

3. N. E. Kolfschoten, P.J. Marang van de Mheen, G.A. Gooiker, E.H. Eddes, J. Kievit, R.A.E.M Tollenaar, M.W.J.M. Wouters, on behalf of the Dutch Surgical Colorectal Audit group Variation in case-mix between hospitals treating colorectal cancer patients in the Netherlands, Eur J Surg Oncol. 2011 Nov;37

4. N.E. Kolfschoten, M.W.J.M. Wouters, G.A. Gooiker, E.H. Eddes, J. Kievit, R.A.E.M Tollenaar, P.J. Marang-van de Mheen, on behalf of the Dutch Surgical Colorectal Audit group, Nonelective colon cancer resections in elderly patients: results from the Dutch Surgical Colorectal Audit. Dig Surg. 2012;29(5):412-9.

5. N.E. Kolfschoten, M. Wouters, P.J. Marang van de Mheen, G.A.Gooiker, N.J. van Leersum, E.H. Eddes, J. Kievit, R. Brand, P. Tanis, W.A. Bemelman. R.A.E.M Tollenaar, J. Meijerink, on behalf of the Dutch Surgical Colorectal Audit group, Successful and safe introduction of laparoscopic colorectal cancer surgery in Dutch hospitals. Ann Surg. 2013 May;257(5):916-21

6. N.E. Kolfschoten, G.A. Gooiker, E. Bastiaannet, N.J. van Leersum, C.J.H. van de Velde, E.H. Eddes, P.J. Marang-van de Mheen, J. Kievit, E. van der Harst, T. Wiggers, M.W.J.M. Wouters, R.A.E.M Tollenaar, on behalf of the Dutch Surgical Colorectal Audit group Combining process indicators to evaluate quality of care for surgical colorectal cancer patients; are scores consistent with short-term outcome? BMJ Qual Saf. 2012 Jun;21(6):481-9.

7. Gooiker GA, Kolfschoten NE, Bastiaannet E, van de Velde CJ, Eddes EH, van der Harst E, Wiggers T, Rosendaal FR, Tollenaar RA, Wouters MW; Dutch Surgical Colorectal Audit group. Evaluating the validity of quality indicators for colorectal cancer care. J Surg Oncol. 2013 Dec;108(7):465-71.

8. N.E. Kolfschoten, J. Kievit, G.A. Gooiker, N.J. van Leersum, H.S. Snijders, E.H. Eddes, R.A.E.M. Tollenaar M.W. Wouters, P.J. Marang van de Mheen. Focusing on desired outcomes of care after colon cancer resections, hospital variations in 'textbook outcome'. Eur J Surg Oncol. 2013 Feb;39(2):156-63.

9. N.E. Kolfschoten, P.J. Marang-van de Mheen, M.W.J.M Wouters, E.H.Eddes, R.A.E.M. Tollenaar, T. Stijnen, J. Kievit, on behalf of the Dutch Surgical Colorectal Audit group. A combined volume and outcome threshold to reliably assess hospital performance. Plos One. 2014

## Other puplications

### Quality of Care

10. Henneman D, Snijders HS, Fiocco M, van Leersum NJ, Kolfschoten NE, Wiggers T, Wouters MW, Tollenaar RA. Hospital variation in failure to rescue after colorectal cancer surgery: results of the Dutch Surgical Colorectal Audit. Ann Surg Oncol. 2013 Jul;20(7):2117-23.

11. van Leersum N, Martijnse I, den Dulk M, Kolfschoten NE, Le Cessie S, van de Velde C, Tollenaar R, Wouters M, Rutten HJ. Differences in Circumferential Resection Margin Involvement After Abdominoperineal Excision and Low Anterior Resection No Longer Significant. Ann Surg. 2013 Oct 3. [Epub ahead of print]

12. Snijders HS, Wouters MW, van Leersum NJ, Kolfschoten NE, Henneman D, de Vries AC, Tollenaar RA, Bonsing BA. Meta-analysis of the risk for anastomotic leakage, the postoperative mortality caused by leakage in relation to the overall postoperative mortality. Eur J Surg Oncol. 2012 Nov;38(11):1013-9.

13. Dekker JW, Gooiker GA, van der Geest LG, Kolfschoten NE, Struikmans H, Putter H, Wouters MW, Tollenaar RA. Use of different comorbidity scores for risk-adjustment in the evaluation of quality of colorectal cancer surgery: does it matter? Eur J Surg Oncol. 2012 Nov;38(11):1071-8.

## Outdoor medicine

14. M.P. Grocott, D.S. Martin, D.Z. Levett, R. McMorrow, J, Windsor,. N.E. Kolfschoten, H.E. Montgomery; Caudwell Xtreme Everest Research Group. Arterial blood gases and oxygen content in climbers on Mount Everest. New England Journal of Medicine. 2009 Jan 8; 360(2):140-9.

15. De behandeling van Koudeletsel Berendsen RR, Kolfschoten NE, de Jong VM, Frima H, Daanen HA, Anema HA. Ned Tijdschr Geneeskd. 2012;156(25)

16. Wilson MH, Davagnanam I, Holland G, Dattani RS, Tamm A, Hirani SP, Kolfschoten N, Strycharczuk L, Green C, Thornton JS, Wright A, Edsell M, Kitchen ND, Sharp DJ, Ham TE, Murray A, Holloway CJ, Clarke K, Grocott MP, Montgomery H, Imray C; Birmingham Medical Research Expeditionary Society and Caudwell Xtreme Everest Research Group. Cerebral venous system and anatomical predisposition to high-altitude headache. Ann Neurol. 2013 Mar;73(3):381-9

## Other projects

17. J.A. Jansen, J.M. Mens, F.J. Backs, N.E. Kolfschoten, H.J. Stam. Treatment of longstanding groin pain in athletes: a systematic review. Scand J Med Sci Sports. 2008 Jun; 18(3):263-74

18. N.E. Kolfschoten, R. Metz, E.J. Verleisdonk. De acute achillespees ruptuur. Ned Tijdschr Geneeskd. 2009;153:A235

19. N.E. Kolfschoten, A. Pronk. Presacrale Epidermoid cyste. Ned Tijds voor Heelkunde 2009;18:279281.

20. R. Metz , G.J. van der Heijden, E.J. Verleisdonk, N.E. Kolfschoten, M.H. Verhofstad, C. van der Werken; Effect of complications after minimally invasive surgical repair of acute achilles tendon ruptures: report on 211 cases. Am J Sports Med. 2011 Apr;39(4):820-4.

# Nederlandse samenvatting en discussie

Recente rapporten over verschillen in behandeling en uitkomsten van zorg tussen ziekenhuizen, in combinatie met de steeds verder stijgende zorgkosten, hebben er toe geleid dat zorgprofessionals meer en meer verantwoordelijk worden gesteld voor de kwaliteit van de door hun geleverde zorg en voor de kosten daarvan. Om grip te houden op de kwaliteit en de kosten van zorg in ziekenhuizen is valide informatie over ziekenhuisprestaties nodig.

Om aan deze toegenomen behoefte aan valide informatie over ziekenhuisprestaties te voldoen, is een nieuwe methode geïntroduceerd voor het systematisch registreren van, en feedback geven over behandeling en uitkomsten van zorg: clinical auditing. In 2009 introduceerden de Nederlandse chirurgen clinical auditing in Nederland met de Dutch Surgical Colorectal Audit (DSCA). Hun doel was de kwaliteit van zorg te verbeteren door:

- ziekenhuizen een instrument te bieden voor het *verbeteren* van hun eigen *prestaties,* waarmee zij inzicht krijgen in de verschillen in behandeling, richtlijnnaleving en uitkomsten van zorg tussen ziekenhuizen en deze verschillen kunnen verkleinen;
- inzicht te krijgen in *nationale* trends in behandelstrategieën en *prestaties*, op basis hiervan 'benchmarks' te definiëren, verbeterpunten te identificeren en 'best practices' te implementeren;
- de *implementatie* van *nieuwe technieken* te kunnen monitoren en evalueren;
- te kunnen voorzien in de maatschappelijke behoefte aan *transparantie* over kwaliteit van zorg.

