

BENCHMARKING COMPREHENSIVE CANCER CARE



Anke Wind

BENCHMARKING COMPREHENSIVE CANCER CARE

By

Anke Wind

Address of correspondence

Anke Wind

Groningerstraat 156

3812 EG, Amersfoort, the Netherlands

ankewind@gmail.com

Cover design: Daan van der Wal

Printing: Gildeprint, Enschede

The printing of this thesis was supported by The Netherlands Cancer Institute

BENCHMARKING COMPREHENSIVE CANCER CARE

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof. dr. T.T.M. Palstra,
volgens besluit van het College voor Promoties
in het openbaar te verdedigen
op donderdag 13 april 2017 om 16.45 uur

door

Anke Wind
geboren op 8 mei 1988,
te Hengelo

Dit proefschrift is goedgekeurd door de promotor:
Prof. dr. W.H. van Harten

© Copyright 2017: Anke Wind, Amersfoort, The Netherlands

ISBN: 978-90-365-4315-6

DOI: 10.3990/1.9789036543156

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system of any nature, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without written permission of the holder of the copyright.

Promotiecommissie

Voorzitter/secretaris

Prof. dr. T.A.J. Toonen, Universiteit Twente

Promotor

Prof. dr. W.H. van Harten, Universiteit Twente

Leden

Prof. dr. J.A.M. van der Palen, Universiteit Twente

Prof. dr. ir. E.W. Hans, Universiteit Twente

Prof. dr. R. Torenvlied, University of Twente

Prof. dr. N.S. Klazinga, Academisch Medisch Centrum/UvA

Dr. M.W.J.M. Wouters, Nederlands Kanker Instituut

Paranimfen

Melanie Lindenberg

Daan van der Wal

Voor mijn opa

Contents

1	General introduction and outline of the dissertation	9
2	Benchmarking specialty hospitals, a scoping review on theory and practice	27
3	Quality assessments for cancer centers in the European Union	67
4	Piloting a Generic Cancer Consumer Quality Index in six European countries	89
5	Development of a Benchmark tool for Cancer Centers; results from a pilot exercise	111
6	Benchmarking cancer centers: from care pathways to Integrated Practice Units	137
7	Management and performance features of cancer centers in Europe: a fuzzy-set analysis	163
8	Actual costs of cancer drugs in 15 European countries	179
9	General discussion and conclusions	191
10	Summary/Samenvatting	205
11	Acknowledgments/Dankwoord	219
12	Curriculum vitae	225

PART 1

General introduction

The aim of this thesis is to contribute to the knowledge on how to develop and use benchmarking for quality improvement of cancer care. In this introduction I will explore the concepts of quality improvement in healthcare and the principles of benchmarking. Furthermore, this chapter describes the research scope, the research methods, and the outline of this thesis.

General introduction

In 2012 there were approximately 3.45 million new cases of cancer (excluding non-melanoma skin cancer) in Europe¹; In that same year 1.75 million people died from cancer¹. The number of cancer patients and survivors is steadily increasing and despite or perhaps because of rapid improvements in diagnostics and therapeutics, important inequalities in cancer survival exist within and between different countries in Europe. Studies indicate that the differences in cancer survival are largely attributable to: socioeconomic factors, inequalities in quality of care and screening, inequalities in diffusion and adherence to clinical guidelines, and inequalities in access to high quality radiotherapy equipment and cancer drugs². Improving the quality of care is part of the approach to reduce suboptimal cancer survival and minimize inequalities in Europe.

Quality Improvement in healthcare

Quality Improvement (QI) is an essential part of healthcare management and can be sought on macro-, meso- and microlevels of the health care system. Governments and payer agencies typically play a role on macro level with general regulations, reporting systems and reviews and inspections. On European level we see activities that are directed towards reducing inequalities in service provision and outcomes and to facilitate cross border treatment. Especially professional societies put effort in educating the individual provider and by defining guidelines for conduct and treatment in specific areas. Institutions (the meso level) have to deal with both the official regulations and many professional guidelines, but also have to ensure the quality on an organizational level. They engage in implementing quality and risk management systems, multidisciplinary guideline systems and a range of other quality management and assurance activities.

It is difficult to define quality of healthcare and definitions often leave room for interpretation. A commonly used definition is from the Institute of Medicine (IoM)³: "Quality of care is the degree to which health services for individuals and populations increase the likelihood of desired outcomes and are consistent with current professional knowledge".

A more practical definition was given earlier by Donabedian⁴. Within this definition quality is described with regard to structure, process and outcomes. Structure measures refer to the availability of for example resources, management systems and guidelines. Process measures correspond to the processes necessary for daily healthcare delivery. Outcomes can contain medical indicators (e.g. mortality ratios, complication rates) as well as patient experience and satisfaction data.

Healthcare institutions are pressured by payers, patients and society to strive for continuous improvement⁵ which has led to a growing need for reliable performance evaluation tools⁶. Measurement is essential for hospital quality improvement⁷; “it provides a means to define what hospitals actually do, and to compare that with the original targets in order to identify opportunities for improvement” (pp.4). There are in principle five different types of measurement of hospital performance⁷:

- Regulatory inspection
- Surveys of consumers’ experiences
- Third-party assessments
- Statistical indicators
- Internal assessments.

(1) Regulatory inspection (most countries have statutory inspectorates) of hospitals causes conformity, and measures performance related to minimal requirements for safety.

(2) Standardized surveys of patients and relatives can measure hospital performance. Advantages of this method are that it identifies what is valued by patients and the general public, and standardized surveys can be tailored to measure specific domains of experience and satisfaction. (3) Third-party assessment includes for example: Peer review (a closed system for professional self-assessment and development); and Accreditation (programs that measure hospital performance in terms of compliance with published standards of organizational – and, increasingly, clinical – processes and results). (4) Statistical indicators can be used to identify issues for performance management, quality improvement and further scrutiny. (5) Internal assessments or self-assessment is used by hospitals to assess and analyze weaknesses and strengths inside the hospital⁸. Benchmarking, which can both be a third-party assessment as an internal assessment, is a common and effective method for measuring and analyzing performance⁹.

Benchmarking

The quality improvement approach examined in this thesis is benchmarking which focuses on learning from others and setting realistic performance targets. The Joint commission⁹ defines benchmarking as:

A systematic, data-driven process of continuous improvement that involves internally and/or externally comparing performance to identify, achieve, and sustain best practice. It requires measuring and evaluating data to establish a target performance level or benchmark to evaluate current performance and comparing these benchmarks or performance metrics with similar data compiled by other organizations, including best-practice facilities” (pp.1).

For healthcare Mosel and Gift¹⁰ provided the following definition: “... *benchmarking is the continual and collaborative discipline of measuring and comparing the results of key work processes with those of the best performers. It is learning how to adapt these best practices to achieve breakthrough process improvements and build healthier communities*”. In general the objectives of benchmarking are: (1) to determine what and where improvements are needed, (2) to analyze how comparable organizations achieve their own high performance levels, and (3) to use this information to improve performance. Benchmarking in the healthcare sector has undergone several modifications¹¹. Initially, benchmarking was essentially the comparison of performance outcomes to identify disparities. In the mid-1990s, it developed to a structured method with the imperative of comparing hospital outcomes to rationalize their funding^{12,13}. It was expanded to include the analysis of processes and success factors for producing higher levels of performance. The most recent modifications to the concept of benchmarking relate to the need to meet patients’ expectations¹⁴.

There are several classifications of benchmark types and models. Benchmarking can be internal (comparing between different groups or teams within an organization) or external (comparing with other organizations in a specific industry or across industries). The most commonly cited typology of benchmarking is Camp’s¹⁵ differentiation between internal, competitive, functional and generic benchmarking. Within these broader categories Bhutta and Huq¹⁶ identified three types of benchmarking: performance, process and strategic. These categories have been expanded upon by other researchers: for example, ‘best in class benchmarking’¹⁷, used to emphasize the organization-independent nature of generic benchmarking. McGonagle and Fleming¹⁸ identified so called “shadow benchmarking”, industrial benchmarking that compares similar organizations, but not exactly the same functions within the same industry/sector, often against the industry leaders (it is similar to functional benchmarking, focusing on a single function to improve the operation of

1 that particular function) and international benchmarking that analyzes processes, comparison with any industry and with world leading organizations (similar to world class or generic benchmarking). To implement benchmarking, there is a need for useful, reliable and up-to-date information. This process of ongoing information management is called surveillance¹¹. Surveillance is the first basis of benchmarking and facilitates and accelerates the benchmarking process. A second basis is learning, information sharing and the adoption of best/good practices to improve performance. This project covered the first base (surveillance) and part of the second base (learning and information sharing). Good practices were identified, however their adoption was not in the scope of this study. This study focused on combining two types of benchmarking: operations and clinical practice (process) benchmarking¹⁹ and associated patient experience²⁰; and performance benchmarking²¹. It compared relative performance drawing on quantitative and qualitative data. First experiences with these in health care generally and oncology specifically, shows that involving comparable centers and services lead to fruitful suggestions for improvement. Brucker et al.²² showed that a nationwide benchmarking system has proved a clinically oriented, practical, flexible, adaptable and extensible tool for measuring and improving the quality of for example breast cancer care. The National Practice Benchmark described by Barr et al.²³ showed that the Oncology community is changing in orderly ways moving toward gains in efficiency as assessed by a variety of measures. Brann et al.²⁴ reported that Benchmarking has the potential to illuminate intra- and inter-organizational performance.

The need for a cancer care benchmarking tool

As mentioned in the general introduction significant inequalities in cancer survival exist within Europe. There seems to be a gap between the potential to provide innovative high quality cancer care and the actual situation in the provision of oncologic care²⁵. The increasing complexity in both multidisciplinary cancer care and translational research requires a new and closer collaboration between cancer centers (CCs). The Stockholm declaration^{25,26} has highlighted the importance of this collaboration between CCs to facilitate high quality care. Identifying what works can assist hospitals in improving their services and reduce inequalities in care provision. It has the potential to raise the level of oncologic services across Europe. Benchmarking is a tool to facilitate the identification of what works and in 2013, the Organisation of European Cancer Institutes (OECI)²⁷ launched the BENCH-CAN project²⁸, aiming at benchmarking comprehensive cancer care to reduce health inequalities in Europe and improve interdisciplinary cancer care by yielding best practice examples. The aim of one of the work packages of this project (work package 4, led by the Netherlands Cancer Institute) was to develop benchmarking tools for comprehensive cancer care (benchmark tool 1) and cancer care pathways (benchmark tool 2) using both qualitative and quantitative approaches.

Outline and aim of this thesis

This thesis is divided in 6 parts: Part 1 Introduction; Part 2 Current situation; Part 3 Patient perspective; Part 4 Qualitative and quantitative benchmarking; Part 5 International financial and quantitative data comparison; and Part 6 Retrospect & prospect describing the results and conclusions and providing a discussion on methodological issues, further research and policy consequences. The aim of this thesis is to present tools to benchmark comprehensive cancer care and cancer care pathways/tumor services. Linked to this aim are several sub-objectives (see Figure 1). They include: (I) assessing the current situation of benchmarking in specialty hospitals and existing quality assessments (Chapter 2 and Chapter 3); (II) measuring patient's perspectives on the quality of care at CCs (Chapter 4); (III) developing and piloting two extensive benchmark tools for comprehensive cancer care (Chapter 5 and Chapter 6); and finally (IV) to investigate the use of quantitative benchmarking and (financial) performance features (Chapter 7 and Chapter 8). This section presents the research objectives, their rationale and the methods.

Research objective 1: Assessing the current situation of benchmarking in specialty hospitals and existing quality assessments for cancer centers

Before embarking on the development and pilot a benchmarking tool for quality assessment of comprehensive cancer care and cancer care pathways/tumor services there was a need to know the state of the art of benchmarking approaches in this field to inform our own approach. To prevent reinventing the wheel, it was assessed which indicators are already used to measure quality in cancer centers.

Chapter 2 Benchmarking specialty hospitals, a scoping review on theory and practice

A scoping literature review was conducted with the following objectives: (i) provide an overview of research on benchmarking in specialty hospitals and care pathways, (ii) describe study characteristics such as method, setting, models/frameworks, and outcomes, and (iii) verify the quality of benchmarking as a tool to improve quality in specialty hospitals and identify success factors.

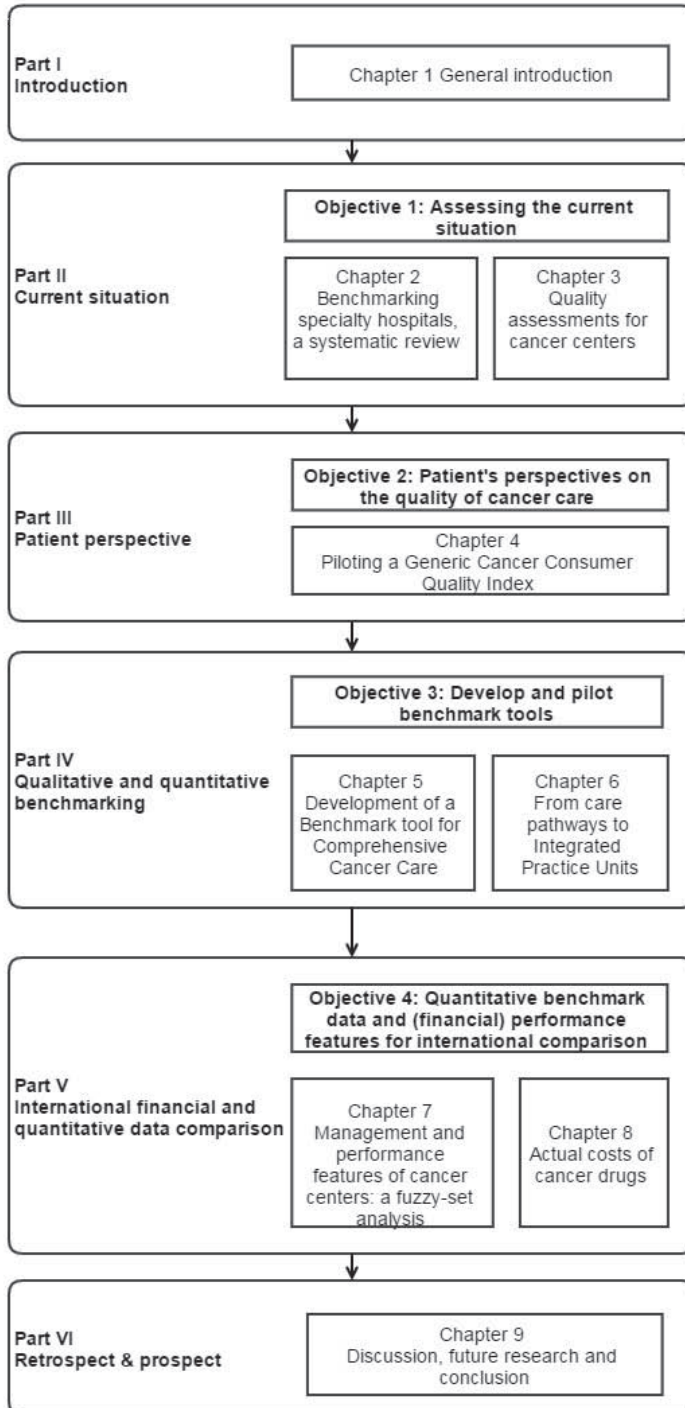


Figure 1 Thesis outline

Chapter 3 Quality assessments for cancer centers in the European Union

European cancer centers go through several assessments at regional, national and international levels. However, many things regarding these assessments remain unclear such as the type of assessments being conducted, who conducts them and with what frequency, and are these assessments focused on assessing research, patient care or both. The goal was to obtain an overview of existing assessments in terms of whether they are: mandatory or voluntary; focused on evaluating research or patient care or both; regional, national and/or international. Data on existing assessments was collected through a survey with the quality managers from CCs in 28 EU member states. Purposive sampling was employed. One CC per member state was contacted. Responses from all CCs were analyzed thematically and verified with the respondents for validity.

Research objective 2: Measuring patient's perspectives on the quality of care at cancer centers

Accounting for patients' perspective has become increasingly important in healthcare quality evaluation. It was therefore decided to develop a patient experience and satisfaction tool as part of the BENCH-CAN project.

Chapter 4 Piloting a Generic Cancer Consumer Quality Index in six European countries

Based on the Consumer Quality Index method (founded on Consumer Assessment of Healthcare Providers and Systems) a questionnaire was recently developed for Dutch cancer patients assessing their experience and satisfaction with care received. As a next step, this study aimed to adapt and pilot this questionnaire for international comparison of cancer patients experience and satisfaction with care in six European countries. We identified two research questions:

1. What are the differences in patient experience and satisfaction between countries and/or patient characteristics?
2. What is the validity and internal consistency (reliability) of the European Cancer Consumer Quality Index?

The Consumer Quality Index was translated into the local language at the participating pilot sites using cross-translation. A minimum of 100 patients per site were surveyed through convenience sampling. Data from seven pilot sites in six countries was collected through an online and paper-based survey. Internal consistency was tested by calculating Cronbach's alpha and validity by means of cognitive interviews. Demographic factors were compared as possible influencing factors.

Research objective 3: Develop and pilot two extensive benchmark tools for comprehensive cancer care

Benchmarking has the potential to illuminate inter-organizational performance differences and facilitate quality improvement. In order to benchmark comprehensive cancer care two tools were developed. One tool looking at cancer centers as a whole, the institutional tool and one tool focusing on cancer care pathways/tumor services. In the benchmarking literature various methods can be found. We used the stepwise approach that was developed Van Lent et al.¹⁹ based on a series of benchmarking pilots in various care organizations (Table 1).

Table 1 13 step benchmarking method developed by van Lent et al.¹⁹

Step	Action
1	Determine what to benchmark
2	Form a benchmarking team
3	Choose benchmarking partners
4	Define and verify the main characteristics of the partners
5	Identify stakeholders
6	Construct a framework to structure the indicators
7	Develop relevant and comparable indicators
8	Stakeholders select indicators
9	Measure the set of performance indicators
10	Analyse performance indicators
11	Take action: results are presented in a report and recommendations are given
12	Develop relevant plans
13	Implement the improvement plans

The indicators that were needed to generate data within this approach were structured (step 6) within a framework based on the European Foundation for Quality Management (EFQM) Model²⁹ and the Institute of Medicine (IOM) domains of Quality³⁰. The European Foundation for Quality Management published a model for performance-assessment and identification of key strengths and improvement areas. It includes 9 criteria in which the organizational structure and processes (enablers) are considered as well as the results, which can be demonstrated by outcome measures. The categories show the various aspects of an organization. Good performance in the enabler's domains is expected to lead to good performance in the results domain³¹. For the results domains the IOM domains of quality were used. For the benchmark tool the domains of quality are adapted into effective, efficient, safe, responsive and personalized, integration, and timely as shown in Figure 2.

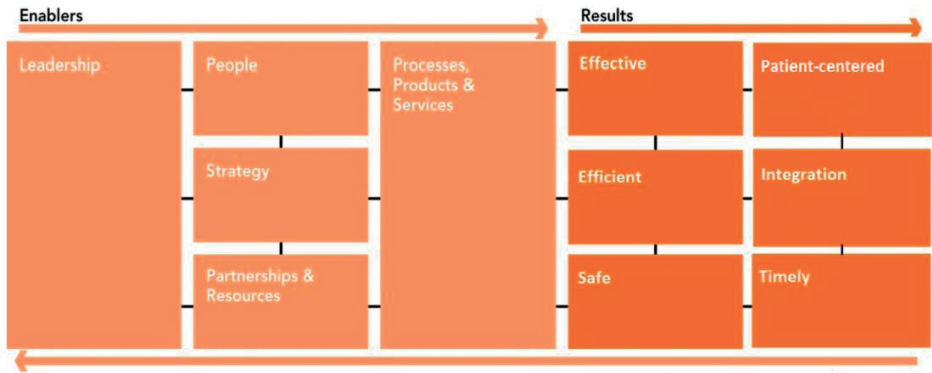


Figure 2 The BENCH-CAN Framework

Chapter 5 Development of a Benchmark tool for Cancer Centers; results from a pilot exercise

Although the method for developing the benchmark tools was largely the same some differences can be found. This paragraph describes the specifics for the institutional tool (Benchmark tool 1, BT1). A comprehensive international benchmarking tool was developed covering all relevant care related and organizational fields. Related to this we identified the following research objectives: (i) develop and pilot an extensive benchmark tool with both qualitative and quantitative indicators, (ii) identify performance differences between cancer centers, and (iii) identify improvement opportunities. Eight cancer centers throughout Europe were selected as pilot sites. The benchmark indicators were tested and pre-piloted in three centers to see whether the definitions were clear and the indicators would yield interesting, discriminative information. After the indicators were adapted the tool was tested in five other centers. The collected data was used to identify improvement suggestions and good practices.

Chapter 6 Benchmarking cancer centers: from care pathways to Integrated Practice Units

Care pathways are often used as a tool to manage the quality in healthcare. It has been shown that their implementation reduces the variability in clinical practice and improves outcomes³², but the European Pathway Association³³ identified a need for international benchmarking. A further step is organizing care according to Integrated Practice Units (IPU), encompassing the whole pathway and all relevant organizational aspects³⁴. Research on this topic is however limited. This study aimed at firstly describing the development and outcomes of benchmark for care pathways (Benchmark tool 2, BT2). The second aim of this study was assessing the degree of development towards an IPU. The benchmark data

1 was used to produce suggestions for pilot sites to improve the organization of cancer care pathways towards the development of IPUs.

Research objective 4: Investigate the use of quantitative benchmarking data and (financial) performance features for international comparison

Quantitative benchmark data can be used to map efficiency in cancer centers and compare costs. Besides simply comparing, the data can be used to identify the importance of quantitative performance features and how they relate to outcomes by means of the fuzzy-set qualitative comparative analysis (fsQCA)³⁵. Quantitative data can also be used to highlight inequalities and call for changes.

Chapter 7 Management and performance features of cancer centers in Europe: a fuzzy-set analysis

Data collected through the quantitative/financial benchmark indicators were used to test a relatively new method within health service research, the fuzzy-set Qualitative Comparative Analysis (fsQCA). In contrast to other quantitative methods, such as a regression analysis, fsQCA can be used for small sample sizes (5–50 cases). The fsQCA method represents cases (cancer centers) as a combination of explanatory and outcome conditions. This study uses the net income and productivity as the outcome conditions and five explanatory conditions: level of dedication to R&D, annual budget, size, type, and whether the center is a Comprehensive Cancer Center (CCC).

Chapter 8 Actual costs of cancer drugs in 15 European countries

International comparison can also be used to identify problems and inequalities in for example the pricing of cancer drugs and stimulate efforts towards joint action. A word-based survey was emailed to all full members of the OECD (n=51), both European Union (EU) members and non-EU members, and to the non-OECD member of Cancer Core Europe. The centers were asked to provide list or official and actual prices, corrected for VAT differences, and asked for information about central or government coordinated purchasing. The actual price was defined as the net price—ie, as price per one dose to allow for a comparison in case of different pack sizes.

Table 2 Summary research objectives, chapters and research methods

Objective	Chapter	Research method
I Assessing the current situation of benchmarking in specialty hospitals and existing quality assessments for cancer centers	Chapter 2	Scoping literature review
	Chapter 3	Survey
II Measuring patient's perspectives on the quality of care at cancer centers	Chapter 4	Patient experience and satisfaction questionnaire (ECCQI)
III Develop and pilot two extensive benchmark tools for comprehensive cancer care	Chapter 5	Multi-center benchmark pilot study
	Chapter 6	Multi-center benchmark pilot study
IV Investigate the use of quantitative benchmarking data and (financial) performance features for international comparison	Chapter 7	fussy-set Qualitative Comparative Analysis
	Chapter 8	Survey

References

1. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer* 2013;49:1374-1403.
2. Verdecchia A, Francisci S, Brenner H, Gatta G, Micheli A, Mangone L, Kunkler I. EURO-CARE-4 Working Group. Recent cancer survival in Europe: a 2000-02 period analysis of EUROCARE-4 data. *Lancet Oncol* 2007;8(9):784-96.
3. Institute of Medicine Committee on Quality of Health in A. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington DC: National Academies Press Report, 2001 Contract No.: Report.
4. Donabedian A. Evaluating the quality of medical care. *The Milbank Memorial Fund quarterly* 1966;44(3):Suppl:166-206.
5. Mohr JJ, Batalden P, Barach P. Integrating patient safety into the clinical microsystem. *Qual Saf Health Care* 2004;13:ii34-8.
6. Swaminathan S, Chernew M, Scanlon DP. Persistence of HMO Performance Measures. *HSR* 2008;43(6): 2033-2049.
7. Shaw C (2003) How can hospital performance be measured and monitored? Copenhagen, WHO Regional Office for Europe. Health Evidence Network report <http://www.euro.who.int/document/e82975.pdf>
8. Joson R. Internal Situation / Environment Assessment and Analysis for Hospitals. <https://journeytowardexcellenceforhospitals.wordpress.com/2012/11/07/internal-situation-environment-assessment-and-analysis-for-hospitals/>
9. Joint Commission. *Benchmarking in Health Care*. Joint Commission Resources; 2011
10. Mosel D, Gift B. Collaborative benchmarking in health care. *Joint Commission Journal on Quality Improvement* 1994, 20:239.
11. Ettorchi-Tardy A, Levif M, Michel P. Benchmarking: a method for continuous quality improvement in health. *Healthcare policy* 2012; 7(4): e101.
12. Camp RC. *Global Cases in Benchmarking: Best Practices from Organizations Around the World*. Milwaukee: American Society for Quality Control Quality Press, 1998.
13. Dewan NA, Daniels A, Zieman G, Kramer T. The National Outcomes Management Project: A Benchmarking Collaborative. *J Behav Health Ser R* 2000; 27(4): 431-36.
14. Ellis J. All Inclusive Benchmarking. *J Nurs Manag* 2006; 14(5): 377-83.
15. Camp RC. *Benchmarking: The Search for Industry Best Practices that Lead to Superior Performance*. Milwaukee,(USA) 1989.
16. Bhutta KS, Huq F. Benchmarking - best practices: an integrated approach. *BIJ* 1999; 6(3): 254.
17. Spendolini M.J. The benchmarking process. *CBR* 1992; 24(5): 21-29.
18. McGonagle JJ, Fleming D. New options in benchmarking. *The Journal for Quality and Participation* 1993; 16(4): 60.
19. van Lent W, de Beer R, van Harten W. International benchmarking of specialty hospitals. A series of case studies on comprehensive cancer centres. *BMC Health Services Research* 2010; 10: 253.
20. Haines S, Warren T. Staff and patient involvement in benchmarking to improve care. *Nurs Manag* 2011; 18(2): 22-25.

21. Duckett SJ, Ward M. Developing 'robust performance benchmarks' for the next Australian Health Care Agreement: the need for a new framework. *Aust New Zealand Health Policy* 2008; 5(1), 1.
22. Brucker S, Schumacher C, Sohn C, Rezai M, Bamberg, Wallwiener D, et al. The Steering Committee: Benchmarking the quality of breast cancer care in a nationwide voluntary system: the first five-year results (2003–2007) from Germany as a proof of concept. *BMC Cancer* 2008; 8: 358.
23. Barr T, Towle E. Oncology Practice Trends From the National Practice Benchmark. *J Oncol Pract* 2012; 8: 292-7.
24. Brann P, Walter G, Coombs T. Benchmarking child and adolescent mental health organizations. *Australas Psychiatry* 2011; 19: 125-32.
25. Ringborg U. The stockholm declaration. *Mol Oncol*. 2008; 2(1): 10-11.
26. Brown H. Turning the Stockholm Declaration into reality: Creating a world-class infrastructure for cancer research in Europe. *Mol Oncol* 2009; 3: 5-8.
27. OECI accessed through <http://www.oeci.eu/>
28. BENCH-CAN accessed through <http://www.oeci.eu/Benchcan/>
29. THE EFQM EXCELLENCE MODEL. <http://www.efqm.org/the-efqm-excellence-model>
30. Institute of Medicine. *Crossing the quality chasm: a new health system for the 21st century*. Washington DC: National Academy Press, 2001.
31. Nabitz U, Klazinga N, Walburg JAN. The EFQM excellence model: European and Dutch experiences with the EFQM approach in health care. *Int J Qual Health Care* 2000; 12(3): 191-202.
32. Panella M. Reducing clinical variations with clinical pathways: do pathways work?. *Int J Qual Health Care* 2003; 15 (6): 509–521.
33. Vanhaecht K, Bollmann M, Bower K, et al. Prevalence and use of clinical pathways in 23 countries—an international survey by the European Pathway Association. *Journal of Integrated Pathways* 2006; 10(1): 28-34.
34. Porter ME, Lee TH. The strategy that will fix health care. *Harv Bus Rev* 2013; 91(12): 24.
35. Ragin CC. *Redesigning social inquiry: Fuzzy sets and beyond*. Chicago: University of Chicago Press, 2008.

PART 2

Current situation

CHAPTER 2

Benchmarking specialty hospitals, a scoping review on theory and practice

**Anke Wind
Wim van Harten**

Abstract

Background: Although benchmarking may improve hospital processes, research on this subject is limited. The aim of this study was to provide an overview of publications on benchmarking in specialty hospitals and a description of study characteristics.

Methods: We searched PubMed and EMBASE for articles published in English in the last ten years. Eligible articles described a project stating benchmarking as its objective and involving a specialty hospital or specific patient category; or those dealing with the methodology or evaluation of benchmarking.

Results: Of 1,817 articles identified in total, 24 were included in the study. Articles were categorized into: pathway benchmarking, institutional benchmarking, articles on benchmark methodology or evaluation and benchmarking using a patient registry. There was a large degree of variability: (1) study designs were mostly descriptive and retrospective; (2) not all studies generated and showed data in sufficient detail; and (3) there was variety in whether a benchmarking model was just described or if quality improvement as a consequence of the benchmark was reported upon. Most of the studies that described a benchmark model described the use of benchmarking partners from the same industry category, sometimes from all over the world.

Conclusions: Benchmarking seems to be more developed in eye hospitals, emergency departments and oncology specialty hospitals. Some studies showed promising improvement effects. However, the majority of the articles lacked a structured design, and did not report on benchmark outcomes. In order to evaluate the effectiveness of benchmarking to improve quality in specialty hospitals, robust and structured designs are needed including a follow up to check whether the benchmark study has led to improvements.

Background

Healthcare institutions are pressured by payers, patients and society to deliver high-quality care and have to strive for continuous improvement. Healthcare service provision is becoming more complex, leading to quality and performance challenges¹. In addition, there is a call for transparency on relative performance between and within healthcare organizations². This pushes providers to focus on performance and show the added value for customers/patients^{3,4}.

Without objective data on the current situation and comparison with peers and best practices, organizations cannot determine whether their efforts are satisfactory or exceptional, and specifically, what needs improvement. Benchmarking is a common and effective method for measuring and analyzing performance. The Joint commission defines benchmarking as:

A systematic, data-driven process of continuous improvement that involves internally and/or externally comparing performance to identify, achieve, and sustain best practice. It requires measuring and evaluating data to establish a target performance level or benchmark to evaluate current performance and comparing these benchmarks or performance metrics with similar data compiled by other organizations, including best-practice facilities⁵.

Benchmarking may improve hospital processes, though according to Van Lent et al.⁶, benchmarking as a tool to improve quality in hospitals is not well described and possibly not well developed. Identifying meaningful measures that are able to capture the quality of care in its different dimensions remains a challenging aspiration⁷. Before embarking on an international project to develop and pilot a benchmarking tool for quality assessment of comprehensive cancer care (the BENCH-CAN project⁸) there was a need to establish the state of the art in this field, amongst others to avoid duplication of work. The BENCH-CAN project⁸ aims at benchmarking comprehensive cancer care and yield good practice examples at European cancer centers in order to contribute to improvement of multi-disciplinary patient treatment. This international benchmark project included 8 pilot sites from three geographical regions in Europe (North-West (n=2), South (n=3), Central-East (n=3)). The benchmarking study was executed according to the 13 steps developed by van Lent et al.⁶, these steps included amongst others the construction of a framework, the development of relevant and comparable indicators selected by the stakeholders and the measuring and analysing of the set of indicators. Accordingly, we wanted to obtain an overview on benchmarking of specialty hospitals and specialty care pathways. Schneider et al.⁹ describe specialty hospitals as hospitals “that treat patients with specific medical

conditions or those in need of specific medical or surgical procedures” (pp.531). These are standalone, single-specialty facilities.

The number of specialty hospitals is increasing⁹. Porter¹⁰ suggests that specialization of hospitals improves performance; it results in a better process organization, improved patient satisfaction, increased cost-effectiveness and better outcomes. Specialty hospitals represent a trend; however, according to van Lent et al.⁶ the opinions about the added value are divided. More insight into the benchmarking process in specialty hospitals could be useful to study differences in organization and performance and the identification of optimal work procedures⁶. Although specialty hospitals may differ according to discipline they have similarities such as the focus on one disease category and the ambition to perform in sufficient volumes. The scope of the BENCH-CAN⁸ project was on cancer centers and cancer pathways, however, we did not expect to find sufficient material on these specific categories and thus decided to focus on specialty hospitals in general. Against this background, we conducted a scoping review. A scoping review approach provides a methodology for determining the state of the evidence on a topic that is especially appropriate when investigating abstract, emerging, or diverse topics, and for exploring or mapping the literature¹¹ which is the goal of this study. This study had the following objectives: (i) provide an overview of research on benchmarking in specialty hospitals and care pathways, (ii) describe study characteristics such as method, setting, models/frameworks, and outcomes, and (iii) verify the quality of benchmarking as a tool to improve quality in specialty hospitals and identify success factors.

Method

Scoping systematic review

There are different types of research reviews which vary in their ontological, epistemological, ideological, and theoretical stance, their research paradigm, and the issues that they aim to address¹². Scoping reviews have been described as a process of mapping the existing literature or evidence base. Scoping studies differ from systematic reviews in that they provide a map or a snapshot of the existing literature without quality assessment or extensive data synthesis¹². Scoping studies also differ from narrative reviews in that the scoping process requires analytical reinterpretation of the literature¹¹. We used the framework as proposed by Arksey and O’Mally¹³. This framework consist of 6 steps: (i) identifying the research question, (ii) identifying relevant studies, (iii) study selection, (iv) charting the data, (v) collecting, summarizing and reporting the results, (vi) optional consultation. Step 6 (optional consultation) was ensured by asking stakeholders from the BENCH-CAN

project for input. Scoping reviews are a valuable resource that can be of use to researchers, policy-makers and practitioners, reducing duplication of effort and guiding future research.

Data sources and search methods

We performed searches in PubMed and EMBASE. To identify the relevant literature, we focused on peer-reviewed articles published in international journals in English between 2003 and 2014. According to Saggese et al.¹⁴ “this is standard practice in bibliometric studies, since these sources are considered ‘certified knowledge’ and enhance the results’ reliability” (pp.4). We conducted Boolean searches using truncated combinations of three groups of keywords and free text terms in title/abstract (see Figure 1). The first consists of keywords concerning benchmarking and quality control. The second group includes key words regarding type of hospitals. All terms were combined with group 3: organization and administration. Different combinations of keywords led to different results, therefore five different searches in PubMed and four in EMBASE were performed. To retrieve other relevant publications, reference lists of the selected papers were used for snowballing. In addition stakeholders involved in the BENCH-CAN project⁸ were asked to provide relevant literature.

Selection method/ article inclusion and exclusion criteria

Using abstracts, we started by excluding all articles that clearly did not meet the inclusion criteria, which covered topics not related to benchmarking and specialty hospitals. The two authors independently reviewed the remaining abstracts and made a selection using the following criteria: the article had to discuss a benchmarking exercise in a specialty hospital either in theory or in practice and/or the article had to discuss a benchmark evaluation or benchmark tool development. Only studies including organizational and process aspects were used, so studies purely benchmarking clinical indicators were excluded. At least some empirical material or theory (or theory development) on benchmarking methodology should be present; essays mainly describing the potential or added value of benchmarking without proving empirical evidence were thus excluded. The articles also had to appear in a peer-reviewed journal. The full texts were reviewed and processed by the first author. Only papers written in English were included.

Data extraction

General information was extracted in order to be able to provide an overview of research on benchmarking in specialty hospitals and care pathways. The following information was extracted from the included articles: first author and year of publication, aim, and area of

2 practice. The analytical data were chosen according to our review objective. They included the following: (I) study design, (II) benchmark model and/or identified steps, (III) type of indicators used, (IV) study outcome, (V) the impact of the benchmarking project (measured by the identified improvements achieved through the benchmark or suggestions for improvements), and (VI) success factors identified. The first author independently extracted the data and the second author checked 25% of the studies to determine inter-rater reliability.

Classification scheme benchmark models

At present, there is no standard methodology to classify benchmark models within healthcare in general and more specifically within specialty hospitals and care pathways. Therefore we looked at benchmark classification schemes outside the healthcare sector, especially in industry. A review of benchmarking literature showed that there are different types of benchmarking and a plethora of benchmarking process models¹⁵. One of these schemes was developed by Fong et al.¹⁶ (box 1). This scheme gives a clear description of each element included in the scheme and will therefore be used to classify the benchmark models described in this paper. It can be used to assess academic/research-based models. These models are developed mainly by academics and researchers mainly through their own research, knowledge and experience (this approach seems most used within the healthcare sector). This differs from Consultant/expert-based models (developed from personal opinion and judgment through experience in providing consultancy to organizations embarking on a benchmarking project) and Organization-based models (models developed or proposed by organizations based on their own experience and knowledge. They tend to be highly dissimilar, as each organization is different in terms of its business scope, market, products, process, etc.)¹⁶.

Classification	Type	Meaning
Nature of benchmarking partner	Internal	Comparing within one organization about the performance of similar business units or processes
	Competitor	Comparing with direct competitors, catch up or even surpass their overall performance
	Industry	Comparing with company in the same industry, including non-competitors
	Generic	Comparing with an organization which extends beyond industry boundaries
	Global	Comparing with an organization where its geographical location extends beyond country boundaries
Content of benchmarking	Process	Pertaining to discrete work processes and operating systems
	Functional	Application of the process benchmarking that compares particular business functions at two or more organizations
	Performance	Concerning outcome characteristics, quantifiable in terms of price, speed, reliability, etc.
	Strategic	Involving assessment of strategic rather than operational matters
Purpose for the relationship	Competitive	Comparison for gaining superiority over others
	Collaborative	Comparison for developing a learning atmosphere and sharing of knowledge

Box 1 Classification scheme for benchmarking by Fong et al.¹⁶

Results

Review

The search strategy identified 1,817 articles. The first author applied the first review eligibility criteria, the topic identification (Figure 1), to the titles and abstracts. After this initial examination 1,697 articles were excluded. Two authors independently reviewed the abstracts of 120 articles. Snowballing identified three new articles that were not already identified in the literature search. Sixty articles were potentially eligible for full text review. The full text of these 60 publications were reviewed by two authors, resulting in a selection of 24 publications that met all eligibility criteria (see Figure 1 and 2)

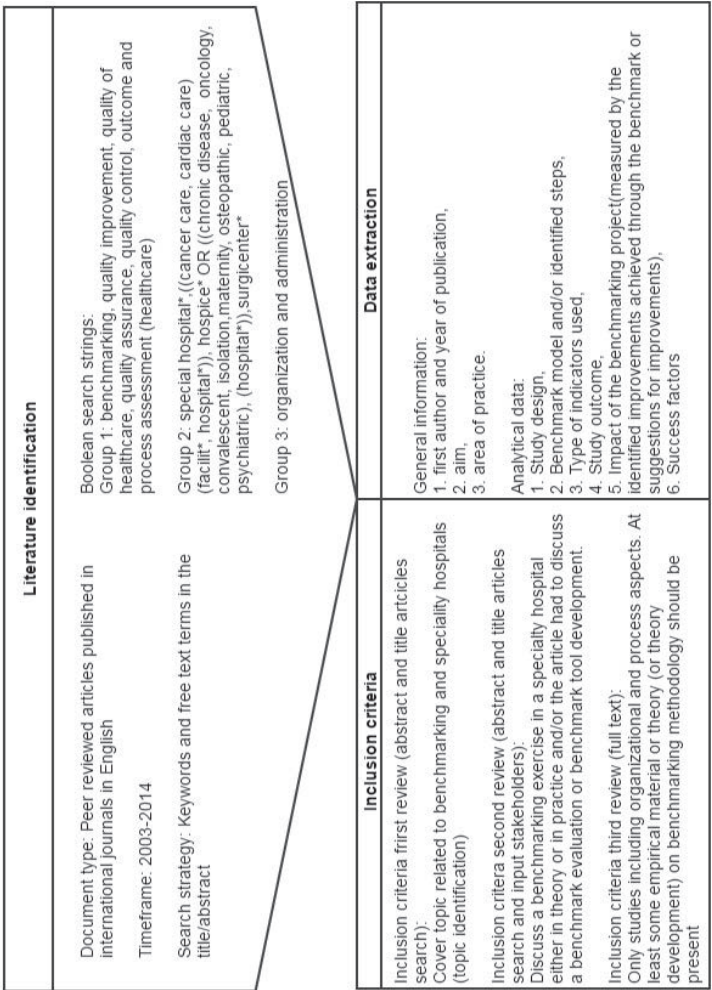


Figure 1 Research Design

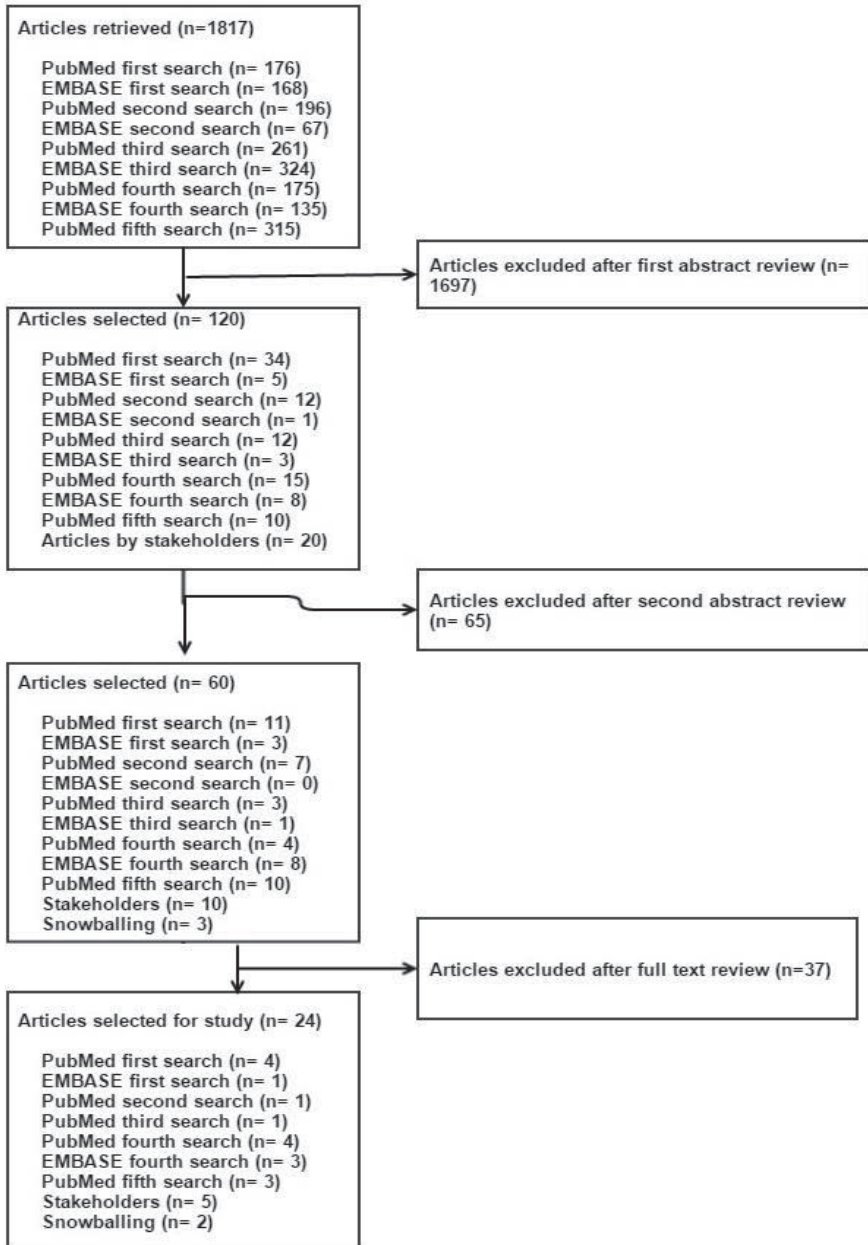


Figure 2 Article selection process

Study characteristics

Table 1 provides an overview of the general information of the included articles. To assist in the analysis, articles were categorized into: pathway benchmarking, institutional

2

benchmarking, benchmark evaluation/methodology and benchmarking using a patient registry (see Figure 3). For each category the following aspects will be discussed: study design, benchmark model and/or identified steps, type of indicators used, study outcome, impact of the benchmarking project (improvements/improvement suggestions) and success factors. The benchmark model and/or described steps will be classified using the model by Fong¹⁶.

