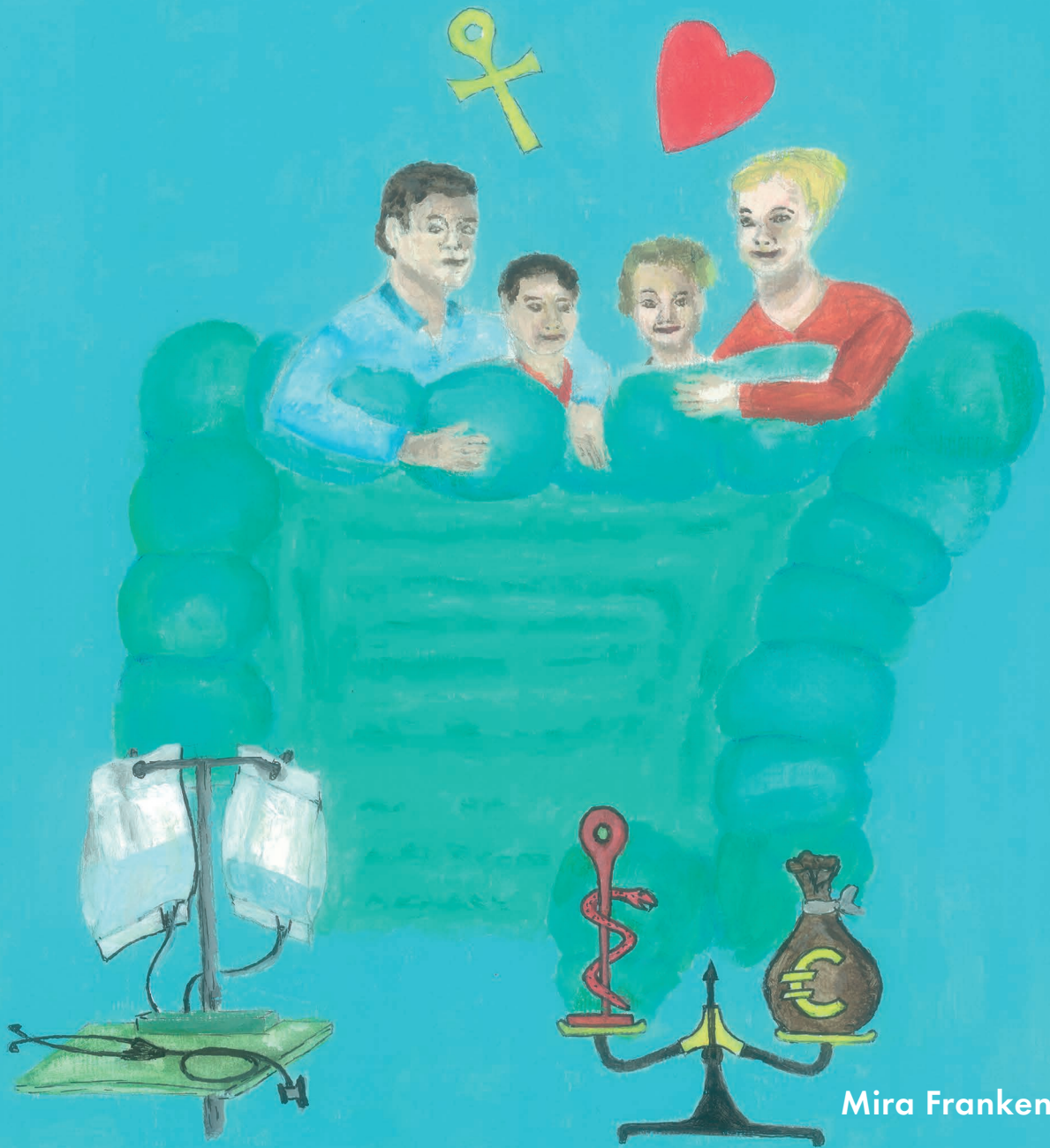
COST-EFFECTIVENESS AND WORK ABILITY IN COLORECTAL CANCER

Towards maintaining sustainable cancer healthcare



Mira Franken

COST-EFFECTIVENESS AND WORK ABILITY IN COLORECTAL CANCER

Towards maintaining sustainable cancer healthcare

Mira D. Franken

Cost-effectiveness and work ability in colorectal cancer

Towards maintaining sustainable cancer healthcare

Thesis with a summary in Dutch, Utrecht University

Proefschrift met een samenvatting in het Nederlands, Universiteit Utrecht

© Mira Franken, 2021, Utrecht, the Netherlands

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means without written permission of the author or the publisher holding the copyright of the published articles.