Echter, een uniforme methodologie om data van clinical audits te gebruiken voor het realiseren van deze doelen was niet beschikbaar. Een van de meest urgente wetenschappelijke vragen op dit gebied is wat de definitie en inhoud van 'valide en betekenisvolle feedback informatie over ziekenhuisprestaties' moet zijn. Het doel van dit proefschrift was dan ook te onderzoeken hoe, op basis van data van clinical audits, valide en betekenisvolle feedback informatie kan worden gegenereerd, die gebruikt kan worden voor het verbeteren van kwaliteit van zorg.

In **Deel 1** wordt besproken hoe *clinical auditing kan worden gebruikt als een instrument voor kwaliteitsverbetering in de chirurgie*. **Deel 2** laat zien hoe data van *clinical audits gebruikt kunnen worden om de zorg op nationaal niveau te evalueren en verbeteren*, in het bijzonder voor hoog-risico patiënten. In **Deel 3** wordt onderzocht hoe *clinical auditing kan worden gebruikt voor de monitoring van de implementatie van nieuwe technieken*. **Deel 4** bespreekt hoe *clinical auditing kan worden ingezet om de kwaliteit van zorg te evalueren en transparant te zijn naar alle belanghebbenden.*
In dit proefschrift worden de data van de DSCA als voorbeeld gebruikt, maar de methodologie kan als een blauwdruk dienen voor andere clinical audits.

# Deel 1. Het gebruik van clinical auditing als een instrument voor kwaliteitsverbetering in de chirurgie

Vele studies hebben reeds een positief effect van clinical auditing op de kwaliteit van zorg gerapporteerd.[1-3] Wij hebben met een systematische review bevestigd dat clinical auditing een positief effect heeft op de uitkomsten van chirurgische zorg. [**Hoofdstuk 2**] Een recente Cochrane review beschreef een vergelijkbaar resultaat voor kwaliteit van niet-chirurgische zorg.[4]
Echter, het blijft lastig te bewijzen dat de gevonden kwaliteitsverbeteringen volledig kunnen worden toegerekend aan clinical auditing omdat ook andere mechanismen bij kunnen dragen aan de kwaliteitsverbetering. Hieronder geven we daar drie voorbeelden van.

- Het eerste mechanisme is 'regression to the mean': een ziekenhuis met een hoge mortaliteit zal, als gevolg van toevalsvariatie, naar alle waarschijnlijkheid het volgende jaar een lagere mortaliteit hebben, onafhankelijk van een eventueel verbeter initiatief. Deze toevalsvariatie is groter wanneer het ziekenhuisvolume kleiner is, of de uitkomstmaat weinig voorkomt.[5] Echter de continue, nationale verbetering zoals beschreven na implementatie van nationale clinical audits, kan niet verklaard worden door dit mechanisme.[1,6]

- Het tweede mechanisme is de autonome kwaliteitsverbetering die vaak wordt gezien tijdens deelname aan een studie, door de toegenomen aandacht voor uitkomsten: het Hawthorne effect. Dit effect kan echter ook gezien worden als het werkingsmechanisme van clinical auditing: door constant de focus op uitkomsten van zorg te stimuleren wordt een soort continu Hawthorne effect gecreëerd. Hoewel dit effect in de bedrijfspsychologie veelvuldig gebruikt wordt als strategie om prestaties te verbeteren,[7,8] is het idee om het Hawthorne effect bewust in te zetten voor het stimuleren van kwaliteitsverbetering in de zorg relatief nieuw. Door tijdig feedback te geven over de mate van verbetering die nodig is en over het effect van verbeter initiatieven, voegt clinical auditing een controle systeem toe aan het Hawthorne effect. Waar het Hawthorne effect, in zijn traditionele vorm (zonder feedback), vaak afneemt met de tijd, kan met clinical auditing steeds de focus op een nieuwe uitkomstmaat worden gelegd. Daarmee wordt steeds een hernieuwde motivatie en verbeter impuls gegeven.[9,10]

- Het derde mechanisme is de perverse prikkel. Alhoewel clinical auditing is ontworpen om het gehele proces van behandeling en uitkomst te monitoren zijn er altijd mazen in het net: processen die niet worden geregistreerd, maar die de uitkomsten op papier verbeteren, terwijl ze geen positief effect hebben op de kwaliteit van zorg. Zo kan bijvoorbeeld de ziekenhuissterfte verlaagd worden door patiënten snel naar een hospice te ontslaan. Enkele ziekenhuizen zouden deze mazen in het net kunnen gebruiken om hun

uitkomsten op papier te verbeteren.[11] Een gedegen validatie met andere databronnen zou dit effect kunnen verminderen.

## The 'missing link'

Waar het in de meeste industrieën zeer gewoon is om het productieproces frequent te auditen om hiermee de efficiëntie te vergroten, is dit idee relatief nieuw in de geneeskunde. De meeste industrieën werken met een Plan-Do-Check-Act (PDCA) cirkel, waarmee een constante verbetercyclus wordt uitgevoerd. Tot nog toe werd binnen de geneeskunde echter alleen het PD deel van de cirkel uitgevoerd: een innovatie werd geïmplementeerd, maar het effect van de implementatie werd niet systematisch gecheckt. Ook een regelmatige audit om zwakke plekken in het proces te identificeren was ongewoon. Clinical auditing is daarom de 'missing link' in de geneeskunde, die zorgprofessionals in de gelegenheid stelt om hun klinische processen en uitkomsten te monitoren en verbeteren. Ook kunnen de data van clinical auditing gebruikt worden ten behoeve van transparantie over de kwaliteit van de geleverde op ziekenhuis, maar ook op nationaal niveau. Een clinical audit is echter alleen succesvol als de data valide zijn en tijdig beschikbaar.

## Een succesvolle audit

De eerste vereiste voor een succesvolle (inter)nationale clinical audit is een up-to-date, complete en valide database, op vijf niveaus.

*Deelname*: alle ziekenhuizen in het beoogde gebied nemen deel aan de audit. In België wordt de behandeling van patiënten met een rectumcarcinoom geregistreerd in de PROCARE database. Echter, slechts 40% van de Belgische ziekenhuizen nemen deel aan deze audit, waardoor de gegevens van deze audit niet representatief zijn voor het gehele land.

- *Volledigheid:* alle patiënten in al deze ziekenhuizen worden geregistreerd. In het Verenigd Koninkrijk is deelname aan de National clinical Bowel Cancer Audit (NBOCAP) vrijwillig. Alhoewel de deelname de afgelopen jaren fors is toegenomen van 44,3% in 2006 tot 90,1% in 2011, is de volledigheid laag: 23,3% van alle ziekenhuizen registreert minder dan 50% van alle patienten.[12] Recente studies hebben uitgewezen dat ziekenhuizen die alle data voor alle patiënten invoerden een significant lagere 30-dagen sterfte hadden dan ziekenhuizen die minder dan 10% van hun patiënten invoerden.[12] Waarschijnlijk wordt dit verschil verklaard doordat de ziekenhuizen die goed registreren ook de meer toegewijde ziekenhuizen zijn, met betere resultaten.[13,14]
- *Compleetheid*: Alle relevante informatie wordt voor iedere patiënt geregistreerd. Wanneer bijvoorbeeld de case-mix informatie ontbreekt, zal de case-mix correctie onvoldoende zijn, en komen de gecorrigeerde uitkomsten van een ziekenhuis mogelijk slechter uit dan ze in werkelijkheid zijn.
- *Tijdigheid*: idealiter wordt de registratie van gegevens voor clinical audits gesynchroniseerd met de klinische verslaglegging, of in ieder geval kort na het beschikbaar komen van de

gegevens uitgevoerd. Dit is van belang omdat onderzoek heeft aangetoond dat tijdige en frequente feedback een belangrijke drijfveer voor kwaliteitsverbetering is.[15]
- *Validiteit*: registratie is juist en kan gevalideerd worden met andere bronnen.