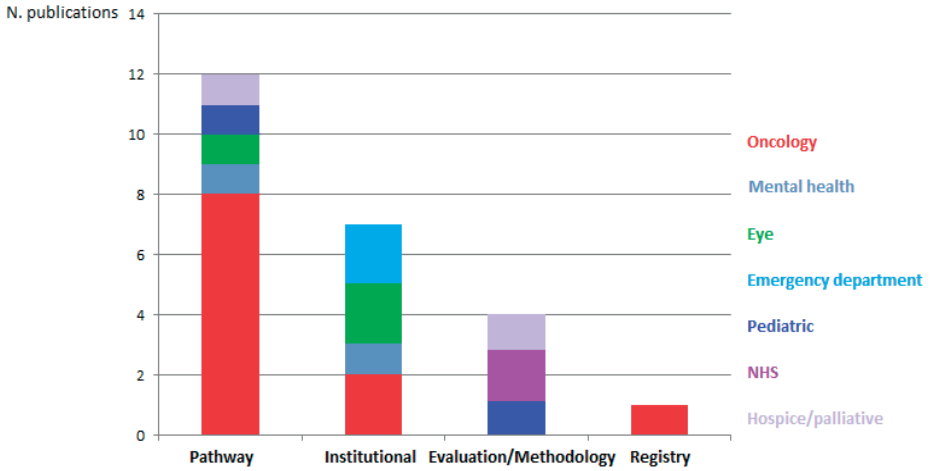


Figure 3 Number of publications per category and area of practice

Table 1 charting categories and associated content for the general information on the benchmarking studies

First author (Year)	Aim	Area of practice
Brucker (2008) [17]	Establish a nationwide network of breast centres; to define suitable quality indicators (QIs) for benchmarking the quality of breast cancer (BC) care; to demonstrate existing differences in BC care quality; and to show that BC care quality improved with benchmarking from 2003 to 2007.	Breast cancer centers Germany
Chung (2010) [18]	Developing organization-based core measures for colorectal cancer patient care and apply these measures to compare hospital performance.	Hospitals registered in the TCDB program in Taiwan
Hermann (2006) [19]	To identify quality measures for international benchmarking of mental healthcare that assess important processes and outcomes of care, are scientifically sound, and are feasible to construct from pre-existing data.	Mental health care professionals from six countries (UK, Sweden, Canada, Australia, Denmark, and the USA) and one international organization, the European Society for Quality in Healthcare (ESQH)
Mainz (2009) [20]	Describing and analyzing the quality of care for important diseases in the Nordic countries (Denmark, Finland, Greenland, Iceland, Norway and Sweden).	Cancer treatment facilities from the different Nordic countries (Denmark, Finland, Greenland, Iceland, Norway and Sweden)
Miransky (2003) [21]	Describing the development of a database for benchmarking outcomes for cancer patients.	A consortium of 12 Comprehensive Cancer Centers in the US
Roberts (2012) [22]	The study had three main aims, to: (i) adapt the acuity-quality workforce planning method used extensively in the UK National Health Service (NHS) for use in hospices; (ii) compare hospice and NHS palliative care staffing establishments and their implications; and (iii) create ward staffing benchmarks and formulae for hospice managers.	Twenty-three palliative care and hospice wards, geographically representing England.
Setoguchi (2008) [23]	Comparing prospectively and retrospectively defined benchmarks for the quality of end-of-life care, including a novel indicator for the use of opiate analgesia.	Seniors with breast, colorectal, lung, or prostate cancer who participated in state pharmaceutical benefit programs in New Jersey and Pennsylvania
Stewart (2007) [24]	Develop tools that lead to better-informed decision making regarding practice management and physician deployment in comprehensive cancer centers and determine benchmarks of productivity using RVUs (Relative value units) accrued by physicians at each institution.	13 major academic cancer institutions with membership or shared membership in the National Comprehensive Cancer Network (NCCN)
Stolar (2010) [25]	Performing a blinded confidential financial performance survey of similar university pediatric surgery sections to start benchmarking performance and define relationships.	19 pediatric surgery sections of university children's hospitals

Table 1 charting categories and associated content for the general information on the benchmarking studies (continued)

First author (Year)	Aim	Area of practice
Van Vliet (2010) [26]	Comparing process designs of three high-volume cataract pathways in a lean thinking framework and to explore how efficiency in terms of lead times, hospital visits and costs is related to process design.	Three eye hospitals in the UK, the USA and the Netherlands
Wallwiener (2011) [27]	Summarize the rationale for the creation of breast centres and discuss the studies conducted in Germany. To obtain proof of principle for a voluntary, external benchmarking programme and proof of concept for third-party dual certification of breast centres and their mandatory quality management systems.	Breast centers in Germany
Wesselman (2014) [28]	Present data from the third annual analysis of the DKG-certified colorectal cancer centers with a particular focus on indicators for colorectal cancer surgery.	Colorectal cancer centers certified by the German Cancer Society (DKG)
Barr (2012) [29]	Revision of 2011 predictions with the use of National Practice Benchmark (NPB) reports from 2011 and development of new predictions. Design of a conceptual framework for contemplating these data based on an ecological model of the oncology delivery system.	Oncology practices in the USA
Brann (2011) [30]	The performance of child and adolescent mental health organizations. To provide an overview of the findings from two projects, undertaken to explore the variability in organizations' performances on particular KPIs (key performance indicators).	Six child and adolescent mental health organizations
De Korne (2010) [3]	The purpose of this study was to evaluate the applicability of an international benchmarking initiative in eye hospitals.	Nine eye hospitals spread over Asia (3), Australia (1), Europe (4), and North America (1).
De Korne (2012) [31]	The aims of this study were to assess the applicability of a benchmarking project in U.S. eye hospitals and compare the results with an international initiative.	Five eye hospitals in the US
Schwappach (2003) [32]	Assess the effects of uniform indicator measurement and group benchmarking. This was followed by hospital-specific activities on clinical performance measures and patients' experiences with emergency care in Switzerland.	Emergency departments of 12 community hospitals in Switzerland, participating in the 'Emerge' project.
Shaw (2003) [33]	To answer basic questions, using precise definitions, regarding emergency department (ED) utilization, wait times, services, and attending physician staffing of representative pediatric EDs (PEDs).	21 Pediatric emergency departments (PED) from 14 states of the USA.
Van Lent (2010) [6]	Examine benchmarking as part of an approach to improve performance in specialty hospitals	International comprehensive cancer centres (CCC) or departments within a CCC in Europe and the US

Table 1 charting categories and associated content for the general information on the benchmarking studies (continued)

First author (Year) [34]	Aim	Area of practice
Ellershaw (2008) [34]	To evaluate the utility of participating in two benchmarking exercises to assess the care delivered to patients in the dying phase using the Liverpool Care Pathway for the Dying Patient (LCP).	Two cancer networks in the northwest of England
Ellis (2006) [35]	Review published descriptions of benchmarking activity and synthesize benchmarking principles to encourage the acceptance and use of Essence of Care as a new approach to continuous quality improvement, and to promote its acceptance as an integral and effective part of benchmarking activity in health services.	NHS (UK)
Matykiewicz (2005) [36]	Introduce Essence of Care, a benchmarking tool for health care practitioners and an integral part of the UK National Health Service (NHS) Clinical Governance agenda	Health care practitioners NHS (UK)
Profit (2010) [37]	To present a conceptual framework to develop comprehensive, robust, and transparent composite indicators of pediatric care quality, and to highlight aspects specific to quality measurement in children.	The Pediatric Data Quality Systems (Pedi-QS) Collaborative Measures Workgroup (consensus panel by the National Association of Children's Hospitals and Related Institutions, Child Health Corporation of America, and Medical Management Planning)
Greene (2009) [38]	Describing The Role of the Hospital Registry in Achieving Outcome Benchmarks in Cancer Care	Carolinas Medical Center (US)

I Pathway benchmarking (PB)

PB Study design

Study design varied across the different pathway studies. Most studies (n=7)^{18-21, 24,25} used multiple comparisons, from which five studies sought to develop indicators. Different methods were used for this indicator development such as a consensus method (Delphi)¹⁸⁻²⁰. In other articles a less structured way of reaching consensus was used such as conference calls²¹ and surveys²⁴. One study used a prospective interventional design¹⁴ while another study²⁶ used a retrospective comparative benchmark study with a mixed-method design. Setoguchi et al.²³ used a combination of prospective and retrospective designs. Existing literature was used in two studies^{27,28}. More information on study design can be found in Table 2.

PB Benchmark model

Eight articles described a benchmarking model and/or benchmarking steps. Applying the classification scheme by Fong et al.¹⁶ most studies used benchmarking partners from the same industry (n=6)^{17,21,23,24,27,28}. Two studies also used partners from the industry but on the global level. A total of 6 studies benchmarked performance^{17,21,23, 27, 28}, one study benchmarked performance and processes¹⁹ and another study used strategic benchmarking²⁶. All studies used benchmarking for collaborative purposes. For more information about the benchmark models see Table 2.

PB Indicators

Most of the pathway studies used outcome indicators (n=7)^{17, 20, 21, 23, 24, 25, 28}. Hermann et al.¹⁹ used a combination of process and outcome indicators e.g. case management and length of stay; and Chung et al.¹⁸ used structure, process and outcome indicators. One study²¹ used a mixture of process and outcome indicators, while another study²⁷ used a combination of structural and process indicators. Most studies used quantitative indicators, such as five-year over-all survival rate¹⁸. Roberts et al.²² described the use of qualitative and quantitative indicators.

PB outcomes

Looking at the outcomes of the different pathway studies it can be seen that these cover a wide range of topics, Brucker¹⁷ for example provided proof of concept for the feasibility of a nationwide system for benchmarking. The goal of establishing a nationwide network of certified breast centres in Germany can be considered largely achieved according to Wallwiener²⁷. Wesselman²⁸ showed that most of the targets for indicators for colorectal care are being better met over the course of time.

Mainz et al.²⁰ reported a major difference between the Nordic countries with regard for 5 years survival for prostate cancer. However, they also reported difficulties such as: threats to comparability when comparing quality at the international level, this is mainly related to data collection. Stolar²⁵ showed that pediatric surgeons are unable to generate sufficient direct financial resources to support their employment and practice operational expenses. Outcomes of the other studies can be found in Table 2.

PB Impact

One article identified improvements in the diagnosis of the patient and provision of care related to participating in the benchmark, for example improvements in the preoperative histology and radiotherapy after mastectomy¹⁷. Three articles identified suggestions for improvements based on the benchmark^{21, 23, 25}, in the provision of care for instance on the use of opiates at the end of life¹⁸ and improvements on the organizational level such as the decrease of the frequency of hospital visits, lead times and costs²⁵. For other improvements see Table 2.

PB Success factors

One study identified success factors. According to Brucker¹⁷ a success factor within their project was the fact that participation was voluntary and all the data was handled anonymous.

Table 2 Summary of the analysis of the pathway benchmarking projects

Author	Study design	Benchmarking model and/or steps	Indicators	Outcomes	Impact (improvements/ improvement suggestions)	Success factors
Brucker [17]	Prospective interventional multi-centre feasibility study.	Partner: Industry Content: Performance Purpose: Collaborative Independent, scientific benchmarking system. Nine guideline-based quality targets serving as rate-based QIs (Quality Indicators) were initially defined, reviewed annually and modified or expanded accordingly. QI changes over time were analyzed descriptively	Quality outcome indicators derived from clinically relevant parameters.	The results from this study provide proof of concept for the feasibility of a novel, voluntary, nationwide system for benchmarking the quality of BC care	Marked QI (Quality Indicators) increases indicate improved quality of BC care.	The project was voluntary and all data was anonymized.
Chung [18]	Multi comparisons study and the development of core measures for colorectal cancer including a modified Delphi method.	N.A.	Quantitative structure, process and outcome indicators	Developing core measures for cancer care was a first step to achieving standardized measures for external monitoring, as well as for providing feedback and serving as benchmarks for cancer care quality improvement.	N.A.	N.A.
Hermann [19]	Multi comparisons study and indicator consensus development process (with elements of the Delphi method).	Partner: Industry/ Global Content: Performance/ Process Purpose: Collaborative Development of indicators for benchmarking.	Process and outcome indicators.	The benchmark was not performed, indicators were developed for a possible benchmark.	N.A.	N.A.
Mainz [20]	Multi comparisons study and the development of indicators based on consensus of a working group	N.A. The results that are available for the prioritized quality indicators cannot really be used for true comparisons and benchmarking	Outcome indicators	A major difference between the Nordic countries has been identified with regard for 5 years survival for prostate cancer.	N.A.	N.A.

Table 2 Summary of the analysis of the pathway benchmarking projects (continued)

Author	Study design	Benchmarking model and/or steps	Indicators	Outcomes	Impact (improvements/ improvement suggestions)	Success factors
Miransky [21]	Multi comparisons study with stakeholder consensus methods. Use of a specialized database for benchmarking outcomes for cancer patients. Conference calls and joint meetings between comprehensive cancer centers and possible benchmark vendors were used to develop this benchmarking database.	Partner: Industry Content: Performance Purpose: Collaborative	Development of a database containing outcome indicators. Benchmarking clinical outcomes and patient	The various databases developed by the collaborative provided the tools through which the group accomplished its goals.	Each consortium member is expected to participate in one quality improvement initiative annually	N.A.
Roberts [22]	Multi comparisons study on staffing and inpatient data at hospices. Study design drew extensively from a UK-wide nursing study (The UK Best practice Nursing Database).	N.A.	Mixture of indicators, both qualitative and quantitative and process and outcome indicators	A broader NHS ward data system, was successfully converted for hospice use. The resultant hospice and palliative care ward data show that, compared to NHS palliative care wards, charitable hospices: (i) look after fewer patients; but generate greater workloads owing to higher patient-dependency and acuity scores; (ii) are much better staffed; and (iii) achieve higher service-quality scores.	N.A.	N.A.

Table 2 Summary of the analysis of the pathway benchmarking projects (continued)

Author	Study design	Benchmarking model and/or steps	Indicators	Outcomes	Impact (improvements/ improvement suggestions)	Success factors
Setoguchi [23]	Retrospective and prospective cohort study.	<p>Partner: Industry</p> <p>Content: Performance measures for the quality of end-of-life cancer care previously developed by Earle et al. New measures were defined for the use of opiate analgesia, which included the proportion of patients who received an outpatient prescription for a long-acting opiate; a short-acting or a long-acting opiate; or both a short acting and a long-acting opiate.</p>	<p>Outcome indicators</p>	Retrospective and prospective measures, including a new measure of the use of opiate analgesia, identified similar physician and hospital patterns of end-of-life care.	Findings suggest that the use of opiates at the end of life can be improved	N.A.
Stewart [24]	Multi comparisons study (clinical productivity and other characteristics of oncology physicians). Data collection by survey	<p>Partner: Industry</p> <p>Content: Performance benchmarks. The clinical characteristics were reviewed of oncology physicians practicing in 13 major academic cancer institutions.</p>	<p>Outcome productivity indicators</p>	Specific clinical productivity targets for academic oncologists were identified. A methodology for analyzing potential factors associated with clinical productivity and developing clinical productivity targets specific for physicians with a mix of research, administrative, teaching, and clinical salary support.	N.A.	N.A.

Table 2 Summary of the analysis of the pathway benchmarking projects (continued)

Author	Study design	Benchmarking model and/or steps	Indicators	Outcomes	Impact (improvements/ improvement suggestions)	Success factors
Stolar [25]	Multi comparisons study using a non-searchable anonymous data capture form through SurveyMonkey. Feedback from stakeholders and availability of information was used to develop indicators. A final questionnaire, containing 17 questions, was send to thirty pediatric surgery practices.	N.A.	Quantitative outcome indicators	A review of the clinical revenue performance of the practice illustrates that pediatric surgeons are unable to generate sufficient direct financial resources to support their employment and practice operational expenses.	The value of the services must accrue to a second party	N.A.
Van Vliet [26]	A retrospective comparative benchmark study with a mixed-method design	Partner: Industry/ Global Content: Strategic Purpose: Collaborative The method comprised of 6 steps: (1) operational focus; (2) autonomous work cell; (3) physical lay-out of resources; (4) multi-skilled team; (5) pull planning and (6) elimination of wastes.	N/A	The environmental context and operational focus primarily influenced process design of the cataract pathways.	When pressed to further optimize their processes, hospitals can use these systematic benchmarking data to decrease the frequency of hospital visits, lead times and costs.	N.A.

Table 2 Summary of the analysis of the pathway benchmarking projects (continued)

Author	Study design	Benchmarking model and/or steps	Indicators	Outcomes	Impact (improvements/ improvement suggestions)	Success factors
Wallwiener [27]	Review of existing literature/data.	<p>Partner: Industry</p> <p>Content: Performance</p> <p>Purpose: Collaborative</p> <p>Phase 1: Benchmarking; Phase 1a: proof of principle; Develop quality indicators; Phase 1b: analysis for a single specific specialty; to demonstrate the feasibility of subgroup analysis.</p> <p>Phase 2: certification of breast centres; to implement a quality management system to assess structural, process and outcome quality.</p> <p>Phase 3: nationwide implementation of certified breast centres.</p>	Structural and process indicators	<p>The voluntary benchmarking programme has gained wide acceptance among DKG/DGS-certified breast centres.</p> <p>The goal of establishing a nationwide network of certified breast centres in Germany can be considered largely achieved.</p>	Improvements in surrogate parameters as represented by structural and process quality indicators suggest that outcome quality is improving.	N.A.
Wesselman [28]	Review of existing literature/data. Analysis of existing benchmarking reports of cancer centers.	<p>Partner: Industry</p> <p>Content: Performance</p> <p>Purpose: Collaborative</p> <p>Analysis of benchmarking reports by the certified centers with the OnkoZert data which reflects the centers' reference results over a period of 3 years.</p> <p>The data for these reports are collected by the centers using an electronic questionnaire and are submitted to OnkoZert. (an independent institute that organizes the auditing procedure on behalf of the DKG)</p>	Respective and guideline-based outcome indicators	<p>The present analysis of the results, together with the centers' statements and the auditors' reports, shows that most of the targets for indicator figures are being better met over the course of time.</p>	There is a clear potential for improvement and the centers are verifiably addressing this.	N.A.

II Institutional benchmarking (IB)

IB Study design

In the two articles by de Korne^{3, 31} mixed methods were used to develop an evaluation frame for benchmarking studies in eye-hospitals. Barr et al.²⁹ used the National Practice Benchmark to collect data on Oncology Practice Trends. Brann³⁰ developed forums for benchmarking child and youth mental-health. Van Lent et al.⁶ conducted three independent international benchmarking studies on operations management of comprehensive cancer centers and chemotherapy day units. Schwappach³² used a pre–post design in two measurement cycles, before and after implementation of improvement activities at emergency departments. Shaw³³ used a questionnaire with 10 questions to collect data on pediatric emergency departments. More information on study design can be found in Table 3.

IB Benchmark model

Characterizing the benchmark models and/or steps with the scheme by Fong¹⁶ it can be seen that all studies used partners from the industry, in two studies these partners were global. Two articles benchmarked performance^{6, 31} while two other articles benchmarked both processes as performance^{3, 32} and one article reported the benchmarking of performance and strategies²⁹. More detailed information on the benchmark models can be found in Table 3.

IB Indicators

Most of the studies used outcome indicators (n=6)^{3, 6, 30-33}. Schwappach et al.³² for example used indicators to evaluate speed and accuracy of patient assessment, and patients' experiences with care by emergency departments. Van Lent⁶ described the use of indicators that differentiated between the organizational divisions of cancer centers such as diagnostics, radiotherapy and research. Brann³⁰ used Key Performance Indicators such as 28-day readmissions to inpatient settings, and cost per 3-month community care period.

IB Outcomes

Different outcomes were mentioned in the study by de Korne³ and on different aspects of operations management by van Lent⁶. However van Lent also showed that the results on the feasibility of benchmarking as a tool to improve hospital processes are mixed. The National Practice Benchmark (NPB)²⁹ demonstrated that the adaptation of oncology practices is moving toward gains in efficiency. Outcomes of the study by Schwappach³² showed that improvements in the reports provided by patients were mainly demonstrated in structures of care provision and perceived humanity. Shaw³³ showed that benchmarking of staffing and performance indicators by directors yields important administrative data. Brann et al.³⁰

presented that benchmarking has the potential to illuminate intra- and inter-organizational performance.

IB Improvements

Improvements mentioned due to participating in the benchmark (Table 3) were a successful improvement project⁶ leading to a 24% increase in bed utilization and a 12% increase in productivity in cancer centers and investments in Emergency Department (ED) structures, professional education and improvement of the organization of care³¹.

IB Success factors

Almost all institutional benchmarking articles identified success factors (n=7). Frequently mentioned factors were commitment of management^{6, 30} and the development of good indicators^{3, 6, 31}.

Table 3 Summary of the analysis of the institutional benchmarking projects

Author	Study design	Benchmarking model and/or steps	Indicators	Outcome	Impact (improvements/ improvement suggestions)	Success factors
Barr [29]	Multi comparisons study using the National Practice Benchmark.	<p>Partner: Industry</p> <p>Content: Performance/ Strategic</p> <p>Purpose: Collaborative National Practice Benchmark survey</p>	N.A.	The National Practice Benchmark reveals a process of change that is reasonably orderly and predictable, and demonstrates that the adaptation of the oncology community is directional, moving toward gains in efficiency as assessed by a variety of measures.	N.A.	To make the survey more accessible, it was stratified into 2 sections (minimum data set and extra).
Brann [30]	Multi comparisons study in which representatives from child and adolescent mental health organizations used eight benchmarking forums to compare performance against relevant KPIs.	N.A.	Key performance indicators looking at outcomes in mental health	Benchmarking has the potential to illuminate intra- and inter-organizational performance.	N.A.	<ol style="list-style-type: none"> 1. Commitment of the management and securing resources. 2. Feeding back benchmarking data to data interpretation. clinical staff to maintain their motivation to the project. 3. Forums for participants to provide them with the opportunity to discuss the performance of their organisation and draw lessons from other organisations.

Table 3 Summary of the analysis of the institutional benchmarking projects (continued)

Author	Study design	Benchmarking model and/or steps	Indicators	Outcome	Impact (improvements/ improvement suggestions)	Success factors
De Korne [3]	Mixture of methods: a systematic literature review and semi-structured interviews. An evaluation frame (based on a systematic literature review) was applied longitudinally to a case study of nine eye hospitals that used a set of performance indicators for benchmarking.	Partner: Industry/ Global Content: Process/ Performance Purpose: Collaborative 4P model : 1) the purposes of benchmarking; 2) the performance indicators used; 3) the participating organizations; and 4) the organizations' performance management systems.	Performance outcome indicators	The benchmarking indicators were mostly used to initiate and to facilitate discussions about management strategies. The eye hospitals in this study were not successful in reaching the goal of quantifying performance gaps or identifying best practices.	Indicators for benchmarking were not incorporated in management system in any of the hospitals, nor were results discussed with employees; only the strategic level was involved.	Performance indicators should; 1. Represent strategically important items; 2.the indicators have to be specific, measurable, acceptable, achievable, realistic, relevant, and timely (SMART); 3. Data have to be converted into measurable quantities; 4. the indicator information has to be comparable to those of other organizations; 5. selected indicators must be relevant to the benchmarking purposes; 6. the indicators should have validity with respect to performance and participants and should also discriminate.
De Korne [27]	Mixture of methods: quantitative analysis included (i) analysis of fiscal year 2009 benchmarking performance data and (ii) evaluation of multiple cases by applying an evaluation frame abstracted from the literature to five U.S. eye hospitals that used a set of 10 indicators for efficiency benchmarking. Qualitative analysis of interviews, document analyses, and questionnaires.	Partner: Industry Content: Performance Purpose: Collaborative 4P model : 1) the purposes of benchmarking; 2) the performance indicators used; 3) the participating organizations; and 4) the organizations' performance management systems.	Efficiency outcome indicators	The benchmark initiative fulfilled many of its purposes, namely, identifying performance gaps, implementing best practices, and stimulating exchange of knowledge.	Case studies showed that, to realize long-term efforts, broader cooperation is necessary.	1. the 4P model suggests that reliable and comparable indicators are a precondition for a successful benchmark, 2. case studies suggest that the development process is an important part of benchmarking. 3. homogeneity in language, reimbursement systems, and administrations

Table 3 Summary of the analysis of the institutional benchmarking projects (continued)

Author	Study design	Benchmarking model and/or steps	Indicators	Outcome	Impact (improvements/ improvement suggestions)	Success factors
Schwappach [28]	Prospective and retrospective mixed methods: Questionnaires, Demographic, clinical, and performance data collected via specific data sheets; systematic data controlling.	<p>Partner: Industry</p> <p>Content: Process/ Performance</p> <p>Purpose: Collaborative EMERGE:</p> <p>(1) selection of interested hospitals, participating on a voluntary basis; (2) joint development of a set of clinical performance indicators agreed upon by all parties; (3) establishment of a measurement system, development of measurement tools and design of data collection instruments; (4) data collection in a first measurement cycle; (5) benchmarking of results and definition of shared, quantitative targets; (5) initialization of hospital-specific improvement activities; (6) data collection in a second measurement cycle; and (7) benchmarking of results.</p>	<p>Outcome</p> <p>Indicator set including two main components: objective measures that evaluate clinical performance in terms of speed and accuracy of patient assessment, and patients' experiences with care provided by Eds.</p>	<p>Concordance of prospective and retrospective assignments to one of three urgency categories improved significantly by 1%, and both under- and over-prioritization, were reduced. Significant improvements in the reports provided by patients were achieved and were mainly demonstrated in structures of care provision and perceived humanity.</p>	<p>A number of improvement activities were initiated in individual hospitals covering a wide range of targets, from investment in ED structures to professional education and organization of care.</p>	<p>Interpretation of results should be guided by a culture of organisational learning rather than individual blame.</p>

Table 3 Summary of the analysis of the institutional benchmarking projects (continued)

Author	Study design	Benchmarking model and/or steps	Indicators	Outcome	Impact (improvements/ improvement suggestions)	Success factors
Shaw[29]	Multi comparisons study with the use of questionnaire containing ten questions.	N.A.	10 'questions' regarding ED patient utilization, wait times, services, and attending physician staffing of the nation's PEDs. Indicators qualified as outcome indicators.	Benchmarking of PEM staffing and performance indicators by PEM directors yields important administrative data. PEDs have higher census and admission rates compared with information from all EDs, while their attending staffing, wait times, and rate of patients who leave without being seen are comparable to those of general EDs.	In larger departments, the opening of fast tracks during high census times has allowed for shorter disposition of lower acuity patients with good success, this has been recommended as one of the solutions to better ED throughput.	N.A.

Table 3 Summary of the analysis of the institutional benchmarking projects (continued)

Author	Study design	Benchmarking model and/or steps	Indicators	Outcome	Impact (improvements/ improvement suggestions)	Success factors
Van Lent [6]	Multi comparisons study internationally benchmarking operations management in cancer centres.	<p>Partner: Industry/ Global</p> <p>Content: Performance</p> <p>Purpose: Collaborative. Spendolini's method and a new '13' step:</p> <ol style="list-style-type: none"> 1. Determine what to benchmark; 2. Form a benchmarking team; 3. Choose benchmarking partners; 4. Define and verify the main characteristics of the partners; 5. Identify stakeholders; 6. Construct a framework to structure the indicators; 7. Develop relevant and comparable indicators; 8. Stakeholders select indicators; 9. Measure the set of performance indicators; 10. Analyze performance differences; 11. Take action: results were presented in a report and recommendations were given; 12. Develop improvement plans; 13. Implement the improvement plans 	<p>Outcome indicators containing a numerator and a de-numerator</p> <p>The selected indicators distinguished between the total organization level, diagnostics, surgery, medication related treatments, radiotherapy and research.</p>	<p>The results on the feasibility of benchmarking as a tool to improved hospital processes are mixed. Success factors identified are a well-defined and small project scope, partner selection based on clear criteria, stakeholder involvement, simple and well-structured indicators, and analysis of both the process and its results.</p>	<p>All multiple case studies provided areas for improvement and one case study presented the results of a successful improvement project based on international benchmarking.</p>	<ol style="list-style-type: none"> 1. Internal stakeholders must be convinced that others might have developed solutions for problems that can be translated to their own settings. 2. Management must reserve sufficient resources for the total benchmarks. 3. Limit the scope to a well-defined problem. 4. Define criteria to verify the comparability of benchmarking partners based on subjects and process. 5. Construct a format that enables a structured comparison. 6. Use both quantitative and qualitative data for measurement. 7. Involve stakeholders to gain consensus about the indicators. 8. Keep indicators simple so that enough time can be spent on the analysis of the underlying processes. 9. For indicators showing a large annual variation in outcomes, measurement over a number of years should be considered. 10. Adapt the identified better working methods so that they comply with other practices in the organisation.

III Benchmarking evaluation/methodology (BEM)

BEM Study Design

Ellershaw³⁴ assessed the usefulness of benchmarking using the Liverpool Care Pathway in acute hospitals in England with the use of a questionnaire. Ellis³⁵ performed a review of benchmarking literature. Matykiewicz³⁶ evaluated the Essence of Care as a benchmarking tool with a case study approach and qualitative methods.

Profit³⁷ used a review of the scientific literature on composite indicator development, health systems, and quality measurement in pediatric healthcare. More information on study design can be found in Table 4.

BEM Benchmark model steps

Three studies describe a benchmark model. They all describe industry partners and process benchmarking (see Table 4).

BEM Indicators

One article described the use of indicators, though very minimally. Matykiewicz³⁶ described benchmarking against best practice indicators, but specific indicators are not mentioned. Profit et al.³⁷ developed a model for the development of indicators of quality of care.

BEM Outcomes

The study by Ellershaw³⁴ displayed that almost three quarters of respondents in the hospital sector felt that participation in the benchmark had had a direct impact on the delivery of care. The outcomes of the study by Ellis³⁵ was that Essence of Care benchmarking is a sophisticated clinical practice benchmarking approach which needs to be accepted as an integral part of health service benchmarking activity. Matykiewicz³⁶ showed that whilst raising awareness is relatively straightforward, putting Essence of Care into practice is more difficult. Profit et al.³⁷ concluded that the framework they presented offers researchers an explicit path to composite indicator development.

BEM Improvements

Improvements due to the benchmark exercise that were identified included specific improvements in levels of communication between health professionals and relatives, within multidisciplinary teams and across sectors³⁴ and that through self-assessment against best practice problems could be identified and solved³⁶.

BEM Success factors

Three articles mentioned success factors, both Ellershaw³⁴ and Matykiewicz³⁶ mentioned the organization of a workshop, while Ellis³⁵ identified reciprocity as an important factor for success.

Table 4 Summary of the analysis of the benchmarking evaluation/methodology studies

Author	Study design	Benchmarking model and/or steps	Indicators	Outcome	Impact (improvements/ improvement suggestions)	Success factors
Ellershaw [34]	Survey to assess the usefulness of benchmarking with the Liverpool Care Pathway	Partner: Industry Content: Process Purpose: Collaborative	N.A.	Whilst almost three quarters of the respondents in the hospital sector felt that participation in the benchmark had had a direct impact on the delivery of care, only around a third in the other two sectors felt the same(hospice and community).	Specific improvements in levels of communication between health professionals and relatives, within multidisciplinary teams and across sectors occurred as a result of participation in the benchmarking exercise.	Holding a workshop for participants to reflect on data, enhances understanding and learn from others.
Ellis [35]	Literature review to encourage the acceptance and use of Essence of Care as a new benchmarking approach.	Partner: Industry Content: Process Purpose: Collaborative Evaluation of a benchmark with the use of Essence of Care	N.A.	Essence of Care benchmarking is a sophisticated clinical practice benchmarking approach which needs to be accepted as an integral part of health service benchmarking activity to support improvement in the quality of patient care and experiences.	N.A.	1. Reciprocity

Table 4 Summary of the analysis of the benchmarking evaluation/methodology studies (continued)

Author	Study design	Benchmarking model and/or steps	Indicators	Outcome	Impact (improvements/improvement suggestions)	Success factors
Matytkiewicz [36]	Case study approach and qualitative methods namely interviews and focus groups	<p>Partner: Industry</p> <p>Content: Process</p> <p>Purpose: Collaborative</p> <p>The Essence of Care process includes: 1) Agree best practice; 2) Assess clinical areas against best practice; 3) Produce/implement action plan aimed at achieving best practice; 4) Review achievement of best practice; 5) Disseminate improvement and/or review action plan; 6) Agree best practice.</p>	Best practice indicators	Whilst raising awareness is relatively straightforward, putting Essence of Care into practice is more difficult to achieve, especially when happening at a time of significant organizational change.	Through self-assessment against the best practice indicators, a problem was identified which, if not dealt with, could have escalated to a more serious situation. The manager saw this as an opportunity to learn from mistakes and initiated a service review that has since resulted in the service being redesigned.	1. Workshops (successful in raising awareness, help people to understand how to apply the benchmarking process in practice)
Profit [37]	Literature review on composite indicator development, health systems, and quality measurement in the pediatric healthcare setting.	N.A.	No indicators were mentioned, however a conceptual framework to develop comprehensive, robust, and transparent composite indicators of pediatric care quality was developed. The model proposed identifying structural, process, and outcome metrics for each of the Institute of Medicine's six domains of quality.	The combination of performance metric development methodology with the quality matrix framework may result in a unique approach for quality measurement that is fair, scientifically sound, and promotes the all-important provider buy-in. The framework presented offers researchers a path to composite indicator development.	N.A.	N.A.

IV Benchmark using patient registry data

The only benchmark study³⁸ using patient registry data originated in oncology practice in the US (see Table 5). For this study National Cancer Database (NCDB) reports from the Electronic Quality Improvement Packet (e-QUIP) were reviewed ensuring all network facilities are in compliance with specific outcome benchmarks. Outcome indicators such as local adherence to standard-of-care guidelines were used. A review of the e-QUIP-breast study at Carolinas Medical Center (CMC) showed that treatment methods could be improved. No improvements were reported. At CMC, the registry has been a key instrument in program improvement in meeting standards in the care of breast and colon cancer by benchmarking against state and national registry data.

Table 5 Summary of the analysis of Benchmark study using patient registry data

Author	Study design	Benchmarking model and/or steps	Indicators	Outcome	Impact (improvements/ improvement suggestions)	Success factors
Greene [38]	Development of a cancer committee; review of the NCDB reports from the Electronic Quality Improvement Packet (e-QUIP) and CP3R ensuring all network facilities are in compliance with specific outcome benchmarks.	N.A.	Outcome indicators	In addition to a role in benchmarking, registry data may be used to assist in establishing new research protocols and in determining market share by the hospital administration. The registry identified several issues which included the lack of physician office contact information, and time lapse for treatment completion.	Two potential issues were identified. With instruction for the pathologists and surgeons regarding these issues, this rate is expected to improve.	N.A.

Discussion

There is a growing need for healthcare providers to focus on performance. Benchmarking is a common and supposedly effective method for measuring and analyzing performance². Benchmarking in specialty hospitals developed from the quantitative measurement of performance to the qualitative measurement and achievement of best practice³⁹.

In order to inform the development of a benchmark tool for comprehensive cancer care (the BENCH-CAN project) we assessed the study characteristics of benchmarking projects in specialty hospitals, avoid duplication and identified the success factors to benchmarking of specialty hospitals. This scoping review identified 24 papers that met the selection criteria which were allocated to one of four categories. Regarding our first two research objectives: (i) provide an overview of research on benchmarking in specialty hospitals and care pathways, (ii) describe study characteristics such as method, setting, models/frameworks, and outcomes, we reviewed the first three categories against a common set of five issues that shape the following discussion. The fourth category (*Benchmark using patient registry data*) had only a single paper so could not be appraised in the same way.

I. Area of practice

In terms of study settings, we were interested in the areas where benchmarking would be most frequently used. Our review identified seven types of specialty hospitals. Most studies were set in oncology specialty hospitals. The majority (n=12) of the articles described projects in which part of a specialty hospital or care pathway was benchmarked. This could be due to the fact that one of the success factors of a benchmarking project defined by van Lent et al.⁶ is the development of a manageable-sized project scope. This can be an identified problem in a department or unit (part of a specialty hospital), or a small process that involves several departments (care pathway).

II. Study design

Looking at the different study designs, both quantitative as qualitative methods can be found. All institutional articles except Schwappach³¹ (retrospective and prospective) made use of a prospective research design while most pathway articles used a retrospective multi-comparison design. Stakeholders often played an important role in the benchmarking process and consensus methods such as the Delphi method were frequently used to develop the benchmark indicators.

III. Benchmark model

Fifteen articles described a benchmark model/steps. All studies that described a benchmarking study made use of partners from the industry, in four articles these were from different countries, e.g. global. Most benchmarks were on performance (n=8), others used a combination of performance and process benchmarking (n=3) or performance and strategic benchmarking (n=1). Three studies described a process benchmark and one benchmarking on strategies. The classification scheme was not developed for healthcare benchmarking specifically. This is shown by the definition of competitor. Some of the described partners in the benchmarking studies fit the first part of the definition: in business, a company in the same industry or a similar industry which offers a similar product or service⁴⁰ for example breast cancer centers or eye hospitals. However there is not always competition between these centers (second part definition). A healthcare specific scheme for benchmarking models would be preferred, this was however not found.

In some cases, a model has been uniquely developed—possibly using field expertise— for performing a particular type of benchmarking, which means that there was no evidence of the usability of the model beforehand. In their article on 'Benchmarking the benchmarking models' Anand and Kodali¹⁵ however identify and recommend some common features of benchmarking models. Their cursory review of different benchmarking process models revealed that the most common steps are: "identify the benchmarking subject" and "identify benchmarking partners"¹⁵. The purpose of the benchmarking process models should be to describe the steps that should be carried out while performing benchmarking. Anand and Kodali¹⁵ recommend that a benchmark model should be clear and basic, emphasizing logical planning and organization and establishing a protocol of behaviors and outcomes. Looking at the models described in this review it shows that only five articles describe models that have all the features described by Anand and Kodali^{3,6, 31, 32, 36}.

IV. Registry

The article about the use of a registry differed in the sense that no benchmark model or benchmarking steps were described. Instead it focused on the usefulness of using a registry for benchmarking. According to Greene et al.³⁸ a registry is a valuable tool for evaluating quality benchmarks in cancer care. Sousa et al.⁴¹ showed the general demands for accountability, transparency and quality improvement make the wider development, implementation and use of national quality registries for benchmarking, inevitable. Based on this we had expected to find more articles describing the use of the registry for benchmarking, these were however not identified through our search.

V. Indicators

Currently, it seems that the development of indicators for benchmarking is the main focus of most benchmarking studies. The importance of indicator development is highlighted by Groene et al.⁴² who identified 11 national indicator development projects. Papers included in this study showed a wide array of approaches to define and select indicators to be used in the projects, such as interviews, focus groups, literature reviews and consensus surveys (Delphi method and others).

A review by Nolte⁴³ shows that there is an ongoing debate about the usefulness of process versus outcome indicators to evaluate healthcare quality. In most papers included in this study outcome indicators were used, especially in the pathway benchmarking papers. This seems contradictory to findings by Mant⁴⁴ who noted that the relevance of outcome measures is likely to increase towards macro-level assessments of quality, while at the organizational or team level, process measures will become more useful. Based on this one would expect the use of process indicators for especially the pathway articles.

Benchmarking as a tool for quality improvement and success factors

Regarding our third objective: “verify the quality of benchmarking as a tool to improve quality in specialty hospitals and identify success factors” we found the following. Only six articles described improvements related to the benchmark. Specific improvements were described in the level of communication between health professionals and relatives, within multidisciplinary teams and across sectors; service delivery and organization of care; and pathway development. Only three articles actually showed the improvement effects of doing a benchmark in practice. This could be linked to the fact that almost no benchmark model described a last step of evaluation of improvement plans as being part of the benchmark process. Brucker¹⁷ showed that nationwide external benchmarking of breast cancer care is feasible and successful. Van Lent⁶ however showed that the results on the feasibility of benchmarking as a tool to improve hospital processes were mixed. This makes it difficult to assess whether benchmarking is a useful tool for quality improvement in specialty hospitals.

Within the pathway studies only one paper mentioned success factors, in contrast with almost all institutional and benchmark evaluation- and methodology papers. Based on our review we’ve come up with a list of success factors for benchmarking specialty hospitals or care pathways (box 2). One article exploring the benchmarking of Comprehensive Cancer Centres⁶ produced a detailed list of success factors for benchmarking project (see box 2), such as a well-defined and small project scope and partner selection based on clear criteria.

This might be easier for specialty hospitals due to the specific focus and characteristics than for general hospitals. Organizing a meeting for participants, either before or after the audit visits, was mentioned as a success factor^{34, 36}. Those workshops or forums provided the opportunity for participants to network with other organizations, discuss the meaning of data and share ideas for quality improvements and best practices. Especially the development of indicators was mentioned often, corresponding to our earlier observation about the emphasis that is put on this issue.

Although this scoping review shows that the included studies seem to focus on indicator development rather than the implementation and evaluation of benchmarking, the characteristics described (especially the models) can be used as a basis for future research. Researchers, policy makers or other actors that wish to develop benchmarking projects for specialty hospitals should learn lessons from previous projects to prevent the reinvention of the wheel. The studies in this review showed that ensuring the commitment to the project by the management team of hospitals participating and the allocation of sufficient resources for the completion of the project is paramount to the development of a benchmarking exercise. The information found in combination with the provided success factors may increase the chance that benchmarking results in improved performance in specialty hospitals like cancer centers in the future.

1. Voluntary participation
2. Anonymous participation
3. Internal stakeholders must be convinced that others might have developed solutions for problems of the underlying processes that can be translated to their own settings.
4. Verify homogeneity participant group to ensure the comparability of benchmarking partners
5. Ensure commitment of the management and secure resources
6. Limit the scope of the project to a well-defined problem
7. Involve stakeholders to gain consensus about the indicators
8. Develop indicators that are specific, measurable, acceptable, achievable, realistic, relevant, and timely (SMART)
9. Use simple indicators so that enough time can be spent on the analysis
10. Measure both qualitative and quantitative data
11. Stratify survey into minimum data set and additional extra's
12. For indicators showing a large annual variation in outcomes, measurement over a number of years should be considered
13. Feed benchmarking data back to clinical staff to maintain their motivation to the project
14. Organize forums and workshops for participants to discuss performance of their organization and learn from other organizations
15. Convert data into measurable quantities
16. Homogeneity in language, reimbursement systems, and administrations
17. Interpretation of results should be guided by a culture of organisational learning rather than individual blame.

Box 2 Success factors benchmarking projects specialty hospitals and pathways

Limitations

A potential limitation is that by searching the titles and abstracts we may have missed relevant papers. The articles included in this review were not appraised for their scientific rigor, as scoping reviews do not typically include critical appraisals of the evidence. In

deciding to summarize and report the overall findings without the scrutiny of a formal appraisal, we recognize that our results speak to the extent of the setting and model of the benchmark study rather than provide the reader with support for the effectiveness of benchmarking.

Conclusion

Benchmarking in specialty hospitals developed from simple data comparison to quantitative measurement of performance, qualitative measurement and achievement of best practice. Based on this review it seems however that benchmarking in specialty hospitals is still in development. Benchmarking seems to be most reported up on and possibly developed in the field of oncology and eye hospitals, however most studies do not describe a structured benchmarking method or a model that can be used repeatable. Based on our study we identified a list of success factors for benchmarking specialty hospitals. Developing 'good' indicators was mentioned frequently as a success factor. Within the included papers there seems to be a focus on indicator development rather than measuring performances, which is an indication of development rather than implementation. Further research is needed to ensure that benchmarking in specialty hospitals fulfills its objective, to improve the performance of healthcare facilities. Researchers wishing –as a next step- to evaluate the effectiveness of benchmarking to improve quality in specialty hospitals, should conduct evaluations using robust and structured designs, focusing on outcomes of the benchmark and preferably do a follow up to check whether improvement plans were implemented.