Financial support for the work in this thesis was kindly provided by the department of Medical Oncology of the University Medical Center Utrecht, Servier Nederland Farma, Chipsoft.

ISBN: 978-94-93197-53-4

Cover: mijn vader - Wouter Franken

Lay-out and printing: Off Page, Amsterdam

COST-EFFECTIVENESS AND WORK ABILITY IN COLORECTAL CANCER

Towards maintaining sustainable cancer healthcare

KOSTENEFFECTIVITEIT EN WERKVERMOGEN BIJ DIKKE DARM- EN ENDELDARMKANKER

Op weg naar het behoud van duurzame kankergezondheidszorg

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof. dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

donderdag 20 mei 2021 des middags te 4.15 uur

door

Mira Desirée Franken
geboren op 1 november 1979
te Delft

PROMOTOREN

Prof. dr. M. Koopman
Prof. dr. C.J.A. Punt

COPROMOTOREN

Dr. M.G.H. van Oijen

TABLE OF CONTENTS

1

INTRODUCTION AND OUTLINE OF THIS THESIS

INTRODUCTION

Colorectal cancer (CRC) is the third most common cancer type worldwide and the median age of patients is about 70 years.[1,2] Although there are hereditary traits which might attribute to 35% of CRC cases, only 5% are due to hereditary forms such as familial adenomatous polyposis and non-polyposis colon cancer (Lynch syndrome).[2] Over the last decade, the incidence of CRC in Western countries has stabilized and CRC death has declined due to advances in detection and treatment.[3–5] However, even though the overall incidence of CRC has stabilized, an increase in incidence in Western countries has been reported in patients younger than 50 years of age.[3]

In the Netherlands, CRC has an incidence of approximately 14,000 patients/year and an annual death rate of around 5,100 patients.[6] Approximately 20-25% of patients have (synchronous) metastatic colorectal cancer at the time of diagnosis[7] and about 20% of patients with non-metastatic disease at the time of diagnosis, will eventually develop (metachronous) metastatic disease.[8]

The development of novel treatment strategies for CRC has resulted in an overall decline in mortality.[4,7] However, healthcare expenses for cancer treatment have increased dramatically over the last decades.[9,10] For instance, in 2009 European healthcare costs of CRC (stage I-IV disease) were estimated at €5.57 billion of which €4.04 billion was spent on in-patient care and €565 million was spent on drugs.[9] We should therefore also consider the impact of treatment choices on healthcare resources while retaining the best possible clinical outcome.

LOCAL TREATMENT FOR COLORECTAL CANCER

Upon CRC diagnosis, treatment decisions are dependent of disease staging. Malignant polyps and some stage I tumours, in which the tumour invades the submucosa, can be curatively treated with endoscopic resections techniques (endoscopic mucosal resection, endoscopic submucosal dissection, or endoscopic full-thickness resection).[11] For more advanced local disease stages (stage I-III), surgery is the standard of care to achieve cure. Adjuvant chemotherapy may increase the survival rate. Patients with pT4N0 high-risk stage II microsatellite stable (MSS*) colon cancer are offered adjuvant systemic therapy with a fluoropyrimidine plus oxaliplatin. Currently, DNA mismatch repair (MMR) deficiency status is the only biomarker with a clinical implication for patients with stage II disease.