De DSCA werd ontwikkeld naar voorbeeld van vele internationale clinical audits. Binnen 2 jaar namen alle Nederlandse ziekenhuizen deel aan de audit. De volledigheid was 92% in 2010 en 95% in 2011, en de audit was 100% compleet voor vrijwel alle items. Feedback gegevens werden iedere week op de website van het ziekenhuis teruggekoppeld en gegevens werden gevalideerd met de database van de Nederlandse Kanker Registratie. [**Hoofdstuk 3**] Wij onderzochten wat de 'drijvende' factoren waren voor het succes van deze registratie en beschreven de kwaliteitsverbeteringen geobjectiveerd na introductie van deze audit. De 'drijvende' factoren waren:
- een leidende rol van de beroepsvereniging in de ontwikkeling van de dataset en de uitkomstmaten;
- integratie van de audit in het nationale kwaliteitsbeleid van de beroepsvereniging;
- een web-based registratie systeem en registratie door de medisch specialist zelf;
- wekelijks online valide en betekenisvolle feedback voor alle deelnemers;
- jaarlijkse externe data verificatie met andere bronnen;
- kwaliteit standaarden gesteld door de beroepsvereniging en introductie van verbeterprojecten om deze standaarden te halen, waarbij de eerste standaard het behalen van een volledige deelname en volledigheid in de audit was. [**Hoofdstuk 3**]

Deze factoren zijn ook recent in een systematische review naar voren gekomen als belangrijke factoren voor het succes van een clinical audit.[16] De DSCA was ontwikkeld als een blauwdruk voor clinical audits in Nederland en vele nieuwe audits zijn reeds gestart naar dit voorbeeld.

## Resultaten van de DSCA

Ziekenhuizen die deelnemen aan de DSCA ontvangen wekelijks feedback informatie over hun prestaties in vergelijking met de andere Nederlandse ziekenhuizen. Deze feedback data bevatten gegevens over de karakteristieken van de behandelde patiënten en tumoren, de behandeling en uitkomsten. Zo wordt feedback informatie gegeven over de mate van richtlijn navolging. Van 2009 tot 2011 werd een significante toename van richtlijn navolging gezien: zo werd een toename gezien van het percentage patiënten dat preoperatief werd besproken in een multidisciplinair overleg, een toename van het gebruik van Magnetic Resonance Imaging (MRI) in de work-up van patiënten met een rectumcarcinoom en een toename in het beschrijven van de circumferentiële resectie marge (CRM). Bovendien werd een duidelijke afname van ziekenhuisvariatie gezien. [**Hoofdstuk 3**] Na 2011 is de richtlijnnaleving nog verder verbeterd.[17]

Ook wordt er feedback informatie gegeven over uitkomsten van de chirurgische behandeling van colorectaal carcinoom, waaronder postoperatieve complicaties, opnameduur en sterfte.

Het percentage postoperatieve complicaties en sterfte na colon en rectumkanker chirurgie in Nederland is significant afgenomen van 2009 tot 2011. Ook de opnameduur nam af met 2 dagen. [**Hoofdstuk 3**] Na 2011 zijn de uitkomsten nog verder verbeterd.[17]

Zoals hierboven reeds beschreven zijn er ook andere mechanismen dan clinical auditing die bijgedragen kunnen hebben aan deze kwaliteitsverbeteringen. Echter de continue verbetering op verschillende onderdelen van richtlijn naleving, de uitkomsten en de afname van de ziekenhuisvariatie laten zien dat clinical auditing een goed instrument is om de kwaliteit van zorg in de oncologische chirurgie te verbeteren.

## Deel 2. Het gebruik van data van clinical audits om de zorg op nationaal niveau te evalueren en verbeteren

Data van clinical audits kunnen ook op nationaal niveau worden geanalyseerd om de behandeling en uitkomsten van specifieke groepen patiënten te bestuderen. Zo kunnen bijvoorbeeld hoog-risico patiënten geïdentificeerd worden om te zien hoe deze patiënten behandeld en eventueel verwezen worden. Er is veel onderzoek gedaan naar risicofactoren in de colorectale kankerchirurgie en verschillen in case-mix tussen ziekenhuizen. Echter, de spreiding van hoog-risico patiënten over de Nederlandse ziekenhuizen en het effect van bekende risicofactoren op uitkomsten van zorg in Nederland zijn niet eerder beschreven. Omdat hoog-risico patiënten vaak niet worden geïncludeerd in klinische studies is er weinig bekend over de beste behandeling en de uitkomsten van deze patiënten. Data van clinical audits zouden dit gat in informatie kunnen opvullen.

Wij hebben laten zien dat de 'verwachte sterfte': het voor een patiënt voorspelde sterfterisico tijdens en na darmkankerchirurgie in Nederland, gebaseerd op case-mix factoren, inzicht kan geven in hoe hoog-risico patiënten zijn verdeeld over de Nederlandse ziekenhuizen. [**Hoofdstuk 4**] We vonden dat er grote verschillen in 'verwachte sterfte' waren tussen de individuele ziekenhuizen, maar ook tussen verschillende typen ziekenhuizen. Zo bleek dat hoog-risico colon kanker patiënten vaker in laag-volume ziekenhuizen of streekziekenhuizen of opleidingsziekenhuizen werden behandeld, in plaats van in academische ziekenhuizen. De hogere verwachte sterfte in deze ziekenhuizen werd voornamelijk verklaard door meer patiënten met een hogere ASA classificatie, meer patiënten met comorbiditeit en meer urgente resecties. We vonden geen verschil in case-mix gecorrigeerde uitkomsten tussen verschillende typen ziekenhuizen. In Nederland worden darmkanker patiënten doorverwezen op basis van het tumorstadium of de aanwezigheid van ingewikkelde comorbiditeit. Echter dit verwijssysteem resulteert dus niet in een hogere 'verwachte sterfte' in de verwijscentra. [**Hoofdstuk 4**]

Dit komt omdat de oudere patiënten, met veel 'niet-ingewikkelde' comorbiditeit, vaak in de laag-volume streek- of opleidingsziekenhuizen worden behandeld. Zeker wanneer deze

patiënten in de urgent worden geopereerd, zijn dit hoog-risico procedures waarbij het sterfterisico tot 41% kan oplopen. [**Hoofdstuk 5**] Mogelijk hebben deze urgente patiënten niet de tijd om naar een hoog-volume, gespecialiseerd ziekenhuis te gaan of verwezen te worden, en worden zij daarom vaker in het lokale ziekenhuis behandeld. Alhoewel wij geen verschil vonden in gecorrigeerde sterfte tussen de verschillende typen ziekenhuizen, [**Hoofdstuk 4**] heeft eerder onderzoek wel uitgewezen dat een hoger ziekenhuis volume voor colorectale kankerchirurgie resulteert in betere uitkomsten. Echter het minimum of maximum volume voor betere uitkomsten is nog nooit wetenschappelijk vastgesteld.[18] Daarnaast is eerder gebleken dat de aanwezigheid van een gespecialiseerde chirurg tijdens een urgente operatie resulteert in betere uitkomsten.[19,20] Alhoewel het verwijzen van hoog-risico patiënten, vooral wanneer zij urgent geopereerd moeten worden, een logistieke uitdaging is, zal het verwijzen van dergelijke patiënten naar gespecialiseerde hoog-volume centra, waar de perioperatieve zorg goed georganiseerd is en altijd een gespecialiseerde colorectale chirurg beschikbaar is, de uitkomsten zeer waarschijnlijk verbeteren. De Nederlandse Vereniging voor Heelkunde (NVVN) is momenteel bezig met een statement over de minimale voorwaarden voor ziekenhuizen die patiënten met darmkanker behandelen, waarin ook de beschikbaarheid van een colorectaal gespecialiseerd chirurg tijdens diensturen wordt beschreven.