References

1. Plsek P and Greenhalgh T. Complexity science: The challenge of complexity in health care. *BMJ* 2001;323:625.
2. Leape L, Berwick D, Clancy C, et al. Transforming healthcare: a safety imperative. *Qual Saf Health Care* 2009;18:424–28.
3. de Korne D, Sol J, van Wijngaarde J, van Vliet EJ, Custers T, Cubbon M, et al. Evaluation of an international benchmarking initiative in nine eye hospitals. *Health Care Manage R* 2010;35:23-35.
4. Campbell S, Braspenning J, Hutchinson, et al. Research methods used in developing and applying quality indicators in primary care. *BMJ* 2003;326:816-19.
5. Joint Commission: Benchmarking in Health Care. Joint Commission Resources; 2011.
6. van Lent W, de Beer R, van Harten W. International benchmarking of specialty hospitals. A series of case studies on comprehensive cancer centres. *BMC Health Serv Res* 2010;10:253.
7. Pringle M, Wilson T, Grol R. Measuring “goodness” in individuals and healthcare systems. *BMJ* 2002;325:704-7.
8. BenchCan <http://www.oeci.eu/Benchcan> (2013). Accessed 20 Feb 2015.
9. Schneider JE, Miller TR, Ohsfeldt RL, Morrisey MA, Zelner BA, Li P. The Economics of Specialty Hospitals. *Med Care Res Rev* 2008;65(5): 531.
10. Porter ME, Teisberg EO. Redefining health care: creating value-based competition on results. Boston, Massachusetts: Harvard Business School Press; 2006.
11. Constand MK, MacDermid JC, Dal Bello-Haas V, Law M. Scoping review of patient-centered care approaches in healthcare *BMC Health Serv Res* 2014;14:271
12. Levac D, Colquhoun H, O’Brien KK. Scoping studies: advancing the methodology. *Implement Sci* 2010;5:69
13. Arksey H, O’Malley L. Scoping studies: Towards a Methodological Framework. *Int J Soc Res Methodol* 2005, 8:19-32.
14. Saggese S, Sarto F, Cuccurullo C. Evolution of the Debate on Control Enhancing Mechanisms: A Systematic Review and Bibliometric Analysis. *IJMR* 2015;00:1-23.
15. Anand G & Kodali R. Benchmarking the benchmarking models. *BIJ* 2008;15:257 – 91.
16. Fong SW, Cheng EWL and Ho DCK. Benchmarking: a general reading for management practitioners. *Manage Decis* 1998;36:407-18.
17. Brucker S, Schumacher C, Sohn C, Rezai M, Bamberg, Wallwiener D, et al. The Steering Committee: Benchmarking the quality of breast cancer care in a nationwide voluntary system: the first five-year results (2003–2007) from Germany as a proof of concept. *BMC Cancer* 2008;8:358.
18. Chung KP, Chang YJ, Lai MS, Nien-Chen Kuo R, Cheng SH, Chen LT, 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* 2010;10:27.
19. Hermann R, Mattke S, Someth D, et al. Quality indicators for international benchmarking of mental health care. *Int J Qual Health C* 2006;31–38.
20. Mainz J, Hjulsgaard M, Thorup Eriksen Og M, Burgaard J. National Benchmarking Between the Nordic Countries on the Quality of Care. *J Surg Oncol* 2009;99:505–507.

21. Miransky J. The Development of a Benchmarking System for a Cancer Patient Population. *J Nurs Care Qual* 2003;18:38-42.
22. Roberts D and Hurst K. Evaluating palliative care ward staffing using bed occupancy, patient dependency, staff activity, service quality and cost data. *Palliat Med* 2012;27:123-130.
23. Setoguchi S, Earle C, Glynn R, Stedman M, Polinski JM, Corcoran CP, et al. Comparison of Prospective and Retrospective Indicators of the Quality of End-of-Life Cancer Care. *J Clin Oncol* 2008;26:5671-78.
24. Stewart M, Wasserman R, Bloomfield C, Petersdorf S, Witherspoon RP, Appelbaum FR, et al. Benchmarks in Clinical Productivity: A National Comprehensive Cancer Network Survey. *J Oncol Pract* 2007;3:2-8.
25. Stolar C, Alapan A, Torres S. University pediatric surgery: benchmarking performance. *J Pediatr Surg* 2010;45:28-37.
26. van Vliet E, Sermeus Q, Kop L, Kop LM, Sol JCA, van Harten WH. Exploring the relation between process design and efficiency in high-volume cataract pathways from a lean thinking perspective. *Int J Qual Health C* 2011;23:83-93.
27. Wallwiener M, Brucker S, Wallwiener D. The Steering Committee: Multidisciplinary breast centres in Germany: a review and update of quality assurance through benchmarking and certification. *Arch Gynecol Obstet* 2012;285:1671-83.
28. Wesselman S, Winter A, Ferencz J, Seufferlein T, Post S. Documented quality of care in certified colorectal cancer centers in Germany: German Cancer Society benchmarking report for 2013. *Int J Colorectal Dis* 2014.
29. Barr T, Towle E. Oncology Practice Trends From the National Practice Benchmark. *J Oncol Pract* 2012;8:292-7.
30. Brann P, Walter G, Coombs T. Benchmarking child and adolescent mental health organizations. *Australas Psychiatry* 2011;19:125-32.
31. de Korne D, van Wijngaarden J, Sol J, Betz R, Thomas RC, Schein OD, et al. Hospital benchmarking: Are U.S. eye hospitals ready? *Health Care Manage Rev* 2012;37:187-198.
32. Schwappach D, Blaudszun A, Conen D, Ebner H, Eichler K, Hochreutener MA. 'Emerge': benchmarking of clinical performance and patients' experiences with emergency care in Switzerland. *Int J Qual Health C* 2003;15:473-85.
33. Shaw K, Ruddy R, Gorelick M. Pediatric emergency department directors' benchmarking survey: Fiscal year 2001. *Pediatr Emerg Care* 2003;19:143-7.
34. Ellershaw J, Gambles M, McGinchley T. Benchmarking: a useful tool for informing and improving care of the dying? *Support Care Cancer* 2008;16:813-9.
35. Ellis J. All inclusive benchmarking. *J Nurs Manag* 2006;14:377-83.
36. Matykiewicz L & Ashton D. Essence of Care benchmarking: putting it into practice. *BIJ*2005;12:467-81.
37. Profit J, Typpo K, Hysong S, Woodard LD, Kallen MA, Petersen LA. Improving benchmarking by using an explicit framework for the development of composite indicators: an example using pediatric quality of care. *Implement Sci* 2010;5.
38. Greene F, Gilkerson S, Tedder P, Smith K. The Role of the Hospital Registry in Achieving Outcome Benchmarks in Cancer Care. *J Surg Oncol* 2009;99:497-9.
39. Zairi M, Leonard P. *Practical Benchmarking. The Complete Guide*. Londen: Chapman & Hall, 1994

40. Business Dictionary: definition competitor. <http://www.businessdictionary.com/definition/competitor.html> Accessed 16 June 2015.
41. Sousa P, Bazely M, Johansson S, Wijk H. The use of national registries data in three European countries in order to improve health care quality. *Int J Health Care Qual Assur* 2006;19:551-60.
42. Groene O, Skau JKH, Frolich A. An international review of projects on hospital performance assessment. *Int J Qual Health Care* 2008;20:162 –71.
43. Nolte E. *International benchmarking of healthcare quality; A review of the literature*. Londen: RAND and London School of Hygiene and Tropical Medicine; 2010.
44. Mant J. Process versus outcome indicators in the assessment of quality of health care. *Int J Qual Health Care* 2001;13:475-80.

CHAPTER 3

Quality assessments for cancer centers in the European Union

Anke Wind
Abinaya Rajan
Wim van Harten

Abstract

Background: Cancer centers are pressured to deliver high-quality services that can be measured and improved, which has led to an increase of assessments in many countries. A critical area of quality improvement is to improve patient outcomes. An overview of existing assessments can help stakeholders (e.g. healthcare professionals, managers and policy makers) improve the quality of cancer research and care and lead to patient benefits. This paper presents key aspects of assessments undertaken by European cancer centers, such as: are assessments mandatory or voluntary? Do they focus on evaluating research, care or both? And are they international or national?

Methods: A survey was sent to 33 cancer centers in 28 European Union member states. Participants were asked to score the specifics for each assessment that they listed.

Results: Based on the responses from 19 cancer centers from 18 member states, we found 109 assessments. The numbers have steadily increased from the 1990s till 2015. Although, a majority of assessments are on patient-care aspects (n=45), it is unclear how many of those include assessing patient benefits. Only few assessments cover basic research. There is an increasing trend towards mixed assessments (i.e. combining research and patient-care aspects).

Conclusions: The need for assessments in cancer centers is increasing. To improve efforts in the quality of research and patient care and to prevent new assessments that “reinvent the wheel”, it is advised to start comparative research into the assessments that are likely to bring patient benefits and improve patient outcome. Do assessments provide consistent and reliable information that create added value for all key stakeholders?

Background

Cancer centers (CCs) in Europe, are located in complex organizational and regulatory environments and are increasingly under pressure to deliver high-quality services and be transparent about it¹. As a consequence of this, there is an increasing emphasis on quality and safety improvement initiatives². Patients and payers increasingly demand proof of guaranteed safety and quality of services. Cancer care activities lead to a steadily growing financial claim on national and regional health systems leading to concerns on sustainability and value for money, especially at a time of austerity measures and deficits in public budgets^{3,4}. This has led to additional need for transparency on quality matters and performance issues⁵.

Determining what is quality and safety of care is complex as it can reflect the combined perspectives of policy makers, purchasers, payers, healthcare professionals, researchers and patients⁶. The complexity of healthcare systems and the unpredictable nature of health care add to this difficulty⁷. Setting and applying clear performance standards through regulatory mechanisms, such as licensing, certification, and accreditation, is crucial to ensure patient safety⁸.

CCs go through several assessments on their performance and quality, assessments being defined as: "A system for evaluating performance, as in the delivery of services or the quality of products provided to consumers, customers, or patients"⁹. Its nomenclature extends to accreditation, certification, performance review, (performance) evaluation and others. This study uses the term assessments as it includes all of the above mentioned terms. So far an overview of the assessments on a European level does not exist. A recent study among Canadian Oncologists by Lim et al.¹⁰ shows that one of the reasons for them not participating in this type of Quality Improvement initiatives is the lack of knowledge about on-going initiatives. This example shows the relevance of obtaining an overview of assessments.

This article presents key findings from a survey that was conducted with CCs in the European Union. The goal was to obtain an overview of existing assessments in terms of whether they are: mandatory or voluntary; focused on evaluating research or patient care or both; regional, national and/or international. An example from the Netherlands¹¹ shows that hospitals spend between 40.1 to 82.3 million Euros on quality assessments in 2014. This study shows that much of the information gathered through these assessments is, however, recorded twice, inefficiently and is accompanied by bureaucracy. Unfortunately there is limited evidence on the added value of these (organizational) assessments for patient care or patient outcomes, primarily due to methodological issues related to limited

insight into the mechanisms through which these exert their effects. Though very relevant, that is not the object of this overview.

The rationale for this study was originally to provide input for the BENCH-CAN project¹². The BENCH-CAN project¹² aims at benchmarking comprehensive cancer and yield best practice example at eight European CCs in order to contribute to improvement of multidisciplinary patient treatment. One of the objectives of the BENCH-CAN project is: To collect, compare and align, by consensus formation, the standards, recommendations and accreditation criteria of comprehensive cancer care adopted in selected European countries representatives of different geographic areas (North-Western Europe; Southern Europe; Central-Eastern Europe). Because of the potential to inform decision makers about existing assessments so that they can take some steps towards regulating these as well as minimizing the related bureaucracy, it was decided to expand the study to other CCs than just the BENCH-CAN pilot sites. Organizations conducting these assessments and (also non EU) CCs can gain better understanding of what type of assessments are currently undertaken in view of growing interest in cooperation in international research consortia^{13,14}.

The context of European cancer centers

Assessments are contextual, and so, first there is a need to understand the type of health system in which the CCs operate. Health systems in the EU can be described in different ways. For this article, the typology developed by Rothgang et al.¹⁵ and Wendt et al.¹⁶ was used, which suggests four types of health systems: the National Health Service (NHS), National Health Insurance (NHI), Social Health Insurance (SHI) and the Etatist Social Health Insurance (ESHI). Three dimensions distinguish each of these systems: financing, service provision, and regulation¹⁷. According to this classification scheme each dimension can be dominated by state (government), societal (for example NGOs, consultancy agencies or research institutes), or private actors (see Figure 1). The US system has a mix of characteristics of those systems; however, unique about the US system in the world is the dominance of the private for profit actors in all three dimensions over the public sector (state/government and societal/ non-governmental)¹⁸.

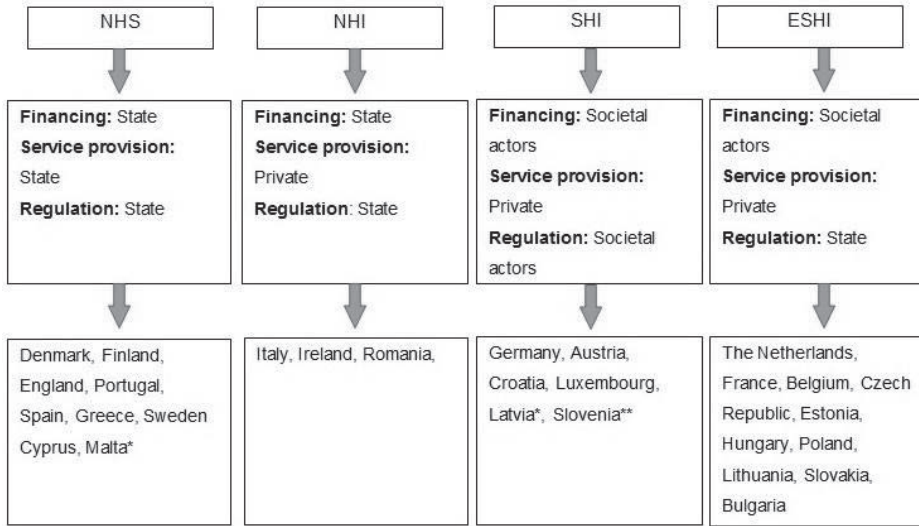


Figure 1. Overview of typology of health systems in the EU

Legend. * Malta and Latvia have mixed public/private service provision

** Slovenia conflicts with the logic of the RW typology as societal actors are in charge of regulation and financing, but service provision lies predominantly in the hands of state actors. Slovenia is, however, gradually evolving into a SHI.

Methods

Survey

A survey was sent initially to the BENCH-CAN pilot sites. After the decision to expand the study, the survey was sent to one cancer center in each of the EU member states with the exception of Belgium, Austria and the UK where 2 cancer centers were contacted. This was due to the lack of response within the given time-frame from the first contacted center. A second center was contacted in each of these countries. In total the survey was sent to 33 cancer centers in the 28 EU countries. Ethics committee approval was deemed irrelevant for this study. For some member states, CCs could not be easily identified and so, other organizations dealing with cancer care and/or research were contacted. CCs were identified through the European Society for Medical Oncology (ESMO)¹⁹ and the Organisation of European Cancer Institutes (OECI)²⁰. The survey was addressed to the lead administrative person in each institute. Participants were asked to describe several topics for each assessment that they listed: (i) the name of the assessment body (i.e. organization that performed the assessment); (ii) whether the body was public or private; (iii) if the assessment was mandatory or voluntary; (iv) the level (i.e. regional/national/international) at which the assessment was performed; (v) if the assessment focused on research, patient

care aspects or a mix of standards; (vi) the frequency of the assessment; (vii) if the assessment led to keeping/losing operating license and/or public funding and (viii) the year in which the assessment was first performed.

Data management and inclusion/exclusion criteria

3 Among the 28 EU member states in which CCs were asked to participate, data were received from 18 member states i.e. one cancer institution per member state (64%), with the exception of Italy (two cancer institutions). Not all surveys were filled out correctly, and some were missing data. So, a follow up was done by e-mail or phone with all respondents to clarify the answers. Two researchers inspected the data and excluded the listed assessments that did not fit the inclusion criteria. The inclusion criteria for the assessments were: the assessment had to assess cancer care, cancer research or a combination of both. All assessments that did not fit these criteria were excluded from the study. Eligible assessments were divided into three categories: clinical/patient care oriented assessments; research oriented assessments; and assessments that are oriented at a combination of care and research. Clinical/Patient care oriented assessments focus on the care delivered by cancer centers, Research oriented assessments focus solely on research performed at cancer centers, while combination oriented assessments focus on a comprehensive assessment of both the care delivered as the research performed in the CC (only applicable in centers where both activities are fully developed). A content analysis was performed. This method enables a more objective evaluation than comparing content based on the impressions of a reader and simplifies the detection of trends. This analysis was executed by converting the different items of the survey (public or private; mandatory or voluntary; regional/national/international; focused on research, patient care aspects or both; the frequency; if the assessment led to keeping/losing operating license and/or public funding and the year in which the assessment was first performed) into variables in excel. By dividing the data into the variables, a structured overview of assessment characteristics was obtained. This overview enabled the researchers to investigate trends in assessments and possible relationships between types of assessment and health systems. Two researchers independently examined the data to check for Inter Rater Reliability. The analysis of findings includes only programs that completed the survey. Validity of the data was verified by checking the assessment body in an online search and by asking the participating CCs to double check the data provided. The full list of included assessments was circulated amongst the respondents for final data validation.

Results

Nature and scope of assessments

Based on the responses, we found 109 known cancer related quality assessments in total in 19 EU member states. The majority of the assessments focus on patient-care aspects (n=45), such as waiting and throughput times, patient participation and patient satisfaction followed by the mixed assessments that focus on patient care as well as research aspects (n=37). In those mixed assessment especially organizational aspects of care and research such as multidisciplinary harmonization/integrated care and scientific interaction and integration receive emphasis, whereas pure research oriented assessments, which are the least in number (n=27), are directed towards research outcomes such as number of publications. The majority of patient care oriented assessments are reported to be mandatory. Mixed assessments are more voluntary.

The majority of assessments (n=62) is done at the national level (performed by a national government sponsored federal agencies or performed by national 'bodies' unaffiliated with governments but with assessment authority), followed by thirty-four assessments that are known to be operational at an international level (performed by international assessment agencies). Some assessments are implemented at a national level, but are also operational at an international level, these have been counted as national. There are only a handful of regional assessments (n=9) such as in Estonia and in Finland (see Table 1). Almost all mandatory assessments are national and are mainly related to keeping license and/or receiving public funding. In contrast, most voluntary assessments are international, and rather aim at quality improvement and are seldom directly tied to licensing or funding.

The majority of assessments (n=62) is done at the national level (performed by national government sponsored federal agencies or performed by national 'bodies' unaffiliated with governments but with assessment authority), followed by thirty-four assessments that are known to be operational at an international level (performed by international assessment agencies). Some assessments are implemented at a national level, but are also operational at an international level, these have been counted as national. There are only a handful of regional assessments (n=9) such as in Estonia and in Finland (see Table 2). Almost all mandatory assessments are national and are mainly related to keeping license and/or receiving public funding. In contrast, most voluntary assessments are international, and rather aim at quality improvement and are seldom directly tied to licensing or funding.

Table 1 Scope and nature of assessments

EU member state	Pure research assessments			Patient care assessments			Research + Patient care assessments				
	Mandatory	Average frequency	Voluntary	Mandatory	Average frequency	Voluntary	Mandatory	Average frequency	Voluntary	Average frequency	Programs operational
Austria	1	2 years	0	0	-	4	2,7 years	0	1	2 years	6
Croatia*	0	-	0	1	1 year	0	-	0	0	-	1
Czech Republic	1	-	1	0	Irregular	1	3 years	2	3	3 years	8
Denmark	0	-	1	0	1 year	0	-	0	0	-	1
Estonia	2	2 years	1	3	2,7 year	1	3 year	0	0	-	7
Finland	0	-	5	0	4,3 years	1	2 years	1	2	4 years	9
France	1	2 years	0	1	4 years	0	-	0	1	4 years	3
Germany	1	5 years	0	1	NC ¹	0	-	1	0	3 years	3
Hungary	0	-	0	1	OTE ³	0	-	1	1	3 years	3
Ireland*	0	-	0	0	-	1	3 years	0	0	-	1
Italy	0	-	1	1	0,3 year	1	3 years	0	4	2,5 years	7
Lithuania	0	-	0	5	3 years	0	-	1	1	4 years	7
Netherlands	0	-	0	2	1 year	1	3 years	0	2	3 years	5
Poland	0	-	0	2	MTOY ²	0	-	3	1	1,3 years	6
Portugal	0	-	0	1	MTOY ²	0	-	3	1	3 years	5
Slovenia	3	1 year	2	5	2,6 years	6	1,5 years	2	2	3 years	20
Spain	1	MTOY ²	1	3	1,5 years	1	MTOY ²	0	1	4 years	7
United Kingdom	5	4,25 years	0	1	2 years	1	2 years	1	2	2,5 years	10
SUBTOTAL	15		12	27		18		15	22		109
TOTAL	27			45				37			109

¹ NC: Not Clear; ² MTOY: More than Once a Year; ³ OTE: One Time Event

Table 2 Level of assessments per country

EU member state	International	National	Regional
Austria	-	6	-
Croatia	-	1	-
Czech Republic	2	6	-
Denmark	1	-	-
Estonia	-	5	2
Finland	4	3	2
France	1	2	-
Germany	-	3	-
Hungary	2	1	-
Ireland	-	1	-
Italy	6	-	1
Lithuania	1	6	-
Netherlands	3	2	-
Poland	-	4	2
Portugal	2	3	-
Slovenia	12	8	-
Spain	2	3	2
United Kingdom	2	8	-
TOTAL	38	62	9

Trend of assessments

Respondents were asked in which year the first assessment for the assessments began. For some this can be easily identified, but for a majority it is difficult to date precisely. The graph in Figure 2 shows a cumulative presentation of the trends in the number and types of assessments. It suggests that:

- The numbers of assessments have steadily increased from the 1990s till 2015.
- In the past two decades, there has been most increase in patient care assessments, followed by the mixed assessments of patient care and research aspects. The rise in pure research assessments has been the least.

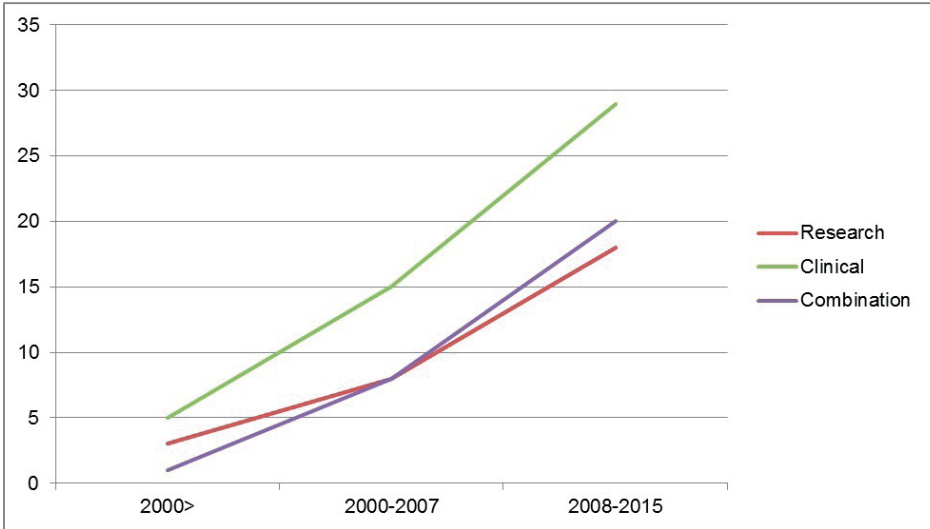


Figure 2 Trends in the number and types of assessments

Discussion

Continuous growth of assessments in Europe and how it compares to the US

The number of assessments in the EU has tripled since 2000. This shows that quality assessment in all its forms is a growing industry. It is particularly interesting to note a steady rise from 2000-2007, and especially following the economic crisis (2007/2008) more assessments seem to have cropped up. Whether this steep rise is related to the need for more accountability during and post financial crunch situations is hard to say. Although the emphasis on mandatory assessments will remain for the purpose of funding and licensing health services, voluntary assessments are equally gaining in popularity. In fact, most of the new assessments are voluntary, however, this does not exclude the pressure on CCs to participate in them. This shows that most assessments seem to be in a transition, moving from a friendly tool of self-assessment and development to a governing tool that agencies use for various purposes.

Regarding CCs, in the US there are at least three main assessments: The Joint Commission accreditation²¹ for healthcare organizations and programs as a whole; The Commission on Cancer (CoC) of the American College of Surgeons for the quality of cancer care delivery²²; and The National Cancer Institute (NCI) designation²³ for assessing excellent multidisciplinary translational cancer research programs, in which almost all leading CCs in the US

participate. Europe is gradually moving towards common European assessment frameworks in order to benchmark and improve cancer research and patient care activities across the EU, but this has not been as developed as it is in the United States. It is with this intention that European Commission is allocating more budgets for research and innovation (e.g. through specific funding programs such as Horizon2020)²⁴ with the idea of improving EU competitiveness in excellent science²⁵. However, the challenges that arise from health care being under national jurisdiction and individual responsibility of each EU member state has meant that only gradual steps towards harmonization towards EU assessments have been seen so far. As healthcare is a major component of national economies (as a user of public funds but also as an investment that generates jobs, taxes and procurement opportunities for Small and Medium Enterprises) within a monetary union, increasing steps towards EU influence on these issues seems inevitable²⁶.

The link between health system type and nature and scope of assessments

A link between the type of health system and the nature of the assessments is visible only in some member states. For example, in the United Kingdom where a National Health Service is being used (regulation, financing and provision by the state, see Figure 1) a lot of mandatory, national assessments can be found. The same goes for Spain. In other countries that have an NHS model, e.g. Finland and Portugal assessments seem to be more voluntary than mandatory. Within the National Health Insurance system (regulation by the state) one would again expect a lot of mandatory and national assessments, but the opposite is the case in Italy, where a lot of international voluntary assessments are performed e.g. the Joint Accreditation Committee- International Society for Cellular Therapy and European Society for Blood and Marrow Transplantation (JACIE)²⁷ and International Organization for Standardization (ISO)²⁸ and the European accreditation by the Organization of European Cancer Institutes (OECI)²⁰. But these initially voluntary assessments are sometimes mandatory for either keeping license and/or are demanded by government to maintain Comprehensive Cancer Center status, such as in Italy. So, the voluntary assessments end up being mandatory at some level.

In the Social Health Insurance (SHI) type, societal actors dominate healthcare regulation and financing, which is reflected in the assessments listed by the centers from Germany and Croatia e.g. in Germany accreditation of cancer care is performed by the German Cancer society²⁹ which is a societal actor dominated by physicians. In most Central and Eastern European countries that have an Etatist Social Health Insurance system, there is a tendency for more mandatory national assessments, while in the majority of Western Europe and Nordic countries there is tendency to participate in more voluntary international assessments. Only in few member states, did we notice regional level assessments e.g.

Italy, Finland and Estonia. This can be partly explained by decentralization/devolution of powers to regions in some EU member states³⁰. Evidence suggests that mandated external quality assessments are less effective than voluntary assessments because the effectiveness of accreditation is dependent on its voluntary nature, non-threatening process, and interactive process with external reviewers as a means of effecting and speeding up quality improvements³¹.

Traditional view of assessments and shifting focus

Assessments focused on research performed by CCs (such as LabQuality which checks the quality of Laboratories and BASG/AGES that looks at the quality of clinical drug trials) are still limited in Europe when compared to patient care assessments. The NCI designation program²³ in the US is one of the anchors of the nation's cancer research efforts. In order to be designated, CCs must meet specific criteria for: breadth and depth of basic cancer research; clinical cancer research; and prevention, control and population/behavioral sciences research in cancer; and strength of interaction among these three major research areas. A European version of the NCI designation was not found in our study.

One of the possible reasons why patient care assessment (such as accreditation) are performed more often than research focused assessments is that, being an accredited center in cancer care could attract patients³². Additionally, in some countries, accreditation is being used as an extension of statutory licensing for institutions³³. Therefore, care assessments such as accreditation seem of more direct importance than assessing research. Another reason is that assessing impact of research on healthcare outcomes is more difficult than assessing care outcomes³⁴. In research, metric-driven indicators such as impact factors are often criticized³⁵ and consensus on value-based indicators is still evolving e.g. how to define success in translational research (bench to bedside and back) in terms of practice-changing innovations³⁶. The awareness that alignment between research and clinical areas is essential in successful translational research³⁷ can explain why more mixed assessments are being introduced in the EU. This is comparable to the SPORE³⁸— the Specialized Programs of Research Excellence — a cornerstone of National Cancer Institute's efforts dedicated to capitalize on research opportunities that have the potential to change the current paradigm in the prevention, detection, diagnosis, and/or treatment of human cancer. Given the amount of funding that goes into research in the EU as in the US³⁹, evaluating research becomes necessary. More specifically, comparative research assessments are needed to make evidence based decisions on most suitable therapies in clinical practice⁴⁰.

Transparency

A review of accreditation and quality systems by the World Health Organization⁴¹ shows that “the move towards statutory and governmental endorsement is associated with freer access by the public to the standards, processes and findings of accreditation”. Half of the assessment bodies make the standards/indicators used for the assessment available at little or no cost. One-third also make full reports of individual assessments publicly available. However, several organizations are unwilling to give away their standards and norm descriptions as this serves also as a source of income and intellectual property. Other difficulties are for example the fact that in many EU member states, the assessment reports as well as the program standards are in the local language. It takes time, money and effort to accurately translate the reports into English. This makes it hard to judge assessments in terms of how each assessment can bring added value to the different stakeholders⁴¹. The first step in deciding the value of assessments is to make their outcomes publicly available and accessible⁴² although this assumption has been challenged⁴³. Next, public consultation must occur with key stakeholders to decide the parameters to assess the added value of assessments⁴⁴. Another issue related to this is also whether the data, if made publicly available, are good enough to actually promote quality improvement and helping consumers make choices⁶. Evidence on whether the assessments undergone by CCs actually provide added value for patient care or patient outcomes is limited. Although most assessments focus on patient care aspects, it is unknown whether patient outcomes are actually improved through these assessments⁴⁵. Evidence shows that for example patient safety can be improved if a healthcare organization undergoes licensing, certification and accreditation⁷ but this is unknown for patient outcomes. Although there is no decisive evidence on the direct impact on patient outcomes, there is some indication that quality assessments such as accreditation could contribute to health outcomes. This is the case if these assessments strengthen interdisciplinary team effectiveness, communication, and enhanced use of indicators leading to evidence-based decision making⁶. This evidence is however limited and study designs are weak. A study focusing on accreditation specifically shows that a lot of information on the added value of assessments is unknown and future research should focus on: determining the impact of accreditation on patient care and outcomes; determining how best to research the validity, impact and value of accreditation processes in health care; determination of value for time and money; and determining the reliability of accreditation surveys to truly assess the quality of organizations⁴⁵.

Strengths and Limitations

This study describes the type and number of assessments at 19 cancer centers in 18 out of 28 member states of the European Union. This is the first systematic European attempt

to gather data on assessments for cancer centers. The results were validated with study participants by asking them not just to confirm the data for their own cancer center but also giving them an opportunity to comment on assessments that were listed by other cancer centers in Europe. This study gives sufficient base data to start thinking about how to reduce the burden of assessments for cancer centers and how to make them more transparent and effective.

3 Content of these assessments (e.g. assessment reports, outcomes) were not easy to access due to language barriers (each cancer center has it in its local European language and is not always translated in English) and/or lack of publicly available information. The individuals from cancer centers who provided the data were quality managers (and/or research directors/senior executive managers) who are usually responsible for organizing and implementing assessments in their center. However, many assessments are multidisciplinary in nature, involving a wide range of staff, therefore future research should focus on validating the responses beyond quality managers. Our assumption is that non-responses may have been the result of not identifying or contacting the appropriate people, rather than reluctance to provide data and/or that formalized assessments do not exist in some member states. Another limitation regarding the year in which the assessment started is the fact that, first assessments may be considered as pilot testing rather than becoming operational. It is therefore difficult in some cases to identify the year in which the actual assessment started.

Conclusion

There seem to be 109 assessments that CCs currently undergo in 19 EU states and the numbers keep increasing. Although there are benefits of assessments, more robust research is needed to understand their value in terms of how they improve patient quality and safety. CCs go through frequent assessments, sometimes as often as more than once a year, this can be a very time consuming as well as expensive for those organizations. Rapid uptake of voluntary assessments is associated with direct financial incentives (such as linkage to core funding or reimbursement) and government encouragement. However, decision makers should regulate assessments to reduce unnecessary assessments that do not bring benefits or added value, that are bureaucratic, time-consuming and/or unaffordable by CCs. This article shows that demand for assessments is increasing and changing rapidly in terms of international assessments as well as mixed assessments of cancer research and care. Assessments must be transparent to bring credibility and accountability among stakeholders. Given the importance of quality of care, patient safety and outcome improvement in cancer care, it would be desirable to evaluate the impact of assessments in these areas. We

recommend future research to go deeper into understanding process and outcome related issues; how much time does each assessment take to prepare and implement, people and money consumed, who are the peer-reviewers and what are their backgrounds, how are standards developed and revised, sources of income for assessment bodies, and last but not least does the exercise meet its objectives?

References

1. Mohr JJ, Batalden P, Barach P. Integrating patient safety into the clinical microsystem. *Qual Saf Health Care*. 2004;13.
2. Institute of Medicine. "Crossing the Quality Chasm: A New Health System for the 21st Century". National Academy of Sciences (2001).
3. Mladovsky P, Srivastava D, Cylus J, et al. "Health policy responses to the financial crisis in Europe. Policy Summary 5". World Health Organisation on behalf of the European Observatory on Health Systems and Policies (2012).
4. Giovannella L and Stegmüller K. The financial crisis and health care systems in Europe: universal care under threat? *Trends in health sector reforms in Germany, the United Kingdom, and Spain*. *Cad. Saúde Pública*. 2014; 30(11): 2263-2281.
5. OECD. "Health care systems: Getting more value for money". OECD Economics Department Policy Notes, No. 2 (2010).
6. Miller MR, Pronovost P, Donithan M, et al. Relationship Between Performance Measurement and Accreditation: Implications for Quality of Care and Patient Safety. *Am J Med Qual*. 2005;20(5):239-52.
7. Hughes RG. Chapter 44 Tools and Strategies for Quality Improvement and Patient Safety. In: Hughes RG. *Patient Safety and Quality: An Evidence-Based Handbook for Nurses*. Rockville (MD): Agency for Healthcare Research and Quality (US); 2008.
8. Institute of Medicine. "To Err is Human: Building a Safer Health System". National Academy of Sciences (1999).
9. Quality assessment. (n.d.) American Heritage® Dictionary of the English Language, Fifth Edition. (2011). <http://www.thefreedictionary.com/Quality+assessment> Accessed 02 May, 2016
10. Lim C, Cheung MC, Franco B, et al. Quality Improvement: An Assessment of Participation and Attitudes of Medical Oncologists. *J Oncol Pract*. 2014; 10(6):e408-14.
11. Ikkersheim D and van der Avoort J. "Onderzoek kosten kwaliteitsmetingen". KPMG Advisory N.V (2015). http://www.nvzziekenhuizen.nl/_library/31906/NVZ%20rapportage%20KPMG%20onderzoek%20kosten%20kwaliteitsmetingen.pdf Accessed April 24, 2016.
12. BENCH-CAN [p://www.oeci.eu/Benchcan/](http://www.oeci.eu/Benchcan/)
13. Trimble LE, Abrams SJ, Meyer MR, et al. Improving Cancer Outcomes Through International Collaboration in Academic Cancer Treatment Trials. *J Clin Oncol*. 2009;27(30): 5109-5114.
14. Rowland JH, Kent EE, Forsythe LP, et al. Cancer Survivorship Research in Europe and the United States: Where have we been, where are we going, and what can we learn from each other? *Cancer*. 2013;119(011):2094–2108.
15. Rothgang H, Cacace M, Grimmeisen S, Wendt C. The changing role of the State in Healthcare systems. *European Review*. 2005;13:187–212.
16. Wendt C, Frisina L, Rothgang H. Healthcare system types: a conceptual framework for comparison. *Soc Policy Adm*. 2009;43(1):7090.
17. Böhm K, Schmid A, Götz R, Landwehr C, Rothgang H. Five types of OECD healthcare systems: Empirical results of a deductive classification. *Health Policy*. 2013;113:258-269.

18. Chua KP. Overview of the U.S. Health Care System. American Medical Student Association(2006).
19. The European Society for Medical Oncology <http://www.esmo.org> Accessed April 30, 2014
20. The Organization for European Cancer Institutes <http://www.oeci.eu> Accessed April 30, 2014
21. The Joint Commission: Accreditation, healthcare, certification <http://www.jointcommission.org> Accessed November 12, 2014
22. Commission on Cancer –American College of Surgeons <https://www.facs.org/quality-programs/cancer> Accessed November 12, 2014
23. National Cancer Institute at the National Institutes for Health <http://www.cancer.gov> Accessed November 12, 2014
24. Horizon2020- the EU framework programme for Research and Innovation <http://ec.europa.eu/programmes/horizon2020/> Accessed December 12, 2014
25. European Commission Directorate General for Research and Innovation, Innovation Union Competitiveness Report 2013 http://ec.europa.eu/research/innovation-union/pdf/competitiveness_report_2013.pdf Published January 2014. Accessed December 12, 2014
26. Fierlbeck K. The changing contours of experimental governance in European healthcare. *Soc Sci Med.* 2014;108:89-96.
27. Joint Accreditation Committee Isct. Embt <http://www.jacie.org> Accessed December 12, 2014
28. International Organization for Standardization <http://www.iso.org/iso/home.html> Accessed December 12, 2014
29. The German Cancer Society <http://www.krebsgesellschaft.de/german-cancer-society.html> Accessed December 12, 2014
30. Progress Consulting S.r.l. and Living Prospects Ltd . 2012. The management of health systems in the EU member states – the role of local and regional authorities ISBN: 978-92-895-0717-2
31. Montagu D. Accreditation and other external quality assessment systems for healthcare, Review of Experiences and lessons learned May 2003, prepared by the Health Systems Resource Center on behalf of Department for International Development. <https://www.wbginvestmentclimate.org/toolkits/public-policy-toolkit/upload/Accreditation-Review-Montagu-2003.pdf> Accessed December 12, 2014
32. Kowalski C, Wesselmann S, Kreienberg R, Schulte H, Pfaff H. The Patients' View On Accredited Breast Cancer Centers: Strengths and Potential for Improvement. *Geburtshilfe und Frauenheilkunde.* 2012;72(2):137-143.
33. Shaw D. Evaluating accreditation. *Int J Qual Health Care.* 2003;15:455–456.
34. Krzyzanowska MK, Kaplan R, Sullivan R. How may clinical research improve healthcare outcomes? *Ann Oncol.* 2011;22(7):10-15.
35. Saha S, Saint S, Christakis AD. Impact factor: a valid measure of journal quality? *J Med Libr Assoc.* 2003;91(1):42–46.
36. Rajan A, Caldas C, van Luenen H, Saghatchian M, van Harten WH. Assessing excellence in translational cancer research: a consensus based framework. *J Transl Med.* 2013;11:274.

37. Rajan A, Sullivan R, Bakker S, van Harten WH. Critical appraisal of translational research models for suitability in performance assessment of cancer centers. *Oncologist*. 2012;17(12):48-57.
38. Translational Research Programmes, National Cancer Institute <http://trp.cancer.gov> Accessed December 12, 2014
39. Woolf HS. The meaning of translational research and why it matters. *JAMA*. 2008;299(2):211-213.
40. Howie L, Hirsch B, Locklear T, Abernethy AP. Assessing The Value Of Patient-Generated Data To Comparative Effectiveness Research *Health Aff* July 2014;33:71220-1228.
41. World Health Organization. "Quality and accreditation in health care services: A global review". Evidence and Information for Policy Department of Health Service Provision (OSD) (2003).
42. Totten AM, Wagner J, Tiwari A, O'Haire C, Griffin J, Walker M. Public Reporting as a Quality Improvement Strategy. Closing the Quality Gap: Revisiting the State of the Science. Evidence Report No. 208. (Prepared by the Oregon Evidence-based Practice Center under Contract No. 290-2007-10057-I.) AHRQ Publication No. 12-E011-EF. Rockville, MD: Agency for Healthcare Research and Quality. July 2012. www.effectivehealthcare.ahrq.gov/reports/final.cfm.
43. Werner RM and Asch DA. The unintended consequences of publicly reporting quality information. *JAMA*. 2005;293(10):1239-44.
44. Deverka PA, Lavalley DC, Desai PJ, et al. Stakeholder participation in comparative effectiveness research: defining a framework for effective engagement. *J Comp Eff Res*. 2012;1(2):181-194.
45. Nicklin W and Dickson S. "The Value and Impact of Accreditation in Health Care: A Review of the Literature". Accreditation Canada (2009).

PART 3

Patient perspective

CHAPTER 4

Piloting a Generic Cancer Consumer Quality Index in six European countries

**Anke Wind
Mark Patrick Roeling
Jana Heerink
Herman Sixma
Pietro Presti
Claudio Lombardo
Wim van Harten**

Abstract

Background: Accounting for patients' perspective has become increasingly important. Based on the Consumer Quality Index method (founded on Consumer Assessment of Healthcare Providers and Systems) a questionnaire was recently developed for Dutch cancer patients. As a next step, this study aimed to adapt and pilot this questionnaire for international comparison of cancer patients experience and satisfaction with care in six European countries.

Method: The Consumer Quality Index was translated into the local language at the participating pilot sites using cross-translation. A minimum of 100 patients per site were surveyed through convenience sampling. Data from seven pilot sites in six countries was collected through an online and paper-based survey. Internal consistency was tested by calculating Cronbach's alpha and validity by means of cognitive interviews. Demographic factors were compared as possible influencing factors.

Results: A total of 698 patients from six European countries filled the questionnaire. Cronbach's alpha was good or satisfactory in 8 out of 10 categories. Patient satisfaction significantly differed between the countries. We observed no difference in patient satisfaction for age, gender, education, and tumor type, but satisfaction was significantly higher in patients with a higher level of activation.

Conclusion: This European Cancer Consumer Quality Index (ECCQI) showed promising scores on internal consistency (reliability) and a good internal validity. The ECCQI is to our knowledge the first to measure and compare experiences and satisfaction of cancer patients on an international level, it may enable healthcare providers to improve the quality of cancer care.

Introduction

The organization of care for cancer patients is complex and multifaceted, cancer can cause a great deal of distress for patients. A study among lung-cancer patients showed that 27 percent mentioned healthcare experiences as an important cause of distress. Waiting times, and lack of information are some mentioned experiences¹. Different healthcare providers are engaged in prevention, diagnosis, treatment and follow up. This requires a high degree of coordination and if inadequately organized, can result in fragmented and discontinued care². The Institute of Medicine (IOM) proposed patient centeredness as a way how healthcare system could improve patients' experience³. Patient centeredness is defined as: care that respects and responds to individual patient's preferences, needs, and values and involves clinical decisions guided by patients³; and is associated with better treatment adherence and improved health outcomes⁴. Healthcare professionals and patients do not always agree on what is important in patient centered care. Wessels et al.⁵ reported that expertise and attitude of healthcare providers as well as accessibility were more important to cancer patients than healthcare professionals expected. This underlines the importance of questionnaires that actually reflect the perspective of the patient. Patient experience and satisfaction are increasingly seen as a quality outcome for health-system or –provider performance, by consumers, practitioners and governing agencies⁶.