* In CRC, two phenotypes can be discerned: microsatellite stable (MSS) and microsatellite instability (MSI). MSI is characterized by mismatch repair deficiency, resulting in many microsatellite mutations and is demonstrated in about 15 % of patients. MSS is characterized by chromosomal changes and demonstrated in 85% of patients.[2]

For all stage III patients, adjuvant systemic therapy with a combination of fluoropyrimidine and oxaliplatin, or fluoropyrimidine monotherapy - in case of contraindication for oxaliplatin and MSS tumour-, is standard.[11–14] Historically, the duration of adjuvant treatment is 6 months. For capecitabine as fluoropyrimidine in combination with oxaliplatin, the treatment duration has recently been limited from 6 to 3 months resulting in less comorbidities without compromising the disease-free survival rates.[13,15] In patients with intermediate to high-risk rectal cancers, neoadjuvant (chemo)radiotherapy has been shown to decrease the local recurrence rate.[11]

Of note, only a subset of patients with stage III disease benefit from adjuvant therapy, while all patients are exposed to systemic treatment and their adverse events. Around 50% of patients with stage III disease achieve long term survival with surgery only and only 20% of patients achieve long term survival following adjuvant systemic treatment.[16,17] There is therefore an urgent need for biomarkers so that patients who might benefit of adjuvant treatment can be identified and patients who do not benefit do not receive unnecessary potentially harmful treatment.[17]

SYSTEMIC TREATMENT FOR METASTATIC COLORECTAL CANCER

Patients with metastatic CRC (mCRC) can be subdivided in 2 categories. For patients with limited metastatic disease, primary tumour resection along with local treatment options with curative intent exist (such as liver surgery, radiofrequency ablation and stereotactic radiotherapy).[11,18] However, the majority of mCRC patients have disseminated permanently unresectable disease for which palliative systemic treatment is indicated.

The standard of care for first line systemic treatment with mCRC is a combination of chemotherapy (fluoropyrimidine, oxaliplatin and/or irinotecan) and targeted therapy against Vascular Endothelial Growth Factor (VEGF), such as bevacizumab.[18] Patients with left-sided RAS and $BRAF^{V600E}$-wild type mCRC may also be treated in first line with chemotherapy in combination with cetuximab or panitumumab, antibodies targeting the Endothelial Growth Factor Receptor (EGFR), or should receive anti-EGFR antibody therapy in subsequent line. Patients with right-sided mCRC do not benefit from anti-EGFR containing therapies, irrespective of RAS/BRAF mutation status.[11] Recently, bevacizumab containing maintenance treatment compared to observation was evaluated in first line treatment. Based on large clinical randomized controlled trials (RCTs) for patients with mCRC, this treatment strategy has resulted in improvements in progression-free survival after first-line treatment without affecting their quality of life.[19–21] Overall, the availability of novel agents and improvements in treatment strategies has resulted in a median overall survival of up to 30 months for mCRC patients participating in clinical trials.[22] It has previously been demonstrated that clinical trial results only have external validity if projected on patients which meet the original study eligibility criteria. If

these eligibility criteria were not met, survival was significantly worse in the general patient population compared to the patients enrolled in clinical trials.[23] Further caution in the extrapolation from trial results to daily practice was recently evident from data showing that the median overall survival for mCRC patients in daily practice has not improved in the Netherlands in the last decade. Only mCRC patients who received metastasectomy and/ or cytoreductive surgery with hyperthermic intraperitoneal chemotherapy (HIPEC) demonstrate a clinically relevant improvement in survival over time.[24]

Lack of survival improvements might be a result of a mostly uniform treatment approach for all patients despite the heterogeneous nature and the complex molecular pathogenesis of CRC. A major challenge for patient tailored therapy, is the lack of sufficient predictive markers for treatment efficacy. Only a limited set of predictive biomarkers, such RAS/BRAF status, currently influence treatment choices. Further knowledge on predictive and prognostic biomarkers, using techniques such as next-generation sequencing along with phenotypic functional assays (i.e. organoids) combined with clinical data, is necessary for personalized treatment to come into reach.[17]

QUALITY OF LIFE

Even though quality of life (QoL) should be an important parameter in treatment choices, only about half of clinical cancer trials specify inclusion of QoL outcomes in their protocol. About 20% of clinical trials eventually report on the quality-of-life outcomes.[25] Quality of life can be measured using health-related generic quality of life questionnaires (i.e. EQ-5D [26]) or disease specific questionnaires (i.e. QLQ-C30[27]).