Door gegevens van clinical audits op nationaal niveau te analyseren krijgen we meer inzicht in de behandeling en uitkomsten van hoog risico patiënten in Nederland. Deze gegevens kunnen gebruikt worden om het klinisch besluitvormingsproces te ondersteunen. Ook kunnen de gegevens van clinical audits helpen om 'best practices' te identificeren die leiden tot betere uitkomsten voor deze patiëntengroep die weinig bestudeerd wordt in klinische studies.

## Deel 3. Het gebruik van clinical auditing voor de monitoring en evaluatie van de implementatie van nieuwe technieken

Clinical audits kunnen ook gebruikt worden als platform voor kwaliteitsbewaking en evaluatie tijdens de implementatie van nieuwe (chirurgische) technieken. Alhoewel nieuwe technieken vaak uitgebreid worden onderzocht in gerandomiseerde studies voordat ze in het hele land worden ingevoerd, worden de resultaten na implementatie vaak niet geëvalueerd. Echter, het is bekend dat de patiëntenpopulatie die in klinische studies wordt geïncludeerd, over het algemeen een jongere, gezondere subgroep is, waarvan de uitkomsten een vertekend beeld zouden kunnen geven voor de gehele populatie. Daarom wordt het bestuderen van resultaten van een nieuwe techniek na brede implementatie ook wel gezien als de belangrijke 'check' stap in PDCA cirkel. Daarnaast kunnen ook verschillen

in gebruik en ervaring met de techniek worden geëvalueerd aan de hand van de data van een nationale clinical audit.

Een voorbeeld voor een dergelijke kwaliteitsbewaking is de Dutch Total Mesorectal Excision (TME) trial waarin de nieuwe TME techniek vergeleken werd met de standaard operatie techniek op dat moment. Om de verschillen in ervaring en interpretatie van de techniek tussen chirurgen te minimaliseren, werden chirurgen getraind middels workshops, videotaping en supervisie tijdens de eerste procedures. Ook kregen chirurgen direct feedback op hun techniek van de patholoog. De nieuwe TME techniek resulteerde in een reductie van het lokaal recidief percentage met 50%.[21] Echter, nadat de studie was afgesloten nam de focus op de TME resectie af, wat resulteerde in dat voor slechts 48% van alle patiënten met een rectumcarcinoom een CRM was beschreven in de DSCA in 2009. [**Hoofdstuk 3**]. Dit illustreert het belang van een constant feedback systeem.

Een ander voorbeeld van goede kwaliteitsbewaking met behulp van clinical auditing is de introductie van laparoscopische chirurgie in Nederland. In 2007 heeft de Nederlandse Vereniging voor Endoscopie een kwaliteitsbewaking programma ontwikkeld voor de introductie van nieuwe laparoscopische technieken. Dit systeem was gebaseerd op een PDCA cirkel waarbij richtlijnen voor het gebruik en onderhoud van apparatuur, een gestructureerd training en certificeringssysteem en een registratie en evaluatie systeem werden ontwikkeld. Wij hebben laten zien dat dit kwaliteitsbewakingssysteem geresulteerd heeft in de veilige en succesvolle introductie van laparoscopische colorectale kankerchirurgie in Nederland. Het percentage laparoscopische resecties in Nederland is hoog, met een zeer acceptabel conversiepercentage. Uitkomsten van laparoscopische resecties in Nederland zijn beter dan na open resecties, zelfs na correctie voor verschillen in case-mix. Uitkomsten na conversie waren vergelijkbaar met uitkomsten na open resecties. [**Hoofdstuk 6**] Dit staat in contrast met eerdere studies, die geen betere uitkomsten na laparoscopische chirurgie vonden.[22] Een mogelijke verklaring is dat deze gerandomiseerde studies vooral laag-risico patiënten bevatten, terwijl het voordeel van laparoscopische chirurgie groter zou kunnen zijn voor de hoog-risico patiënten,[23] die vooral in nationale clinical audits zoals de DSCA worden beschreven. Alhoewel we grote verschillen in het percentage laparoscopische resecties zagen tussen ziekenhuizen, vonden we geen bewijs dat deze verschillen van invloed waren op de uitkomsten. [**Hoofdstuk 6**]

## Deel 4. Het gebruik van clinical auditing om de kwaliteit van zorg te evalueren en transparant te zijn naar alle belanghebbenden

Overweldigende media aandacht voor ziekenhuisvariaties in kwaliteit van zorg heeft geleid tot een sterke maatschappelijke behoefte aan transparantie over kwaliteit van zorg.

Gegevens over ziekenhuisprestaties op het gebied van behandeling en uitkomsten van zorg kunnen gebruikt worden als 'kwaliteitsindicatoren' om deze vraag naar transparantie voor alle belanghebbenden te beantwoorden. Kwaliteitsindicatoren worden gedefinieerd als meetbare aspecten van zorg die de kwaliteit van zorg in zijn geheel reflecteren. Transparantie over dergelijke data zou de kwaliteit van zorg ook kunnen verbeteren via twee wegen[24].

- *Selectie*: data kunnen gebruikt worden door patiënten om een ziekenhuis te kiezen voor hun behandeling, door verzekeraars om ziekenhuizen op te selecteren of door de inspectie om slecht presterende ziekenhuizen te identificeren. De gedachte is dat door een dergelijke 'vrije markt' strategie de beter presterende ziekenhuizen de meeste patiënten behandelen, en de slecht presterende ziekenhuizen hiermee uitgeselecteerd worden.
- *Verbetering*: transparantie over behandeling en uitkomsten van zorg kan ook als stimulans werken voor ziekenhuizen om de kwaliteit van zorg te verbeteren en de kosten te reduceren. [**Hoofdstuk 2 en 3**]

Daarbij is de keuze van de kwaliteitsindicatoren cruciaal. Alhoewel kwaliteitsindicatoren al uitgebreid gebruikt worden in Nederland is de ideale kwaliteitsindicator, die bruikbaar is voor alle belanghebbenden, nog niet beschreven. In dit deel gaan we in op de keuze en definitie van kwaliteitsindicatoren, een zinvolle presentatie van kwaliteitsindicatoren en de noodzaak van het toepassen van correcties daarbij.

## Kwaliteitsindicatoren

Aan een goede kwaliteitsindicator worden een aantal eisen gesteld met betrekking tot *relevantie*, *wetenschappelijke waarde* en *bruikbaarheid* [25].