The Consumer Quality Index (CQI) used in this study is based on the American CAHPS (Consumer Assessment of Healthcare Providers and Systems)⁷. The CAHPS is one of the most well-known initiatives to measure quality of care from the healthcare user's perspective. CAHPS is widely used in the United States and translated and used in the Netherlands. The CQI is also based on the Dutch QUOTE (Quality Of care Through the patient's Eyes)⁸. Many researchers have designed instruments to measure patient experience and satisfaction that are specific to a country's health system or individual hospital⁹⁻¹⁴. In order to compare performance across health systems and providers, standardized and comparable measures of patient experience and satisfaction are necessary, to our knowledge there is no such instrument yet. Our objective was to adapt and test the psychometric properties of a generic questionnaire that measures the actual experiences and satisfaction of cancer patients with care in different countries in Europe based on the Dutch version of the CQI. A generic questionnaire has advantages: it can be used for patients with all tumor types, which makes developing different tumor-specific questionnaires redundant⁴. Questions regarding actual experiences tend to reflect the quality of care better and are more interpretable and actionable for quality improvement purposes, while satisfaction ratings shows whether expectations were met¹⁵. In order to get a comprehensive picture, both satisfaction and experience are measured. Our research questions were:

1. *What are the differences in patient experience and satisfaction between countries and/or patient characteristics?*
2. *What is the validity and internal consistency (reliability) of the European Cancer Consumer Quality Index?*

Materials and Methods

2.1 Questionnaire

4

To use the existing CQI in an international context, questions related specific to the Dutch system were removed based on expert opinion. The updated questionnaire was sent to the European Cancer Patient Coalition and a patient representative at each of the pilot sites to check for appropriateness for international measurement. Patient representatives were asked to judge whether their patients would be able to read and comprehend the questions. Twelve institutes across Europe were invited to participate of which seven institutes in six countries (two in Italy) responded positively. These countries were: Hungary (HUN), Portugal (PRT), the Netherlands (NLD), Romania (ROM), Lithuania (LIT), and Italy (ITA). The CQI was translated into the local language at the pilot sites and translated back into English, to ensure that no information was lost in translation, so called cross-translation. Cross-translation was used to ensure the translated instruments are conceptually equivalent in each of the target countries/cultures¹⁶. The CQI used in this study will be referred to as European Cancer Consumer Quality Index (ECCQI) and consists of 65 questions/items divided into 13 categories. The three categories with demographic or disease specific information were used as background and were not part of the analysis which therefore includes 10 categories (45 items). Participants were given the opportunity to comment on the questionnaire.

2.2 Data collection

The target response was a minimum of 100 respondents per institute. Every institute assigned a person who ensured the distribution and collection of the questionnaires. In the Netherlands, data were collected through an online survey tool¹⁷, in other institutes (n=6) the questionnaire was paper-based due to the fact that internet coverage was not sufficient in these countries. Respondents were selected by convenience sampling. This study was performed in agreement with the declaration of Helsinki. Approval by a medical ethics committee was not required. All participants consented to the use of the data provided by them. Data from interviews and questionnaires were analyzed anonymously.

2.3 Inclusion criteria

The following criteria were used for inclusion of the questionnaires: (1) Patients had to be 18 years or older, (2) patients had to be examined, treated or had after-care for cancer within the last two years in the examined center, (3) gender, age and level of education had to be known, (4) 50% of the questions answered.

2.4 Cognitive interviews

Cognitive interviews were performed in order to measure the face validity of the ECCQI and to identify problems in the wording or structure of questions which might lead to difficulties in question administration, miscommunication, etc. Face validity is the extent to which a test is subjectively viewed as covering the concept it is supposed to measure, which in this study, is the experience of and satisfaction of cancer patients with care received at the cancer center. Both 'thinking aloud' and 'verbal probing'¹⁸, were used in this study. When thinking aloud, respondents are asked to read the questions out loud and to verbalize their thoughts as they fill out the questionnaire. With verbal probing, the interviewer asks follow-up questions to understand a participant's interpretation more clearly and precisely. The cognitive interviews were conducted in the Netherlands, Romania (with interpreter), and Portugal (with interpreter). Data collected through the cognitive interviews were analyzed by means of the Question Appraisal System (QAS-99)¹⁹. The QAS-99 consists of seven elements: (i) Determine if it is difficult to read the question uniformly for all respondents; (ii) Look for problems with any introductions, instructions, or explanations from the respondent's point of view; (iii) Identify problems related to communicating the intent or meaning of the question to the respondent; (iv) Determine if there are problems with assumptions made or the underlying logic the questions; (v) Check whether respondents are likely to not know or have trouble remembering information; (vi) Assess questions for sensitive nature or wording, and for bias; (vii) Assess the adequacy of the range of responses to be recorded.

2.5 Recoding

Data were recoded in order to be analyzed. Almost all categories of the CQI consist of questions with four response options: never = 1, sometimes = 2, usually = 3 and always = 4. For the categories that did not consist of those four response options, the options were recoded into one of the four options above. Response codes of the questions about demographic characteristics were also recoded; (i) Age: 18 – 34, 35 – 64, and 65 or older; (ii) Years of education: low (1-8 years), moderate (9-13 years), and high (14 and higher). The answers 'I don't know/I no longer remember' and 'Not applicable' were scored as missing.

2.6 Analyses

For descriptive analyses we used SPSS v.22. To aid future comparison of samples and normalization, descriptive statistics involved calculating the weighted mean for each scale and country. In line with the instructions²⁰, patient's scores were only valid if 50% or more questions within a scale were answered. We performed a chi-square test to determine whether the distribution of patient characteristics such as age differed between countries. For every category the weighted mean was calculated per country, where the weight depended on the number of items rated by the patient. We summed the scale scores and calculated the weighted mean of overall patient experience and satisfaction for every patient. The possible effects of demographic characteristics on ECCQI score were examined with one way Analyses Of VAriance (ANOVA) analysis (95% Confidence Intervals: CI).

To estimate the internal consistency (reliability) of each scale, we calculated the Cronbach's alpha²¹ (Cronbach, 1951; α) for ordinal items. In short, we followed the method from Gardemann et al.²², where α was calculated on the polychoric correlation matrix (calculated with the *psych* package available in the R programming language), instead of the normal Pearson correlation. Acceptable α scores fall between 0.5 to 0.7 and α is considered good if higher than 0.7²³.

The ECCQI presented here is based on the factor structure of the CQI. We tested the structural validity of the ECCQI in our data with Confirmatory Factor Analysis (CFA). The rationale behind applying CFA is that a predefined measurement model can be tested with Structural Equation Modeling, where CFA provides insight into the fit of the model on the current data. CFA analyses were conducted in Mplus v.7²⁴ fitted using the Weighted Least Squares Mean Variance adjusted (WLSMV). As general measures of fit, the Root Mean Square Error of Approximation (RMSEA) and Comparative Fit Index (CFI) were evaluated. The RMSEA provides an indication of how well the model fits in the population. Values $>.10$ indicate poor model fit, values between $.08$ and $.05$ indicate adequate model fit, and values of $.05$ or below indicate good fit of the model to the data²⁵. The CFI ranges from zero to one and higher values indicate better fit. It has been shown to be an adequate fit statistic for ordinal data²⁶ with values larger than $.90$ indicating moderate fit and $.95$ indicating good fit. Also, we fitted all models using the Weighted Least Squares Mean Variance adjusted (WLSMV).

2.7 Patient activation

To investigate relationships between level of patient activation and ECCQI score the Patient Activation Measure (PAM) was administered^{27, 28}. The PAM was included later on in the

study. It was only send to institutes in the Netherlands, Romania and Italy, since these countries indicated that they still could implement the PAM at that time, however not all patients filled out the PAM.

Results

3.1 Response

Initially 958 questionnaires were collected. After application of the inclusion criteria 698 questionnaires were included in this study (see Figure 1). Respondent characteristics can be found in Table 1. In order to ensure anonymity data are presented by country and not by individual institute (the Italian institutes are combined).

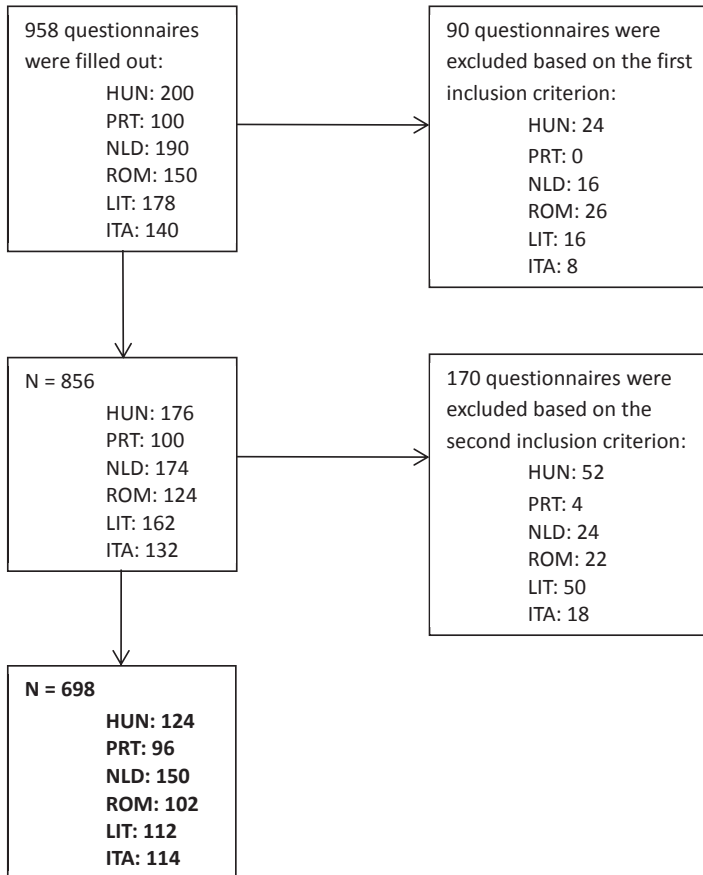


Figure 1 Flow-chart of sample size ECCQI

Table 1 ECCQI Respondent characteristics. Percentage and absolute numbers

	HUN	PRT	NLD	ROM	LIT	ITA	Total
Age, % (#)							
18 – 34	3,2 (4)	3,1 (3)	5,3 (8)	5,9 (6)	3,6 (4)	7,0 (8)	4,7 (33)
35 – 64	57,3 (71)	59,4 (57)	68,7 (103)	78,4 (80)	74,1 (83)	71,1 (81)	68,1 (475)
> 65	39,5 (49)	37,5 (36)	26,0 (39)	15,7 (16)	22,3 (25)	21,9 (25)	27,2 (190)
Gender, % (#)							
Male	22,6 (28)	39,6 (38)	44,0 (66)	27,5 (28)	34,8 (39)	29,8 (34)	33,4 (233)
Female	77,4 (96)	60,4 (58)	56,0 (84)	72,5 (74)	65,2 (73)	70,2 (80)	66,6 (465)
Education, % (#)							
Low	7,3 (9)	56,3 (54)	1,3 (2)	7,8 (8)	13,4 (15)	23,0 (26)	16,4 (114)
Moderate	41,9 (52)	29,2 (28)	24,7 (37)	43,1 (44)	29,5 (33)	54,0 (61)	36,6 (255)
High	50,8 (63)	14,6 (14)	74,0 (111)	49,0 (50)	57,1 (64)	23,0 (26)	47,1 (328)
Activation, % (#)							
Level 1			32,4 (11)	8,6 (5)		27,7 (18)	18,6 (34)
Level 2			47,1 (16)	12,1 (7)		16,9 (11)	18,6 (34)
Level 3			8,8 (3)	53,4 (31)		40,0 (26)	43,7 (80)
Level 4			11,8 (4)	25,9 (15)		15,4 (10)	19,1 (35)
Experienced health % (#)							
Excellent	5,6 (7)	1,0 (1)	4,3 (6)	0	1,8 (2)	5,3 (6)	3,2 (22)
Very good	15,3 (19)	1,0 (1)	17,0 (24)	10,9 (11)	3,6 (4)	13,3 (15)	10,8 (74)
Good	36,3 (45)	37,5 (36)	55,3 (78)	42,6 (43)	42,8 (48)	44,3 (50)	43,7 (300)
Moderate	30,6 (38)	48 (46)	21,3 (30)	36,6 (37)	50,0 (56)	30,1 (34)	35,1 (241)
Poor	12,2 (15)	12,5 (12)	2,1 (3)	9,9 (10)	1,8 (2)	7,0 (8)	7,2 (50)
Type of cancer % (#)							
Digestive organs	6,4 (8)	27,1 (26)	8,7 (13)	16,7 (17)	24,1 (27)	12,3 (14)	15,0 (105)
Lung	4,8 (6)	4,2 (4)	7,3 (11)	6,9 (7)	6,2 (7)	5,3 (6)	5,9 (41)
Breast	40,3 (50)	29,2 (28)	32,0 (48)	39,2 (40)	16,1 (18)	43,9 (50)	33,5 (234)
Male reproductive organs	6,4 (8)	8,3 (8)	13,3 (20)	6,9 (7)	10,7 (12)	2,6 (3)	8,3 (58)
Skin	9,7 (12)	1,0 (1)	8,0 (12)	1,0 (1)	0,9 (1)	2,6 (3)	4,3 (30)
Blood, bone marrow and lymph nodes	4,0 (5)	2,1 (2)	2,0 (3)	1,0 (1)	0	9,6 (11)	3,1 (22)
Urinary tract	3,2 (4)	1,0 (1)	4,0 (6)	1,0 (1)	7,1 (8)	0,9 (1)	3,0 (21)
Female reproductive organs	8,9 (11)	6,2 (6)	2,7 (4)	19,6 (20)	23,2 (26)	3,5 (4)	10,2 (71)
Head and neck area	3,2 (4)	4,2 (4)	1,3 (2)	1,0 (1)	2,7 (3)	3,5 (4)	2,6 (18)
Central nervous system	1,6 (2)	0	1,3 (2)	1,0 (1)	0	0	0,7 (5)
Bone or soft tissue	1,6 (2)	1,0 (1)	2,7 (4)	1,0 (1)	0,9 (1)	2,6 (3)	1,7 (12)
Endocrine glands	4,8 (6)	0	0,7 (1)	1,0 (1)	0	0	1,1 (8)
Eye or eye socket	0,8 (1)	0	0	0	0	0	0,1 (1)
Other	0,8 (1)	1,0 (1)	5,3 (8)	2,0 (2)	0,9 (1)	0	1,9 (13)
Multiple forms	3,2 (4)	14,6 (14)	10,7 (16)	2,0 (2)	7,1 (8)	13,1 (15)	8,4 (59)

Table 1 ECCQI Respondent characteristics. Percentage and absolute numbers (continued)

	HUN	PRT	NLD	ROM	LIT	ITA	Total
Years of cancer % (#)							
< 1	14,6 (18)	6,5 (6)	2,0 (3)	28,9 (28)	27,9 (29)	27,4 (31)	16,9 (115)
1 – 2	58,5 (72)	54,3 (50)	86,7 (130)	54,6 (53)	62,5 (65)	26,5 (30)	58,9 (400)
3 – 5	13,8 (17)	16,3 (15)	10,0 (15)	12,4 (12)	4,8 (5)	21,2 (24)	13,0 (88)
6 – 10	8,1 (10)	13,0 (12)	0	4,1 (4)	4,8 (5)	13,3 (15)	6,8 (46)
> 10	4,9 (6)	9,8 (9)	1,3 (2)	0	0	11,5 (13)	4,4 (30)
Treatment received %* (#) (more than 1 answer possible)							
Examinations	68,8 (84)	69,8 (67)	93,6 (132)	59,4 (60)	73,2 (82)	91,1 (103)	76,9 (528)
Operation	56,6 (69)	53,6 (37)	70,2 (99)	44,5 (45)	67,0 (75)	33,6 (38)	52,8 (363)
Radiotherapy	33,6 (41)	28,1 (27)	50,3 (71)	38,6 (39)	14,3 (16)	16,8 (19)	31,0 (213)
Chemotherapy	39,3 (48)	80,2 (77)	39,7 (56)	75,2 (76)	36,6 (41)	68,1 (77)	54,6 (375)
Hormone therapy	9,8 (12)	12,5 (12)	19,1 (27)	8,9 (9)	0,9 (1)	12,4 (14)	10,0 (75)
Immunotherapy	1,6 (2)	8,33 (8)	6,4 (9)	2,0 (2)	1,8 (2)	8,0 (9)	4,7 (32)
Aftercare	15,6 (19)	3,13 (3)	75,9 (107)	1,0 (1)	8,0 (9)	2,7 (3)	20,7 (142)
Stage of treatment % (#)							
Tests to ascertain diagnosis	2,5 (3)	0	0,7 (1)	2,0 (2)	6,4 (7)	0	1,9 (13)
Diagnosis known, will be treated soon	6,7 (8)	2,1 (2)	0,7 (1)	4,0 (4)	9,1 (10)	2,7 (3)	4,1 (28)
Treatment that is intended to cure	37,0 (44)	60,0 (57)	14,7 (21)	66,7 (68)	59,1 (65)	68,7 (77)	48,8 (332)
No further treatment possible	0,8 (1)	0	0	3,9 (4)	0	0	0,7 (5)
Non-curative treatment	5,0 (6)	32,6 (31)	11,2 (16)	10,8 (11)	10,9 (12)	10,7 (12)	12,9 (88)
Check-ups or treatments of the symptoms	39,5 (47)	5,3 (5)	68,5 (98)	11,8 (12)	13,6 (15)	16,1 (18)	28,6 (195)
Finished with treatments and check-ups	8,4 (10)	0	4,2 (6)	1,0 (1)	0,9 (1)	1,8 (2)	2,9 (20)

* percentages indicate percentage of total patients that received that type of treatment

Results of the chi-square test showed a significant difference in the distribution of the patient characteristics such as level of education ($\chi^2(10) = 210.315, p < 0.001$) and perceived overall health ($\chi^2(20) = 77.641, p < 0.001$).

3.2 Results of the ECCQI per country.

Table 2 shows the descriptive statistics of the ECCQI. The weighted mean of the summed scale scores was 3.35, ranging from 2.05 to 4, being slightly skewed (skewness = .871). Comparison between countries revealed a significant difference in experience and satisfaction [$F(5,692) = 5.337, p < 0.001$]. Post hoc comparisons indicated that this overall effect was predominantly influenced by a significant ($p < 0.001$) mean difference between Hungary (mean = 3.29, Standard Deviation (StDev) = .34) and the Netherlands (mean =



3.46, StDev = .33), and Italy (mean = 3.28, StDev = .33) and the Netherlands. Looking more specifically Portugal (mean = 3.11 and StDev = .97) scored fairly low on 'own inputs' as does Italy (mean = 3.09 and StDev = .89). 'Coordination' is scored quite low by Italian patients (mean = 3.03 and StDev = .54), whereas Hungarian patients give a relatively low score to 'rounding of the treatment' (mean = 2.99 and StDev = .53). However, for none of the categories significant differences were found between highest scoring country and the lowest scoring country. Looking at some specific questions about practical experiences it was found that patients in Hungary, Romania and Lithuania found it difficult to park at the institute (average score of 1). In all countries except Romania the majority of the patients received their diagnosis when expected (in Romania a majority, 47.5%, received it sooner). Looking at the satisfaction questions specifically (Table 3) it can be seen that all patients give a higher grade to the likeliness of recommending the center than to how they experienced the center themselves.

3.3 Patient characteristics

When looking at the division by age it can be seen that patients who are 65 or older report the highest score at half of all categories. The total scale score increased with age, being 3.27 (StDev = .39) in patients aged 18-34, 3.34 (StDev = .33) in patients 35-64 and 3.39 (StDev = .32) in patients aged >65. The age differences were not significant [$F(2,692) = 2.68, p = .069$]. Stratification by gender shows that females scored lower (mean = 3.34, StDev = .33) compared to males (mean = 3.38, StDev = .34), but this difference is not significant [$F(1,696) = 1.828, p = 0.177$]. Also, quality of care was not reported differently by patients with a higher/longer education [$F(5,694) = 0.093, p = .911$]. When we clustered the patients on tumor type, we observed no significant differences [$F(14,683) = 1.297, p = 0.204$]. A representative subset of 172 patients (score 1 believing the patient role is important N= 31; score 2 having the confidence and knowledge necessary to take action N= 32; score 3 actually taking action to maintain and improve one's health N= 76; and score 4 staying the course even under stress N=33) also completed the PAM, which revealed that reported quality of care significantly differs across PAM level [$F(3,168) = 2.362, p < 0.034$]. Post hoc comparisons showed that this effect is mainly driven by patients in the highest level of activation scoring higher (mean = 3.26, StDev = .36) than respondents with the lowest level (mean = 3.48, StDev = .26) of activation.

Table 2 Results of the ECCQI per country and category, mean and median score 4 point scale and range, StDev = Standard Deviation

Category		HUN	PRT	NLD	ROM	LIT	ITA	Total
Accessibility	Mean	3.03	3.39	3.79	2.84	3.03	3.58	3.32
	Median	3.00	3.50	4.00	2.67	3.00	3.67	3.50
	Range	3.00	2.33	3.00	3.00	3.00	2.33	3.00
	StDev	.81	.54	.42	.76	.66	.57	.72
Organization	Mean	2.21	2.13	2.35	2.43	2.32	3,25	2.29
	Median	2.33	2.00	2.20	2.33	2.33	2.17	2.33
	Range	2.33	2.33	2.27	2.00	2.40	2.33	2.60
	StDev	,50	.50	,53	,52	,48	,52	,51
Hospitalization	Mean	3,33	3,36	3,39	3,17	3,23	3,01	3,25
	Median	3.50	3.50	3.50	3.25	3.33	3.00	3.33
	Range	2.00	1.50	1.00	1.67	1.47	2.00	2.33
	StDev	,37	,40	,28	,41	,35	,39	,38
Safety	Mean	3.68	3.93	3.81	3.71	3.61	3.90	3.77
	Median	4.00	4.00	4.00	4.00	3.50	4.00	4.00
	Range	3.00	1.00	2.00	3.00	2.00	1.00	3.00
	StDev	.57	.20	.45	.53	.48	.27	.46
Attitude of HP	Mean	3,39	3,55	3,57	3,69	3,70	3,45	3,55
	Median	3.50	3.67	3.80	3.83	3.92	3.55	3.67
	Range	2.20	2.00	2.00	1.83	2.00	2.20	2.20
	StDev	,57	,52	,48	,41	,47	,49	,51
Communication and information	Mean	3,49	3,68	3,62	3,65	3,56	3,52	3,59
	Median	3.75	4.00	3.75	3.88	3.75	3.67	3.75
	Range	3.00	2.00	2.00	2.00	2.50	2.25	3.00
	StDev	,60	,51	,48	,53	,50	,49	,52
Own input	Mean	3,34	3,11	3,54	3,33	3,45	3,08	3,33
	Median	3.50	3.50	4.00	3.50	4.00	3.00	3.50
	Range	3.00	3.00	2.50	3.00	3.00	3.00	3.00
	StDev	,75	,97	,60	,80	,75	,89	,80
Coordination	Mean	3,46	3,20	3,27	3,43	3,50	3,03	3,31
	Median	3.50	3.25	3.25	3.50	3.50	3.00	3.33
	Range	2.50	2.25	2.75	2.00	2.50	2.50	2.75
	StDev	,49	,59	,62	,52	,48	,54	,57
Supervision and support	Mean	3,26	3,29	3,32	3,20	3,45	3,15	3,28
	Median	3.33	3.40	3.56	3.30	3.60	3.30	3.40
	Range	2.40	2.30	3.00	2.70	2.78	2.33	3.00
	StDev	,58	,62	,75	,61	,57	,63	,63
Rounding off the treatment	Mean	2,99	3,05	3,10	3,23	3,25	3,29	3,11
	Median	3.17	3.25	3.33	3.33	3.33	3.00	3.33
	Range	2.67	1.67	2.33	2.50	2.07	2.50	2.83
	StDev	,53	,58	,63	,57	,40	,61	,54
Mean of all categories	Mean	3,29	3,36	3,46	3,35	3,35	3,28	3,35
	Median	3.37	3.43	3.56	3.40	3.40	3.34	3.41
	Range	1.69	1.42	1.49	1.85	1.85	1.67	1.95
	StDev	,34	,34	.32	,32	,33	,33	,33



Table 3 Overall opinion absolute numbers, mean and median scale 1-10 and range, StDev=Standard Deviation

Category		HUN	PRT	NLD	ROM	LIT	ITA
Hospital score	Mean	8,91	8,91	9,11	9,24	8,78	8,57
	Median	9,00	10,00	10,00	10,00	9,00	10,00
	Range	5,00	5,00	3,00	4,00	7,00	6,00
	StDev	1,15	1,27	,87	,98	1,37	1,25
Likeliness to recommend	Mean	9,46	9,42	9,53	9,65	9,02	9,10
	Median	10,00	10,00	10,00	10,00	10,00	10,00
	Range	9,00	7,00	3,00	4,00	7,00	6,00
	StDev	1,19	1,27	,81	,82	1,39	1,12

3.4 Validity and evaluation of the questions

Fourteen cognitive interviews were conducted. Patients felt that in general the questionnaire was appropriate to measure patient satisfaction and experience. However, in 18 questions at least one problem was identified based on the QAS-99 method¹⁹. Most problems concerned the interpretation of questions. The most frequently mentioned comment was that the questionnaire does not differentiate between nurses and doctors (n=7), whereby patients could not give a nuanced answer. CFA revealed that the ECCQI measurement model had a moderate to good fit on our data (RMSEA = 0.039, CFI = 0.943).

3.5 Internal consistency

Seven categories ('attitude of the healthcare professional', 'communication and information', 'coordination', supervision and support' and 'rounding off the treatment') represent a good level of internal consistency ($\alpha > 0.7$) for all countries and overall (see Table 4). In three categories ('organization', 'hospitalization' and 'own inputs) level of internal consistency was acceptable (α between .5 and .7) to good. The alphas in the categories 'accessibility' and 'safety' were lower and represented an unacceptable internal consistency ($\alpha > 0.5$) in three countries (accessibility), possibly due to a low number of variables (accessibility = 3, safety = 2) and a smaller sample size after splitting the data to country specific. With the exemption of the Dutch population, removing the question: "Is it difficult to get to the this hospital (either by your own transport, by public transport or by taxi)" could increase α , but the correlational stability of this item increased with sample size.

Table 4 Ordinal Cronbach's alpha(α) score per ECCQI category and country and number of respondents (N) per ECCQI category and country

Category (N items)		HUN	PRT	NLD	ROM	LIT	ITA	Total
Accessibility (3)	α	.37	.23	.73	.54	.49	.71	.70
	Valid N	59	69	120	66	76	95	485
Organization (5)	α	.68	.58	.63	.58	.63	.60	.68
	Valid N	41	27	24	46	48	36	222
Hospitalization (5)	α	.72	.77	.81	.68	.62	.59	.73
	Valid N	88	53	84	87	102	69	483
Safety (2)	α	.64	.43	.66	.78	.54	.90	.65
	Valid N	96	92	95	99	97	107	586
Attitude of HP (6)	α	.91	.91	.91	.91	.93	.86	.91
	Valid N	61	62	31	63	48	82	347
Communication and information (4)	α	.90	.87	.88	.90	.84	.81	.88
	Valid N	102	82	113	96	92	100	585
Own inputs (2)	α	.65	.76	.87	.80	.78	.81	.78
	Valid N	87	60	111	80	76	85	499
Coordination (4)	α	.81	.76	.84	.71	.84	.70	.78
	Valid N	103	86	109	92	93	104	587
Supervision and support (10)	α	.90	.90	.96	.90	.92	.91	.93
	Valid N	40	45	9	37	39	51	221
Rounding off the treatment (4)	α	.82	.95	.77	.86	.96	.93	.78
	Valid N	10	7	11	17	36	14	95

Discussion

We developed a questionnaire that measures patient experiences and satisfaction with cancer care in hospitals in European countries for patients with all types of cancer. It measures a broad array of topics capturing specific needs and wishes of cancer patients. We found no significant differences between tumor types, supporting the use of a generic questionnaire⁴.

With regard to our first question: What are the differences in patient experience and satisfaction between countries and/or patient characteristics?- we found that patient experience and satisfaction is scored different between countries, with significant differences ranging from an average of 3.27 to 3.46 on a 4-point scale. Patient experience and satisfaction is scored, on average, the lowest in Italy and the highest in the Netherlands. Using one questionnaire for different cultural groups (different nationalities) could lead to measurement bias which could be an explanation for the differences between countries.



4 Looking at possible effects of cultural differences applying Hofstede's cultural dimension theory^{29, 30}, possible explanatory factors for the difference in patient satisfaction between countries can be found. High masculine societies (Hungary and Italy) had significantly lower satisfaction scores than low masculine societies (the Netherlands). According to Hofstede, Hofstede & Minkov³¹, a high masculine score indicates an assertive judgmental behavior without having much concern for the feelings of others, which could result in lower satisfaction scores. A low masculine score indicates more tenderness and sympathy for others, resulting in less willingness to provide criticism and therefore higher satisfaction scores. Previous studies on ethnic groups^{32, 33} showed however that differences in satisfaction with care should not be ascribed to measurement bias but should be viewed as arising from actual differences in experiences. Evaluation of the measurement equivalence across race and ethnicity on the CAHPS shows that that measurement bias does not substantively influence conclusions based on patients' responses³³. A study amongst 15 countries performed by Ipsos³⁴ showed that Italy scores low on patient experience which corresponds to our findings. Another population survey conducted in 2010³⁵ showed a high degree of satisfaction with health-care services and access to health care in both outpatient and inpatient setting in Lithuania.

Regarding the second question: What is the validity and internal consistency (reliability) of the ECCQI? - the cognitive interviews showed problems with different questions. Most problems concerned the interpretation of questions. These questions will be reviewed in order to make them more clear and understandable. The structural validity of the ECCQI measurement model was moderate to good. Given the relative large number of items and scales, versus the number of respondents, the fit could be improved by including more persons to increase the person vs. item ratio. Also, the fit of the model was evaluated for all six countries combined and it is possible that the ECCQI is not measurement invariant across countries or cultures. With more data, it would be possible to investigate whether the measurement model (and thus the latent constructs of the scales) are identical across nations³⁶. The validity of the ECCQI could be also be increased with more specificity in the questions, for example by dividing healthcare professionals into doctors and nurses. Regarding internal consistency, alpha was satisfactory to good in eight out of ten categories. Lack of questions in the categories with a low alpha are most likely the reason for the low internal consistency score. It is recommended to investigate whether reliable scales could be created by means of creating other sub-scales, or replace these scales by single-item questions.

The small differences between countries could be attributed to the difference in completing the questionnaire. In the Netherlands the questionnaires were Internet-based, while in other countries they were paper-based. Studies investigating the equivalence between

Internet and paper-based questionnaires are conflicting. Fang³⁷ indicated that differences were apparent when analyzing data from distinct survey modes (Internet and paper-based). On the other hand, other studies provided results which support the measurement equivalence of survey instruments across Internet and paper-based surveys³⁸⁻⁴⁰.

Age does not significantly influence the results. For the total satisfaction score in all countries, differences between the highest scoring age group and lowest scoring group were not significant. This finding contrasts other studies^{41, 42} showing that age needs to be considered when looking at patient experience and satisfaction data. In addition, results show that males were more positive than women which corresponds to results from other studies⁴¹, this difference was however not significant. Further, level of activation seems to have a significant influence, since low activated patients reported lower scores and highly activated patients reported higher scores. It can be seen that all patients give a higher mark to likelihood that they would recommend the hospital to other patients than that they rate the hospitals for themselves. Our results indicate that when measuring patient experience and satisfaction results need to be adjusted for nationality and level of activation but not for age or other demographic characteristics. Based on this research, the current questionnaire should be further tested for its ability to discriminate between hospitals and countries.

A possible limitation of this study design is the sampling method. With convenience sampling the chance of selection bias is high which could have influenced the outcomes. For example, in education level a majority of the Portuguese patients had a low education level, a majority of the Italian patients had a moderate education while in the other countries the majority had a high education level. Regarding physical health, patients in Portugal were more negative giving a moderate score, while in the other countries most patients rated their physical health as good or excellent. Analysis of the total study population however showed no influence of demographic characteristics.

The real value of these studies lies in their use to stimulate quality improvements. Even though the centers studied are not necessarily representative of all cancer centers in the study countries, the results indicate areas of improvement and might provide evidence about how organizations and providers could meet patients' needs more effectively.

Conclusions

To our knowledge, the questionnaire used in this study is the first that measures the experiences and satisfaction of cancer patients with care provided by cancer centers in Europe.

Our results show that patient satisfaction is scored significantly different between countries. We showed that differences exist in experiences and satisfaction between people with different characteristics such as activation levels. After testing for discriminatory power our questionnaire can be used Europe-wide to measure quality of cancer care from the patient perspective and to identify differences in the experiences of patients in different hospitals. This ECCQI is a first step towards the international comparison of patient experience and satisfaction, which could enable healthcare providers and policy makers to improve the quality of cancer care.

References

1. Tishelman C, Lovgren M, Broberger E, Hamberg K, Sprangers MA. Are the most distressing concerns of patients with inoperable lung cancer adequately assessed? a mixed-methods analysis. *J Clin Oncol* 2010;28:1942–1949.
2. Ouwens M, Hermens R, Hulscher M, Vonk-Okhuijsen S, Tjan-Heijnen V, Termeer R, et al. Development of indicators for patient-centred cancer care. *Support Care Cancer* 2010;18:121–130.
3. Institute of Medicine, Committee on Health Care in America: Crossing the quality chasm: a New health system for the 21st century. Washington, DC: National Academy Press, 2001.
4. Booij JC, Zegers M, Evers PM, Hendriks M, Delnoij DM, Rademakers JJ. Improving cancer patient care: development of a generic cancer consumer quality index questionnaire for cancer patients. *BMC Cancer* 2013;13:203.
5. Wessels H, de Graeff A, Wynia K, de Heus M, Kruitwagen CL, Teunissen SC, et al. Are health care professionals able to judge cancer patients' health care preferences correctly? A cross-sectional study. *BMC Health Serv Res* 2010;10:198.
6. Institute of Medicine. *Crossing the Quality Chasm*. Washington, DC: Institute of Medicine, 1999.
7. Hargraves JL, Hays RD, Cleary PD. Psychometric properties of the consumer assessment of health plans study (CAHPS) 2.0 Adult core survey. *Health Serv Res* 2003; 38:1509–1527.
8. Sixma HJ, Kerssens JJ, Campen CV, Peters L. Quality of care from the patients' perspective: from theoretical concept to a new measuring instrument. *Health Expect* 1998;1:82–95.
9. Draper M, Cohen P, Buchan H. Seeking consumer views: what use are results of hospital patient satisfaction surveys. *Int J Qual Health Care* 2001; 13:463–468.
10. Cheng S, Yang M, Chiang T. Patient satisfaction with and recommendation of a hospital: effects of interpersonal and technical aspects of care. *Int J Qual Health Care* 2003; 15:345–355.
11. Perneger TV, Kossovsky MP, Cathieni F, di Florio V, Burnand B. A randomized trial of four patient satisfaction questionnaires. *Med Care* 2003;41:1343–1352.
12. Boulding W, Glickman SW, Manary MP, Schulman KA, Staelin R. Relationship between patient satisfaction with inpatient care and hospital readmission within 30 days. *Am J Manag Care* 2011;17:41–48.
13. Fujimura Y, Tanii H, Saijoh K. Inpatient satisfaction and job satisfaction/stress of medical workers in a hospital with the 7:1 nursing care system (in which 1 nurse cares for 7 patients at a time). *Environ Health Prev Med* 2011;16:113–122.
14. Tataw DB, Bazargan-Hejazi S, James FW. Health services utilization, satisfaction, and attachment to a regular source of care among participants in an urban health provider alliance. *J Health Hum Serv Adm* 2011;34:109–141.
15. Sixma HJ, Calnan S, Calnan M, Groenewegen PP. User involvement in measuring service quality of local authority occupational therapy services: a new approach. *Int J Consum Stud* 2001;25(2):150-159.
16. World Health Organization. Process of translation and adaptation of instruments.; 2015. http://www.who.int/substance_abuse/research_tools/translation/en/.
17. Surveymonkey <https://www.surveymonkey.com/>

18. Willis GB. Cognitive Interviewing. A “how to” guide. Rockville: Research Triangle Institute, 1999.
19. Willis GB, Lessler JT. Questionnaire Appraisal System QAS-99. Rockville: Research Triangle Institute, 1999.
20. Sixma HJ, Boer De D, Delnoij D. Handboek CQ-index ontwikkeling: richtlijnen en voorschriften voor de ontwikkeling van een CQ-index meetinstrument. Utrecht: NIVEL, 2008.
21. Cronbach L J. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16,297-334.
22. Gadermann AM, Guhn M, Zumbo BD. Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Pract assess res eval* 2012;17(3):1-13.
23. Streiner LD, Norman GR. *Health Measurement Scales: A Practical Guide to Their Development and Use*. Fourth edition. Oxford: Oxford University Press, 2008.
24. Muthén LK, Muthén BO. *Mplus User’s Guide*. Seventh Edition. Los Angeles, CA: Muthén & Muthén, 1998-2012.
25. Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the fit of structural equation models: test of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research* 2003;8:23-74.
26. Yu CY. Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes.;2002 <http://www.statmodel.com/download/Yudissertation.pdf>:
27. Hibbard JH, Mahoney ER, Stockard J, Tusler M. Development and testing of a short form of the patient activation measure. *Health Serv Res* 2005;40(6):1918–1930.
28. Hibbard JH, Stockard J, Mahoney ER, Tusler M. Development of the Patient Activation Measure (PAM): conceptualizing and measuring activation in patients and consumers. *Health Serv Res* 2004;39(4):1005–1026.
29. Hofstede, G. *Dimensionalizing Cultures: The Hofstede Model in Context*. ORPC 2011, 2(1).
30. Meeuwesen L, van den Brink A, Hofstede, G. Can dimensions of national culture predict cross-national differences in medical communication? *Patient Educ Couns* 2009, 75(1):58-66.
31. Hofstede G, Hofstede GJ, Minkov M. *Cultures and organizations: software of the mind: intercultural cooperation and its importance for survival*. Third edition. New York: McGraw Hill, 2010.
32. Na’poles AM, Gregorich SE, Santoyo-Olsson J, O’Brien H, Stewart AL. Interpersonal processes of care and patient satisfaction: do associations differ by race, ethnicity, and language? *Health Serv Res* 2009;44:1326–1344.
33. Morales LS, Reise SP, Hays RD. Evaluating the equivalence of health care ratings by whites and Hispanics. *Med Care* 2000;38(5):517-527.
34. Ipsos. *International Healthcare Report Card: Citizen-Patients in 15 Countries Assess Improvement to Healthcare.*; 2013 <http://www.ipsos-na.com/download/pr.aspx?id=12775>
35. Social Information Centre and European Research. *Final report on the project on survey of patients and health care service providers*. Vilnius: Ministry of Health, 2012.

36. Rescorla L, Ivanova MY, Achenbach TM, Begovac I, Chahed M, Drugli MB, Emerich DR, et al. International epidemiology of child and adolescent psychopathology II: integration and applications of dimensional findings from 44 societies. *J Am Acad Child Adolesc Psychiatry* 2012;51(12):1273-1283.
37. Fang J, Wen C, Prybutok VR. An assessment of equivalence between Internet and paper-based surveys: Evidence from collectivistic cultures. *Qual Quant* 2014;48(1):493-506.
38. De Beuckelaer, A, Lievens, F. Measurement equivalence of paper-and-pencil and internet organisational surveys: A large scale examination in 16 countries. *Appl Psychol* 2009;58(2):336-361.
39. Cole MS. The Measurement Equivalence of Web-Based and Paper-and-Pencil Measures of Transformational Leadership: A Multinational Test Organ Res Meth 2006; 9(3):339-368.
40. Van De Looij-Jansen PM, De Wilde EJ. Comparison of web-based versus paper-and-pencil self-administered questionnaire: Effects on health indicators in Dutch adolescents. *Health Serv Res* 2008;43(5):1708-1721.
41. Jaipaul CK, Rosenthal GE. Are Older Patients More Satisfied With Hospital Care Than Younger Patients? *J Gen Intern Med* 2003;18(1):23-30.
42. Hargraves, JL, Wilson IB, Zaslavsky A, James C, Walker JD, Rogers, G, et al. Adjusting for patient characteristics when analyzing reports from patients about hospital care. *Med Care* 2001;39(6):635-641.

PART 4

Qualitative and quantitative benchmarking

CHAPTER 5

Development of a Benchmark tool for Cancer Centers; results from a pilot exercise

Anke Wind
Joris van Dijk
Isabelle Nefkens
Wineke van Lent
Peter Nagy
Ernestas Janulions
Tuula Helander
Francisco Nuno Rocha-Gonçalves
Wim van Harten

Submitted

Abstract

Background: Differences in cancer survival exist between countries in Europe. Benchmarking of good practices can assist cancer centers to improve their services aiming for reduced inequalities. The aim of the BENCH-CAN project was to develop a cancer care benchmark tool and yield good practice examples, contributing to improving the quality of interdisciplinary care. This paper describes the development of this benchmark tool and its validation in cancer centers throughout Europe.

Methods: A benchmark tool was developed and executed according to a 13 step benchmarking process. Indicator selection was based on literature, existing accreditation systems, and expert opinions. A final format was tested in eight cancer centers. Center visits were performed to verify information, grasp context and answer additional questions. Based on the visits, the benchmark methodology identified opportunities for improvement.

Results: The final tool consisted of 61 qualitative and 141 quantitative indicators, which were structured in an evaluative framework. Data from all eight participating centers showed inter-organization variability on many indicators, such as bed utilization and provision of survivorship care. Next, improvement suggestions for centers were made; 85% of which were agreed upon. Overall centers provided positive feedback regarding the benchmark tool.

Conclusion: A benchmarking tool for cancer centers was successfully developed and tested and is available in an open format. The tool allows comparison of inter-organizational performance. Improvement opportunities were successfully identified for every center involved and the tool was positively evaluated.

Introduction

The number of cancer patients is steadily increasing and, despite rapid improvements in therapeutic options, inequalities in access to quality cancer care and thus survival exist between different countries¹. These inequalities indicate room for improvement in quality of cancer care. Identifying good practices can assist cancer centers (CCs) in improving their services and can ultimately reduce inequalities. Benchmarking is an effective method for measuring and analyzing performance and its underlying organizational practices². Developed in industry in the 1930s, benchmarking made its first appearance in healthcare in 1990². Benchmarking involves a comparison of performance in order to identify, introduce, and sustain good practices. This is achieved by collecting, measuring and evaluating data to establish a target performance level, a benchmark. This performance standard can then be used to evaluate the current performance by comparing it to other organizations, including good-practice facilities³. Due to globalization, absence of national-comparators, and the search for competitive alternatives, there is an increasing interest in international benchmarking⁴. However, a study by Longbottom⁵ on 560 healthcare benchmarking projects, showed only 4% of the projects involved institutions from different countries. In literature, relatively few papers are published on healthcare benchmarking methods⁶. Moreover, to the best of our knowledge, there is no confirmed indicator set for benchmarking comprehensive cancer care. In 2013, the Organization of European Cancer Institute (OECI)⁷ therefore launched the BENCH-CAN project⁸, aiming at reducing health inequalities in cancer care in Europe and improving interdisciplinary comprehensive cancer care by yielding good practice examples. In view of this aim, a comprehensive international benchmarking tool was developed covering all relevant care related and organizational fields. In this study comprehensive refers to thorough, broad, including all relevant aspects - which is also a means to describe interdisciplinary, state of the art, holistic cancer care.

The aims of this study were (i) to develop and pilot a benchmark tool for cancer care with both qualitative and quantitative indicators, (ii) identify performance differences between cancer centers, and (iii) identify improvement opportunities.

Method

Study design and sample

This multi-center benchmarking study involved eight cancer centers (CCs) in Europe, six of which designated as a comprehensive cancer center (encompassing care, research and education) by the OECI⁹. Centers were located in three geographical clusters: North/

Western-Europe (n=2), Southern-Europe (n=3) and Central/Eastern-Europe (n=3). The benchmark tool was developed and executed according to the 13 step method by van Lent et al.⁶ (see Table 1). In short, the first five steps involve the identification of the problem, forming the benchmarking team, choosing benchmark partners and define their main characteristics, and identify the relevant stakeholders. Step 6 to 12 will be explained in more detail in the following paragraphs.

Table 1 Benchmarking steps developed by van Lent and application in this study

13 steps by van Lent	Application of the steps in this study
1 Determine what to benchmark	Comprehensive cancer care, structured through the domains of the BENCH-CAN framework such as People, Process, Product & Services, and Efficient (step 6).
2 Form a benchmarking team	International consortium existing of representatives from cancer centers, health R&D organisation, biomedical innovations consultancy company, and OECl.
3 Choose benchmarking partners	Cancer centers in Europe.
4 Define and verify the main characteristics of the partners	A mapping exercise of the external environment in which the cancer centers are located was performed.
5 Identify stakeholders	Four stakeholder groups were identified: patients, management, clinicians and researchers.
6 Construct a framework to structure the indicators	The framework is based on the European Foundation for Quality Management (EFQM) Excellence Model [10] and the adapted six domains of quality of the Institute of Medicine [11].
7 Develop relevant and comparable indicators	Indicators were retrieved from literature [12] and expert opinion.
8 Stakeholders select indicators	Stakeholders from the BENCH-CAN project and other experts from cancer centers provided feedback on the indicators.
9 Measure the set of performance indicators	Indicators were first pre-piloted in three centers to check clarity of the definitions and whether indicators would yield interesting information. Data collection phase was three months. Next, the three month during data collection phase was repeated for the other centers. A team performed a center visit to each pilot center to verify the data, to grasp the context and clarify any questions arising from the provided data.
10 Analyse performance indicators	The researchers compared the performance of the pilot cancer centers. Reports of this comparison were checked by the other members of the center visit team.
11 Take action: results are presented in a report and recommendations are given	For each participating cancer centre, a report was made containing the outcomes of the benchmark for all centers. Data was anonymized. Improvement recommendations were sent in a separate document.
12 Develop relevant plans	Pilot centers were asked to develop improvement plans for recommendations that they agreed with.
13 Implement the improvement plans	Outside the scope of this study.

Framework and indicators

As described in step 6 we developed a framework to structure the indicators. The European Foundation for Quality Management (EFQM)¹⁰ Excellence Model (comparable to the Baldrige model¹³) was used for performance-assessment and identification of key strengths and improvement areas¹⁴. Apart from the enabler fields, we adapted the Institute of Medicine domains of quality¹¹ for outcomes or results: effective, efficient, safe, patient-centered, integration and timely (Figure 1).

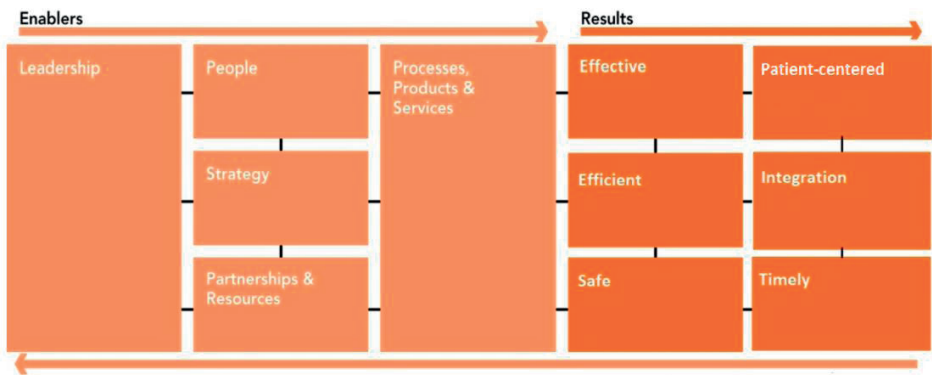


Figure 1 the BENCH-CAN framework.

Note: The enabler domains from the EFQM model describe factors that enable good quality care. The results domains adapted from the IOM domains of quality describe how good quality care can be measured.

Indicators (step 7) were derived from literature¹² and expert opinion. Existing assessments were used as basis for the benchmark tool¹⁶. Stakeholders of the BENCH-CAN project such as representatives from the European Cancer Patient Coalition, and clinicians and experts (such as quality managers) from cancer centers (OEI member centers, n=71) provided feedback to reach consensus on the final set of indicators to be used in the benchmark (step 8). For the financial and quantitative indicators this included the standardization of data collection to allow comparison between pilot centers and determining the level of detail for cost accounting.

Reliability and validity

A priori stakeholder involvement was used to ensure reliability and validity⁶. After collecting the indicators in step 9, the validity of the indicators was checked using feedback from the pilot centers based on three criteria^{16, 17}: 1) definition clarity, 2) data availability and reliability, 3) discriminatory features and usability for comparisons.

Indicator refinement and measurement

The indicators were pre-piloted in three centers to see whether the definitions were clear and the indicators would yield relevant, discriminative information. Based on this pilot, we decided to add and remove indicators, and refine definitions of some indicators. After refinement, the resulting set of 63 qualitative indicators and 193 quantitative indicators was measured in the five remaining centers. The pre-pilot centers submitted additional information on the added indicators in order to make all centers comparable.

We collected data from the year 2012 and each pilot center appointed a contact person who was responsible for the data collection within the institute and the delivery of the data to the research team. After an initial analysis of the data, a one-day visit to each pilot center was performed to verify the data, grasp the context and clarify questions arising from the provided data. The visits were also used to collect additional information through semi-structured interviews and to acquire feedback on the benchmark tool.

Analysis

Two methods were used to compare the qualitative and quantitative data. A deductive form of the Qualitative Content Analysis was used to analyze the qualitative data¹⁸. This method contains eight steps which are described in Table 2.

Table 2 steps Qualitative Content Analysis

Step	Action
1	Read through the benchmark data (transcripts) and make notes
2	Go through the notes and list the different types of information found
3	Read through the list and categorize each item (domains of the framework were used as main categories)
4	Repeat the first three stages again for each data transcript
5	Collect all of the categories or themes and examine each in detail and consider it's fit and its relevance
6	Categorize all data (all transcripts together) into minor and major categories/themes
7	Review all categories and ascertain whether some categories can be merged or sub-categorized
8	Return to original transcripts and ensure that all the information has been categorized

Quantitative data was first checked for consistency and correctness, and all cost data was converted into euros and adjusted for Purchasing Power Parity¹⁹. In addition, data was normalized when necessary to be able to compare different types and sizes of centers. Used normalizations were: 1) openings hours of departments, 2) number of inpatient beds,

3) number of inpatient visits, and 4) number of full-time equivalent (FTE). All data was summarized and possible outliers were identified. Outliers were discussed with the relevant centers to elaborate on the possible reasons for the scores.

To ensure validity, a report with all data (qualitative and quantitative) was sent to the pilot centers for verification. Not all centers were able to provide all data, so for some indicators centers are missing, as we did not use imputation. Data is structured according to the adapted domains of quality from the IOM; effective, efficient, safe, patient-centered, and timely.

Improvement suggestions

After comparison of all quantitative and qualitative data, three researchers independently identified improvement opportunities for each center. Improvement suggestions or opportunities (at least three per center) were only mentioned for those areas where the researchers felt the center could actually make the improvement without being restricted by for example regulations. Based on these improvement suggestions, if in agreement, pilot centers developed improvement plans.

Results

Reliability and validity

Ten indicators were deemed irrelevant (such as sick leave) and were removed after the pre-pilot. Nineteen indicators were added based on evaluation criteria and feedback. Several indicator definitions were clarified. The final pilot-list contained 63 qualitative indicators and 193 quantitative indicators. After the pilot data collection, a secondary evaluation of the definition clarity, data availability, data reliability and discriminative value was performed. This re-valuation resulted in a final set of 61 qualitative indicators and 141 quantitative indicators that were deemed suitable for wider use in benchmarking cancer centers.

Results of the benchmark

The performances of the participating centers varied on many indicators, of which a selection is shown in Table 3 and described below. Organizations are anonymized. The results are structured according to the adapted domains of quality¹¹.

I Effective

The majority of centers register crude mortality rates of their patient groups (n=6) as shown in Table 3. Only Institute A publishes this rate. Another type of mortality, 30-day surgical mortality, was not registered in center B, C and G. Centers also reported difficulties with providing novel technologies and therapies limiting their ability to provide the optimal care for patients.

II Efficient

Medical efficiency

The medical efficiency, defined as the use of medical production factors to gain desired health outcome with a minimum waste of time, effort, or skills, greatly varies between the participating centers as shown in Figure 2. Center G scores high (ratio of 7), whereas center C has a low number of daycare treatments (ratio 0.3) in relation to their inpatients visits compared to the other centers.

5

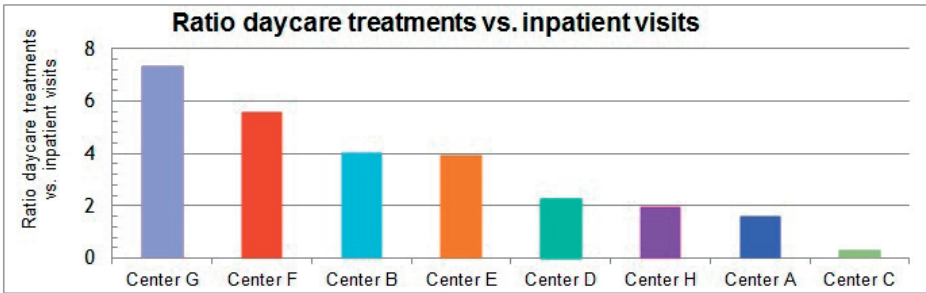


Figure 2 Number of daycare treatments in relation to the number of inpatient visits.

The utilization of beds differs between centers, as shown in Figure 3. Especially center C, G and H have a relatively low inpatient bed utilization. Similarly, a large variation in utilization of the daycare beds is observed. Center E has a high daycare bed utilization, but scores average in the ratio between daycare treatments/inpatient visits. In contrast, center G also had a relatively high number of daycare treatments but a lower utilization.

Input efficiency

Number of scans per radiology device varies between centers, as shown in Figure 4. Center D scores high on the efficiency of MRI (4462 scans per MRI) X-ray (7703 scan per X-ray machine), and CT (13,836 scans). Center H scores high on the efficiency of MRI and CT. Center E has outsourced their MRI and no data was available from center G considering X-rays.

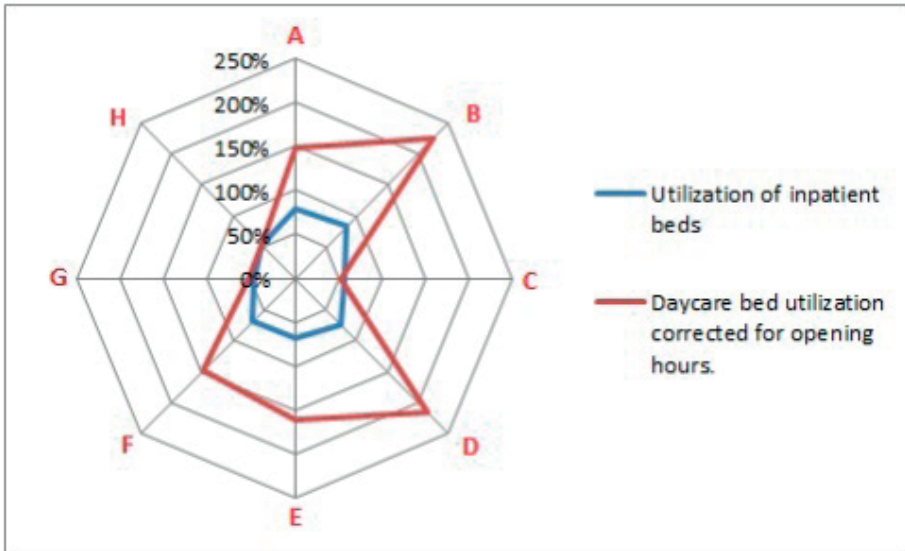


Figure 3 Inpatient and day-care bed utilization

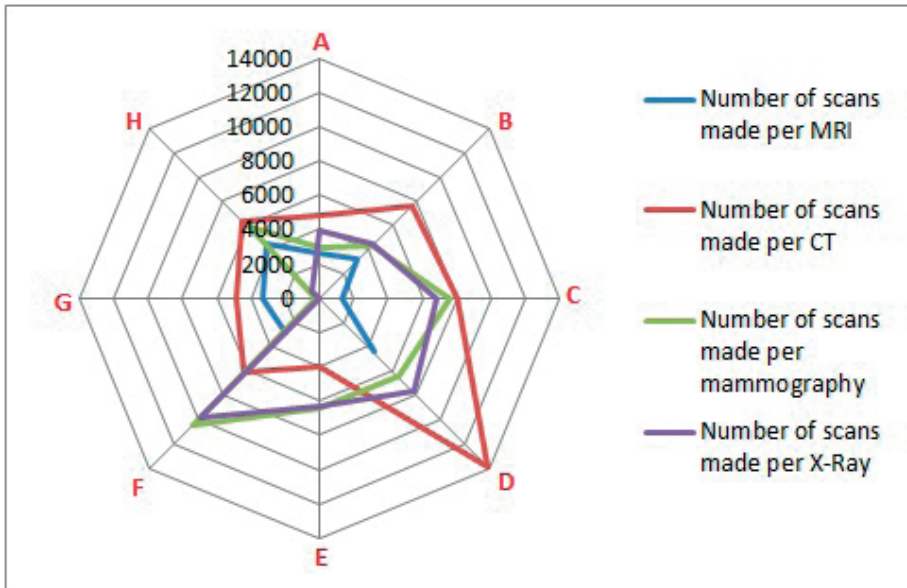


Figure 4 Total number of scans made per device in one year

III Safe

Center A has a safety management system which is audited annually by an independent external agency. Prospective risk assessments are performed in center A before implementing new treatments, new care pathways or critical changes in key processes. Center B divided risk management into general risk management (e.g. risks of fire) and clinical risk management (e.g. transfusion risks and medication errors). Institute H adopted the “International Patient Safety Goals” (IPSG) issued by the Joint Commission International²⁰. Most centers (n=7) have an institution-wide reporting systems that registers different types of adverse events: near miss; incident; adverse event; sentinel event. Only doctors can make official notifications of a medical error in institute E and nurses cannot report an incident directly. Center G uses a system that generates reports for patient satisfaction, patient safety and patient complaints. Near misses should be reported in institute H according to their procedures, but in practice only actual events are reported. For more information on the domain of safety see Table 3.

IV Patient-centered

Although all center have some type of contact-person for patients, none had an official case-manager for all patient pathways. In institute A and D a formalized inclusion of patients in the strategy development is present. Other centers reported to collaborate with external patient organizations to represent patients. All centers provide some care for cancer survivors, however, only center A has an extensive survivorship program in-house with a dedicated budget. Center G also reports to have a budget for survivorship care (e.g. Psychosocial support). For more information on patient centeredness see Table 3.

V Timely

For seven centers the waiting times are set by the government (see Table 3). Institute A indicated that they encountered difficulties in meeting the maximum waiting time for some types of surgeries. The maximum waiting times are input for negotiations with healthcare insurers, and have potential influence on the funding for center A. Center H reports waiting times to the regional government who uses this data to adjust the amount of services offered by the regional healthcare-system. Possible reasons mentioned for long waiting times are high demand of patients for diagnostic tests and insufficient staff. The largest variation between institutes occurred in overall waiting time before first visit, which varied between 1.5 and 21.8 days.

Table 3 Profiles of the cancer centers against a selection of indicators. For each domain a selection of indicators and their outcomes is presented.

Topic	Center							
	A	B	C	D	E	F	G	H
Type of center	Comprehensive CC	Comprehensive CC	Clinical CC	Comprehensive CC	x	Comprehensive CC	Comprehensive CC	Comprehensive CC
<i>Effective</i>								
Survival/mortality-registration	Gross mortality ratios (2.08%) per tumor type and -stage are published on the website.	Gross mortality ratio per number of discharges is registered (9.5%) and Risk-Adjusted Mortality (e.g. age, sex, tumor etc.) (0,8836)	Gross mortality rate is registered	Gross inpatients mortality ratio (2.0%) is registered.	Gross mortality ratio is registered	Mortality ratios are collected only for patients who had surgery	Mortality ratios are registered by the Cancer Registry	Not recorded at institutional level
Colo-rectal surgery mortality (within 30 days)	0	Not registered	Not registered	0	Unknown	Unknown	Not registered	Unknown
Innovative technology and therapies	Budgets are not keeping up with the developments in technology's and therapies and the increase in costs that comes with this	Targeted biological treatments are not covered. New therapies will lead to a negative balance. This will cause a big challenge in the future.	It is hard to introduce the latest technologies because they are not reimbursed. It might take 2-3 years to arrange reimbursement.	- New treatments are very expensive, not always paid for by insurance - Healthcare services are not able to provide state of the art treatment to all patients	Not foreseen as a major challenge.	Biosimilar, and generics are a challenge. Several expensive drugs come to the market. Need for biomarkers	All drugs are reimbursed, but there is a 6 month wait before the reimbursement comes. The national system is slower. ESMO guidelines are followed.	Unknown

Table 3 Profiles of the cancer centers against a selection of indicators. For each domain a selection of indicators and their outcomes is presented. (continued)

Topic	Center							
Type of center	A	B	C	D	E	F	G	H
Risk management	<ul style="list-style-type: none"> - Quality, Occupational health and environment service - Prospective risk assessments by MDT - Staff training - Dedicated emergency managers, available at the institute 24/7. 	<ul style="list-style-type: none"> - Occupational health and Overall Risk Management Service - Risk management system: general risk management and clinical risk management 	<ul style="list-style-type: none"> - Annual risk factors analysis and prevention action plan. - Evaluation of implementation end each year. 	<ul style="list-style-type: none"> - Several protocols for risk management - New employees undergo a thorough medical examination to decide if they are fit for employment. - Emergency Assessments - External Risk Assessments 	<ul style="list-style-type: none"> - Department for prevention and control of nosocomial infections. - Staff involved is formally trained and informed about the regulations 	<ul style="list-style-type: none"> - Health and Safety function deals with prevention and management of staff safety - The medical physics unit manages the prevention and control of radioactive risks for patients and staff. 	<ul style="list-style-type: none"> - Strategy and program measurements are part of the institute and available on the website 	<ul style="list-style-type: none"> - Risk management plan overseeing clinical management and risk - Medical Directorate coordinates all activities of preventive medicine and environmental health - Protection and Prevention Service.

Table 3 Profiles of the cancer centers against a selection of indicators. For each domain a selection of indicators and their outcomes is presented. (continued)

Topic		Center							
Type of center	A	B	C	D	E	F	G	H	
Adverse events	<ul style="list-style-type: none"> - Safe incident reporting system - Every employee of the institute has access to this system and can report incidents - Incidents are analyzed by relevant departments and feedback is given to the detector and an advice for improvement measures to the manager - Patient safety committee monitors hospital-wide 	<ul style="list-style-type: none"> - Everybody is allowed to make a notification of an adverse event - Form allows notifier to make recommendations to prevent this situation - All adverse events are collected by the CG service - Quarterly reports on website 	<ul style="list-style-type: none"> - Every staff member can inform the head of clinical department about potential threats. - Medical staff fill a Report of discrepancies or adverse event and inform Head department - Risk and prevention factors evaluation group prepare annual Risk factors and prevention actions plan for the Institute 	<ul style="list-style-type: none"> - Adverse events and near misses reported on departmental level and institutional level - Radiotherapy department has an IT system for reporting irregularities and adverse events - Analysis of irregularities is reported on the shared drive and as well as patient complaints analysis 	<ul style="list-style-type: none"> - Patient incidents are reported to the complaints registry thereafter addressed to the Management, and after evaluation addressed to the Ethics Council. - Every incident is electronically recorded and sent to the public health authorities quarterly. - In every department there is a local registry for incidents, reporting on paper. - Medical errors: if nurse is involved she reports it to the doctor, if the doctor finds it serious, sends it to the medical committee. Nurses cannot report an incident directly. 	<ul style="list-style-type: none"> - Institutional Incident Reporting program. The events are reported on a voluntary and not-anonymous base; - Summary report with the strategic directorate. - Reporting usually on departments have it electronically - Patients can report incidents as well 	<ul style="list-style-type: none"> - System that registers and generates reports for patient satisfaction, patient safety, patient complaints. 	<ul style="list-style-type: none"> - The "incident reporting module" is available for staff in the institutional intranet and a counseling service is available. - Sheets are anonymous. - In practice only the actual events are reported. 	<ul style="list-style-type: none"> Comprehensive CC

Table 3 Profiles of the cancer centers against a selection of indicators. For each domain a selection of indicators and their outcomes is presented. (continued)

Topic	Center	A	B	C	D	E	F	G	H	
Type of center		Comprehensive CC	Comprehensive CC	Clinical CC	Comprehensive CC	x	Comprehensive CC	Comprehensive CC	Comprehensive CC	
<i>Patient centered</i>										
Case manager		<p>Case managers for head and neck cancer, breast cancer and melanoma</p> <ul style="list-style-type: none"> - Case manager usually nurse specialist or physician assistant - Contact person for questions about their treatment etc. - Case managers redirects patients to supportive services - All patient receive leaflet with all information on how to reach their case manager. 	<p>The patient can contact:</p> <ul style="list-style-type: none"> - The physician at an appointment; - A member from the nursing staff that follows the patient; - The social worker and/ or the psychologist that follows the patient - An administrative office that is called "Patient Support Office" 	<ul style="list-style-type: none"> - Patients are informed by treating physicians - Nurses are also indicated as contact persons 	<ul style="list-style-type: none"> - Contact person for each patient, by law, is their responsible specialized MD - a designated Case Manager exists at the Same Day Surgery Unit - Tasks of the Case Manager: - Schedule procedure and inform patient in writing - Collect the necessary paperwork and provide them to the anesthesiologist - Provide patients with discharge records 	x	<p>The contact person for each patient is the physician (medical oncologist, surgeon etc.). The discharge letter contains all the information the patient needs, including a telephone number for the doctor in charge of the case, and in some cases, the number of the nurse of the day hospital clinic.</p>	<p>There is a contact person/team – usually nurses – for each person. A contact number and all the needed pieces of information are given to the patient before discharge.</p> <p>In the breast unit there is one case manager to act as a contact person.</p>	<ul style="list-style-type: none"> - "personal support nurse" there in practice, not officially - when needed there is option contact physician in charge 	<p>Basically there is a physician as contact person for patients but a specific case manager is not appointed for each patient.</p> <p>In clinical trials a case manager, often a research nurse, is appointed to coordinate appointments (exams, visits, follow-up) as primary contact for patients.</p>

Table 3 Profiles of the cancer centers against a selection of indicators. For each domain a selection of indicators and their outcomes is presented. (continued)

Center		A	B	C	D	E	F	G	H
Topic	Type of center	Comprehensive CC	Comprehensive CC	Clinical CC	Comprehensive CC	x	Comprehensive CC	Comprehensive CC	Comprehensive CC
Patient involvement (care)		<ul style="list-style-type: none"> - Patient portal: access to their full medical file - Patients can choose their own doctor - Patient participation in decision-making regarding treatment depends on type of health insurance. Not all treatments covered by all types of health insurance 	<ul style="list-style-type: none"> - Patients participate in the multidisciplinary appointments - Physician discusses alternative treatments with the patient - Patients can choose their own doctor 	<ul style="list-style-type: none"> - Patients receive information about their diagnostic and treatment processes and must sign Statement of Faith. - Information to patients and other entities is provided according to procedure - Patients can choose their own doctor 	<ul style="list-style-type: none"> - Patients can choose their own doctor - Patients have the right to a Patient Representative - Patients have the right to review their patient documentation and pose any questions 	<ul style="list-style-type: none"> - Patients are explained the purpose of every diagnostic or therapeutic procedure and alternatives where available. The patient signs informed consent forms at every major step of the care pathway. - Patients have the right to review their patient documentation and pose any questions - Patients can choose their own doctor 	<ul style="list-style-type: none"> - Oncologist and/or the multidisciplinary team propose the diagnostic and treatment processes. If the patient decides to accept the proposal, he/she has to sign the informed consent. - Patients have the right to review their patient documentation and pose any questions 	<ul style="list-style-type: none"> - MDT makes a recommendation, and physician decides with the patient. Patient receives most of the patient documentation automatically (and largely in e-format). - All patients (citizens) have access to their own patient records 	<ul style="list-style-type: none"> - Patients are involved in the discussion of their diagnostic and treatment plan during visits performed by physicians. Signature of the informed consent is intended as the acceptance of the patient to the proposed treatment.

Table 3 Profiles of the cancer centers against a selection of indicators. For each domain a selection of indicators and their outcomes is presented. (continued)

Topic	Center							
	A	B	C	D	E	F	G	H
Type of center	Comprehensive CC	Comprehensive CC	Clinical CC	Comprehensive CC	x	Comprehensive CC	Comprehensive CC	Comprehensive CC
Patient involvement (strategy)	- Patients are represented by the Patient Advisory Board (PAB). - PAB promotes the common interests of patients and gives solicited and unsolicited advice to the Board of Directors	- Patient Support Office to establish connection between users and Board and streamline communication of the patient and the professionals	- Patient representatives are not involved that much due to the lack of legislation and to the fact that representatives are not very active.	- The Patient Representative participates in the Board of Directors' meeting each week and communicates directly with the Board.	- Patients can participate in the institute through offering suggestions or filing complaints/suggestions during their care in the institute.	- Patients' representatives are part of the Patient Education working group and proactively propose improvements to services	- For some brochures patients are asked for opinion on design. - Patient Advisory Board, heavily involved in developing services and care. - There is a good relationship with the national Cancer Society.	- Collaboration with external patient organizations that address patients' priorities and needs.
Survivors	- The institute offers multidisciplinary rehabilitation - General rehabilitation and special rehabilitation program for patients with head and neck cancer. - % Annual budget for survivorship programmes: 0.6%	- Website under development: "I Have Cancer" - Groups related with the institute that give support to patients. - % Annual budget for survivorship programmes: 0	-Rehabilitation treatment -Cancer Information Center -Patient's school -Social services and other rehabilitation services. - % Annual budget for survivorship programmes: 0	- The survivorship programs are organized in collaboration with the League Against Cancer. - % Annual budget for survivorship programmes: 0	- Psychological support is available in the Institute for patients and their families. - Follow up program for survivors - % Annual budget for survivorship programmes: 0	- The institute has a specific clinic for long term cancer survivors and cancer-free patients and a dedicated patient organization. - % Annual budget for survivorship programmes: 0	- Support unit. - Physiotherapy, Psychosocial support, Nutrition, Peer support, Sexuality, social support, support for families with children - % Annual budget for survivorship programmes: 0.4%	- Chaplain and/or social workers and/or psychologists -Multidisciplinary collaboration to support cancer patients and survivors - % Annual budget for survivorship programmes: 0

Table 3 Profiles of the cancer centers against a selection of indicators. For each domain a selection of indicators and their outcomes is presented. (continued)

Topic	Center							
	A	B	C	D	E	F	G	H
Type of center	Comprehensive CC	Comprehensive CC	Clinical CC	Comprehensive CC	x	Comprehensive CC	Comprehensive CC	Comprehensive CC
<i>Timely</i>								
Waiting and throughput registration	- Access and throughput standards (maximum times) set by government - Dashboard quality and safety reports quarterly on waiting times - Waiting times published on website - Pass standards entry level requirement for negotiations with insurance companies	- Maximum set by government. - Reported in annual report	- Maximum waiting times are available in Institute information board.	- Waiting times are kept according to government decree	- Maximum set by government. - There are currently no standard means of measuring waiting and throughput times.	- The center records waiting times at two institutional levels: 1- Regional: 2- National: Ministry of Health oversight.	-Waiting times on the website for several tumors: breast, prostate and bowel cancers (more to be added)	- Waiting and throughput times are recorded and published on website. -Regional level maximum waiting times. -No consequences regarding reimbursement.
Average overall waiting time before first visit	9.1 days	1.54 days	12 days	7 days	not available	21.8 days	Unknown	9.6 days
Average waiting time from first visit to diagnosis	9.1 days.	Majority patients already have a diagnosis when they come to the first visit.	20 days	14 days	Average not available, depends on pathology result (availability for pathology result is established by law and is maximum 30 days).	Not available	Unknown	Not available

Table 3 Profiles of the cancer centers against a selection of indicators. For each domain a selection of indicators and their outcomes is presented. (continued)

Topic	Center							
	A	B	C	D	E	F	G	H
Type of center	Comprehensive CC	Comprehensive CC	Clinical CC	Comprehensive CC	x	Comprehensive CC	Comprehensive CC	Comprehensive CC
Average waiting time diagnosis-start treatment	Not available	6.53 days	Not available	7 days	Not available	19.7 days	Unknown	12.6 days

Improvement suggestions

Table 4 describes examples of improvement suggestions per pilot center and resulting improvement plans. Improvement suggestions varied from broader processes such as the involvement of patients in the care process, to specific recommendations (e.g. measure staff satisfaction). Adoption of case managers was a frequently mentioned improvement suggestion. Regarding the suggestion to improve patient participation in the organization, center C only partially agreed as they stated “not all patients want to be involved”. Center A felt a complication registry was mainly useful per discipline and therefore partly agreed with the suggestion to implement an institution-wide complications registry. Out of the total improvement suggestions, pilot centers agreed with 85% and partially agreed with 15%. For center G improvement suggestions were given, however, no improvement plan was received.

Table 4 Improvement suggestions, response and planned actions.

Suggestions	Institute/ Agreement	Comments	Actions to be taken identified by pilot centers
Case managers for (all) patients/all tumor types	A/ Agree	This is important but requires specialized staff, currently shortage of this specialized staff	Currently there are official case managers for 5 tumor types, development of case managers for other tumor types will follow these examples
	B/ Agree	Case managers are an important tool in patient treatment so we want to improve this area.	Already part of the strategic vision so no extra actions need to be taken
	C/ Agree	It would be good to have case manager-the process has to be more organized, more patient oriented	Educate the right staff and dedicate them as case-manager
	F/ Agree	A case manager for each pathway will be formally identified	Define clear role and responsibility for the case manager for each pathway/tumor type
Develop more support for survivors	B/ Agree	With the increase of the survival rates in cancer patients we recognize that this is an area that we must improve.	A website where survivors can exchange information and experiences was already launched. A portal for survivors, amongst others, is under development.
	D/ Agree	Survivorship programs are provided mostly by the patient organizations.	Develop own survivorship program for the institute and further formalize the collaboration with patient organizations in survivorship programs.

Table 4 Improvement suggestions, response and planned actions. (continued)

Suggestions	Institute/ Agreement	Comments	Actions to be taken identified by pilot centers
Increase patient participation in the care process	B/ Agree	We are already working in this area.	An area on the website is under development where patient can access: future appointments, exams results and requisitions, among other clinical and administrative information. The portal that is under development will have one tab containing the patients targeted information.
Improve patient participation in the organization/strategy development	C/ Partially agree	Patients have to be involved. However not all patients want to be involved.	All patients have to pass the MDT. And after discussion-take a decision on whether to participate. This participation has to be organized.
Develop a structured, institute wide adverse events analysis system	C/ Agree	Its absolutely necessary to check and register these events. Important for the quality of care.	Depends on the staff. Sometimes they hide the information
Measure staff satisfaction	C/ Agree	Staff has to be honest and not just provide the socially accepted answers	Regular discussions with staff, improve existing questionnaires
Central complication registry may be useful	A/ Partially agree	Complication registration is mainly useful for healthcare professionals, current registration system allows health professionals to see the data important for them, per discipline. Central registration could be useful to annually analyze the results and look at the trends compared to trends in for example new patients. The national institute for Clinical Auditing registers complications as well on a national level	Create system that can extract data from existing system or develop new registration system
Implement Computerized Physician Order Entry	E/ Agree		Electronic prescriptions are currently being implemented: in the short term there will be 2 pilot actions for 2 departments. It is currently planned to include treatment details (chemotherapy data), transfusions and clinical trial participation.
	F/ Partially agree	This is an important and urgent objective, but unfortunately due to regional restrictions the institute cannot be proactively proceed	

Table 4 Improvement suggestions, response and planned actions. (continued)

Suggestions	Institute/ Agreement	Comments	Actions to be taken identified by pilot centers
Improve patient transition protocol	H/ Agree	We should improve the network with other hospitals/institutes, care facilities and general practitioners (GPs) as well	Improvement of the electronic chart (e-chart): at regional level, the first attempted has been made within the region
Assess and improve inpatient bed utilization	H/ Partially agree	Inpatient bed utilization is planned and regulated at regional level	

Discussion

In this study, we developed a benchmark tool to assess the quality and effectiveness of comprehensive cancer care consisting of 61 qualitative indicators and 141 quantitative indicators. The tool was successfully tested in eight cancer centers to assess its suitability for yielding improvement suggestions and identifying good practices.

The benchmark data showed performance differences between cancer centers which led to improvement suggestions/opportunities for all participating centers. In general, the indicators revealed well-organized centers. However, there were indicators on which centers performed less. For example, not all centers register mortality rates and it is unclear whether these rates, when registered, are made public. Nevertheless, there is broad consensus that public reporting of provider performance can be an important tool to drive improvements in patient care²¹. An indicator on which only two centers performed well was the offering of in-house survivorship care by having a dedicated budget. An advantage of follow-up taking place in cancer centers is that it is comfortable for patients and provides continuity of care²². However, it is debatable whether offering this kind of care should be the responsibility of cancer centers, as multiple pilot centers already indicated to have tight budgets.

Large variety existed in the domain of efficiency between centers. This variety was only partly related to differences in healthcare systems, leading to multiple improvement suggestions. For example, center C, G and H had a relatively low inpatient bed utilization, which is likely to be less cost-efficient. Center G had a high number of daycare treatments but a lower bed utilization, possibly indicating a utilization loss. A higher ratio indicates efficient use of beds and chairs and, hence, most likely also staff use. Centers C and D might have a surplus of daycare beds and chairs. Wind et al.²³ showed that having fewer beds has no association with low financial performance and could indeed improve efficiency.

Another important improvement area was patient-centeredness. Specifically in the area of case management for which all centers agreed that it was necessary to implement or expand. Case management is an organizational approach used to optimize the quality of treatment and care for individuals within complex patient groups²⁴. However, centers indicated that implementing or extending these case managers will take a long time and therefore categorized this as mid-term (2-5 years) or long-term (6-10 years) goals.

5 Several assumptions underpinned this study. First, although we thoroughly searched the literature and existing quality assessments to identify indicators for the initial list, some suitable indicators may have been missed. Identifying suitable outcome indicators was more challenging than for example process indicators due to the difference in case-mix and healthcare system and financing. We tried to minimize this influence by including a large group of experts from various fields who had affinity with development and management of cancer centers and quality assessment in cancer care. We continuously modified the set of indicators in response to feedback on their relevancy, measurability and comparability by the pilot centers. An advantage of this approach is that the indicators benchmark what the cancer centers want to know, which can increase adoption of the benchmark format as a tool for future quality improvement.

Second, the tool was only tested once in eight European cancer centers. This makes it impossible to say whether the benchmark actually led to quality improvements. Consequently, future research should evaluate the implementation of improvement plans to investigate whether the benchmark actually leads to quality improvement. In addition, future inclusion of more centers will allow to assess the actual discriminative capabilities of the indicator set. The benchmark tool was successfully applied in eight European countries with different wealth status. Although differences in healthcare systems and social legislation unavoidably led to differences in nature and availability of data, comparison still revealed relevant and valuable recommendations for all centers. We mainly achieved this by correcting for size, case-mix and type of healthcare reimbursements.

Finally, due to the extensive scope of indicators, it was difficult to go into detail for each topic. A benchmark focused on a single domain would allow to yield more profound information and more specific improvement suggestions and good practices. Future research is therefore advised to focus on specific domains of the BENCH-CAN framework, such as strategy and effectiveness, to gain a more profound understanding of the processes behind the performance differences, enabling a better comparison and more applied improvement recommendations.

Multiple lessons were learned from benchmarking cancer care in specialized centers throughout Europe. First, representatives of the pilot centers indicated that international projects such as these can increase awareness that performance can be improved and promote the notion that countries and centers can learn from each other. Identifying successful or good-practice approaches can assist hospitals in improving their services, and reduce inequalities in care provision raising the level of oncologic services across countries. Pilot centers did however indicate not to be able to implement all suggestions or good practices due to socio-economic circumstances. Second, learning through peers enabled cancer centers to improve their performance and efficiency without investing in developing these processes separately. A frequently mentioned comment was the casual, non-competitive atmosphere which led to an open collaboration. Involvement of key stakeholders from the centers at the start of the benchmark is highly recommended to develop interest, strengthen commitment, and ensure sufficient resources which not only accommodates a successful benchmark but also ensures implementation of the lessons learned.

In conclusion, we successfully developed and piloted a benchmark tool for cancer centers. This study generated more insight into the process of international benchmarking, providing cancer centers with common definitions, indicators and a tool to focus, compare and elaborate on organizational performance. Results of the benchmark exercise highlight the importance of an accurate description of underlying processes and understanding the rationale behind these processes. The tool allowed comparison of inter-organizational performance in a wide range of domains, and improvement opportunities were identified. The tool and the thereof derived improvement opportunities were positively evaluated by the participating cancer centers. Our tool enables cancer centers to improve on quality and efficiency by learning from good practices from their peers instead of reinventing the wheel.

References

1. De Angelis R, Sant M, Coleman MP, et al. Cancer survival in Europe 1999–2007 by country and age: results of EURO-CARE-5—a population-based study. *The lancet oncology* 2014; 15(1): 23-34.
2. Ettorchi-Tardy A, Levif M, Michel P. Benchmarking: a method for continuous quality improvement in health. *Healthcare policy*. 2012; 7(4): e101.
3. Joint Commission: Benchmarking in Health Care. Joint Commission Resources; 2012.
4. Gudmundsson H, Wyatt A, Gordon L. Benchmarking and sustainable transport policy: learning from the BEST network. *Transport Reviews*. 2005; 25(6): 669-690.
5. Longbottom D. Benchmarking in the UK: an empirical study of practitioners and academics. *Benchmarking. An International Journal* 2000; 7(2): 98-117.
6. van Lent W, de Beer R, van Harten W. International benchmarking of specialty hospitals. A series of case studies on comprehensive cancer centres. *BMC Health Services Research* 2010; 10: 253.
7. www.oeci.eu Accessed 14 May, 2013
8. www.oeci.eu/benchcan Accessed 18 December, 2013
9. <http://oeci.selfassessment.nu/cms/> Accessed 7 August, 2013
10. <http://www.efqm.org/the-efqm-excellence-model> Accessed 15 October, 2013
11. Institute of Medicine. Crossing the quality chasm: a new health system for the 21st century. Washington DC: National Academy Press, 2001.
12. Thonon F, Watson J, Saghatchian M. Benchmarking facilities providing care: An international overview of initiatives. *SAGE open medicine*. 2015; 3
13. MALCOLM BALDRIGE NATIONAL QUALITY AWARD (MBNQA) <http://asq.org/learn-about-quality/malcolm-baldrige-award/overview/overview.html> Accessed 28 December, 2016
14. Hakes C. The EFQM Excellence Model for assessing organizational performance. Van Haren, 2007.
15. Wind A, Rajan A, van Harten W. Quality assessments for cancer centers in the European Union. *BMC Health Services Research* 2016; 16: 474.
16. De Korne D F, Sol K J C A, van Wijngaarden J D H et al. Evaluation of an international benchmarking initiative in nine eye hospitals. *Health Care Management Review*. 2010; 35: 23
17. Cowper J, Samuels M. Performance benchmarking in the public sector: the United Kingdom experience. *Benchmarking, Evaluation and Strategic Management in the Public Sector*. Paris: OECD, 1997.
18. Zhang Y, Wildemuth BM. Qualitative Analysis of Content. *Applications of Social Research Methods to Questions in Information and Library Science*, 2016 318. https://www.ischool.utexas.edu/~yanz/Content_analysis.pdf Accessed September 15, 2016
19. Purchasing Power Parities - Frequently Asked Questions OECD. <http://www.oecd.org/std/prices-ppp/purchasingpowerparities-frequentlyaskedquestionsfaqs.htm> Accessed January 6, 2017
20. "International Patient Safety Goals" (IPSG) issued by the Joint Commission International. <http://www.jointcommissioninternational.org/improve/international-patient-safety-goals/> Accessed 16 November, 2016

21. Joynt KE, Orav EJ, J Zheng, Jha AK. Public Reporting of Mortality Rates for Hospitalized Medicare Patients and Trends in Mortality for Reported Conditions. *Ann Intern Med.* 2016; 165(3): 153-60
22. Models of Long-Term Follow-Up Care. American Society for Clinical Oncology. <https://www.asco.org/practice-guidelines/cancer-care-initiatives/prevention-survivorship/survivorship/survivorship-3> Accessed 28 August, 2016
23. Wind A, Lobo MF, van Dijk J, et al. Management and performance features of cancer centers in Europe: A fuzzy-set analysis. *JBR.* 2016; 69(11): 5507–5511
24. Wulff CN, Thygesen M, Søndergaard J, Vedsted P. Case management used to optimize cancer care pathways: a systematic review. *BMC health services research.* 2008; 8(1): 1

CHAPTER 6

Benchmarking cancer centers: from care pathways to Integrated Practice Units

Anke Wind

Francisco Nuno Rocha-Gonçalves

Edit Marosi

Lucia Da Pieve

Monica Groza

Marco Asioli

Marco Albini

Wim van Harten

Submitted

Abstract

Background: Structuring cancer care in pathways can reduce variability in clinical practice and improve patient outcomes. International benchmarking can aid centers with regard to development, implementation and evaluation. A further step in the development of multidisciplinary care is organizing care in Integrated Practice Units (IPU), encompassing the whole pathway and relevant organizational aspects. Research on this topic is however limited. This paper describes the development and results of a benchmark tool for cancer care pathways and exploration of IPU development in cancer centers.

Methods: The benchmark tool was developed according a 13 step benchmarking method. The tool was piloted in seven cancer centers in Europe. Centers provided data and site visits were performed to grasp the context and clarify additional questions. Benchmark data was structured into pathway development and evaluation and assessed against key IPU features.

Results: Benchmark results showed that most centers have formalized multidisciplinary pathways, teams differed in composition and almost twofold differences were found in mammography use efficiency. Improvements suggestions included to position pathways formally and structurally evaluate outcomes in a sufficient high frequency. Based on the benchmark three centers indicating to have a breast IPU were scored differently on implementation. Overall we found centers in Europe are in various stages of development of pathways and IPUs, varying from an informal pathway structure to a full IPU type of organization.

Conclusion: A benchmark tool for care pathways was successfully developed and tested and is available in an open format. The tool allows the assessment of pathway organization and can be used to assess the status of IPU development. Improvement opportunities were identified for the organization of care pathways and development towards IPUs. Three centers are in varying degrees of implementation and can be characterized as breast IPU. Organizing cancer care in an IPU could yield multiple performance improvements.

1. Background

Healthcare-systems struggle with rising costs of cancer care¹ and are increasingly under pressure to deliver high-quality services². Cancer care often takes place across different settings and involves health-professionals from multiple disciplines³. High-quality multi-disciplinary care can be compromised by inadequate coordination between health-professionals⁴. It has been shown that the implementation of care pathways reduces variability in clinical practice and improves outcomes⁵. Clinical/Care pathways, it's nomenclature varying from critical pathways, integrated care pathways, case management plans to care maps, are used to systematically manage a patient focused care program⁶. Although there are divergent views on purpose, content and implementation⁷, consensus exists on pathway characteristics: strong multidisciplinary character, aimed at improving quality and efficiency; and strong emphasis on alignment issues throughout the care chain for a homogeneous patient group. Vanhaecht et al.⁸ showed that improvement in care pathway concepts and methodology demands international knowledge sharing. International benchmarking can enable knowledge sharing, not only regarding integration of evidence-based guidelines in pathways, but also on specific topics like IT integration⁸. According to Polite et al.⁹ oncology pathways offer many potential advantages, nevertheless, several issues must be addressed such as who should control the development of pathways.