- *Relevantie*: is de indicator relevant voor een grote populatie met een aanzienlijk risico, en is er ruimte voor verbetering? Om een goede ziekenhuisvergelijking mogelijk te maken is het van belang dat de indicator voldoende vaak voorkomt,[5,26] echter hoe vaak is niet bekend.
- *Wetenschappelijke waarde*: is de indicator betrouwbaar en valide?.
  - Betrouwbaar betekent dat de data reproduceerbaar zijn en dat alle patiënten geïncludeerd zijn. Data moeten op een uniforme manier worden geregistreerd en verzameld, en definities moeten helder zijn. Voor een betrouwbare ziekenhuisvergelijking moeten gegevens ook gecorrigeerd worden voor verschillen in *case-mix* en voor *toevalsvariatie*.
  - Validiteit betekent dat de indicator meet wat hij zou moeten meten, en dat er een duidelijke en consistente relatie is tussen een goed resultaat op de kwaliteitsindicator en goede kwaliteit van zorg. Nu wordt kwaliteit van zorg vaak beoordeeld op basis van procesindicatoren, waarbij wordt aangenomen dat ziekenhuizen die patiënten volgens de richtlijnen behandelen ook goede uitkomsten hebben. Echter, voor de meeste procesindicatoren voor darmkanker is een relatie met uitkomsten van zorg niet onderzocht of niet bewezen.[27] Bovendien is het bekend dat ziekenhuizen vaak wisselend scoren op verschillende indicatoren,[28,29] wat het lastig maakt een oordeel te vormen over het ziekenhuis.

- *Bruikbaarheid*: de kwaliteitsindicatoren moeten begrijpelijk en bruikbaar zijn voor het bedoelde publiek. Verschillende belanghebbenden zullen geïnteresseerd zijn in verschillende aspecten van kwaliteit van zorg: zo zal de inspectie voor de gezondheidszorg misschien meer geïnteresseerd zijn in veiligheid en effectiviteit, zullen verzekeraars efficiëntie waarschijnlijk interessanter vinden, terwijl voor patiënten veiligheid, effectiviteit, tijdigheid en patiëntgerichtheid misschien het belangrijkst zijn. Daarom is het van belang de bruikbaarheid van de indicator voor het bedoelde publiek te evalueren.

## Clinical audits en kwaliteitsindicatoren.

In Nederland heeft de maatschappelijke behoefte aan transparantie over kwaliteit van zorg tot een enorme lijst van kwaliteitsindicatoren geleid, die in de media worden gebruikt om ziekenhuizen te rangschikken en beoordelen. Zij spelen ook een grote rol spelen in de inkoopstrategie van zorgverzekeraars. Echter, voor de meeste van deze indicatoren zijn relevantie, betrouwbaarheid, validiteit en bruikbaarheid nooit onderzocht, of bewezen dubieus.[30,31] Dit is niet verrassend gezien er voor de implementatie van clinical auditing in Nederland geen registratie systeem was voor kwaliteitsinformatie, waarmee deze kwaliteitsindicatoren bestudeerd konden worden. Clinical audits bevatten gedetailleerde, uniform geregistreerde informatie over zorgprocessen en uitkomsten, waarmee gefundeerd onderzoek naar goede kwaliteitsindicatoren voor de Nederlandse ziekenhuiszorg mogelijk gemaakt wordt.

## Procesindicatoren en uitkomstindicatoren

De meeste kwaliteitsindicatoren voor het colorectaal carcinoom zijn procesindicatoren, omdat deze minder beïnvloed worden door *case-mix* en *toevalsvariatie* dan uitkomstindicatoren. Echter valide proces indicatoren moeten een bewezen associatie hebben met relevante uitkomsten. Wij bestudeerden de associatie tussen proces- en uitkomstindicatoren in de DSCA, en vonden dat voor sommige procesindicatoren een goede ziekenhuisscore geassocieerd was met goede scores op de uitkomstindicatoren. Dit gold echter niet voor alle procesindicatoren. [**Hoofdstuk 7**] Dit zou verklaard kunnen worden doordat de ziekenhuisscores voor verschillende procesindicatoren inconsistent waren: ziekenhuizen met een goede score op de ene indicator hadden niet per se een goede score op de andere indicator. Daarom werden de procesindicatoren gecombineerd tot een samengestelde maat voor richtlijnnaleving. Wij lieten zien dat een goede score op richtlijnnaleving geassocieerd was met goede postoperatieve uitkomsten. [**Hoofdstuk 8**]. Gezien de meeste processen die in de richtlijn beschreven worden vooral effect hebben op de lange termijn (ziekte vrije) overleving, maar niet op het postoperatieve beloop, is het niet verrassend dat voor de individuele patiënt het behandeld worden via de richtlijn de postoperatieve uitkomsten niet beïnvloedde. Echter, op ziekenhuisniveau werd een sterke associatie gevonden tussen een goede score op richtlijnnaleving en goede postoperatieve uitkomsten. [**Hoofdstuk 8**] Alhoewel er altijd gevallen zullen zijn waarin er goede redenen zijn om af te wijken van de richtlijnen, zal dit de minderheid van de patiënten betreffen. Om alle andere patiënten

volledig volgens de richtlijn te behandelen is een logistieke uitdaging. Daarom beschrijft de gevonden associatie tussen richtlijnnaleving en een goede uitkomst waarschijnlijk ook niet een oorzakelijk verband, maar juist het 'ziekenhuis effect': de invloed van een toegewijd team en goed georganiseerd zorgproces.

## Samengestelde maten

Voor patiënten is het lastig om ziekenhuizen te beoordelen aan de hand van een lijst met uitkomsten: hoe kies je tussen een hoge kans op complicaties, maar een lage kans op postoperatieve sterfte? Echter, eerder onderzoek heeft laten zien dat patiënten graag een samengestelde maat voor uitkomsten zouden gebruiken voor hun ziekenhuiskeuze: de 'textbook outcome' maat die het percentage patiënten voor wie alle gewenste uitkomsten gehaald zijn weergeeft.[32,33] Deze maat is simpel en *bruikbaar* , niet alleen voor patiënten maar voor alle belanghebbenden. Voor patiënten beschrijft het hun kansen op de beste uitkomst na opname in een ziekenhuis. Voor zorgprofessionals kan deze maat feedback geven over hoe vaak de behandeling volledig verloopt zoals bedoeld. Voor verzekeringsmaatschappijen vat het de ziekenhuisresultaten op het gebied van veiligheid, effectiviteit en efficiëntie weer, en zou het een bruikbare indicator kunnen zijn voor selectieve inkoop. De inspectie zou deze maat kunnen gebruiken als surveillance maat. Door het combineren van verschillende uitkomsten van zorg in een begrijpelijke uitkomstmaat, voorkomt de 'textbook outcome' maat defensieve en indicator gedreven praktijkvoering. Een goede klinische besluitvorming zal daarom lijden tot een optimale 'textbook outcome' score. Bovendien bleek deze maat onderscheidend te zijn: gemiddeld hadden 49% van alle patiënten een 'textbook outcome', maar dit percentage verschilde tussen ziekenhuizen van 27 tot 71%. Zelfs na correctie voor case-mix bleken er ziekenhuizen te zijn waar slechts half zoveel patiënten een 'textbook outcome' had als verwacht, maar ook waar 1,4 keer zoveel patiënten een 'textbook outcome' hadden. Wij konden acht negatieve 'outliers' identificeren: ziekenhuizen waarin de kwaliteit slechter was dan gemiddeld op verschillende uitkomsten. [**Hoofdstuk 9**]