A recent development in healthcare is the transformation from volume-based healthcare to value-based healthcare¹⁰. Most current models, lack on the aspect of value (i.e. patients' health benefits per healthcare dollar spent)¹¹: 1) an ability to measure outcomes that matter to patients, 2) transparency around measured clinical and financial outcomes, and 3) care coordination across all providers in the care pathway. According to Porter and Lee¹², transformation to value-based healthcare requires a shift from silos organized by specialty, to organizing care around patient's medical condition, including the whole pathway. They propose Integrated Practice Units (IPU), in which a team of clinical and nonclinical personnel provides the full care pathway¹². Wherever IPUs exist, consistent results are found: faster treatment, better outcomes, and lower costs¹². Integration of medical practice across the whole care pathway has therefore been identified as high-value healthcare delivery, particularly in the management of complex diseases such as cancer¹². An example of IPU development in cancer care are the Multidisciplinary-Care-Centers at MD Anderson¹³. However research regarding this subject is limited. This study aimed at firstly describing the development and outcomes of benchmark for oncology care pathways. Benchmarking is defined as: continual and collaborative measuring and comparing results of key work processes with those of the best performers. Learning how to adapt these best practices to achieve breakthrough process improvements and build healthier communities¹⁴. The second aim of this study was exploring the degree of development towards an IPU, based on the benchmark data.

2. Methods

2.1 Study design and sample

This international benchmarking study (part of the BENCH-CAN project¹⁵, a European benchmark project aiming at benchmarking cancer care to contribute to improving the quality of interdisciplinary patient treatment) involved seven cancer centers across Europe (South Europe n= 4 and Central/East Europe n=3). The participating organizations were all members of the Organisation of the European Cancer Institutes (OECI)¹⁶ and five were designated as Comprehensive Cancer Center by the OECI¹⁷. The benchmark tool used to collect data was developed and executed according to the 13 steps developed by van Lent et al.¹⁸ as depicted in Table 1. In the following paragraph step 7 to 12 are further elaborated on.

Table 1 13 step benchmarking method developed by van Lent et al.¹⁸

Step	Action	Application in this study
1	Determine what to benchmark	Organizational aspects of pathways and tumor services
2	Form a benchmarking team	BENCH-CAN consortium
3	Choose benchmarking partners	Seven cancer centers in North-West, South and Central-Easter Europe
4	Define and verify the main characteristics of the partners	mapping exercise of the external environment of the cancer centers and the influence this could have on the pathway and tumor service development
5	Identify stakeholders	Patients, Clinicians, Administrators, Researchers
6	Construct a framework to structure the indicators	This was not specifically done for this study, the framework of the BENCH-CAN project was used
7	Develop relevant and comparable indicators	Based on literature and expert opinion (cancer centers and oncology care pathway expert from the Netherlands Comprehensive Cancer Organisation (IKNL))
8	Stakeholders select indicators	Stakeholders from the BENCH-CAN project and other experts from cancer centers provided feedback on the indicators
9	Measure the set of performance indicators	Data collection phase was three months. A team (consisting of the first author, other members of the BENCHCAN consortium and a member of European Cancer Patient Coalition (ECPC)) performed a site visit to each pilot center to verify the data, grasp the context and clarify any questions arising from the provided data.
10	Analyze performance differences	Deductive Qualitative Content Analysis ¹⁷ performed
11	Take action: results were presented in a report and recommendations were given	For each participating cancer centre, a report was made containing the anonymized outcomes of the benchmark for all centers. Improvement recommendations were sent in a separate document.
12	Develop improvement plans	If in agreement pilot sites developed improvement plans
13	Implement the improvement plans	Outside the scope of this study

2.2 Indicator development and collection

Indicators (step 7) were derived from literature and expert opinion. Experts included a representative of the IKNL researching pathways in Oncology^{20,21}. Stakeholders of the BENCH-CAN project (such as researchers, clinicians and (quality) managers) and experts from other cancer centers (OEI members, n=71) were asked for feedback, and consensus was reached on the indicator set (step 8) containing 51 qualitative indicators and 193 quantitative indicators. Indicators assessed multiple topics concerning the organization of cancer care pathways, more specifically the pathway for breast cancer and colo-rectal cancer. Focus were organizational aspects of care pathways, not the clinical interpretation. Pilot sites appointed a team responsible for the data collection covering multiple departments. The team existed of members from different stakeholder groups: patients, clinicians, researchers and management. The data was collected for the year 2012. After a quick data scan, a one day visit to each pilot center was performed to verify the data, grasp the context and clarify questions arising from the data. The visit consisted of semi-structured interviews and a tour of the cancer center with a specific focus on the breast unit (if available). The visits were also used to collect additional information and to acquire feedback on the benchmark tool. The validity of the indicators was checked using feedback from pilot sites based on three criteria^{22, 23}: 1. Clear definition, 2. Data availability and reliability, 3. Discriminatory features .



2.3 Analyzes

After the completion of all site visits data for each indicator was compared. A deductive form of the Qualitative Content Analysis¹⁹ was used to report on the collected data. This method contains 9 steps which are described in Table 2.

Table 2 steps in the deductive Qualitative Content Analysis

Step	Action
1	Reading through the (benchmark) data transcript and making notes in the margins
2	Assessing the notes made in the margins and listing the different types of information found
3	Reading through this list and categorizing each item
4	Repeat the first three stages again for each data transcript
5	Collect all of the categories or themes and examine each in detail and consider the fit and relevance
6	Categorize all data (all transcripts together) into minor and major categories/themes
7	Review all categories and ascertain whether some categories can be merged or if some need to be sub-categorized
8	Return to the original transcripts and ensure that all the information has been categorized
9	To ensure the validity, send a report containing all data to the pilot sites for verification

Data in this paper is anonymized. The description of pathways is structured according to the criteria by Kinsman et al.²⁴. Data analysis focused mainly on the pathway for breast cancer. The indicators developed for the pathway tool were similar to the IPU criteria as developed by the Harvard Business School²⁵ such as organization of care multidisciplinary teams (MDTs) and description of the steps taken in the care pathway. To explore the degree of IPU development every center was scored against the IPU criteria²⁵. The key-IPU criteria described by the Harvard Business School²⁵ are depicted in Box 1.

1. Organized around the patient medical condition or set of closely related conditions;
2. Involves a dedicated, multidisciplinary team who devotes a significant portion of their time to the condition;
3. Providers involved are members of or affiliated with a common organizational unit;
4. Care is led by a physician team captain and a care manager who oversee each patient's care process;
5. Providers function as a team, meeting formally and informally on a regular basis to discuss patients, processes and results;
6. Takes responsibility for the full cycle of care for the condition, encompassing outpatient, inpatient, and rehabilitative care as well as supporting services (e.g. nutrition, social work, behavioral health);
7. Incorporates patient education, engagement, and follow-up as integral to care
8. Measures outcomes, costs, and processes for each patient using a common information platform;
9. Accepts joint accountability for outcomes and costs
10. Utilizes a single administrative and scheduling structure;
11. Co-located in dedicated facilities;

Box 1 key-IPU criteria²⁵

3. Results

3.1 Indicators

After the data collection, definition clarity, data availability, data reliability and discriminative value of the indicators were evaluated with the pilot centres. Seven indicators were deemed irrelevant and removed. One indicator regarding the minimal volume of surgeries was added. This evaluation resulted in a final set of 46 qualitative indicators and 141 quantitative indicators that were considered suitable for wider use in benchmarking care pathways and exploring IPUs.

3.2 Pathway benchmark

An overview of the current status of pathways in the pilot sites can be found in Table 3. Most cancer centers only started using official pathways quit recently, and in center D and E have not implemented and formalized all pathways and were recommended to do so. All centers base their pathways on guidelines, either national or international and have a clear “director” to guide the pathway development. All centers perform at least mammography, ultrasound and physical examination before breast surgery. Furthermore, all centers perform annual mammography’s in the first five year of follow up for all patients. However, institute E indicates they perform the follow-up for about 60% of their patients. As mammography plays an important role in both diagnosis as follow-up efficient use of these machines is important. The number of scans made per device per year varied from 4125 to 10444 scan. Besides the breast cancer pathway centers have developed pathways for several other tumors such as lung cancer (4 centers), and Melanoma (4 centers).

Pathway characteristics

An overview of pathway characteristics can be found in Table 4. Pathways are developed for multidisciplinary teams (MDTs), roles of various healthcare professionals are described (see Table 4). Center D was least structured and was recommended to include discussion by an MDT in the pathway for all patients. The care pathway describes a logical sequence from diagnostics, through treatment onto either survivorship care or palliative care. Centers E and F were recommended to evaluate the transition protocol for patients transferring from hospitals to other healthcare facilities as this was lacking as part of their formal pathway. In most countries we found maximum waiting and throughput times set by government for the different steps in the pathway.

Table 3 Overview of the pathway for breast cancer described per institute

Institute	Pathway used since	Basis	Pathway director	Diagnosis (% patients pre-surgical mammography, ultrasound and physical examination) and follow up (% patients routine mammographic screening)	Other pathways
A	2007	The pathways were developed using national and international guidelines and also based in the long experience of professionals in cancer treatment.	the (Breast) Clinic Coordinator	Diagnosis: 100% of patients Follow-up: 100% of patients Efficiency mammography (scans/device/year): 4341	Head and Neck; Endocrine Tumours; Nervous System; Lung; Skin/Melanoma, Soft Tissues and Bone; Paediatrics; Onco-Haematology; Breast; Gynaecology; GI; and Urology
B	2011	Pathways are part of Quality management system documentation. Treatment guidelines are reviewed and updated by a commission at least once a year. Guidelines are based on the Eusoma and SIS guidelines	Head of the Breast Surgery Department	Diagnosis: 90-95% of patients Follow-up: Unknown Efficiency mammography (scans/device/year): 7617	Lung, breast, gynecological, urological, head and neck, GI.
C	2010	The patient pathways were developed by the leadership of the institute and members of the relevant MDT based on international and national guidelines, relevant legislation, and clinical practices.	The director of the breast cancer pathway is the head of the breast cancer MDT. The director of each pathway is the head of the relevant MDT.	Diagnosis: 100% of patients Follow-up: 100% of patients Efficiency mammography (scans/device/year): 6490	Lung, Breast, GI, Urological, and Melanoma.

Table 3 Overview of the pathway for breast cancer described per institute (continued)

Institute	Pathway used since	Basis	Pathway director	Diagnosis (% patients pre-surgical mammography, ultrasound and physical examination) and follow up (% patients routine mammographic screening)	Other pathways
D	2014	From international and national guidelines, the institute has developed institutional protocols and working procedures. These are periodically revised by MDTs and the management. Some pathways are not formalized, but they exist for years as there are multidisciplinary teams for years.	The director of the cancer care pathway is the chief of the department.	Diagnosis: 100% of patients Follow-up: 100% of patients Efficiency mammography (scans/device/year): 6427	Gynecologic cancer, hematology, pediatric oncology, digestive cancer, Urological cancer, lung cancer, colorectal cancer, breast cancer
E	2014	The clinical pathways were developed in accordance with the regional and national legislation, as well as clinical practice. At present the pathways are partially implemented	There are clear directors within the development of pathways. For breast cancer it is the Head of the Breast Surgical Oncology Division.	Diagnosis: 100% of patients Follow-up: ± 60% receive follow up in this institute. Of these 100% receive mammography. Efficiency mammography (scans/device/year): 10444	Breast, colon, melanoma, ovary, lymphoma, gastric, sarcoma, head and neck
F	2005	Clinical pathways are based on international and national guidelines and conferences to continually stay up to date with the clinical literature and in order to keep the pathways current.	Cancer care pathways are developed by the Clinical and Surgery Departments with the supervision of the Healthcare Directorate and in close collaboration with the Medical Directorate.	Diagnosis: 100% of patients Follow-up: 100% of patients Efficiency mammography (scans/device/year): 5852	Thyroid and parathyroid, Prostate, Neuroendocrine (NETs), Pancreatic, hepatic, and biliary tumors, Sarcomas, Head and neck, Breast

Table 3 Overview of the pathway for breast cancer described per institute (continued)

Institute	Pathway used since	Basis	Pathway director	Diagnosis (% patients pre-surgical mammography, ultrasound and physical examination) and follow up (% patients routine mammographic screening)	Other pathways
G	2011	PDSA cycles are used for developing, implementing and reviewing Clinical Pathways which are based on national and international guidelines.	Disease Units (Integrated Practice Units) and MDTs are responsible for defining and implementing evidence-based pathways and algorithms as a support for their activities. The Cancer Center Managing Director and the Clinical and Scientific Director are responsible for facilitating, overseeing and approving clinical pathways.	Diagnosis: % of patients Follow-up: % of patients Efficiency mammography (scans/device/year): Unknown	Urology, Cervix, Chronic Lymphatic Leukemia, Chronic myeloid leukemia, Colorectal, Endometrium, Liver, Lung, Melanoma, Mesothelioma, Multiple Myeloma, Onco-thrombosis, Ovary, Pancreas, Pituitary, Prostate, Sarcoma, Supportive Treatments, Thymoma, Upper GI, Women Fertility and Sexuality Preservation



Table 4 Description of breast cancer pathway characteristics described per institute

Institute	Multidisciplinary	Guidelines- or evidence base	Content of the pathway description	Criteria-for waiting and throughput times
A	<p>The care pathway describes the roles of the different disciplines involved. Formalized position</p> <p>Breast MDT core members: Surgical Oncology, Medical Oncology, Radiotherapy, Imaging, Pathology and Nursing. Non-core members: Plastic Surgery, Psycho-Oncology, Medical Physics and Rehabilitation, Genetics, Neurosurgery, Orthopedics and Palliative Care</p>	<p>The breast cancer pathway was developed using national and international guidelines.</p>	<p>All steps, from the patient admission process and patient referral requirements till the follow-up are described in the pathway.</p>	<p>The pathway contains maximum waiting times between the various steps.</p>
B	<p>The care pathway describes the roles of the different disciplines involved.</p> <p>Breast MDT members: breast surgeons, radiologists, medical oncologists, radiotherapist, pathologist, geneticist, nurse, administrator.</p>	<p>The breast pathways are developed based on well-known international guidelines, for breast: Eusoma and SIS guidelines</p>	<p>The pathway contains information about Epidemiological, Scientific, Genetic, Clinical steps</p>	<p>The government sets maximum waiting times for certain steps of the pathway</p>
C	<p>The care pathway describes the roles of the different disciplines involved in the MDT.</p> <p>Breast MDT members: breast surgeon, clinical oncologist, plastic surgeon, psychologist imaging diagnostician pathologist</p>	<p>The pathways were developed based on international guidelines such as the NCCN and the ESMO guidelines, as well as national guidelines.</p>	<p>All steps from patient admission to follow-up of care are detailed in the pathway, including all clinical treatment algorithms.</p>	<p>The government does not limit maximum waiting times in general.</p>

Table 4 Description of breast cancer pathway characteristics described per institute (continued)

Institute	Multidisciplinary	Guidelines- or evidence base	Content of the pathway description	Criteria-for waiting and throughput times
D	The care pathway describes the roles of the different disciplines involved and which cases should be presented in the MDT meeting.	Breast MDT members: medical oncologist, radiotherapist, radiologist, surgeon, pathologist.	The pathways are based on international and national guidelines.	There government only limits the time from referral by the general practitioner to the time the specialized physician sees the patient. There are currently no standard means of measuring waiting and throughput times between the steps in the pathway.
E	A multidisciplinary team for each pathway	Breast MDT members: preoperative setting: radiologist, surgeon, pathologist, breast nurse postoperative setting: surgeon, medical oncologist, radiotherapist, pathologist, nuclear medicine doctor, breast nurse	Pathways are based on national and international guidelines	The government decides on the maximum waiting between the steps within the pathway
F	The care pathway describes the roles of the different disciplines involved.	Breast MDT members: breast and plastic surgeons, medical oncologists, radiotherapists, geneticists and pathologists	Pathways are based on international and national guidelines, clinical study protocols and conferences.	There are governmental set maximum waiting times between the various steps in the pathway. Patient access to the hospital is described and monitored in the pathway.

Table 4 Description of breast cancer pathway characteristics described per institute (continued)

Institute	Multidisciplinary	Guidelines- or evidence base	Content of the pathway description	Criteria-for waiting and throughput times
G	<p>All the participants in MDTs were involved in developing and reviewing the Clinical Pathways. Therefore Clinical Pathways are integrated in the activities of MDTs and, vice versa, Clinical Pathways include MDTs as a crucial step.</p> <p>Breast MDT members: Surgeons, Oncologists, Radiotherapists, Pathologists, nurses. On call: Radiologists, Plastic surgeons, Psychologists</p>	<p>Pathways are based on international and national guidelines and conferences.</p>	<p>All steps from patient registration to follow up including a detailed description are described in the pathway. This includes all clinical diagnosis and treatment algorithms and follow-up</p>	<p>There are set waiting times within pathway. Inside the regional portal these are monitored.</p>

Evaluation

From Table 5 can be seen that the extent of evaluation methods vary from an informal evaluation by MDTs to an extensive internal and external evaluation. The Clinical Governance at center A makes a patient pathway audit for every Pathology Clinic (IPU). This audit focuses on two areas: compliance with therapeutic guidance documents and evaluation of waiting times. The breast cancer pathway in institute B was evaluated through external accreditation. In center D evaluation is performed internally and externally. In institute E the pathways were not systematically evaluated. An indicator matrix containing indicators from various sources was in the making for this purpose. In center F pathways were collectively discussed through MDT meetings in close collaboration with the Healthcare Directorate and the Medical Directorate that are in charge of supervising cancer care pathways. Center G organizes an extensive internal and external evaluation (by a regional and national agency) of the pathway. Overall pathway evaluation seems to focus mainly on waiting and throughput times and less on quality performance.

Table 5 General Pathway evaluation policy described per institute

Institute	Evaluation method	Examples of type of measurements	Frequency	Outcomes evaluation
A	The Clinical Governance makes a patient pathway audit for every Pathology Clinic, where times between referral/ appointment/ diagnosing / treatment are evaluated, including reasons for delay.	Efficiency	Every trimester	Quarterly reports and an annual report . There's also an annual report with clinical indicators that compares the results of the last two years.
B	Evaluation of the breast pathway was performed by an external evaluation body by means of accreditation	Efficiency, each year the pathway leadership conducts or participates in two or more studies (e.g. CP3R) that measure quality and/or outcomes	Once a year	The outcome of the evaluation was a structured pathway
C	Members of the MDTs evaluate the applicable pathways.	Evaluation of patient flow efficiency	Every 3 years.	The outcomes of the pathway evaluation are discussed by members of the relevant MDT and are included in the annual report.
D	Pathways are evaluated through the evaluation of patients' files. There are no goals set beforehand. Internal evaluation is performed by two means: quality of care control team and internal audit. The external evaluation is performed by the National Healthcare Insurance.	Monitoring patient complaints suggestions, diagnosis and treatment guidelines adherence check, evaluation quality of services	Internally every 3 months, externally quarterly	The quality of care team and the external evaluation reports are presented to the Institute's Medical Council and to the general manager. As a result of the evaluation the responsible parties are informed in regards to their individual performance and provided with suggestions for improvement.

Table 5 General Pathway evaluation policy described per institute (continued)

Institute	Evaluation method	Examples of type of measurements	Frequency	Outcomes evaluation
E	The pathways were not systematically monitored, they will be monitored through specific indicators. The indicators proposed so far are the EUSOMA's indicators for Breast Unit accreditation, integrated with internal performance indicators and indicators proposed by BENCH-CAN.	EUSOMA's indicators for Breast Unit, BENCH-CAN indicators	Not applicable	Not applicable
F	Medical Directorate coordinates the activities for the evaluation of diagnostic and therapeutic pathways (PDTA). The evaluation process is hence performed inside the institute in close collaboration with the Clinical and Surgery Departments and with the supervision of the Healthcare Directorate.	Unknown	every 6 months	Findings from the evaluation process are implemented by the Clinical and Surgery Departments in order to translate results into practice. Quality, Training and Privacy Office coordinates the quality improvement process of the pathways in close collaboration with the Departments.
G	<p>External evaluation</p> <p>External evaluation and benchmark is provided by two different institutions, regional and national</p> <p>Regional benchmarks the single hospital performance against other regional hospitals. Two kinds of evaluation are provided:</p> <ul style="list-style-type: none"> - general evaluation at departments' level - Evaluation on the pathway' level <p>The national evaluation is performed by a national agency that once a year displays on a Web Portal the outcome measures (such as mortality), benchmarked against all national hospitals.</p> <p>Internal evaluation</p> <p>Internal system that calculates and provides all the performance indicators to the Hospital. Review of Clinical Pathways is based on a periodical review of patient records. It's in plan to include the pathways inside the EPR system, in order to facilitate clinical decisions and to improve standardization.</p>	<p>30 days Mortality, Readmissions, Surgical re-interventions, transfers to other acute hospitals, Efficiency (e.g.% of patients operated within 60 days from mammography)</p>	<p>Continuous improvement</p> <p>Small/relevant every year as evidence changes, formally/big every three years.</p>	<p>Several actions are implemented for improving outcomes and quality provided to our patients, as a result of the pathways' evaluation. For example for Breast Cancer: Starting from the group of evaluation indicators there is a continuous improvement with an annual evaluation visit by EUSOMA.</p>

3.3 Degree of IPU development

By definition cancer centers are organized around a medical condition (i.e. cancer). However differentiating and looking at specific types of cancer, Table 6 shows that this can only be seen at center A (for 11 different cancers including breast cancer), center B (breast cancer), and center G (5 types of cancer including breast). This is linked to the common organization unit where, in all centers providers involved are members of the cancer center but in center A, B and G they are members of a specific organizational unit.

All centers work with MDTs specific to a tumor type, hence dedicating time by their members to the specific condition. Members of the MDT are however not always solely or clearly dedicated to one tumor type, or unit. A physician heads the MDT, though only centers E and G have a dedicated case manager.

Regarding responsibility for the full care cycle, this was the case for MDTs in all centers. Nevertheless, only center A and G indicated that the unit encompassed rehabilitative care as well as supporting services (e.g. psychosocial services). For other centers this is part of services for the whole center. This is the same regarding patient education, engagement, and follow-up which are seen as integral to care however not specifically arranged for a unit (except for center A and G). The fact that units take responsibility for the whole care cycle does not mean all patients complete the whole pathway at the cancer center. Center G mentioned that many patients admitted to the unit already have a diagnosis and many patients also complete the follow-up elsewhere. This compromises' the measurement of outcomes of the care pathway for each patients.

All centers measure outcomes, costs, and processes, however most (n=5) do this for the whole center and not unit specific. Centers are either lacking data platforms that allow for the collection of specific inputs (in terms of HR and finances) and outcomes per unit or do not collect these data for now. Similarly, a specific administrative and scheduling structure is lacking or not fully implemented, as centers C, D, E and F don't have a formalized unit. For center B which indicates to have a breast IPU it was unclear whether they have an accompanying dedicated scheduling structure.

Overall three groups can be identified: centers that have the IPU structure fully implemented (centers A and G); centers that have the IPU partially implemented (center B) and centers that have certain features of the IPU but did not develop or implement the IPU (center C, D, E, F). The evaluation of the implementation outcomes of the breast unit in center A is described in box 2.

Table 6 IPU development for breast cancer assessed against IPU criteria²⁵

Topic	Institute						
	A	B	C	D	E	F	G
Organized around a medical condition	+++	+++	++	++	++	++	+++
Multidisciplinary team with dedicated time to the condition	+++	+++	++	++	++	++	+++
Providers involved are members of or affiliated with a common organizational unit	+++	++	++	++	++	++	+++
Physician team captain or clinical care manager oversees care process	++	++	+	+	++	+	+++
Providers function as a team, meeting formally and informally on a regular basis	+++	+++	+++	+++	+++	+++	+++
Responsibility for the full cycle of care for the condition, encompassing outpatient, inpatient, and rehabilitative care as well as supporting services (e.g. nutrition, social work, behavioral health)	+++	++	++	+	+	+	+++
Incorporates patient education, engagement, and follow-up as integral to care	+++	+	+	+	+	+	+++
Measures outcomes, costs, and processes for each patient using a common information platform	+++	+	+	+	+	+	+++
Accepts joint accountability for outcomes and costs	+++						+++
Single administrative and scheduling structure	+++	+	+	+	+	+	+++
Co-located in dedicated facilities	+++	+++	+++	+++	+++	+++	+++
Total degree of implementation out of 33	32	21	19	18	19	18	33

Note: Columns display the institutes development on the IPU criteria. A triple plus symbol (+++) indicates that institute has this feature fully developed, a double plus symbol (++) indicates that the feature is mostly developed by the center, a plus symbol (+) indicates the feature is partially implemented while a minus symbol (-) indicates the absence of this feature in the center. If a cell is left blank this could not be assessed.

Evaluation of the breast IPU developed in 2007 over a seven year period (2007-2014).

Number of patients

For patients being admitted for the first time in the breast clinic there was a sustained improvement, valued at 25% increase from 2007 to 2014. There was also an increase of 44% in medical appointments without resourcing to more staff.

Costs

Gains in efficiency and in scale, reflected in a reduction of costs of an appointment by 68%. The organization of the Breast IPU resulted in 41% reduction of the cost of medication supplied to an average patient, due to purchasing power; clinical trials; risk sharing agreements; and also better protocols.

Organization

The activities of Clinical Governance and of Management Control were much improved after the implementation of the Pathology Clinics. In fact, staffing, resource planning and auditing became clearer after the new procedures were in place. Pathways within the IPU were developed around the patient needs, and also allow for interface with therapeutic protocols with drugs and radiotherapy treatments.

Box 2 Outcomes of an IPU implementation evaluation in center A

Discussion

In this study, we developed a benchmark tool to assess development, implementation and evaluation of cancer care pathways. The tool was successfully tested in seven cancer centers to assess its suitability for yielding improvement suggestions regarding pathway organization and data to explore the status of IPU organization.

The director concerned with pathways is usually a medical specialist which is in accordance with Polite et al.⁹ who argues that responsibility for oncology pathway development must always lie primarily with clinicians. Based on our data it seems that centers have no clear strategy when developing pathways. This is in line with a study by Vanhaecht et al.⁶, in which a minority of sampled countries (43%) used a systematic approach to develop, implement and evaluate care pathways. Regarding pathway evaluation, none of the centers work with beforehand set goals for the evaluation. In-built continuous evaluation and follow-up should guarantee the effectiveness of care pathways²⁶, this was therefore recommended as an improvement opportunity for cancer centers.

Porter²⁷ states that care pathways are beneficial, but not sufficient for value-based high qualitative care and recommends the establishment of IPUs. In this explorative study assessing the centers' benchmark data against the criteria of an IPU²² we found three groups. Group 1, centers fully meeting all criteria (center A and G), and group 2, partially meeting the criteria (center B) both indicated to have a breast IPU. Based on our data we agree with center A and G having an IPU for breast cancer. However for center B an improvement in four criteria is required to have a fully implemented IPU. Group 3, centers meeting only some of the criteria did not indicate to have this type of organization implemented. The fact they do meet some of the criteria is likely due to the pathway organization and the organization of MDTs that has been developed for years²⁸. Dedicating specific resources such as staff, measuring outcomes specific to the IPU and IPU specific administrative and scheduling systems were lacking.

Sarai et al.²⁹ identify IPUs as units in which providers commit a substantial portion of their time to treating a focused set of care pathways, implying pathways are IPU building blocks. We identify pathways as tools to map the current organizational processes and improve these processes where necessary (value and quality improvement on process level). IPUs are structured organizational units where the process identified through the pathway takes place. IPU organization seems to require an organizational change from a focus on discipline based departments (e.g. Radiotherapy) to those departments facilitating the IPUs which are pathway based. In this study IPUs are therefore seen, as tools for quality and value improvement on a strategic level.

The evaluation of the breast IPU in center A showed that improved efficiency led to more time available for patients, through volume, economies of scale and improvement in quality. Although based on a small number, this corresponds to findings by Low et al.³⁰ who found that the IPU resulted in reduced readmissions in patients who are at highest risk of readmission. There is no structured information on how patients feel about this development which is something that needs to be known in order to facilitate true patient-centered care. In a fee-for-service system measuring patient-reported-outcomes (PRO) before and after interventions provides healthcare-providers with the possibility to market their results to patients, and payers and increase market share³¹. Enthoven et al.³² however state that for patients with multiple morbidities the IPU organization is unfavorable. This stems from the original ideas of Porter and Teisberg¹⁰ that IPUs should be entities with sufficient degrees of “self organization”.

6 Developing and implementing an IPU has barriers which will vary based on the condition being treated, provider organization, and health-system characteristics. Keswani et al.³¹ divided these barriers into three subcategories: Operational, technology, and payment/contracting. Regarding operational barriers, if cancer centers have multiple IPUs, certain IPUs may draw upon for example physical therapy, psychosocial services and rehabilitation (supportive services). Most centers in this study have supportive services organized as shared services. There are however trade-offs worth considering in organizing these services. Having supportive services embedded within the IPU, as opposed to shared services, enables immediate access to those services for the IPU patients, along with reduced scheduling complexity³¹. However, the major risk is that some IPUs may not have sufficient demand for services, resulting in wasted capacity. Describing and assessing pathways for a specific tumor can aid in the decision regarding the optimal organization model.

There are several technology competences that are essential for ensuring patient-centered value through an IPU. It is important to have a patient-engagement platform that enables two-way communication and sharing of data (such as personalized treatment plans) between the patient and the IPU care team³¹. This can be made available through web and mobile devices and collect outcomes (e.g. PROs and patient satisfaction) longitudinally and display data in real-time for use in and beyond the clinic. This serves two of the IPU criteria, it improves communication within the care team and enables outcome measurement for all patients. Major barriers to obtaining this technology are the upfront costs of customizing the platform specifically for the IPU, training staff and patients and the platform upkeep. Ideally the platform can be shared with other healthcare providers enabling benchmarking.

Barriers related to payment and contracting are health system specific. Through providing personalized treatment to patients who are most likely to benefit, and by addressing

the psychosocial needs, an IPU can achieve improved health outcomes at lower cost per patient. However, given the fact that many supportive psychosocial services are often not (fully) covered by payers integrating all services will likely result in lower or even negative margins relative to current performance. Cancer centers thinking of developing an IPU need to take this into consideration if this is the case in their country. For those countries a change in healthcare financing will be needed in order to implement IPUs. As recommended by Porter²⁷ once IPUs are getting form, it will be easier to lobby for reimbursement through bundled prices for care pathways.

This study holds several implications. Firstly, within this study we did not emphasize the detailed clinical content of the pathway. Future research focusing on the exact pathway content as part of the benchmark will further enable the purpose of international knowledge sharing. Furthermore, as the initial focus of the indicators were pathways, we had to deduct the IPU criteria from the material. This assessment was aided by the focus of the site visit on the breast cancer department which showed whether centers had the IPU organization in place. Indicators that look specifically at IPU organization should be refined and can aid in a more thorough assessment against IPU criteria. Additionally, as little scientific material is available some degree of subjectivity was inevitable with regard to the explorative assessment of the benchmark data against the IPU criteria. Further studies into scale and cut-off points that distinguish between different levels of IPU development are recommended. Future research focusing on more extensive outcome evaluation over multiple years allows to identify the comparison of IPU organized cancer centers and non-IPU organized centers therefore assessing the added value of IPU organization.

Finally, this study focused on a limited number of cancer centers (n=7), and collected data for one year. Future research should include larger series (including general hospitals with an Oncology department) and over multiple years to assess actual discriminative capabilities of the tool and establish the sensitivity for changes over time.

In conclusion, we successfully developed and piloted a benchmark tool to compare and elaborate on organizational performance and suggestions towards improving the organization of cancer care pathways (both implementation as evaluation) were made. The data generated through the benchmark enabled to exploration of the status of IPU development. Pathways can be seen as the process that takes place in the strategic unit of the IPU. Based on our assessments we found that centers varied with regard to fulfilling the IPU criteria, usually lacking specific resources such as staff, measuring IPU specific outcomes and IPU specific administrative and scheduling systems. The development of an IPU requires a strategic organization change with several implications for operations, technology, and

payment and contracting. An evaluation of an IPU for breast cancer showed performance improvements in efficiency and finances.

References

1. Porter ME, Lee TH. The Strategy That Will Fix Health Care. *Harvard Business Review*. October 2013
2. Mohr JJ, Batalden P, Barach P. Integrating patient safety into the clinical microsystem. *Qual Saf Health Care*. 2004;13.
3. Wulff CN. The Effect of Hospital-Based Case Management in Cancer care Pathways. 2012 PhD Thesis: 2-5.
4. Institute of Medicine, Committee on Quality of Health Care in America. *Crossing the quality chasm: a new health system for the 21st century*. Washington, D.C.: National Academy Press; 2001.
5. Preston C, Cheater F, Baker R, Hearnshaw H. Left in limbo: patients' views on care across the primary/secondary interface. *Qual Health Care* 1999; 8(1): 16-21.
6. De Bleser L, Depreitere R, De Waele K, Vanhaecht K, Vlayen J, Sermeus W. Defining pathways. *J Nurs Manag* 2006; 14(7): 553-63.
7. Currie L, Harvey G. The origins and use of Care Pathways in the USA, Australia and the United Kingdom
8. Vanhaecht K, Bollmann M, Bower K, Gallagher C, Gardini A, Guezo J, et al. Prevalence and use of clinical pathways in 23 countries—an international survey by the European Pathway Association. *Journal of Integrated Pathways* 2006; 10(1): 28-34.
9. Polite BN, Page RD, Nabhan C. Oncology Pathways—Preventing a Good Idea From Going Bad. *JAMA oncology* 2016; 2(3),=: 297-298.
10. Porter, M. E., & Teisberg, E. O. (2006). *Redefining health care: creating value-based competition on results*. Harvard Business Press.
11. Keswani A, Koenig KM, Bozic KJ. Value-based Healthcare: Part 1—Designing and implementing integrated practice units for the management of musculoskeletal disease. *Clinical Orthopaedics and Related Research* 2016; 474(10): 2100-2103.
12. Porter ME, Lee TH. The strategy that will fix health care. *Harv Bus Rev* 2013; 91(12): 24.
13. Mosel D, Gift B. Collaborative benchmarking in health care. *Joint Commission Journal on Quality Improvement* 1994, 20:239.
14. BENCH-CAN 2016. <http://www.oeci.eu/Benchcan/>. Accessed 23 March 2016.
15. <http://oeci.selfassessment.nu/cms/> Accessed 7 August, 2016.
16. van Lent W, de Beer R, van Harten W. International benchmarking of specialty hospitals. A series of case studies on comprehensive cancer centres. *BMC Health Serv Res*. 2010; 10: 253.
17. Zhang Y, Wildemuth BM. Qualitative Analysis of Content. *Applications of Social Research Methods to Questions in Information and Library Science*, 2016 318. https://www.ischool.utexas.edu/~yanz/Content_analysis.pdf Accessed September 15, 2016
18. van Hoeve J, de Munk L, Otter R, de Vries J, Siesling S. Quality improvement by implementing an integrated oncological care pathway for breast cancer patients. *Breast* 2014; 23(4):364-70.
19. van Hoeve JC, Elferink MA, Klaase JM, Kouwenhoven EA, Schiphorst PP, Siesling S. Long-term effects of a regional care pathway for patients with rectal cancer. *Int J Colorectal Dis* 2015; 30(6): 787-95.

20. De Korne D F, Sol K J C A, van Wijngaarden J D H et al. Evaluation of an international benchmarking initiative in nine eye hospitals. *Health Care Management Review*. 2010; 35: 23
21. Cowper J, Samuels M. Performance benchmarking in the public sector: the United Kingdom experience. *Benchmarking, Evaluation and Strategic Management in the Public Sector*. Paris: OECD, 1997.
22. Kinsman L, Rotter T, James E, Snow P, Willis J. What is a clinical pathway? Development of a definition to inform the debate. *BMC medicine* 2010; 8(1): 31.
23. President & Fellows of Harvard College. Integrated Practice Units. <http://www.isc.hbs.edu/health-care/vbhcd/pages/integrated-practice-units.aspx> Accessed on 19 August, 2016.
24. de Vries M, van Weert JC, Jansen J, Lemmens VE, Maas HA. Step by step development of clinical care pathways for older cancer patients: Necessary or desirable?. *European journal of cancer* 2007; 43(15): 2170-2178.
25. Porter M. Value-Based Health Care Delivery. HBS Healthcare Initiative & Healthcare Club, 2012. <http://vbhcprize.com/wp-content/uploads/2014/03/Porter-Value-based-health-care-delivery.pdf> Accessed on 14 August 2016
26. Taylor C, Shewbridge A, Harris J, Green JS. Benefits of multidisciplinary teamwork in the management of breast cancer. *Breast cancer* 2013; 5(7): 79-85.
27. Sarai M, Duszak R, Kaufman JA. Value-Based Medicine: Organizing IR Care through Integrated Practice Units. *Journal of vascular and interventional radiology* 2016; 27(3): 336-338.
28. Low LL, Tan SY, Ng MJM, Tay WY, Ng LB, Balasubramaniam K, et al. Applying the Integrated Practice Unit Concept to a Modified Virtual Ward Model of Care for Patients at Highest Risk of Readmission: A Randomized Controlled Trial. *PLoS one* 2017; 12(1): e0168757.
29. Keswani A, Koenig KM, Ward L, Bozic KJ. Value-based healthcare: Part 2—Addressing the obstacles to implementing integrated practice units for the management of musculoskeletal disease. *Clinical Orthopaedics and Related Research* 2016; 474(11): 2344-2348.
30. Enthoven AC, Crosson FJ, Shortell SM. 'Redefining health care': medical homes or archipelagos to navigate?. *Health Affairs* 2007; 26(5), 1366-1372.

PART 5

**International financial and
quantitative data comparison**

CHAPTER 7

Management and performance features of cancer centers in Europe: A fuzzy-set analysis

Anke Wind
Mariana Fernandes Lobo
Joris van Dijk
Isabelle Nefkens-Lepage
José Laranja-Pontes
Vítor da Conceição Gonçalves
Wim van Harten
Francisco Nuno Rocha-Gonçalves

Abstract

The specific aim of this study is to identify the performance features of cancer centers in the European Union by using a fuzzy-set qualitative comparative analysis (fsQCA). The fsQCA method represents cases (cancer centers) as a combination of explanatory and outcome conditions. This study uses data on seven centers from a European benchmarking project: BENCH-CAN. The fsQCA uses the net income and productivity as the outcome conditions and five explanatory conditions: the level of dedication to R&D, annual budget level, size, type, and whether the center is a comprehensive cancer center. Despite the modest number of cases, the study successfully applies the fsQCA. The findings show that public, comprehensive cancer centers with at least two of the three other explanatory conditions (dedication to R&D, annual budget, or size) have an association with high net income and high productivity.

Introduction

With a mortality rate in Europe of 1.75 million in 2012, cancer became the second leading cause of death (Ferlay et al., 2013). Therefore, this disease has an important effect on the healthcare system. The research by Aggarwal and Sullivan (2014) shows that in high income countries the cost of delivering high quality, equitable care is outstripping the present social budgets. Cancer centers are essential actors in the delivery of care in various countries. They are important for scientific discovery and for advancing excellence in cancer care (Simone, 2002). But, society forces cancer centers to provide a high productivity level at lower costs to cope with increasing demand. Across all disciplines, healthcare is becoming more complex, leading to quality and performance challenges (Plsek & Greenhalgh, 2001). Further, society calls for transparency on the relative performance between and within healthcare organizations (Leape et al., 2009). A variety of features both internal and external to a center, such as available resources, level of professional training, and size (Merkow, Chung, Paruch, Bentrem, & Bilimoria, 2014), are possible explanatory conditions for high performance.

Benchmarking is a common and effective method for measuring and analyzing performance in order to identify areas of improvement. Van Lent, de Beer, and van Harten (2010) define benchmarking as “the search for and implementation of best practices” (pp. 253). Their research shows that benchmarking can produce relevant input to improve the operational management of specialty hospitals such as cancer centers.

One method to identify the importance of quantitative performance features and how they relate to outcomes is the fuzzy-set qualitative comparative analysis (fsQCA). This method provides a systematic, transparent, and exhaustive analytical approach in the realm of comparative research (Ragin, 2000). The method essentially uses a qualitative comparative analysis (QCA) for the identification of patterns that hold across the sample of cases (Fiss, 2007; Rihoux & Ragin, 2009) and is becoming increasingly common in organization and management science (Wagemann, Buche, & Siewert, 2015). In contrast to other quantitative methods, such as a regression analysis, fsQCA can use small sample sizes (5–50 cases) (Fiss, 2012). Several examples of studies with ten or fewer cases exist in which fsQCA provides valuable insights (Stokke, 2007; Magnier-Watanabe & Senoo, 2008) and in which statistical tests would be unreliable. Few examples exist of the application of fsQCA within health research (Warren, Wistow, & Bamba, 2013).

This study explores the use of the fsQCA to investigate the association between preselected explanatory conditions and financial performance outcomes (net income, the profit per discharge; and productivity, the total number of patient visits divided by the number of

inpatient beds) that use data from the BENCH-CAN project (2016). BENCH-CAN is an international project that aims to benchmark comprehensive cancer care and yield examples of good practices in a way that contributes to improving the quality of interdisciplinary patient treatment. Following Merkow et al. (2014) and Delgado (2008), this study hypothesizes that a large budget, substantial size, and a significant involvement in research and development (R&D) are important conditions for highly productive and financially sound and profitable centers. Thus, this study tests the following proposition: a highly productive center has large profits if the center has a high budget, is large, and focuses primarily on R&D. Following this introduction, section 2 contains the theoretical framework. Section 3 describes the data used and analyses performed. Section 4 presents the findings. Section 5 offers discussions with limitations and section 6 the conclusion.

Performance measurement

2.1. Healthcare management and benchmarking

Since 1934 when the University of Chicago first offered its seminal program in hospital administration, both the private and public healthcare sectors have established themselves as a specific field of applied management. In fact, this field combines a multidisciplinary approach by demanding skills in leadership, policy, management, finance, and social service delivery that manifest themselves through several distinctive factors: strong information asymmetry; lack of integration; high uncertainty regarding outcomes; highly complex organizations; and specificity in measuring production through financing rules, leadership sources, types of outcomes, and stakeholders' conflicting expectations (Costa, 2008). Some phenomena are at the heart of healthcare managers' concerns today, such as "how to measure and control costs", "how to innovate" or "how to best deliver innovation", "how to finance," or "how to sustain profitability" (Requart, 2015; Devine, O'Clock, & Lyons, 2000). Different studies show that specialty hospitals such as cancer centers might have a financial advantage over general hospitals because they can recap the benefits of profitable services without having to cross-subsidize unprofitable services (Schneider et al., 2007; Choudhry, Choudhry, & Brennan, 2005). However, this financial advantage is only true when all cancer services are profitable, which is not always the case. The innovative financing of cancer care, for example, through comprehensive cancer centers, can drive efforts toward universal health coverage (Gospodarowicz, Trypuc, D'Cruz, Khader, & Knaul, 2015).

A well-known management saying claims that "what is not measured cannot be managed or improved" (Kaplan & Porter, 2011). Thus, benchmarking is a useful technique

for establishing patterns, measuring performance, and performing a comparative analysis (Gonzalez- Padron, Akdeniz, & Calantone, 2014). Although many studies examine the healthcare industry, Nogueira, Lira, Albuquerque, and Linhares (2015) show that these studies are insufficient to evaluate a hospital's financial performance, hence the need for more research. The settings of healthcare organizations cause the insufficient nature of these studies. These organizations often involve developing teaching and research while providing complex health services that require humane and integral care in an imposing technological, financial, legal, and ethical environment.

2.2. The fuzzy-set qualitative comparative analysis

Since Ragin (1987), the QCA has evolved into a robust and valuable alternative method to rather traditional statistical approaches (Ragin, 2008; Woodside, 2013). The QCA uses set theory and Boolean algebra where the variables define the conditions, a condition or a combination of conditions defines a set, and cases are members of the sets if they have any of the conditions. To compare sets, the QCA uses a minimization algorithm and the causality between explanatory and outcome conditions in terms of necessity and sufficiency.

In the early stages, the research used the crisp-set theory to build a QCA. In a crisp set, cases can be either members or non-members (1 or 0). More recently, the research has developed the fsQCA as an alternative to the crisp set. In this approach, cases have membership degrees (a value between 0 and 1), that is, cases can be partially in a set. This new approach represents a qualitative enhancement of the method by allowing better adjustment to the real data, an issue that would otherwise arise when dichotomizing variables that are continuous in nature (Ragin, 2000).