## Case-mix correctie

Ziekenhuis processen en uitkomsten kunnen sterk beïnvloed worden door de complexiteit van de patiënten en ziekten die behandeld worden. Daarom is het voor valide ziekenhuisvergelijkingen essentieel dat er gecorrigeerd wordt voor verschillen in case-mix. Echter, de registratie van deze case-mix gegevens is tijdrovend. Daarom wordt de vraag gesteld of case-mix correctie altijd relevant is voor alle uitkomstmaten. Dimick liet eerder zien dat voor coronaire bypass operaties gecorrigeerde en ongecorrigeerde uitkomsten sterk gerelateerd waren. Beide voorspelden de ziekenhuisprestaties van het volgende jaar even goed. Hieruit werd geconcludeerd dat de verschillen tussen ziekenhuizen in uitkomsten van coronaire bypass chirurgie niet beïnvloed werden door case-mix, mogelijk omdat de case-mix niet erg verschilde tussen ziekenhuizen.[34] Ook Snijders et al. lieten zien dat de grote variatie in het percentage naadlekkage tussen Nederlandse ziekenhuizen niet verklaard werd door verschillen in case-mix. Zij suggereerden dat andere factoren dan de meegenomen case-mix

de ziekenhuisvariatie verklaarden.[35] Echter, wij hebben laten zien dat de 'verwachte sterfte' na darmkanker resecties, berekend op basis van patiënt- en tumorkarakteristieken, varieerde tussen ziekenhuizen van 1,5 tot 14%. Deze resultaten laten duidelijk zien dat case-mix correctie essentieel is wanneer ziekenhuizen vergeleken worden op basis van mortaliteit na darmkanker resecties. [**Hoofdstuk 4**] Alhoewel de variatie in uitkomstindicatoren tussen ziekenhuizen voor een deel verklaard kan worden door verschillen in case-mix, zal dit deel niet even groot zijn voor alle uitkomsten. Echter case-mix correctie blijft belangrijk voor de geloofwaardigheid van ziekenhuis vergelijkingen op basis van uitkomstindicatoren.[34,35]

## Relevantie, het minimale ziekenhuisvolume voor het beoordelen van kwaliteit

Een ander probleem bij het vergelijken van ziekenhuisprestaties is de steekproefgrootte: wanneer het aantal procedures laag is of de incidentie van de uitkomstindicator laag is, zijn de betrouwbaarheidsintervallen rond de indicatoren erg groot. Door dit statistische fenomeen kunnen laagvolume ziekenhuizen een 5 keer slechtere prestatie leveren dan verwacht, zonder dat bewezen kan worden dat zij significant slechter dan gemiddeld presteren. [**Hoofdstuk 10**] Daarom stellen wij een gecombineerde maat voor volume en uitkomst voor, waarmee ziekenhuizen geïdentificeerd kunnen worden die bewijs leveren van goede kwaliteit van zorg. Deze ziekenhuizen hebben goede uitkomsten bij een voldoende aantal patiënten om te kunnen garanderen dat, als de prestaties gelijk blijven, dat de uitkomsten het volgende jaar niet slechter zullen zijn dan een vooraf bepaald niveau van 'ondermaatse zorg'. Wij lieten zien dat deze gecombineerde maat voor volume en uitkomst de ziekenhuisprestaties in het volgende jaar beter voorspelde dan volume of uitkomst alleen. Echter, de gecombineerde maat werkte beter voor morbiditeit dan voor mortaliteit. [**Hoofdstuk 10**] Eerdere studies hebben laten zien dat de invloed van toevalsvariatie op ziekenhuisprestaties groter is wanneer de uitkomstindicator weinig voorkomt, terwijl de invloed van systematische variatie, zoals praktijkvoering in een ziekenhuis, groter is wanneer een uitkomstindicator vaker voor komt.[26] Dit zou mogelijk verklaren waarom de gecombineerde maat voor volume en uitkomst beter werkte voor morbiditeit dan voor mortaliteit. Daarom is het belangrijk te kijken naar het vóórkomen van een uitkomstindicator wanneer men ziekenhuizen wil vergelijken op basis van deze indicator.

## Goede kwaliteitsindicatoren voor valide en betekenisvolle feedback informatie over ziekenhuisprestaties

Individuele indicatoren kunnen altijd nuttig zijn, zo kan een procesindicator gebruikt worden om de implementatie van een specifiek proces te monitoren. Echter individuele indicatoren geven niet altijd een goed beeld van de kwaliteit van het gehele zorgproces. Wanneer indicatoren gecombineerd worden tot een samengestelde maat voor richtlijnnaleving, blijkt een goede ziekenhuisscore geassocieerd te zijn met goede uitkomsten. Daarom is deze samengestelde maat voor richtlijnnaleving meer *valide* en *bruikbaar* om de kwaliteit van zorg te beoordelen dan een individuele proces indicator. Wanneer ziekenhuizen worden beoordeeld op uitkomstindicatoren, moet men zich realiseren dat uitkomsten vaak aan

elkaar gerelateerd zijn, en dat een goede kwaliteit van zorg niet alleen over veiligheid gaat, maar ook over effectiviteit en efficiëntie. Bovendien moeten uitkomstindicatoren altijd gecorrigeerd worden voor verschillen in case-mix om de geloofwaardigheid te waarborgen. Als laatste is de betrouwbaarheid van een uitkomstindicator afhankelijk van het voorkomen van de indicator en het ziekenhuisvolume. Daarom stellen wij een gecombineerde maat voor volume en uitkomst voor om ziekenhuizen met goede uitkomsten in voldoende aantallen te selecteren waarvoor betrouwbaar gesteld kan worden dat de kwaliteit van zorg nu èn volgend jaar goed genoeg is. Wanneer deze volume-uitkomst combinatie wordt toegepast op de 'textbook outcome' indicator, kunnen ziekenhuizen geselecteerd worden waar patiënten erop kunnen vertrouwen dat zij een grote kans hebben op een 'textbook outcome'.

## De toekomst

De snelle en succesvolle introductie van de DSCA en de imposante kwaliteitsverbetering die werd gezien na de implementatie hebben geleid tot initiatieven voor vele nieuwe clinical audits. Met de DSCA als blauwdruk kunnen nieuwe audits binnen een jaar ontwikkeld en opgestart worden. Clinical auditing blijft bovendien niet beperkt tot de chirurgische oncologie: in de melanomen audit is ook systemische behandeling opgenomen, de vaatchirurgie heeft twee audits opgestart voor de behandeling van het abdominaal aneurysma en voor carotis chirurgie en er is een audit voor kinderchirurgie en voor de behandeling van cerebrovasculaire accidenten en de ziekte van Parkinson. Wij hopen en verwachten dat binnen enkele jaren alle specialismen en alle veel voorkomende ziektebeelden worden geëvalueerd in een clinical audit. Een recent rapport van de Gezondheidsraad bepleit clinical auditing als de belangrijkste stap richting transparantie en verbetering van kwaliteit van zorg.[36]

## Transparantie in de zorg

Tot voor kort werden gegevens van de DSCA geanonimiseerd gepubliceerd. Angst voor datafraude, perverse prikkels en een 'naming and shaming' cultuur waren de belangrijkste drempels voor de openbaarmaking van gegevens. De Boston Consulting Group beschreef echter recent aan de hand van het Zweedse voorbeeld dat het openbaar maken van valide gegevens over kwaliteit van zorg afkomstig van clinical audits niet resulteerde in nationale media schandalen, maar in een snelle, nationale verbetering van de kwaliteit van zorg.[37] Daarom heeft de Nederlandse Vereniging van Heelkunde in 2012 besloten tot een getrapt transparantie programma: het eerste jaar wordt alleen het ziekenhuis volume gepubliceerd, het volgende jaar worden proces indicatoren openbaar gemaakt, en het derde jaar worden ook uitkomstindicatoren naar buiten gebracht. Een dergelijk getrapt transparantie programma biedt ziekenhuizen de ruimte hun data te controleren en eventuele achterstanden in te halen, maar beantwoordt ook de maatschappelijke behoefte aan transparantie over kwaliteit van zorg. Ieder jaar stellen afgevaardigden van de medische beroepsgroep, samen met verzekeraars en patiëntenverenigingen, de indicatoren vast voor de verschillende audits.

Gezamenlijk besluiten zij welke indicatoren dat jaar leidend zullen zijn voor de beoordeling van de kwaliteit van zorg, en voor de inkoop strategie van verzekeraars dat jaar.

De grootste belanghebber in transparantie in de zorg is waarschijnlijk de patiënt. Echter, eerder onderzoek heeft laten zien dat patiënten publiekelijk beschikbare informatie over kwaliteit van zorg maar zelden gebruiken voor hun keuze voor een ziekenhuis, omdat zij de informatie niet begrijpen of niet vertrouwen.[38,39] Met behulp van clinical audits kunnen relevante, valide en bruikbare gegevens over kwaliteit van ziekenhuiszorg gegenereerd worden, zoals de 'textbook outcome' indicator, die door patiënten gebruikt kunnen worden voor hun keuze voor een ziekenhuis van behandeling.[32] Verder onderzoek zou zich kunnen richten op de vraag hoe de behoefte van patiënten aan informatie over kwaliteit van zorg het beste vervuld kan worden, en op welke manier deze informatie het beste gepresenteerd kan worden.

De Nederlandse Federatie van Kankerpatiëntenorganisaties heeft hierin reeds het voortouw genomen en vier relevante aspecten van kwaliteit van zorg gedefinieerd:
structuur informatie: beschrijft wat er aangeboden wordt in een ziekenhuis, en of het ziekenhuis aan de door de beroepsgroep gestelde normen voldoet;
proces- en uitkomstinformatie van clinical audits: beschrijft of de ziekenhuizen voldoet aan de standaarden voor proces- en uitkomstindicatoren zoals bepaald door de beroepsgroep;
Patiënt Reported Outcome Measures (PROM's): beschrijven de resultaten van de behandeling en de kwaliteit van leven vanuit het perspectief van de patiënt;
Consumer Quality index: beschrijft de subjectieve beleving van patiënten behandeld in dat ziekenhuis.
Het Dutch Institute for Clinical Auditing heeft in samenwerking met het NFK en Zorgverzekeraars Nederland een transparantie project gelanceerd waarin deze vier aspecten van kwaliteit van zorg worden geïntegreerd, met de DSCA als showcase. De verschillende aspecten van kwaliteit van zorg worden op een publiekelijk beschikbare website gerapporteerd. Deze website, volledig ontworpen voor en door patiënten, is de volgende stap in het vergroten van het maatschappelijk vertrouwen in een goede kwaliteit van zorg in Nederland.

## Referenties


1. Khuri SF, Daley J, Henderson WG. The comparative assessment and improvement of quality of surgical care in the Department of Veterans Affairs. Arch Surg. Jan 2002;137(1):20-27.
2. O'Connor GT, Plume SK, Olmstead EM, et al. A regional intervention to improve the hospital mortality associated with coronary artery bypass graft surgery. The Northern New England Cardiovascular Disease Study Group. JAMA. Mar 20 1996;275(11):841-846.
3. Wibe A, Moller B, Norstein J, et al. A national strategic change in treatment policy for rectalcancer--implementation of total mesorectal excision as routine treatment in Norway. A national audit. Dis Colon Rectum. Jul 2002;45(7):857-866.
4. Ivers N, Jamtvedt G, Flottorp S, et al. Audit and feedback: effects on professional practice and healthcare outcomes. Cochrane Database Syst Rev. 2012;6:CD000259.
5. Dimick JB, Welch HG, Birkmeyer JD. Surgical mortality as an indicator of hospital quality: the problem with small sample size. JAMA. Aug 18 2004;292(7):847-851.
6. van Gijn W, Wouters MW, Peeters KC, van de Velde CJ. Nationwide outcomeregistrations to improve quality of care in rectal surgery. An initiative of the European Society of Surgical Oncology. J Surg Oncol. Jun 15 2009;99(8):491-496.
7. Hsueh Y. The Hawthorne experiments and the introduction of Jean Piaget in American industrial psychology, 1929-1932. History of psychology. May 2002;5(2):163-189.
8. Gale EA. The Hawthorne studies-a fable for our times? QJM : monthly journal of the Association of Physicians. Jul 2004;97(7):439-449.
9. Lehmann C, Nowak A. The Hawthorne effect: can it be measured and utilized? British journal of anaesthesia. Apr 2013;110(4):658-659.
10. Usichenko TJ, A; Wendt, M. 'Yes, we can' utilize the Hawthorne effect to improve postoperative analgesia. British journal of anaesthesia. 2013;110(4):659.
11. Pouw ME, Peelen LM, Lingsma HF, et al. Hospital standardized mortality ratio: consequences of adjusting hospital mortality with indirect standardization. PLoS One. 2013;8(4):e59160.
12. Almoudaris AM, Burns EM, Bottle A, Aylin P, Darzi A, Faiz O. A colorectal perspective on voluntary submission of outcome data to clinical registries. Br J Surg. Jan 2011;98(1):132-139.
13. Engel A. This paper is about a rather small colorectal unit trying to come to grips with the frustration produced by the discrepancy between their perceived performance and NBOCAP performance. Colorectal Dis. Aug 2012;14(8):1019.
14. Lewis RS, Graham DG, Watson JD, Lunniss PJ. Learning the hard way: the importance of accurate data. Colorectal Dis. Aug 2012;14(8):1015-1018.
15. Jamtvedt G, Young JM, Kristoffersen DT, O'Brien MA, Oxman AD. Does telling people what they have been doing change what they do? A systematic review of the effects of audit and feedback. Qual Saf Health Care. Dec 2006;15(6):433-436.
16. Dixon N. Proposed standards for the design and conduct of a national clinical audit or quality improvement study. Int J Qual Health Care. May 21 2013.
17. Wouters MWJM. DICA rapportages 2012: Dutch Institute for Clinical Auditing;2013.
18. van Gijn W, Gooiker GA, Wouters MW, Post PN, Tollenaar RA, van de Velde CJ. Volume and outcome in colorectal cancer surgery. Eur J Surg Oncol. Jul 6.
19. Elshove-Bolk J, Ellensen VS, Baatrup G. Logistics and outcome in urgent and emergency colorectal surgery. Colorectal Dis. Nov 10 2009.
20. Biondo S, Kreisler E, Millan M, et al. Impact of surgical specialization on emergency colorectal surgery outcomes. Arch Surg. Jan;145(1):79-86.
21. Peeters KC, van de Velde CJ. Surgical quality assurance in rectal cancer treatment: the key to improved outcome. Eur J Surg Oncol. Aug 2005;31(6):630-635.
22. Schwenk W, Haase O, Neudecker J, Muller JM. Short term benefits for laparoscopic colorectal resection. Cochrane Database Syst Rev. 2005(3):CD003145.
23. McCloskey CA, Wilson MA, Hughes SJ, Eid GM. Laparoscopic colorectal surgery is safe in the high-risk patient: a NSQIP risk-adjusted analysis. Surgery. Oct 2007;142(4):594-597; discussion 597 e591-592.
24 Berwick DM, James B, Coye MJ. Connections between quality measurement and improvement. Med Care. 2003 Jan;41(1 Suppl):I30-8.
25. Dimick JB. What makes a "good" quality indicator? Arch Surg. Mar 2010;145(3):295.
26. van Dishoeck AM, Lingsma HF, Mackenbach JP, Steyerberg EW. Random variation and rankability of hospitals using outcome indicators. BMJ quality & safety. Oct 2011;20(10):869-874.
27. Patwardhan M, Fisher DA, Mantyh CR, et al. Assessing the quality of colorectal cancer care: do we have appropriate quality measures? (A systematic review of literature). J Eval Clin Pract. Dec 2007;13(6):831-845.
28. Habib MR, Solomon MJ, Young JM, Armstrong BK, O'Connell D, Armstrong K. Evidence-based and clinical outcome scores to facilitate audit and feedback for colorectal cancercare. Dis Colon Rectum. Apr 2009;52(4):616-622; discussion 622-613.
29. Lilford R, Mohammed MA, Spiegelhalter D, Thomson R. Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. Lancet. Apr 3 2004;363(9415):1147-1154.