A study by the Agency for Healthcare Research and Quality (Thygeson, Peikes, & Zutshi, 2013) that evaluates the patient-centered medical home (PCMH) model shows that the fsQCA is a powerful approach for studying PCMHs and other health services (Thygeson et al., 2012; Kahwati et al., 2011). The fsQCA assumes that many pathways can lead to the same outcome, equifinality (Ragin, 2008; Fiss, 2011). Second, the fsQCA assumes that each pathway can contain different combinations of explanatory characteristics (Thygeson et al., 2013). Third, if the analysis can link the presence of a combination of conditions to the presence of an outcome, then the absence of the combination does not necessarily lead to the absence of the outcome, the so called asymmetric causality (Fiss, 2011). In sum, rather than quantifying the effect of a variable in the outcome, this method detects which combinations of conditions have a positive association with the outcome and which have a negative effect (Woodside, 2013).

Material and methods

3.1. Data

The study uses data from seven cancer centers in Europe regarding seven variables (Prod — level of productivity; Income — net income; RD — level of dedication to R&D; Budget — annual budget; Size — size of center based on the number of beds; CCC — comprehensive cancer center; and Public — public center. The data comes from a European benchmarking project (BENCH-CAN). The centers within this project were selected through convenience sampling of three geographical regions within the European Union (North-Western, Southern, East-Central). Centers are represented by capital letters to ensure that they remain anonymous.

3.2. Outcome conditions

The outcome conditions considered to assess the financial performance of centers are: Income – net income measured as the profit per discharge (inpatients and daycare patients), and Prod–productivity measured as the total number of patient (inpatients, daycare, and outpatients) visits divided by the number of inpatient beds.

3.3. Explanatory conditions

The study uses five explanatory conditions (R&D, Budget, Size, Public, CCC) regarding the center's (i) level of dedication to R&D (percentage of new patients in clinical trials per annum); (ii) annual budget; (iii) size as measured by the number of inpatient beds; (iv) whether the center is public, non-profit, or private; and (v) whether the center is a CCC. A CCC is a cancer center that has a well-established combination of fundamental and translational cancer research and that has a sufficient portfolio of cancer care services that extend along the total care pathway. The CCCs in this study receive their accreditation from the Organisation of European Cancer Institutes (OEI) Accreditation and Designation program (Saghatchian et al., 2014).

3.4. Analysis

The fsQCA tests whether the empirical data supports a relation between the explanatory and the outcome conditions. Given that the set theory is the basis for the QCA, the analysis compares subsets and super sets. A set is a combination of conditions (conditional pattern), and cases (cancer centers) are members of a set if they present that conditional pattern. When a condition is dichotomous (e.g., CCC and Public), this condition defines

its membership or non-membership. However, the study uses the fuzzy-set calibration to set membership when the conditions are continuous variables. The fsQCA allows partial membership of cases in a set based on three anchor values (percentiles 0.95, 0.5, and 0.05): full membership equal to 0.95 or higher, a crossover point of maximum ambiguity equal to 0.5, and full non-membership equal to 0.05 or lower. Cases on different sides of the crossover point are qualitatively different, while cases with differing memberships on the same side of the crossover point differ in degree (Ragin, 2008). Table 1 depicts the sample after the calibration. In the data, a high score in Budget indicates a high degree of membership in the set defined by a high annual budget. The symbol “~” prior to the condition’s name indicates the negation of a condition. In this case the absence of a high annual budget means the presence of a low annual budget. This study uses consistency and coverage metrics to assess the necessity and sufficiency analyses (Ragin, 2000). Under the fsQCA approach, condition A is necessary to outcome Y if in each case the degree of membership in Y is consistently less than or equal to the degree of membership in A ($Y \leq A$). Condition A is sufficient to Y if across all cases the degree of membership in condition A is consistently less than or equal to the degree of membership in Y (Legewie, 2013).

Table 1
Summary table of membership scores (after implementation of calibration).

Center	Outcome measures	Causal conditions					
	Prod	Income	RD	Budget	Size	CCC	Public
A	0.51 ^a	0.51 ^a	0.98	0.75	0.1	1	1
B	0.81	0.64	0.43	0.98	0.51 ^a	1	1
C	0.04	0.03	0.07	0.06	0.64	1	0
D	0.96	0.78	0.51 ^a	0.04	0.63	1	1
E	0.08	0.11	0.04	0.19	0.98	0	1
F	0.36	0.98	0.73	0.51	0.05	1	1
G	0.94	0.26	0.72	0.51 ^a	0.05	1	1

^a The table uses a score of 0.51 instead of 0.5 in order to avoid scores of maximum ambiguity that reduce the number of cases that define conditional patterns.

The sufficiency analysis constructs a truth table that lists all possible conditional patterns and the respective number of empirical cases (in our case 25 = 32 patterns are possible). The conditional patterns in the truth table are then subject to a minimization algorithm that simplifies the patterns into sufficient solutions. These analyses identify both the presence and the absence of outcomes (Income, ~Income, Prod, ~Prod) in order to address asymmetric causality.



The simplification process yields different solutions depending on the logical remainders (conditional patterns that do not have empirical cases). When limited variability exists in the sample and when many of the conditional configurations (rows in a truth table) are logical remainders, then the research recommends a counter-factual analysis based on theoretic knowledge about the relation between the conditions and the outcome (Ragin, 2008). This study implements the counter-factual analysis by following Fiss (2011). His approach provides a way to identify core and peripheral conditions using the parsimonious and intermediate solutions of the minimization algorithm. Fiss (2011) states that “core elements are those causal conditions for which the evidence indicates a strong causal relationship with the outcome of interest and peripheral elements are those for which the evidence for causal relationship with the outcome is weaker.” For better transparency and reproducibility of the results, the current study can supply, on request, the truth table and sufficiency analysis output table. All analyses use the fsQCA 2.5 software (Ragin & Davey, 2009).

Findings

Tables 2 and 3 show that of the five explanatory conditions, being CCC and Public are necessary conditions for both outcomes (high net income and high productivity). These conditions are necessary because they meet the recommended thresholds of consistency (≥ 0.9) and coverage (≥ 0.5) (Ragin, 2008; Legewie, 2013).

Table 2
Analysis of necessary conditions for the outcomes for high net income and low net income.

Outcome measure	Income		~Income	
	Consistency	Coverage	Consistency	Coverage
CCC	0.97	0.53	0.76	0.47
~CCC	0.03	0.11	0.24	0.89
Public	0.99	0.55	0.74	0.45
~Public	0.01	0.03	0.26	0.97
RD	0.76	0.72	0.52	0.55
~RD	0.53	0.50	0.74	0.77
Budget	0.63	0.69	0.45	0.55
~Budget	0.59	0.49	0.75	0.69
Size	0.45	0.50	0.62	0.77
~Size	0.79	0.65	0.60	0.55

Table 3
 Analysis of necessary conditions for the outcomes for high productivity and low productivity.

Outcome measure	Prod	~Prod		
	Consistency	Coverage	Consistency	Coverage
CCC	0.98	0.60	0.72	0.40
~CCC	0.02	0.08	0.28	0.92
Public	0.99	0.61	0.71	0.39
~Public	0.01	0.04	0.29	0.96
RD	0.71	0.75	0.46	0.44
~RD	0.47	0.50	0.74	0.69
Budget	0.64	0.77	0.47	0.51
~Budget	0.59	0.56	0.79	0.66
Size	0.39	0.49	0.60	0.67
~Size	0.74	0.68	0.55	0.45

Table 4 shows the results from the sufficiency analysis for high net income and low net income, and Table 5 shows the results for high productivity and low productivity. For each solution, the study presents the Centers with a given causal pattern. The study also presents the individual and overall consistency and coverage metrics. According to Ragin (2008), overall consistency scores of at least 0.8 are acceptable.

The results show that three alternative causal patterns exist for centers with a high net income and high productivity. These patterns indicate that Public and CCC with a high level of at least two of the explanatory conditions, dedication to R&D, annual budget, or size, tend to have high net income (Income) and high productivity (Prod) (Tables 4 and 5 show the configurations for high performance). However, the conditional pattern that combines a high level of dedication to R&D with a high annual budget (solution 3 in Table 5) covers more empirical cases (raw coverage 0.50). Centers A, F, and G most consistently show this conditional pattern. These two models have an overall consistency above 0.8, which is acceptable.



Table 4

Configurations of causal conditions for high and low net income including the individual and overall consistency.

Configuration	Configurations for high net income (Income)			Configurations for low net income (~Income)		
	1	2	3	1	2	3
Centers	B	D	A, F, G	E	C	B
Causal conditions						
RD		+	+	--	--	--
Budget	+		+	-	-	
Size	+	+				
CCC	++	++	++		+	+
Public	++	++	++	+		+
Consistency	1.00	1.00	0.78	0.72	0.74	0.55
Raw coverage	0.23	0.34	0.53	0.37	0.40	0.24
Unique coverage	0.02	0.14	0.33	0.22	0.25	0.09
Overall solution consistency:	0.82			0.78		
Overall solution coverage:	0.69			0.72		

Note: Columns display a conditional pattern for the intermediate result (solution). For each pattern, the table depicts the core conditions with large-size symbols and the peripheral conditions with small-size symbols. A plus symbol (+) indicates that the solution includes the presence of the explanatory condition, while a minus symbol (-) indicates that the solution includes the negation of the explanatory condition. Double + or - represent core conditions, and single + and - indicate peripheral conditions. Blank cells indicate that the explanatory condition is not in the conditional pattern of a solution, that is, its presence or absence is irrelevant for the outcome when associated with that solution.

Table 5

Configurations of causal conditions for high and low productivity.

Configuration	Configurations for high productivity (Prod)			Configurations for low productivity (~Prod)	
	1	2	3	1	2
Centers	B	D	A, F, G	E	C
Causal conditions					
RD		+	+	--	--
Budget	+		+	--	--
Size	+	+			
CCC	++	++	++	-	+
Public	++	++	++	+	-
Consistency	1.00	1.00	0.83	1.00	1.00
Raw coverage	0.20	0.31	0.50	0.25	0.28
Unique coverage	0.02	0.13	0.32	0.25	0.28
Overall solution consistency:	0.86			1.00	
Overall solution coverage:	0.65			0.53	

Note: Columns display a conditional pattern for the intermediate result (solution). For each pattern, the table depicts the core conditions with large-size symbols and the peripheral conditions with small-size symbols. A plus symbol (+) indicates that the solution includes the presence of the explanatory condition, while a minus symbol (-) indicates that the solution includes the negation of the explanatory condition. Double + or - represent core conditions, and single + and - indicate peripheral conditions. Blank cells indicate that the explanatory condition is not in the conditional pattern of a solution, that is, its presence or absence is irrelevant for the outcome when associated with that solution.

Analysis on the negation of outcomes shows three alternative conditional patterns that associate centers with low net income and two alternative causal patterns that lead to low productivity. Either low dedication to R&D or a low budget is a core condition for low performance. Each solution covers only one empirical case. The overall consistency of the negation models shows a consistency that is close to the acceptable threshold.

Discussion

This exploratory study successfully applies the fsQCA (which was so far limitedly applied in healthcare research) to study the association between five explanatory conditions and two financial performance outcomes in a healthcare context. Of the five necessary conditions, this study finds that being a CCC and a public center are necessary for both high income and high productivity when in combination with at least two of the three other possible explanatory conditions (dedication to R&D, annual budget, or size).

However, because almost all of the empirical cases in this study are public and comprehensive cancer centers, a possible argument might be whether this relation will hold true when including more private centers that are not CCCs. The literature suggests that CCCs are usually able to focus efforts and organize resources for the efficient and interactive accomplishment of the goals in patient care and research (Simone, 2002). According to Simone (2002), cancer centers enable and catalyze a high level of cancer-focused achievement that would not happen without such a formal organization of staff and programs. Merkow et al. (2014) show that NCI (National Cancer Institute) accredited cancer centers provide more efficient care and that accredited centers offer more structural resources. Further, the NCI accreditation program is the basis for the OEI accreditation and designation program by which CCCs become accredited (van Harten, 2014).

Further examination of the results shows that three alternate causal patterns exist that associate centers with high performance. These patterns indicate that Public and CCC in combination with a high level of at least two explanatory conditions, dedication to R&D, annual budget, or size, tend to have high net income and high productivity. Providing cancer care is expensive. With the fast increase in the costs of medical technologies and cancer drugs (Aggarwal, Ginsburg, & Fojo, 2014), achieving a high net income and high productivity become far more relevant aims. The research by Delgado (2008) supports the finding that size could be an explanatory condition by showing that larger cancer centers have better financial performance than smaller centers. However, research by Litvak and Bisognano (2011) shows that having more beds does not necessarily make a hospital more efficient. The results indicate that having fewer beds has no association with low

performance. However, a lack of dedication to R&D and a low budget do dictate low performance.

The causal pattern that combines the presence of a high level of dedication to R&D and a high annual budget is a more common solution within the analysis. The fact that centers A, F, and G most consistently have this conditional pattern is interesting because these three centers are in three different geographic regions (Western Europe, Northern Europe and Southern Europe). Although the correct terminology for fsQCA is to speak about causal patterns, some cases such as the patterns with a high level of involvement in R&D might better reflect a relational pattern. If a center receives a lot of money from, for example, the government to perform R&D activities, then a relational pattern exists. If a positive relation exists between R&D and the desired outcomes, then that relation is an extension of the relation between internal innovation and external innovation (Hidalgo & D'Alvano, 2014). Organizations that pursue R&D in healthcare can in fact have a positive effect on the care that organization provides to patients. The results of this study support that the presence of these conditions relates to the presence of high (financial) performance, while their absence relates to low (financial) performance. Therefore, these conditions appear to be critical factors for (financial) performance.

7 This study has several limitations. First, the small sample of centers leads to a lack of variability in some causal conditions. Most of the cancer centers are public with comprehensive care. Although the fsQCA is conducive to small samples like this one, the lack of variability affects the association of the presence of these conditions with the conditional pattern for both the presence and absence of the two outcome conditions (high net income and high productivity). Second, the study uses a limited number of medical conditions that is arbitrary (however based on logical thinking that might convey a possible bias). Including more conditions and more cases could lead to a better understanding of what actually makes these cancer centers (financially) perform so highly. The research by Delgado (2008) shows that significant evidence exists that points to high productivity as a causal condition for high financial performance, which this study does not address. Future research should take this evidence into account. Despite the limitations, this study shows that the fsQCA is a promising method to measure the performance in healthcare and more specifically in health service research. Due to its young age, the fsQCA is still under development and improvement, but the launch of new software packages and innovative forms of graphical representation constantly improves the possibilities for its application.

Conclusion

This study shows that the fsQCA is a powerful approach for health service research despite the method's limited use. The study provides a new avenue of research in healthcare services management through the use of this novel technique. From a managerial and healthcare organizational point of view, this study shows that centers that are public and that offer comprehensive cancer care have both a high income and a high productivity. Second, these centers show a high level of dedication to R&D and have a high annual budget. Some of these variables are likely to be actionable from the point of view of the hospital's management or from the perspective of the healthcare system — so the possible choices that exist for these variables could enable better performance for healthcare providers.

References

- Aggarwal, A., & Sullivan, R. (2014). Affordability of cancer care in the United Kingdom—Is it time to introduce user charges? *Journal of Cancer Policy*, 2(2), 31–39.
- Aggarwal, A., Ginsburg, O., & Fojo, T. (2014). Cancer economics, policy and politics: What informs the debate? Perspectives from the EU, Canada and US. *Journal of Cancer Policy*, 2(1), 1–11.
- BENCH-CAN (2016). Accessed 7 January 2016 <http://www.oeci.eu/Benchcan/>
- Choudhry, S., Choudhry, N. K., & Brennan, T. A. (2005). Specialty versus community hospitals: What role for the law? *Health Affairs Web Exclusive*, W5, 361–372.
- Costa, C. (2008). Editorial. *Revista Portuguesa de Saúde Pública* 7.
- Delgado, R. I. (2008). Financial performance drivers and strategic control: The case of cancer treatment centers. *Texas medical center dissertations*.
- Devine, K., O’Clock, P., & Lyons, D. (2000). Health-care financial management in a changing environment. *Journal of Business Research*, 48(3), 183–191.
- Ferlay, J., Steliarova-Foucher, E., Lortet-Tieulent, J., Rosso, S., Coebergh, J. W. W., Comber, H., & Bray, F. (2013). Cancer incidence and mortality patterns in Europe: Estimates for 40 countries in 2012. *European Journal of Cancer*, 49(6), 1374–1403.
- Fiss, P. C. (2007). A set-theoretic approach to organizational configurations. *Academy of Management Review*, 32(4), 1180–1198.
- Fiss, P. C. (2011). Building better causal theories: A fuzzy set approach to typologies in organization research. *Academy of Management Journal*, 54(2), 393–420.
- Fiss, P. C. (2012). A crisp and fuzzy set qualitative comparative analysis. University of Southern California (Accessed 22 December 2015, http://www.indiana.edu/~wim/docs/11_9_2012_Peer%20Fiss_slides.pdf).
- Gonzalez-Padron, T., Akdeniz, M. B., & Calantone, R. J. (2014). Benchmarking sales staffing efficiency in dealerships using extended data envelopment analysis. *Journal of Business Research*, 67(9), 1904–1911.
- Gospodarowicz, M., Trypuc, J., D’Cruz, J., Khader, O. S., & Knaul, F. (2015). Cancer services and the Comprehensive Cancer Center in disease control priorities: Cancer. Washington: World Bank Group.
- Hidalgo, A., & D’Alvano, L. (2014). Service innovation: Inward and outward related activities and cooperation mode. *Journal of Business Research*, 67(5), 698–703.
- Kahwati, L., Lewis, M., Kane, H., Williams, P., Nerz, P., Jones, K., ... Kinsinger, L. (2011). Best practices in the Veterans Health Administration’s MOVE! Weight management program. *American Journal of Preventive Medicine*, 41(5), 457–464.
- Kaplan, R. S., & Porter, M. E. (2011). How to solve the cost crisis in healthcare. *Harvard Business Review*, 46–64.
- Leape, L., Berwick, D., Clancy, C., Conway, J., Gluck, P., Guest, J., & Pinakiewicz, D. (2009). Transforming healthcare: A safety imperative. *Quality and Safety in Health Care*, 18(6), 424–428.
- Legewie, N. (2013). An introduction to applied data analysis with qualitative comparative analysis (QCA). *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, 14(3) (Art. 15).

- Litvak, E., & Bisognano, M. (2011). More patients, less payment: Increasing hospital efficiency in the aftermath of health reform. *Health Affairs*, 30(1), 76–80.
- Magnier-Watanabe, R., & Senoo, D. (2008). Organizational characteristics as prescriptive factors of knowledge management initiatives. *Journal of Knowledge Management*, 12(1), 21–36.
- Merkow, R. P., Chung, J.W., Paruch, J. L., Bentrem, D. J., & Bilimoria, K. Y. (2014). Relationship between cancer center accreditation and performance on publicly reported quality measures. *Annals of Surgery*, 259(6), 1091–1097.
- Nogueira, D. L., Lira, G. V., Albuquerque, I. M. A. N., & Linhares, M. S. C. (2015). Evaluation of Brazil's teaching hospitals: A systematic review. *Revista Brasileira de Educação Médica*, 39(1), 151–158.
- Plsek, P. E., & Greenhalgh, T. (2001). Complexity science: The challenge of complexity in health care. *BMJ: British Medical Journal*, 323(7313), 625–635.
- Ragin, C. C. (1987). *The comparative method. Moving beyond qualitative and quantitative strategies*. Berkeley: University of California Press.
- Ragin, C. C. (2000). *Fuzzy-set social science*. Chicago: University of Chicago Press.
- Ragin, C. C. (2008). *Redesigning social inquiry: Fuzzy sets and beyond* (pp. 190–212). Chicago: University of Chicago Press.
- Ragin, C. C., & Davey, S. (2009). fs/QCA, version 2.5 [computer program]. Tucson, AZ: Department of Sociology, University of Arizona. (Accessed 14 December 2015, <http://www.socsci.uci.edu/%7Ecragin/fsQCA/software.shtml>).
- Requart, H. (2015). An insider's guide to the transformation of health care. BCG perspectives (Accessed 14 December 2015, <https://www.bcgperspectives.com>).
- Rihoux, B., & Ragin, C. C. (2009). *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques*. London: Sage.
- Saghatchian, M., Thonon, F., Boomsma, F., Hummel, H., Koot, B., Harrison, C., ... van Harten, W. (2014). Pioneering quality assessment in European Cancer Centers: A data analysis of the Organization for European Cancer Institutes Accreditation and Designation Program. *Journal of Oncology Practice* (August 12th Publish Ahead of Print).
- Schneider, J. E., Ohsfeldt, R. L., Morrisey, M. A., Li, P., Miller, T. R., & Zelner, B. A. (2007). Effects of specialty hospitals on the financial performance of general hospitals, 1997–2004. *Inquiry*, 44, 321–334.
- Simone, J. V. (2002). Understanding cancer centers. *Journal of Clinical Oncology*, 20(23), 4503–4507.
- Stokke, O. S. (2007). Qualitative comparative analysis, shaming, and international regime effectiveness. *Journal of Business Research*, 60(5), 501–511.
- Thygeson, N. M., Solberg, L. I., Asche, S. E., Fontaine, P., Pawlson, L.G., & Scholle, S.H. (2012). Using fuzzy set qualitative comparative analysis (fs/QCA) to explore the relationship between medical “homeness” and quality. *Health Service Research*, 47(1), 22–45.
- Thygeson, N. M., Peikes, D., & Zutshi, A. (2013). *Fuzzy-set qualitative comparative analysis: A configurational comparative method to identify multiple pathways to improve patient-centered medical home models*. Rockville: U.S. Department of Health and Human Services.
- Van Harten, W. H. (2014). Comprehensive cancer centres based on a network: the OECl point of view. *Ecancer medical science*, 8(43).

- Van Lent, W. A., de Beer, R. D., & van Harten, W. H. (2010). International benchmarking of specialty hospitals. A series of case studies on comprehensive cancer centres. *BMC health services research*, 10(1), 253–264.
- Wagemann, C., Buche, J., & Siewert, M. B. (2015). QCA and business research: Work in progress or a consolidated agenda? *Journal of Business Research* (Nov. 10th, Epub ahead of print).
- Warren, J., Wistow, J., & Bamba, C. (2013). Applying Qualitative Comparative Analysis (QCA) to evaluate a public health policy initiative in the North East of England. *Policy and Society*, 32, 289–301.
- Woodside, A. G. (2013). Moving beyond multiple regression analysis to algorithms: Calling for adoption of a paradigm shift from symmetric to asymmetric thinking in data analysis and crafting theory. *Journal of Business Research*, 66(4), 463–472.

CHAPTER 8

Actual costs of cancer drugs in 15 European countries

Wim van Harten
Anke Wind
Paolo De Paoli
Mahasti Saghatchian
Simon Oberst

Lancet Oncology 2016;17:18-20.

The financial sustainability of cancer services as part of national health systems is a major challenge;¹ oncology consumes up to 30% of total hospital expenditure and the amount spent on expensive cancer drugs is rising fast.² In view of the pipeline of new drugs, these costs are likely to continue to grow.³ Apart from the risk of unequal access between European countries, burdening health systems with fast-growing costs for these drugs means that the sustainability of cancer care could be compromised.

One element increasingly under scrutiny is the pricing policy of pharmaceutical companies. Both in the recent American Society of Clinical Oncology (ASCO) meeting and in various publications,⁴⁻⁶ attention has been drawn to the lack of transparency on the pricing of various drugs. Different models to judge the appropriateness of prescribing drugs have been proposed; the European Society for Medical Oncology (ESMO) and ASCO models try to balance cost and effectiveness. However, a wholehearted decision to include cost effectiveness as a criterion does not yet seem possible.

An overview of actual prices in European countries does, to our knowledge, not exist, and anecdotal evidence has suggested that differences in price levels might be high. We surveyed the prices for several cancer drugs in European countries through the membership of the European Organization of Cancer Institutes (OECI) and Cancer Core Europe.⁷ A word-based survey was emailed to all full members of the OECI (n=51), both European Union (EU) members and non-EU members, and to the non-OECI member of Cancer Core Europe. The survey was sent to the board of the centre; for most centres the leading pharmacist responded, together with an oncologist. We asked centres to provide list or official and actual prices, corrected for VAT differences, and asked for information about central or government coordinated purchasing. The actual price was defined as the net price—ie, as price per one dose (eg, one 100 mg vial or one capsule of 12.5 mg) to allow for a comparison in case of different pack sizes. Any type of discount was taken into account because we asked the centres to provide us with the price they actually pay. We received a response from 21 centres from 15 countries. Most responses were received in June and July 2015. For some countries we received more than one response and from these we present an average.

We noted that official or list prices differ substantially between countries (up to 92% lower than the highest), and actual prices also differ between countries (up to 58% lower; appendix A). Additionally, reductions on list prices were very different between countries. The table shows a selection of prominent examples (the appendix provides a full overview): Lithuania to show the unavailability of drugs in certain countries; Spain to highlight the high discounts some countries get by contrast with others (e.g., France); and the Nether-

lands as an example of a high-income country. Only data retrieved through the survey are included in the table and appendix A.

Country		Lithuania	Spain (n=2)	France (n=2)	The Netherlands (n=3)
GDP/capita (€)		12,400	22,800	32,200	39,300
Perjeta® 1 vial/420 mg *	Actual price	N/A	2,590.18	2,891.10	3,000.00
	List/ official price	N/A	2,910.58	2,891.10	3,000.00
Mabthera® 1 vial/100 mg *	Actual price	N/A	210.56	266.44	271.13
	List/ official price	N/A	238.06	266.44	279.27
Yervoy® 1 vial/ 50 mg *	Actual price	5,500.00	2,338.83	3,536.50	4,144.00
	List/ official price	5,500.00	4,086.54	3,536.50	4,250.00

Prices are ex VAT. N/A = not available. * Prices are in Euro's per lowest dosage

Table Actual and formal list prices of expensive cancer drugs for common minimal dosages as provided by single- or a number of cancer centres in 4 European countries

No correlation of cost with the existence of central purchasing authorities could be found, other than for France, where its central purchasing system realizes quite low prices over a broad range. However, findings from a study by Lopes and colleagues⁸ showed that France limits discounts granted for reimbursable drugs, which could explain that there is almost no difference between the actual price and the official or list price. By contrast, Spain shows quite high discounts. Apart from the extraordinary budget constraints imposed by the Spanish government, we did not find any explanation for this. It seems that especially in Lithuania, several drugs are not available, and drugs that are available are sometimes expensive, by contrast with what would be expected from GDP per person. If a pattern can be detected, we conclude that pricing seems to be highest in the wealthier countries such as the Netherlands, but a strict association between GDP per person and pricing level could not be established.

In some cases only list or official prices were available or only one official price, because of central purchasing. Many institutions report that confidentiality clauses in their contracts make it difficult for them to freely report prices they pay. At the same time, we noted a very positive attitude towards revealing this information because oncologists are motivated to increase awareness of the issue. Some respondents referred to sensitivity around research funding; the centres that treat most patients with cancer are often cooperating closely with the pharmaceutical industry in undertaking translational drug research, and are at the same time confronted with the drug cost issue.

To date, the pharmaceutical industry has not presented transparent explanations or calculations underlying their pricing decisions. In some countries experiments were started with managed entry agreements, such as pay-for-performance arrangements, but no method-

ologically sound reports have been presented that show a decrease in overall costs at a national level. Neither are we aware of agreements whereby the innovation is rewarded and prices reduce relative to sales volume.

In conclusion, we noted substantial price differences in the prices of cancer drugs in this illustrative example from 15 European countries. This calls for joint action by countries and medical societies with the pharmaceutical industry, since fast and equitable access to promising new drugs is important to improving treatment results. The societal challenge is to combine the development and availability of promising new drugs with the sustainability of our system. All parties involved must agree innovative and sustainable business models to ensure fast access to relevant drugs for patients with cancer.

References

1. Sullivan R, Peppercorn J, Sikora K, et al. Delivering affordable cancer care in high-income countries. *Lancet Oncol* 2011; 12: 933–80.
2. Luengo-Fernandez R, Leal J, Gray A, Sullivan R. Economic burden of cancer across the European Union: a population-based cost analysis. *Lancet Oncol* 2013; 14: 1165–74.
3. Dutch Cancer Society (KWF). Effective new anti-cancer drugs, but the funding system is creaking at the seams. 2015. <https://www.kwf.nl/SiteCollectionDocuments/english-summary-signalingreport-expensivecancermedication.pdf> (accessed Sept 2, 2015).
4. Ramsey S. How state and federal policies as well as advances in genome science contribute to the high cost of cancer drugs. *Health Affairs (Millwood)* 2015; 34: 571–75.
5. Schnipper L, Davidson N, Wollins D, et al. American Society for Clinical Oncology Statement: a conceptual framework to assess the value of cancer treatment options. *J Clin Oncol* 2015; 33: 2563–77.
6. Cherny N, Sullivan R, Dafni U, et al. A standardized generic, validated approach to stratify the magnitude of clinical benefit that can be anticipated from anti-cancer therapies: the European Society for medical oncology Magnitude of Clinical Benefit Scale (ESMO-MCBS). *Ann Oncol* 2015; 26: 1547–73.
7. Celis JE, Ringborg U. Cancer Core Europe: a first step towards a virtual cancer institute in Europe? *Mol Oncol* 2014; 8: 1161–62.
8. Lopes S, Marty C, Berdai D. PHIS pharma profile France 2011. http://whocc.goeg.at/Literaturliste/Dokumente/CountryInformationReports/PHIS_Pharma%20Profile%20FR_2011_final.pdf (accessed Oct 30, 2015).

Appendix A

Actual and formal list prices of expensive cancer drugs for common minimal dosages as provided by single- or a number of cancer centers in 15 European countries listed from the lowest GDP per capita ratio. Prices are ex VAT. N/A = not applicable or not available. For the countries that did not send their prices in Euro's the exchange rate applicable for the date of analysis was used (see below). The same rate was used for actual and list/official prices.

*= Original in RON. Exchange rate 15/06/2015 1 euro = 4.47051 RON

** = Original in HUF. Exchange rate 15/06/2015 1 euro= 313,271536 HUF

***= Original in PLN. Exchange rate 19/10/2015 1 euro= 4.24886829 PLN

****= Original in Pounds. Exchange rate 15/06/2015 1 euro = 0,7240 Pound

*****= Original in NOK. Exchange rate 03/07/2015 1 euro = 8.85 NOK

Country		Romania*	Hungary**	Poland***	Lithuania	Czech Republic	Estonia
GDP/capita (€)		7,500	10,500	10,700	12,400	14,700	14,800
Herceptin®	Actual price	563.14	516.47	N/A	476.56	481.82	674.80
1 vial/150 mg (€) ^a	List / official price	585.26	630.38	537.99	476.56	502.73	N/A
Trastuzumab							
Perjeta®	Actual price	2,677.24	2,594.37	N/A	N/A	2,590.91	2,762.00
1 vial/420 mg (€) ^a	List/ official price	2,677.24	N/A	2,419.94	N/A	2,649.10	N/A
Pertuzumab							
Mabthera®	Actual price	238.19	N/A	N/A	N/A	N/A	295.27
1 vials /100 mg(€) ^a	List/ official price	247.96	220.53	287.72	N/A	420.91	N/A
Rituximab							
Avastin®	Actual price	283.67	N/A	176.53	237.12	N/A	345.26
1 vial/100 mg (€) ^a	List/ official price	297.47	274.65	317.73	237.12	261.82	N/A
Bevacizumab							
Yervoy®	Actual price	4,024.29	N/A	N/A	5,500.00	3,150.91	4,237.38
1 vial/50 mg (€) ^a	List/ official price	4,024.29	3,251.40	4,340.73	5,500.00	3,682.73	N/A
Ipilimumab							
Glivec®	Actual price	15.05	N/A	N/A	17.93	15.80	7.06
1 tab/100 mg (€) ^a	List/ official price	15.12	16.46	17.60	17.93	15.86	N/A
Imatinib							
Sutent®	Actual price	37.52	N/A	N/A	38.48	34.03	38.16
1 cap/12,5 mg (€) ^a	List/ official price	37.52	37.02	44.89	38.48	36.18	N/A
Sunitinib							
Zelboraf®	Actual price	38.96	N/A	34.88	27.15	29.22	N/A
1 tab/240 mg (€) ^a	List/ official price	38.96	37.35	37.67	27.15	29.22	N/A
Vemurafenib							
Xtandi®	Actual price	27.77	N/A	19.30	N/A	25.79	N/A
1 vial/40 mg (€) ^a	List/ official price	27.77	N/A	N/A	N/A	25.79	N/A
Enzalutamide							

(€)^a Prices are in Euro's per lowest dosage

Portugal	Spain (N=2)	Italy (N=2)	France (N=2)	Germany (N=2)	United Kingdom****	Belgium (N=2)	The Nether- lands (N=3)	Norway *****
16,600	22,800	26,600	32,200	32,500	34,500	36,000	39,300	73,400
N/A	506.88	557.70	542.43	670.00	562.71	517.92	605.25	497.74
509.13	573.58	913.45	542.43	718.33	562.71	536.76	606.49	523.50
N/A	2,590.18	2,850.81	2,891.10	2,440.00	3,308.01	2,754.72	3,000.00	N/A
2,761.64	2,910.58	4,113.44	2,891.10	2,580.83	3,308.01	2,840.66	3,000.00	2,474.58
N/A	210.56	262.39	266.44	N/A	241.16	210.38	271.13	210.80
228.80	238.06	395.67	266.44	691.62	241.20	218.38	279.27	217.47
N/A	278.01	300.33	250.65	N/A	335.16	312.26	321.54	260.93
266.41	328.67	412.86	250.65	286.88	335.16	322.89	321.08	286.58
N/A	2,338.83	2,684.92	3,536.50	3,200.00	N/A	4,250.00	4,144.00	3,143.06
2,975.00	4,086.54	5,75484	3,536.50	3,273.25	5,179.56	4,380.80	4,250.00	5,191,65
N/A	18.26	16.35	18.23	24.00	19.85	19.37	20.96	17.51
18.05	18.98	25.10	19.02	24.56	19.85	20.31	20.96	17.74
N/A	39.33	42.50	41.44	N/A	N/A	43.38	36.56	37.85
42.38	40.80	66.01	45.39	52.72	38.70	46.45	43.91	38.19
N/A	33.61	34.94	29.01	26.00	N/A	30.21	29.32	23.16
27.92	39.66	55.61	30.48	27.42	43.16	34.82	37.81	34.12
N/A	21.77	21.51	27.75	N/A	N/A	27.50	27.50	N/A
26.98	21.10	45.61	28.62	29.98	33.73	28.35	27.50	29.00

PART 6

Discussion

CHAPTER 9

General discussion and conclusions

The overall aim of this dissertation, as introduced in chapter 1, was to present tools to benchmark comprehensive cancer care and cancer care pathways/tumor services in the European Union. Linked to this aim are several sub-objectives which are addressed in the papers included in this thesis using multiple research methods including a literature review, quantitative and qualitative analyses. In this final chapter, the main findings regarding the research objectives are summarized and discussed, followed by the methodological considerations. Subsequently, the generalizability, areas for future research are defined and implications for policy and practice are discussed. Finally, concluding remarks are made.

Reflection on main findings

Research objective 1: Assessing the current situation of benchmarking in specialty hospitals and existing quality assessments

Our first research objective concerned the evaluation of the current situation of benchmarking in specialty hospitals and existing quality assessments.

Chapter 2

In chapter two we systematically obtained literature to establish the state of the art in the field of benchmarking in specialty hospitals. The results of the 24 included papers showed a focus on indicator development which is usually an indication of development rather than implementation. Overall it seemed that the majority of the studies lacked a structured design, and did not report on benchmark outcomes. Benchmarking seemed to be most reported up on and possibly developed in the field of oncology and eye hospitals, however most studies do not describe a structured benchmarking method or a model that can be repeatedly used. In order to evaluate the effectiveness of benchmarking to improve quality in specialty hospitals, robust and structured designs are needed including a follow up to check whether the benchmark study has led to improvements.

Chapter 3

In chapter 3 we obtained an overview of existing (performance) assessments at European cancer centers to inform the development of the benchmark tool. Based on the responses from 19 cancer centers from 18 member states, we found 109 assessments in Europe. The number of assessments have steadily increased from 1990s till 2015. The number of care focused assessments rose most rapidly, followed by the mixed assessments of patient care and research aspects. We tried to assess the content of the assessment, however, not

all existing assessment reports of EU CCs are available in English and/or easily accessible. Hence, it is hard to know the exact criteria that are being used in these assessments. The criteria that could be retrieved were evaluated for their use in the benchmark tool. The majority of assessments (n=63) are done at the national level, followed by international level assessments (n=38). There are only a handful of regional assessments (n=9). Almost all mandatory assessments are national and are mainly related to keeping license and/or receiving public funding. In contrast, most voluntary assessments seem to be international, and mainly aim at quality improvement. Improving the organization of care by quality assessments seems a justified goal, but it may be questioned whether the efforts are justifiable if no quantifiable clinical benefits can be shown.

Research objective 2: Measuring patient's perspectives on the quality of care at cancer centers

Our second research objective concerned taking into account the perspective of the patient in looking at quality of care in cancer centers.

Chapter 4

In chapter 4, it was described that overall patient satisfaction with care received at cancer centers in Europe on a scale from 1-10 was high. In order to measure patient experience and satisfaction we adapted an existing Dutch Consumer Quality Index. To our knowledge, the questionnaire used in this study is the first that measures the experiences and satisfaction of cancer patients with care provided by cancer centers in Europe. A total of 698 patients from six European countries filled the European Cancer Consumer Quality Index (ECCQI). The questionnaire was tested for internal consistency, Cronbach's alpha was good or satisfactory in 8 out of 10 categories. The cognitive interviews revealed problems in 18 of the 65 items. Patient satisfaction significantly differed between the countries. We observed no difference in patient satisfaction for age, gender, education, and tumor type, but satisfaction was significantly higher in patients with a higher level of activation. We found no significant differences between tumor types, supporting the use of a generic questionnaire. Outcomes of a patient experience and satisfaction survey should not be seen as objective, hard measures of quality of care, but as a measure of whether the care provided lives up to the patients expectations.

Research objective 3: Develop and pilot two extensive benchmark tools for comprehensive cancer care

The third objective of this study, develop and pilot two extensive benchmark tools resulted in two tools.

Chapter 5

Chapter 5 dealt with 1) development and pilot of an extensive benchmark tool with both qualitative and quantitative indicators, 2) identification of performance differences between cancer centers, and 3) the identification of improvement opportunities.

The selected pilot setting consisted of eight (comprehensive) cancer centres, geographically distributed throughout Europe. Regarding the development of the tool we developed 61 qualitative and 141 quantitative indicators based on literature and expert opinions and structured in a framework (BENCH-CAN framework) to produce a comprehensive set and valuable improvement suggestions. Stakeholders were involved in multiple phases (indicator development, indicator refinement, verification of the data).

For the second aim of this study, the identification of performance differences between cancer centers we found on many indicators, such as bed utilization and the provision of survivorship care. These differences can partially be explained by country and health system related differences, but these did not prevent us from providing relevant feedback data to the pilot sites. This does however shows the difficulty of international benchmarking which confirms findings from two recent international benchmarking studies in eye hospitals¹ and comprehensive cancer centers².

Finally, the research team distilled improvement opportunities for each pilot site. The suggestions varied from very concrete suggestions such as the development of a Computerized Physician Order Entry (CPOE) to more general suggestions such as the increasing of patient involvement in the strategy/organization. The most common mentioned suggestion was the implementation of a case manager. There was a high degree of agreement with the suggestions (85%). As most centers identified the improvement suggestions as being mid or long-term the effect of these suggestions is not expected in the coming years. We do expect a difference in the uptake of the suggestions with concrete suggestions such as the development of an CPOE being implemented faster and more successfully.

Chapter 6

In chapter 6 we discussed the development of a tool for the benchmarking of pathways and assessed the degree of development of cancer centers towards Integrated Practice Units (IPUs). The pilot included seven cancer centers in Southern and Central/Eastern

Europe. We found that although cancer centers all have a clear director for the pathway development and a clear bases (international and national guidelines) cancer centers have no clear strategy when developing pathways. The majority of the centers don't have a structured evaluation method. The pathway benchmark data proved to be usable for the assessment of IPU development. Regarding this development we found that three centers stated to have this developed for breast cancer. Assessing the centers against the pathway criteria by the Harvard Business School³ we identified three groups. Group 1, centers fully meeting all criteria (center A and G), and group 2, partially meeting the criteria (center B) all indicated to have a breast IPU. Based on our data we agree with center A and G having an IPU for breast cancer. However for center B an improvement in four criteria is recommended to have a fully implemented IPU. Group 3, centers meeting only some of the criteria did not indicate to have this type of organization implemented. Dedicating specific resources such as staff, measuring outcomes specific to the IPU and IPU specific administrative and scheduling systems were lacking.

Research objective 4: Investigate the use of quantitative benchmarking data and (financial) performance features for international comparison

The fourth objective this study resulted in the application of a fairly new approach in health service research and an overview of cancer drug prices.

Chapter 7

In chapter 7 we examined the use of the fuzzy-set Qualitative Comparative Analysis (fsQCA) method to identify performance features of cancer centers in the European Union. The study successfully applied the fsQCA (which was so far limitedly applied in healthcare research) to study the association between five explanatory conditions (level of dedication to R&D, annual budget, size, type, and whether the center is a comprehensive cancer center) and two financial performance outcomes (net income and productivity) in a cancer center context. Of the five necessary conditions, this study found that being a Comprehensive Cancer Center and a public center are necessary for both high income and high productivity when in combination with at least two of the three other possible explanatory conditions (dedication to R&D, annual budget, or size).

Chapter 8

A focused benchmark was performed on costs for expensive cancer drugs as pharmaceutical drug pricing policies are increasingly under scrutiny. This is enforced by the lack of transparency on the pricing of various drugs. As there was anecdotal evidence that the dif-

ference in prices of cancer drugs in Europe were high, a survey amongst cancer centers was performed assessing the list and actual price for cancer drugs. We noted that substantial differences between countries with regard to, official or list prices (up to 92% lower than the highest), and actual prices (up to 58% lower). Additionally, reductions on list prices differed, where for example Spain shows quite high discounts. Apart from the extraordinary budget constraints imposed by the Spanish government, we did not find any explanation for this. It seems that especially in Lithuania, several drugs are not available, and drugs that are available are sometimes very expensive, by contrast with what would be expected from GDP per person. If a pattern can be detected, we conclude that pricing seems to be highest in the wealthier countries such as the Netherlands, but a strict association between GDP per person and pricing level could not be established.

Methodological considerations

Framework and indicators

A structured and well documented approach was used for the benchmark development². The literature found in chapter 1 and existing assessments identified in chapter 2 were used as a basis for the benchmark tool. The framework developed to structure the indicators was partially based on the EFQM model of Excellence. The EFQM model is both generic as specific⁴. As described by Nabitz et al.⁴ it is derived from theories on organizational change and knowledge management and innovation enabling performance-assessment and identification of key strengths and improvement areas. The indicators were derived from both literature as expert and stakeholder opinion. This resulted in “new” indicators which can be seen as both a strength as a limitation. A strong aspect of this is that by asking cancer centers what they want to know from other cancer centers we increase the likeliness that the benchmark tool will actually be used. As mentioned before the development of “new” indicators could also be a limitation as these indicators have not been tested before. This study could have laid more emphasis on the methodological quality of the indicators. However, the validity of the indicators was checked using feedback from the pilot centers based on three criteria 1) definition clarity, 2) data availability and reliability, 3) discriminatory features and usability for comparisons. In chapter 5 and 6, a draft indicator set was prepared and consensus was achieved using a stakeholder exercise across Europe. It is however inevitable that the drafting of the initial indicator set was biased by the researcher. It is therefore important to keep in mind that indicators in this benchmark tool are not the ultimate or only set to benchmark quality and efficiency in cancer centers in Europe.

We developed this benchmark tool by involving stakeholders from across Europe throughout the development and piloting phases. This has helped ensure acceptability and accountability among stakeholders. The tools contain both qualitative and quantitative indicators encompassing the whole organization of cancer centers. Next to this we developed a tool to measure the patients perspective, chapter 4, therefore adding an extra dimension to quality assessments of cancer centers.

Benchmark pilot/data collection

The focus of this project was learning from peers instead of judging. Centers were hesitant at first because it reminded them of an accreditation, but we succeeded in explaining the differences between accreditation and benchmarking. The casual non-competitive atmosphere lead to an open collaboration in which pilot sites felt they could learn from each other. This corresponds to findings by de Korne¹ who states that including a number of institutes providing similar services (such as cancer care) in a non-competitive environment could provide a fruitful landscape for benchmarking.