30 Gooiker GA, Verbeek L, van der Geest LG, et al. The quality indicator ‚tumour positive surgical margin following breast-conserving surgery' does not provide transparent insight into care. Ned Tijdschr Geneeskd. 2010;154:A1142.

31 Beperkt zicht; Onderzoek naar de betrouwbaarheid, validiteit en bruikbaarheid van prestatie-indicatoren over de kwaliteit van de Nederlandse ziekenhuiszorg. Rapport NFU-consortium Kwaliteit van Zorg. 2012.

32. Marang-van de Mheen PJ, Dijs-Elsinga J, Otten W, et al. The relative importance of quality of care information when choosing a hospital for surgical treatment: a hospital choice experiment. Medical decision making : an international journal of the Society for Medical Decision Making. Nov-Dec 2011;31(6):816-827.

33. Dijs-Elsinga J, Otten W, Versluijs MM, et al. Choosing a hospital for surgery: the importance of information on quality of care. Medical decision making : an international journal of the Society for Medical Decision Making. Sep-Oct 2010;30(5):544-555.

34. Dimick JB, Birkmeyer JD. Ranking hospitals on surgical quality: does risk-adjustment always matter? J Am Coll Surg. Sep 2008;207(3):347-351.

35. Snijders HS, Henneman D, van Leersum NL, et al. Anastomotic leakage as an outcome measure for quality of colorectal cancer surgery. BMJ quality & safety. Sep 2013;22(9):759-767.

36. Gezondheidsraad. Rapport: Publieke indicatoren voor kwaliteit van curatieve zorg. De stand van de discussie. 2013

37. Kuenen JM, R; Larsson, S; Leeuwen, W van. Zorg voor waarde. Boston Consultancy Group. 2011.

38. Marshall MN, Shekelle PG, Leatherman S, Brook RH. The public release of performance data: what do we expect to gain? A review of the evidence. JAMA. Apr 12 2000;283(14):1866-1874.

39. Faber M, Bosch M, Wollersheim H, Leatherman S, Grol R. Public reporting in health care: how do consumers use quality-of-care information? A systematic review. Med Care. Jan 2009;47(1):1-8.

# Curriculum Vitae

Nicoline Kolfschoten was born in Woerden on April 12th 1982. She grew up in the city of the Hague. After graduating from the Hague Montessori Lyceum in 2000, she joined the Royal Navy and studied at the 'Koninklijk instituut voor de Marine' where she graduated for her propedeusis in 2001. In the same year she started Medical School at the University of Utrecht.

While attending medical school she founded a student cover band and sang at many events. Also, she actively participated in the Utrecht Student Alpine Club and further developed her mountaineering skills. During her elective internship in sports medicine, she studied chronic groin pain in athletes together with Jaap Jansen.

Her passion for mountaineering brought her in contact with Prof Dr Mike Grocott and the Caudwell – Xtreme Everest group, with whom she spent 6 weeks in Tibet climbing Cho Oyu, and 3 months at Everest Base Camp to study human adaptation to hypoxia for her final-year's research thesis.

During her final internship in emergency medicine at the Diakonessenhuis in Utrecht, she collaborated in a systematic review after treatment of the acute achilles tendon rupture, together with Roderik Metz and Dr Egbert Verleisdonk. In 2008 she graduated Medical school and started working at the surgical department of the Diakonessenhuis in Utrecht.

In 2009 she started her research at the Leiden University Medical Centre under the supervision of Prof Dr. Job Kievit, Prof Dr Rob Tollenaar, dr Michel Wouters and dr Perla Marang van der Mheen. At a time when the first data of the first Dutch clinical audit, the DSCA came available, she was investigated and developed 'valid and valuable feedback data'. Furthermore she assisted with the developement of the Dutch Institute for Clinical auditing, which was founded by dr Eric Hans Eddes, Prof dr Rob Tollenaar and dr Michel Wouters. She also assisted in the development and implementation of the Dutch Upper-GI Audit.

In january 2012 she started her surgical training at the Leiden University Medical Centre under the supervision of prof. Dr. Jaap Hamming, and in january 2014 she continued her surgical training at the Medisch Centrum Haaglanden under the supervision of dr. Sven Meijlaerts. In July 2014 she decided to switch to emergency medicine and she started her training in emergency medicine in January 2015.

# Dankwoord

Dit proefschrift is het resultaat van mijn noeste arbeid en een fantastische ondersteuning en teamwork, waarvoor ik heel veel mensen wil bedanken, en een paar mensen in het bijzonder.

Allereerst dank ik mijn beide promotoren Job Kievit en Rob Tollenaar, die mij de mogelijkheid hebben gegeven om deze ervaringen op de doen en dit onderzoek te volbrengen. Daarnaast hebben zij mij de kans gegeven om nauw betrokken te zijn bij de opstart van het Dutch Institute for Clinical Auditing, een instituut wat een 'gamechanger' in de chirurgie, en de ziekenhuiszorg in Nederland blijkt te zijn.

Natuurlijk wil ik ook mijn beide co-promotoren Perla Marang-van de Mheen en Michel Wouters bedanken voor hun sturende begeleiding. Michel Wouters wil ik bedanken voor de inhoudelijke sturing, waarmee altijd de juiste insteek werd gevonden om de boodschap beter te 'verkopen'. Perla Marang-van de Mheen wil ik bedanken voor haar statistische en kritische wetenschappelijke begeleiding, waarmee ik niet alleen mijn proefschrift, maar ook mijn werk binnen de DICA inhoudelijk beter invulling kon geven. Ook Eric-Hans Eddes wil ik bedanken voor zijn steun en hulp tijdens mijn promotieperiode, en de kansen en het vertrouwen.

Daarnaast wil ik natuurlijk mijn collega onderzoekers van de heelkunde van het LUMC bedanken voor hun hulp en gezelligheid in deze periode. Ik ben er trots op deel uit te maken van het succes van C-11-14. Specifiek wil ik de kwaliteits-meisjes Gea Gooiker, Nicolien van Leersum en Heleen Snijders bedanken voor de positieve stimulans van het samenwerken, en Daan en Martijn en alle volgende DICA onderzoekers voor het overdragen en verder uitbouwen van wat wij hebben opgebouwd. Ook wil ik Anouck en Duveken bedanken voor alle gezelligheid in en na het LUMC.

De DSCA en DUCA wetenschappelijke commissie en methodologische raad wil ik bedanken voor de enorme inspirerende inzet voor het succes van de audits, voor het vertrouwen en de steun voor mijn carrière en de inhoudelijke input in mijn proefschrift. Daarnaast dank ik alle chirurgen, fellows, assistenten, semi-artsen, co-assistenten, nurse-practicioners, physician assistents en research verpleegkundigen in Nederland die de data van de DSCA invoerden, en op die manier bijdroegen aan het succes van de DSCA en de DICA, en specifiek de chirurgen van het LUMC voor hun ongezouten kritiek vanuit het veld.

Mijn vrienden uit school en studietijd, mijn jaarclub en klimvrienden, en specifiek Lonneke, Iris, Sietske, Saar, Veronique en Janne, voor hun support, (sportieve) afleiding, avonturen en altijd positieve stimulans om door te gaan. En natuurlijk mijn familie, en vooral mijn ouders Richard en Rinske en zus Gwendolyn. Dank voor alle steun en motivatie, dat ik mocht leren van jullie eerdere ervaringen. Ik ben er trots op deel uit te maken van ons gezin.