During the pilot phase (chapter 5 and 6) we found that for some institutes it was difficult to provide all requested data. This was mainly the case for the quantitative indicators where different reimbursement systems led to different recording systems. It seems the confidentiality of some data was the reason for one of the pilot sites dropping out. Dropping of cancer centers resulted in an uneven distribution of centers, only two centers from North/Western Europe were included compared to four from Southern Europe and four from Central/Eastern Europe. Another possible reason for this could be that centers in Western Europe felt that the benchmark tool would not yield sufficient improvement suggestions. This was however not the case for the included center in Western Europe. This lack of geographical distribution could have resulted in the missing of good practices.

Budget Impact Analysis

In order to gain insight into the resources needed from the oncologic centres to perform the benchmarking exercise a Budget Impact Analysis (BIA) was carried out. As such, it provided an estimate of the costs of implementing benchmarking in future institutions by multiplying the hours spend by the weighted expenditures on wages of the staff involved in the benchmark. There was a high variation in the estimated spent hours, which could be explained by the wide variation in data availability between centers due to differences in size, geographical location and country, and difference in reimbursement systems. The most important factor influencing the easiness of data collection seemed the ongoing or recently finished accreditation program by the OECl⁵ and/or dedicated ICT system for

registration and automatic reimbursement. Besides the costs for staff, centers did not have to invest any money to be part of a benchmark while being provided with improvement suggestions and good practice examples from peers. This makes benchmarking a feasible tool for quality improvement which enables centers to learn from peers without having to re-invent the quality improvement wheel. Most good practices did not require major investments but were merely clever ways of organizing care or involving patients. The benchmark can be used as a tool to generate discussions about how to deliver services, from a volume-based organization of care to a value-based organization for example. This increases the understanding and the learning from each other, rather than actual performance comparison.

Good practices

As mentioned in the introduction various good practices were identified in this study. Despite much discussion, there is no universally accepted definition of a good practice. Little was found in literature on good practice identification, nor on specific good practices for cancer care. At a minimum a good or best practice must⁶: 1. Demonstrate evidence of success; 2. Affect something important, that is, contribute to the organization's mission or program goals; 3. Have the potential to be replicated or adapted to other settings. Good practice within this study are defined as: the existence of a good or interesting idea in a given practice, one that deserves attention. Firstly pilot sites were asked what they thought were their good practices, secondly literature was reviewed for good practices and thirdly an expert panel with representatives from European cancer centers identified good practices in the collected data. This triangulation led to the good practices described in this study. Although we tried to identify these good practices as objectively as possible by using data triangulation (institutes own suggestion, literature and expert opinion) some bias was inevitable. We did make sure that all good practices fitted the minimum requirements⁶. Mosel and Gift⁷ state that benchmarking is a continuous process, the institutes in this study were only benchmarked once. The good practices identified are therefore only snapshots. In order to see whether the benchmark actually led to quality improvement it needs to be repeated in a couple of years and good practices need to be updated to the present situation.

Generalizability

The tools developed in this project were piloted in nine cancer centers. As these centers were from different regions in Europe we feel that the results are generalizable to other centers in Europe as well as non-European countries such as the USA, or Australia. The validity and reliability of the results and good practices was checked in multiple phases of the

study, for example by performing a site visit, and sending all results (although anonymized) to all centers so they could not only verify their own data but also see how other centers performed. Applicability in third world countries is probably only feasible for a subset, as for instance a certain level of infrastructure (ICT, technologies) is required. Cancer care is represented in nearly all hospitals and not just cancer centers. As the indicators were only measured in cancer centers and not in general hospitals providing cancer care it is unclear whether the outcomes are generalizable for cancer care in general. As to the applicability of the tool to other healthcare areas than oncology we expect the benchmark tool, to be applicable to other settings as well, because cancer care seems representative for many (multidisciplinary) hospital processes, due to its urgency, complexity and uncertainty. Thus, the benchmark tools seem general applicable, while the scope of the resulting changes is context specific. This context specificity was shown in transfer of good practices and the implementation of the improvement suggestions. This was not always possible mainly due to regulatory and financial restraints.

Future research

The findings of this thesis have implications for future research as well for a wide range of stakeholders (i.e. Cancer Centers and other cancer care providers, policy makers and regulatory agencies). In the following paragraph future research areas are set out.

More rigorous evaluation methods

In order to evaluate whether the benchmark tool developed in this study is an effective tool for quality improvement, a time series design is needed as this allows studying the variability in performance of the cancer centers and the sustainability of improvements⁸. Using the input collected through the earlier described BIA the costs of participating to this benchmark could be related to the outcomes of the improvement opportunities assessing the feasibility of the benchmark tool. Although difficult, it would be beneficial to include more (clinical) outcome indicators to see whether these are affected by the improvement in the organization. Second, from a methodological perspective, the use of a control group would enable the assessment of whether improvements occur due to the benchmark or would have occurred naturally. This is however often impossible due to organization specific characteristics and the willingness of hospitals to participate in this type of research and would require a much longer time frame than was available for the BENCH-CAN project (three years).

Future research should involve larger series of hospitals that include organizations with similar characteristics (cancer centers) and organizations with different characteristics (for example general hospitals) to assess the actual discriminative capabilities of the benchmark tool. This will allow both literal and theoretical reproduction⁸. As this study only used data over the year 2012, the sensitivity for changes over time could not be established, this should be included in future studies.

Inclusion of more and diverse centers is also needed to verify the relationships identified through the application of the fuzzy-set Qualitative Comparative Analysis (chapter 7). Including more conditions (other types of hospitals) and more cases could lead to a better understanding of what actually makes these cancer centers (financially) perform highly.

Identify a suitable and feasible benchmark processes

The research design used in this study was fit for pilot purposes (one researcher working full-time for the project) but will most likely not be applicable in practice. The BENCH-CAN manual⁹ provides two methods to perform the benchmark: i) Health facility (cancer center or hospital) instigated benchmarking, and ii) Third party benchmarking. Healthcare facilities can decide to hire a third party to serve as an objective, evaluating party in the benchmarking process. This approach would however require a bigger investment as this third party would have to be paid. Third parties can be any company or independent professional body with experience in benchmarking. Each benchmarking project is different and can have different processes, steps. Therefore the models presented do not necessarily fit all situations. As the developments steps of the 13 step method by van Lent have already been executed focus should lie with the data collection and analysis. It is recommended to develop an online data collection tool to aid the data collection. Future research should assess whether the proposed methods are feasible in practice and what other methods could be identified.

Implications for policy and practice

Implications for cancer centers and other healthcare facilities

Accounting for the patient's perspective is gaining importance in quality measurement and the organization of cancer centers¹⁰. The tool described in chapter 4 can assist cancer centers to fulfill that requirement. The BENCH-CAN tools, described in chapter 5 and 6, provide cancer centers with a tool to internationally compare performance, identify good practices and improve their organization through the adaption of these good practices.

The tools can be used for self-assessment, benchmarking against the good practices and identifying indicators for which processes or data is lacking. How organizational learning takes place in healthcare is very much dependent on the relationships between the different stakeholders such as clinicians and management. This enforces the importance of having a benchmark team that involves all stakeholders. This will stimulate interaction and thereby learning on tactical and operational levels. This could prove an incentive to attract and motivate personnel.

The developed benchmark tool, depending on the process used (e.g. health facility instigated or third-party) is a relatively cheap (in terms of financial input) tool to see where a center stands compared to peers and where improvements are needed. The good practices identified through the BENCH-CAN project can be starting points for quality improvements. Both benchmark tools offer an extensive range of indicators, but centers could also decide to focus on a certain domain, for example patient centeredness.

The evaluation of an IPU described in chapter 6 showed that the implementation of an IPU led to higher productivity and utilization and reduced costs. A challenge remains for hospitals to link data on outcomes to costs. As the implementation of an IPU holds several barriers the decision to integrate should not be made lightly and health systems factors such as regulatory restrictions and reimbursement issues need to be taken into account. For all developments and suggestions described it is important to keep in mind that these cannot be implemented over night. Healthcare managers need to think on the long-term, anticipating future opportunities and threats based on today's data. Benchmarking should not be a one-time event, but part of continuously quality improvement, repeating the benchmark every few years to measure whether quality improvements were achieved and to make sure good practices remain up-to-date.

Policy makers

The overview about existing assessments given in chapter 3 can assist policy makers towards regulating these as well as minimizing the related bureaucracy. The tools developed were tested in international environment but could be used on a national level as well. Policy makers can use the benchmark to see what services are doing particularly well. We do however advise to use that information to focus resources on centers/areas that require improvement rather than "punishing" these centers for scoring low on the benchmark. As discussed before, policy makers play a role in encouraging transparency of performance outcomes. Within our study we also found that some improvement suggestions could not be implemented due to regulatory restrains. Here lies a responsibility for policy makers

making sure that healthcare facilities such as cancer centers get the opportunity to deliver the best possible care, as that is also what is demanded from them.

As service provision changes, for example by the creation of IPUs financing systems will need to change. The unit of reimbursement needs to be equal to the unit of value, e.g. encompassing the care pathway. Chapter 8 showed that benchmarking cannot only be used to minimize inequalities and differences but also to highlight them and call for action.

Conclusions

As scientific material on healthcare benchmarking methods and results were scarcely available a thorough benchmark development based on exiting quality assessments was conducted. This study showed that benchmarking is a useful instrument to identify (quality) performance differences between cancer centers in Europe. An extensive set of tools was developed and tested to look at quality of care from three perspectives: qualitatively, quantitatively and through the eyes of patient. To further test the usability and feasibility of these tools for quality improvement future research should adopt a time series design including more healthcare organizations.

References

1. de Korne D, Sol J, van Wijngaarde J, van Vliet EJ, Custers T, Cubbon M, et al. Evaluation of an international benchmarking initiative in nine eye hospitals. *Health Care Manage R.* 2010;35:23-35.
2. van Lent W, de Beer R, van Harten W. International benchmarking of specialty hospitals. A series of case studies on comprehensive cancer centres. *BMC Health Services Research* 2010; 10: 253.
3. President & Fellows of Harvard College. Integrated Practice Units. <http://www.isc.hbs.edu/health-care/vbhcd/pages/integrated-practice-units.aspx> Accessed on 19 August, 2016.
4. Nabitz U, Klazinga N, Walburg JAN. The EFQM excellence model: European and Dutch experiences with the EFQM approach in health care. *International Journal for Quality in Health Care.* 2000; 12(3): 191-202.
5. <http://oeci.selfassessment.nu/cms/> Accessed 7 August, 2016
6. National Collaborating Centre for Methods and Tools. Sharing internal best practices. Hamilton, ON: McMaster University, 2005 (Updated 29 March, 2011) <http://www.nccmt.ca/resources/search/84>. Accessed on 21 May, 2015.
7. Mosel D, Gift B. Collaborative benchmarking in health care. *The Joint Commission journal on quality improvement* 1995; 20(5): 239-249.
8. Eccles M, Grimshaw J, Campbell M, Ramsay C. Research designs for studies evaluating the effectiveness of change and improvement strategies. *Quality and Safety in Health Care* 2003; 12: 47.
9. Wind A, Nefkens I, van Dijk J, Thonon F, Sebestyén E, et al. BENCH-CAN implementation manual. http://oeci.eu/Benchcan/Doc/general/Resources/BenchCan_Manual_2016_FINAL.pdf accessed 16 November 2016
10. Wind A, Roeling MP, Heerink J, Sixma H, Presti P, Lombardo C, van Harten W. Piloting a generic cancer consumer quality index in six European countries. *BMC cancer* 2016; 16(1): 711.

CHAPTER 10

Summary/Samenvatting

Summary

Chapter 1 Introduction

The number of cancer patients and survivors is steadily increasing and despite or perhaps because of rapid improvements in diagnostics and therapeutics, important inequalities in cancer survival exist within and between different countries in Europe. Improving the quality of care is part of the approach to reduce suboptimal cancer survival and minimize inequalities in Europe. Quality Improvement (QI) is an essential part of healthcare management and can be sought on macro-, meso- and microlevels of the health care system. The aim of this thesis was to contribute to the knowledge on how to develop and use benchmarking for quality improvement of cancer care. There are several classifications of benchmark types and models. This study focused on combining two types of benchmarking: operations and clinical practice (process) benchmarking and associated patient experience; and performance benchmarking. It compared relative performance drawing on quantitative and qualitative data.

Chapter 2 Benchmarking specialty hospitals, a scoping review on theory and practice

Chapter 2 had the following objectives: (i) provide an overview of research on benchmarking in specialty hospitals and care pathways, (ii) describe study characteristics such as method, setting, models/frameworks, and outcomes, (iii) verify the quality of benchmarking as a tool to improve quality in specialty hospitals and identify success factors. Of 1,817 articles identified in total, 24 were included in the study. Articles were categorized into: pathway benchmarking, institutional benchmarking, articles on benchmark methodology or -evaluation and benchmarking using a patient registry. We found that benchmarking seems to be more developed in eye hospitals, emergency departments and oncology specialty hospitals. Some studies showed promising improvement effects. However, the majority of the articles lacked a structured design, and did not report on benchmark outcomes. In order to evaluate the effectiveness of benchmarking to improve quality in specialty hospitals, robust and structured designs are needed including a follow up to check whether the benchmark study has led to improvements. Several success factors were identified.

Chapter 3 Quality assessments for cancer centers in the European Union

Chapter 3 presents key findings from a survey that was conducted with cancer centers in the European Union. The goal was to obtain an overview of existing assessments in terms of whether they are: mandatory or voluntary; focused on evaluating research or patient care

or both; regional, national and/or international. Based on the responses from 19 cancer centers from 18 member states, there are 109 assessments in Europe. The numbers have steadily increased from 1990s till 2015. The number of patient care assessments have risen most rapidly in Europe, followed by the mixed assessments of patient care and research aspects. The rise in pure research assessments has not increased much. Some mixed assessments (combining research and care elements) may involve assessing translational research, which translates research to practice. However, not all existing assessment reports of EU cancer centers are available in English and/or easily accessible. Hence, it is hard to know the exact criteria that are being used in these assessments. The majority of assessments (n=63) are done at the national level, followed by international level assessments (n=38). There are only a handful of regional assessments (n=9). Almost all mandatory assessments are national and are mainly related to keeping license and/or receiving public funding. In contrast, most voluntary assessments seem to be international, and mainly aim at quality improvement and are seldom directly tied to licensing or funding.

Chapter 4 Piloting a Generic Cancer Consumer Quality Index in six European countries

The research question of chapter 4 was twofold: 1. What are the differences in patient experience and satisfaction between countries and/or patient characteristics? 2. What is the validity and internal consistency (reliability) of the European Cancer Consumer Quality Index? An existing Consumer Quality Index was adapted and translated into the local language at the participating pilot sites using cross-translation. A total of 698 patients from six European countries filled the questionnaire. Patient satisfaction significantly differed between the countries. We observed no difference in patient satisfaction for age, gender, education, and tumor type, but satisfaction was significantly higher in patients with a higher level of activation. This European Cancer Consumer Quality Index (ECCQI) showed promising scores on internal consistency (reliability) and a good internal validity. Cronbach's alpha was good or satisfactory in 8 out of 10 categories.

Chapter 5 Development of a Benchmark tool for Cancer Centers; results from a pilot exercise

Differences in cancer survival exist between countries in Europe. Benchmarking of good practices can assist cancer centers to improve their services aiming for reduced inequalities. The aim of the BENCH-CAN project was to develop a cancer care benchmark tool and yield good practice examples, contributing to improving the quality of interdisciplinary care. The aims of the study described in chapter 5 were (i) to develop and pilot a benchmark tool for cancer care with both qualitative and quantitative indicators, (ii) identify perfor-

mance differences between cancer centers, and (iii) identify improvement opportunities. A benchmark tool was developed and executed according to a 13 step benchmarking process. The final tool consisted of 61 qualitative and 141 quantitative indicators, which were structured in an evaluative framework. Data from all eight participating centers showed inter-organization variability on many indicators, such as bed utilization and provision of survivorship care. The tool allowed comparison of inter-organizational performance. Improvement opportunities were successfully identified for every center involved and the tool was positively evaluated.

Chapter 6 Benchmarking cancer centers: from care pathways to Integrated Practice Units

Structuring cancer care in pathways can reduce variability in clinical practice and improve patient outcomes. International benchmarking can aid centers with regard to development, implementation and evaluation. A further step in the development of multidisciplinary care is organizing care in Integrated Practice Units (IPU), encompassing the whole pathway and relevant organizational aspects. Research on this topic is however limited. This chapter describes the development and results of a benchmark tool for cancer care pathways and exploration of IPU development in cancer centers. The benchmark tool was developed according to a 13 step benchmarking method. The tool was piloted in seven cancer centers in Europe. Benchmark results showed that most centers have formalized multidisciplinary pathways, teams differed in composition and almost twofold differences were found in mammography use efficiency. Improvements suggestions included to position pathways formally and structurally evaluate outcomes in a sufficient high frequency. Overall we found centers in Europe are in various stages of development of pathways and IPUs, varying from an informal pathway structure to a full IPU type of organization. Organizing cancer care in an IPU could yield multiple performance improvements.

Chapter 7 Management and performance features of cancer centers in Europe: A fuzzy-set analysis

The aim of the study in chapter 7 is to identify the performance features of cancer centers in the European Union by using a fuzzy-set qualitative comparative analysis (fsQCA). The fsQCA method represents cases (cancer centers) as a combination of explanatory and outcome conditions and is a method to identify the importance of quantitative performance features and how they relate to outcomes. This method provides a systematic, transparent, and exhaustive analytical approach in the realm of comparative research. This study explored the use of the fsQCA to investigate the association between preselected explanatory conditions (the level of dedication to R&D, annual budget level, size, type, and

whether the center is a comprehensive cancer center) and financial performance outcomes (net income, the profit per discharge; and productivity, the total number of patient visits divided by the number of inpatient beds). This study used data on seven centers from a European benchmarking project: BENCH-CAN. Despite the modest number of cases, the study successfully applied the fsQCA. The findings show that public, comprehensive cancer centers with at least two of the three other explanatory conditions (dedication to R&D, annual budget, or size) have an association with high net income and high productivity.

Chapter 8 Actual costs of cancer drugs in 15 European countries

The financial sustainability of cancer services as part of national health systems is a major challenge; oncology consumes up to 30% of total hospital expenditure and the amount spent on expensive cancer drugs is rising fast. In view of the pipeline of new drugs, these costs are likely to continue to grow. Apart from the risk of unequal access between European countries, burdening health systems with fast-growing costs for these drugs means that the sustainability of cancer care could be compromised. An overview of actual prices in European countries does, to our knowledge, not exist, and anecdotal evidence has suggested that differences in price levels might be high. We surveyed the prices for several cancer drugs in European countries through the membership of the European Organization of Cancer Institutes (OECI) and Cancer Core Europe. We noted substantial price differences in the prices of cancer drugs in this illustrative example from 15 European countries. This calls for joint action by countries and medical societies with the pharmaceutical industry, since fast and equitable access to promising new drugs is important to improving treatment results.

Chapter 9 Discussion

This thesis showed that benchmarking is a useful instrument to identify (quality) performance differences between cancer centers in Europe. An extensive set of tools was developed and tested to look at quality of care from three perspectives: qualitatively, quantitatively and through the eyes of patient. To develop this set a structured and well documented approach was used. The literature found in chapter 1 and existing assessments identified in chapter 2 were used as a basis for the benchmark tool. The framework developed to structure the indicators was partially based on the EFQM model of Excellence. We developed this benchmark tool by involving stakeholders from across Europe throughout the development and piloting phases. This has helped ensure acceptability and accountability among stakeholders. However it is however inevitable that the drafting of the initial indicator set was biased by the researcher. It is therefore important to keep in mind that indicators in this benchmark tool are not the ultimate or only set to benchmark

quality and efficiency in cancer centers in Europe. The tools contain both qualitative and quantitative indicators encompassing the whole organization of cancer centers. Next to this we developed a tool to measure the patients perspective, chapter 4, therefore adding an extra dimension to quality assessments of cancer centers.

The focus of this project was learning from peers instead of judging. Centers were hesitant at first because it reminded them of an accreditation, but we succeeded in explaining the differences between accreditation and benchmarking. The casual non-competitive atmosphere lead to an open collaboration in which pilot sites felt they could learn from each other. In order to gain insight into the resources needed from the oncologic centres to perform the benchmarking exercise a Budget Impact Analysis (BIA) was carried out. As such, it provided an estimate of the costs of implementing benchmarking in future institutions by multiplying the hours spend by the weighted expenditures on wages of the staff involved in the benchmark. There was a high variation in the estimated spent hours, which could be explained by the wide variation in data availability between centers due to differences in size, geographical location and country, and difference in reimbursement systems.

In order to evaluate whether the benchmark tool developed in this study is an effective tool for quality improvement a time series design is needed as this allows studying the variability in performance of the cancer centers and the sustainability of improvements. Future research should involve use larger series of hospitals that include organizations with similar characteristics (cancer centers) and organizations with different characteristics (for example general hospitals) to assess the actual discriminative capabilities of the benchmark tool. Future research should also assess whether the proposed benchmarking methods are feasible in practice and what other methods could be identified.

Samenvatting

Hoofdstuk 1 Introductie

Het aantal kankerpatiënten en overlevenden neemt gestaag toe en ondanks of misschien wel dankzij snelle verbeteringen in de diagnostiek en therapie, zijn er belangrijke verschillen in overleving van kanker binnen en tussen de verschillende landen in Europa. Verbetering van de kwaliteit van de zorg is een onderdeel van het minimaliseren van de ongelijkheid in Europa. Quality Improvement (QI) is een essentieel onderdeel van de gezondheidszorg management en kan worden gezocht op macro-, meso- en microniveau van het zorgstelsel. Het doel van dit proefschrift is een bijdrage te leveren aan de kennis over hoe een benchmarking tool te ontwikkelen en te gebruiken voor de verbetering van de kwaliteit van kankerzorg. Er zijn verschillende classificaties van benchmark-types en modellen. Dit onderzoek richt zich op het combineren van twee soorten van benchmarking: “operations” en de klinische praktijk (proces) benchmarking en de bijbehorende ervaringen van patiënten; en “performance” benchmarking.

Hoofdstuk 2 Benchmarking van gespecialiseerde ziekenhuizen, een literatuur studie naar theorie en praktijk

Hoofdstuk 2 had de volgende doelstellingen: (i) een overzicht geven van het onderzoek naar benchmarking in gespecialiseerde ziekenhuizen en zorgpaden, (ii) studie kenmerken, zoals methode, omgeving, modellen / kaders, en de resultaten beschrijven, (iii) verifiëren van de kwaliteit van benchmarking als een instrument om de kwaliteit van gespecialiseerde ziekenhuizen te verbeteren en het identificeren van succesfactoren. Van 1817 artikelen zijn in totaal 24 opgenomen in deze studie. Artikelen zijn onderverdeeld in: zorgpad benchmarking, organisatie benchmarking, artikelen op benchmark-methodiek of -beoordeling en benchmarking met behulp van een patiënt register. We vonden dat benchmarking meer ontwikkeld lijkt in oogziekenhuizen, spoedeisende hulp afdelingen en oncologisch gespecialiseerde ziekenhuizen. Sommige studies toonden veelbelovende verbeter effecten. Bij de meeste artikelen ontbrak echter een gestructureerd ontwerp en werd niet gerapporteerd over resultaten. Om de effectiviteit van benchmarking om de kwaliteit in gespecialiseerde ziekenhuizen te verbeteren te meten zijn robuuste en gestructureerde studie ontwerpen nodig inclusief een follow-up om te controleren of de benchmark studie heeft geleid tot verbeteringen. Verschillende succesfactoren werden geïdentificeerd.

Hoofdstuk 3 Kwaliteitsbeoordelingen voor kankercentra in de Europese Unie

Hoofdstuk 3 presenteert de belangrijkste bevindingen van een onderzoek dat werd uitgevoerd onder kankercentra in de Europese Unie. Het doel was om een overzicht van de bestaande kwaliteitsbeoordelingen te verkrijgen: verplicht of vrijwillig; gericht op het evalueren van onderzoek of patiëntenzorg of beide; regionaal, nationaal en /of internationaal. Op basis van de reacties van 19 kankercentra uit 18 lidstaten, vonden wij 109 assessments in Europa. De aantallen zijn gestaag toegenomen van 1990 tot 2015. Het aantal patiëntenzorg assessments is het snelst gestegen, gevolgd door de gemengde beoordelingen van de patiëntenzorg en onderzoek aspecten. De stijging van de pure onderzoeksbeoordelingen is niet veel gestegen. Niet alle bestaande evaluatieverslagen van de kankercentra in de EU zijn beschikbaar in het Engels en / of gemakkelijk toegankelijk. Daarom is het moeilijk om de exacte criteria die worden gebruikt in deze beoordelingen te onderzoeken. Het merendeel van de beoordelingen (n=63) worden gedaan op het nationale niveau, gevolgd door internationale beoordelingen (n=38). Er zijn slechts een handvol regionale beoordelingen (n=9). Bijna alle verplichte beoordelingen zijn nationaal en hebben voornamelijk betrekking op het houden van een licentie en / of ontvangen van overheidsfinanciering. In tegenstelling, de meeste vrijwillige evaluaties lijken internationaal te zijn, en met name gericht op kwaliteitsverbetering. Deze worden zelden rechtstreeks gekoppeld aan de licentie en financiering.

Hoofdstuk 4 Het testen van een "Generic Cancer Consumer Quality Index" in zes Europese landen

De onderzoeksvraag van hoofdstuk 4 was tweeledig: 1. Wat zijn de verschillen in de ervaringen en de tevredenheid van patiënten tussen landen en / of patiënt kenmerken? 2. Wat is de validiteit en interne consistentie (betrouwbaarheid) van de European Cancer Consumer Quality Index? Een bestaande Consumer Quality Index werd aangepast en vertaald in de lokale taal van de deelnemende pilot-sites met cross-vertaling. Een totaal van 698 patiënten uit zes Europese landen vulden de vragenlijst in. We vonden dat de tevredenheid van de patiënt aanzienlijk verschilde tussen de landen. Wij namen geen verschil waar in patiëntentevredenheid voor verschillen in leeftijd, geslacht, opleiding en tumortype, maar tevredenheid was significant hoger bij patiënten met een hoger niveau van activering. Deze Europese Cancer Consumer Quality Index (ECCQI) toonde veelbelovende scores op interne consistentie (betrouwbaarheid) en een goede interne validiteit. Cronbach's alpha was goed of bevredigend in 8 van de 10 categorieën.

Hoofdstuk 5 Ontwikkeling van een Benchmark instrument voor kankercentra; resultaten van een proefproject

Er zijn grote verschillen in de overleving van kanker tussen landen in Europa. Benchmarking van goede praktijken/voorbeelden kan kankercentra helpen om hun dienstverlening te verbeteren en streeft naar verminderde ongelijkheid. Het doel van het BENCH-CAN project was om een oncologische zorg benchmark-tool te ontwikkelen en voorbeelden van goede praktijken/voorbeelden te identificeren, die bijdragen aan de verbetering van de kwaliteit van de interdisciplinaire zorg. De doelstellingen van de in hoofdstuk 5 beschreven studie waren (i) het ontwikkelen en testen van een benchmark tool voor kankerzorg met zowel kwalitatieve als kwantitatieve indicatoren, (ii) het identificeren van prestatieverschillen tussen kankercentra, en (iii) het identificeren van verbetermogelijkheden. De benchmark tool was ontwikkeld en uitgevoerd op basis van een 13 stap proces voor benchmarking. Het uiteindelijke instrument bestond uit 61 kwalitatieve en 141 kwantitatieve indicatoren, die werden gestructureerd in een evaluatieve kader. Gegevens van alle acht deelnemende centra toonde inter-organisatie variabiliteit tussen vele indicatoren, zoals bed gebruik en de verstrekking van nazorg. De tool maakt vergelijking van inter-organisatorische prestaties mogelijk. Verbeter mogelijkheden werden met succes geïdentificeerd voor elke betrokken centrum betrokken en de tool werd positief geëvalueerd.

Hoofdstuk 6 Een benchmark tool voor kankercentra; van zorgpaden naar Integrated Practice Units

Structureren van kankerzorg in zorgpaden kan variabiliteit in de klinische praktijk verminderen en patiënt uitkomsten verbeteren. Internationale benchmarking kan centra helpen met betrekking tot de ontwikkeling, implementatie en evaluatie van zorgpaden. Een verdere stap in de ontwikkeling van multidisciplinaire zorg is het organiseren van zorg in Integrated Practice Units (IPU), die het hele traject en de relevante organisatorische aspecten bevatten. Onderzoek naar dit onderwerp is echter beperkt. Dit hoofdstuk beschrijft de ontwikkeling en de resultaten van een benchmark-tool voor kankerzorgpaden en exploratie van IPU ontwikkeling in kankercentra. De benchmark tool is ontwikkeld volgens een 13 stappen benchmarking methode. De tool werd getoetst in zeven kankercentra in Europa. Benchmark resultaten toonden aan dat de meeste centra multidisciplinaire zorgpaden hebben geformaliseerd, multidisciplinaire teams verschillen in samenstelling en bijna tweevoudige verschillen werden gevonden in mammografie efficiëntie. Over het algemeen vonden we dat kankercentra in Europa in verschillende stadia van ontwikkeling van zorgpaden en IPU's zijn, variërend van een informele zorgpad structuur naar een volledige IPU type organisatie. Het organiseren van kankerzorg in een IPU kan meerdere prestatieverbeteringen opleveren.

Hoofdstuk 7 Management en prestatiekenmerken van kankercentra in Europa: een fuzzy-set analyse

Het doel van de studie in hoofdstuk 7 is om prestatiekenmerken van kankercentra in de Europese Unie te identificeren met behulp van een fuzzy-set kwalitatieve vergelijkende analyse (fsQCA). De fsQCA methode representeert gevallen als een combinatie van verklarende en uitkomst omstandigheden en is een methode om het belang van kwantitatieve prestatiekenmerken en hoe zij betrekking hebben op de resultaten te identificeren. Deze methode zorgt voor een systematisch en transparant verloop, en volledige analytische aanpak op het gebied van vergelijkend onderzoek. Deze studie onderzocht het gebruik van de fsQCA om de associatie tussen voorgeselecteerde verklarende omstandigheden (het niveau van toewijding aan R&D, de jaarlijkse begroting niveau, grootte, type, en of het centrum een integraal kankercentrum is) en de financiële prestaties van de resultaten (netto inkomen, de winst per ontslag; en de productiviteit, het totale aantal patiënt bezoeken gedeeld door het aantal intramurale bedden) te bestuderen. Deze studie gebruikte gegevens van zeven centra binnen een Europees benchmarking project: BENCH-CAN. Ondanks het bescheiden aantal gevallen, wist de studie met succes de fsQCA toe te passen. De bevindingen tonen aan dat publieke, integrale kankercentra met ten minste twee van de drie andere verklarende voorwaarden (toewijding aan R&D, de jaarlijkse begroting, of grootte) een associatie tonen met een hoog netto-inkomen en een hoge productiviteit.

Hoofdstuk 8 De werkelijke kosten van kankermedicijnen in 15 Europese landen

De financiële houdbaarheid van kankerzorg als onderdeel van het nationale gezondheidstelsels is een belangrijke uitdaging; oncologie verbruikt tot 30% van het totale ziekenhuis uitgaven en het bedrag besteed aan dure kankermedicijnen stijgt snel. Gezien de nieuwe geneesmiddelen in de pijnlijn, zullen deze kosten waarschijnlijk blijven groeien. Naast het risico van ongelijke toegang tussen de Europese landen, worden gezondheidssystemen belast met snelgroeiende kosten voor deze medicijnen wat betekent dat de duurzaamheid van de kankerzorg kan worden aangetast. Een overzicht van de werkelijke prijzen in Europese landen bestaat, voor zover wij weten, niet, en anekdotisch bewijs suggereert dat de verschillen in prijsniveaus hoog zouden kunnen zijn. We onderzocht de prijzen voor een aantal geneesmiddelen tegen kanker in Europese landen. We vonden aanzienlijke verschillen in de prijzen van geneesmiddelen tegen kanker in dit illustratieve voorbeeld uit 15 Europese landen. Dit vraagt om gezamenlijke actie van landen en medische verenigingen met de farmaceutische industrie, omdat een snelle en rechtvaardige toegang tot veelbelovende nieuwe medicijnen belangrijk is voor de verbetering van behandel resultaten.

Hoofdstuk 9 Discussie

Dit proefschrift toonde aan dat benchmarking een nuttig instrument is voor het identificeren van (kwaliteit) prestatieverschillen tussen kankercentra in Europa. Een uitgebreide set van tools is ontwikkeld en getest om te kijken naar de kwaliteit van de zorg vanuit drie perspectieven: kwalitatief, kwantitatief en door de ogen van de patiënt. Om deze set te ontwikkelen is een gestructureerd en goed gedocumenteerde aanpak gebruikt. De literatuur gevonden in hoofdstuk 1 en bestaande assessments die in hoofdstuk 2 werden geïdentificeerd, zijn gebruikt als basis voor de benchmark tool. Het kader ontwikkeld om de indicatoren te structureren werd gedeeltelijk gebaseerd op het EFQM-model of Excellence. Deze benchmark instrumenten zijn ontwikkeld in samenwerking met stakeholders uit heel Europa, waarbij stakeholders betrokken waren gedurende de hele ontwikkeling. Dit heeft bijgedragen aan de aanvaardbaarheid en de toerekenbaarheid van de tools voor de belanghebbenden. Het is echter onvermijdelijk dat het opstellen van de oorspronkelijke indicator set werd vertekend door de achtergrond onderzoeker. Daarom is het belangrijk om in gedachten te houden dat de indicatoren in deze benchmark tool niet de ultieme of enige indicatoren set is voor het benchmarken van de kwaliteit en efficiëntie bij kankercentra in Europa. De tools bevatten zowel kwalitatieve en kwantitatieve indicatoren en omvat de hele organisatie van kankercentra. Daarnaast ontwikkelden we een instrument (zie hoofdstuk 4) om het patiënten perspectief te meten, hierbij een extra dimensie aan kwaliteitsbeoordelingen van kankercentra toevoegend.

De focus van dit project was het leren van collega's in plaats van oordelen. Centra waren aarzelend in eerste instantie omdat het hen deed denken aan een accreditatie, maar we zijn erin geslaagd een informele niet-competitieve sfeer te creëren. Deze sfeer heeft geleid tot een open samenwerking waarin proeflocaties vonden dat ze kunnen leren van elkaar. Om inzicht te krijgen in welke middelen nodig waren van de oncologische centra om de benchmark uit te voeren middelen werd een Budget Impact Analysis (BIA) uitgevoerd. Als zodanig is het een schatting van de kosten van de uitvoering van benchmarking in toekomstige instellingen door vermenigvuldiging van de uren met de gewogen uitgaven van de lonen van het personeel dat meewerkte bij de benchmark. Deze BIA toonde grote verschillen tussen kankercentra, deze kunnen mogelijk verklaard worden door verschillen in beschikbare data, verschillen in grootte, geografische ligging, en het verschil in financiële systemen.

Om te beoordelen of de benchmark tool ontwikkeld in deze studie een doeltreffend instrument is voor kwaliteitsverbetering is een tijdreeks ontwerp nodig. Dit maakt het mogelijk om de variabiliteit in de prestaties van de kankercentra te bestuderen en de duurzaamheid van de verbeteringen. Toekomstig onderzoek moet grotere series ziekenhuizen en

organisaties met vergelijkbare kenmerken (kankercentra) en organisaties met verschillende eigenschappen bevatten (bijvoorbeeld algemene ziekenhuizen). Dit kan worden gebruikt om de feitelijke discriminerende mogelijkheden van de benchmark tool te beoordelen. Toekomstig onderzoek moet ook beoordelen of de voorgestelde benchmarking methoden haalbaar zijn in de praktijk en welke andere methoden kunnen worden geïdentificeerd.

CHAPTER 11

Acknowledgement/Dankwoord

Mijn boekje is af! Na een drie jaar en een paar maanden durende rollercoaster is het gelukt alle papers af te krijgen, de laatste loodjes wogen verreweg het zwaarst. Dit proefschrift was er niet gekomen zonder de bijdrage en support van velen. Een aantal van hen wil ik hier in het bijzonder voor bedanken.

In de eerste plaats gaat mijn dank uit naar mijn promotor: prof. dr. Wim van Harten. Beste Wim, ik wil je heel graag bedanken voor je altijd kritische blik, je oplossing als ik het niet zag zitten, sturing als ik uit de bocht dreigde te vliegen en je snelle feedback. Je hebt mij gestimuleerd om zowel uit het project als uit mezelf het maximale te halen, het steeds een beetje beter te maken. Het begin van mijn PhD was heel erg wennen, maar dankzij jou heb ik mijn draai gevonden en ik ben heel blij met hoe onze samenwerking is geweest. Ik hoop dat onze wegen elkaar in de toekomst weer zullen kruisen.

Dan de secretaresses van Wim van Harten bij het NKI-AVL, Jorrita en Marianne. Bedankt voor alle hulp en voor de gezellige momenten tijdens het wachten op mijn overleg. Speciale dank aan Marion, Wim van Hartens secretaresse bij het Rijnstate. Beste Marion, we moesten beiden even wennen, maar heel erg bedankt voor je ondersteuning bij het regelen van alle promotiezaken. Mijn dank gaat ook uit naar de leden van de promotiecommissie, prof. dr. N.S. Klazinga, Dr. M.W.J.M. Wouters, Prof. dr. R. Torenvlied, prof. dr. J.A.M. van der Palen, en Prof. dr. ir. E.W. Hans hartelijk dank dat jullie mijn manuscript wilden lezen en beoordelen.

The BENCH-CAN project was a group effort and without the people from the consortium this project would not have been possible. First of all I would like to thank Joris van Dijk and Isabelle Nefkens from PANAXEA for all their feedback and support. Secondly my deepest gratitude goes to Pietro Presti and Edit Marosi, Edit and Pietro thank you for being great companions for all the site visits. I really enjoyed the professional discussions, but also the dinners and the joint site seeing in all the cities that we visited. From NIO I would also like to thank Peter Nagy for being a great colleague. Peter, your support and feedback helped me in times when I was doubting myself and my work. All other people at NIO thank you for your willingness to be a pilot center.

Edit Sebestyen thank you and your colleagues from Health ClusterNet very much for your coordination, and feedback. Jonathan Watson, although we did not work together very often I appreciate your feedback and support throughout the project. Frederique Thonon thank you for our collaboration. Germana Gianquinto and Claudio Lombardo, thank you for making sure everything ran smoothly, Claudio I hope I was able to explain the difference between accreditation and benchmarking.

To all centers that participated in this study, my deepest gratitude and although I am aware that this was a center wide effort there are some specific people I would like to thank. First of all Lucia Da Pieve from CRO, I really enjoyed working together not only on BENCH-CAN but also on other projects. Monica Groza, from IOCN, thank you, I hope the tulips are still doing well! Marco Asioli, INT, thank you for all your feedback and help. Tuula Helander, HUCH, thank you for your warm welcome to Finland. Ernestas Janulionis, NCI, you were always critical but very helpful. I really enjoyed visiting Vilnius again. Special thanks goes to Francisco Nuno Rocha-Gonçalves, IPO, you helped me with my first publication and with two more. Your ideas really helped to sharpen my PhD.

Het NKI-AvL was zowel pilot site als werkplek tijdens mijn PhD. Vanuit de pilot site kant wil ik alle werknemers bedanken die hebben bijgedragen aan de data collectie. In het bijzonder gaat mijn dank uit naar Anke Serrarens, Anke jou altijd positieve feedback heeft mij heel erg op weg geholpen. Daarnaast wil ik graag Wineke van Lent bedanken, Wineke jou kritische blik maakte mijn indicatoren veel scherper.

Ik wil mijn collega's van de afdeling Psychosociaal Onderzoek en Epidemiologie (PSOE) bedanken voor de kritische en opbouwende feedback en de gezelligheid tijdens het jaarlijkse uitje, de afdelingsborrels en het kerstdiner. Melanie, Ann-Jean, Vera en Jacqueline dank voor een super leuk jaar als feestcommissie. Alle andere mensen binnen het NKI, dank voor de gezelligheid tijdens borrels en de goeie gesprekken bij een kop koffie. Aan mijn nieuwe collega's op de gang van de RvB dank voor jullie warme welkom.

The "Wim van Harten group": Abi, Anna, Ann-Jean, Bruno, Laura, Melanie, Wilma, Valesca and Wim. I really appreciated sharing our research experiences and ideas, and of course the dinners and borrels.

Melanie speciale dank aan jou natuurlijk voor het feit dat je mijn paranimf wilde zijn, maar daarnaast voor het zijn van de beste buurvrouw. We hebben niet heel lang naast elkaar gezeten, maar in die tijd hebben onze gesprekken mij heel erg geholpen. Ik waardeer dat we nog steeds samen kunnen lachen, maar ook af en toe even lekker klagen.

Vrienden, friends. Zonder namen te noemen, bedank ik al mijn vrienden die mij in de afgelopen jaren hebben gesteund of juist de druk van de ketel haalden. Weekendjes weg, soms ver weg, soms dichtbij, drankjes drinken, samen eten, al deze momenten hebben mij zeer geholpen om deze reis te volbrengen. For my English friends, I'm not going to name you all separately, but thank you very much for being there, sometimes listening to my frustrations but also to celebrate the happy moments.

Mijn familie. Broertjes en zusjes, oma's en opa's, ooms en tantes, Sandra, Kris en Monica, neefjes ik wil jullie allemaal bedanken voor de interesse en steun die jullie, allemaal op jullie eigen manier, gegeven hebben de afgelopen jaren. Papa, dank voor alle steun en af en toe delen in de PhD frustraties. Onze inhoudelijke gesprekken hebben mij erg geholpen het grotere geheel van de gezondheidszorg en alle processen die daarin spelen te zien. Mama, dank voor alle opbeurende woorden tijdens koffies op de UT en op andere momenten in tijden dat ik het even niet zag zitten. Ondanks dat het antwoord op de vraag: "kan ik nog wat voor je doen" vaak nee was, veel dingen moest ik alleen doen, heeft dit mij erg geholpen.

En dan natuurlijk Daan, mijn paranimf en mijn liefste. Dank voor alle ruimte en steun die je me geeft, zonder jou waren mijn artikelen allemaal te lang geweest en was dit proefschrift er zeker niet zo snel gekomen. Jij hebt verreweg het meeste meegekregen van de hele reis die mijn PhD was, het vieren van de geaccepteerde artikelen, maar ook alle frustratie af en toe, ik had me hier geen betere partner in kunnen wensen.

Ten slotte wil ik dit proefschrift opdragen aan mijn opa die er helaas niet meer is. Vanuit de grond van mijn hart dank ik je voor alle steun gedurende mijn hele leven, de wijze raad en de stimulatie om te studeren en het beste uit mezelf te halen zonder me onder druk te zetten. Jou aanwezigheid zal gemist worden, niet alleen tijdens mijn verdediging, maar voor altijd. Ik weet dat jij ontzettend trots zou zijn geweest.

CHAPTER 12

Curriculum vitae

Anke Wind was born in Hengelo on 8th of May 1988. After graduating high school in Hengelo she completed a Bachelor in Human Geography and Planning from the University of Groningen. She pursued an Erasmus Mundus Masters degree in Sustainable Regional Health Systems and Public Health. This master was provided by 4 EU universities: Deusto University (Spain), Corvinus University (Hungary), Vilnius University (Lithuania) and University of Verona (Italy). She graduated cum laude in 2013. From June 2013 she pursued a PhD at the Netherlands Cancer Institute-Antoni van Leeuwenhoek (NKI-AVL) in Amsterdam in collaboration with the University of Twente in Enschede. The topic of her PhD was the development and pilot of benchmark tools for cancer centers. This project comprised several sub projects such as the adaptation and pilot of a European Cancer Consumer Quality Index, a project that was initiated by Anke. During her PhD she supervised two master students. This PhD was performed as part of the BENCH-CAN project funded by the European Commission. Anke is currently employed as a policy advisor quality and safety at the NKI-AVL. Her professional interest lies in areas of healthcare quality improvement, policy implementation, stakeholder engagement, patient participation and international collaboration.