Generic health-related quality of life questionnaire outcomes are generally used to express health-related quality of life utilities, which are of use in cost-effectiveness studies. A frequently used questionnaire for this purpose, is the EQ-5D questionnaire developed by EuroQoL research foundation. The EQ-5D questionnaire addresses 5 domains: mobility, self-care, daily activities, pain and mood, and also includes a visual analogue scale to rate quality of life on a scale from 0-100.[26] Utility values reflect preferences for different health states, benchmarked towards country specific outcomes. EQ-5D outcomes are used for utility calculations, where utilities are expressed on a scale of 1 – best imaginable health and less than zero reflecting health states worse than death.[28–31] Surprisingly high utilities have been reported in CRC patients, for all disease stages, with utilities ranging from 0.64-0.95[32–34], while the VAS score of the EQ-5D reflected a worse quality of life and seemed incongruent compared to the utilities.[33] Therefore, the EQ-5D questionnaire is regarded to be insufficiently sensitive for disease specific quality of life measurements in cancer patients.

The European Organization for Research and Treatment of Cancer (EORTC) has developed a cancer specific quality of life instrument for use in Clinical Trials, the QLQ-C30.

During the development and validation of the QLC-C30 questionnaire, significant changes in quality of life domains for patients were observed while receiving treatment for non-resectable lung cancer.[27] Treatment-related toxicity for CRC patients surprisingly did not show a negative impact on global health QLQ-C30 score[35,36], while increases in toxicity were seen in two-thirds of the patients.[35] In addition, in an elderly patient population (>70 years) no decline in global health or symptom scores were reported[36] or only short term function decline for patients receiving surgery.[37] Previously, QoL as measured with the QLQ-C30 instrument in CRC patients and compared to the general population, did not demonstrated differences in global health score at 1 and 3 years after CRC diagnosis. However, patients experienced clinically meaningful decrements in the emotional and social functioning domain, and in specific symptoms (dyspnea, constipation and diarrhea) up to 3-15 years after CRC diagnosis.[38,39] Tumour location was of influence on symptoms, for instance patients with rectal cancer experienced more decrements in social functioning, abdominal problems, pain and fatigue.[39,40] For CRC patients <60 years, decrements in role functioning and financial difficulties after 3 years were additionally reported.[38] Elderly patients (>70 years) did not show clinically meaningful decrements on symptom scores or global health score.[38,40] In contrast, a large cross-sectional survey on social distress demonstrated that up to 15% of CRC patients experienced social distress. The strongest risk factors associated with social distress where the presence of ≥3 comorbidities, unemployment and recurrent or metastatic disease.[41] Thus, a global health score alone is insufficient to assess the impact on quality of life, and symptom scores should be evaluated more thoroughly when assessing the impact of treatment strategies on QoL and risk factors for social distress should be identified. Recently, the EORTC has proposed a QLQ-C30 summary score, which might become a more relevant outcome measure.[42]

WORK ABILITY

The ability to work is important for mental wellbeing and quality of life.[43,44] However, an important number of cancer patients fail to return to work, while work absenteeism or reduced work ability has been associated with reduced quality of life.[45,46] Work absenteeism or reduced work ability is influenced by multiple factors. For instance, younger age, higher levels of education, continuity of care, absence of surgery, less physical symptoms, marital status, cancer type, disease stage have been associated with return to work. [46] Cancer did not only reduce employability, but also resulted in a decline in income. More important effects on employment and income were observed in advanced stages of disease.[47]

Data available on work participation for CRC patients show that up to 40 % of patients experience financial distress,[48] 45-85% decrease or cease working following or during CRC treatment [49,50], resulting in reduction of annual labour income.[51] Advanced disease stage, lower level of education, sick leave in the year prior to diagnosis, reduced quality of life in the physical functioning component score and children aged under 18 at home have

been associated with increased risk for receiving a disability pension.[47,50,52,53] Receiving chemotherapy, ongoing chemotherapy treatment and sick leave at the time of diagnosis were predictors for reduced work ability or delayed return to work.[53–55]

As the retirement age currently rises in the Netherlands and the incidence of CRC in patients under 50 years has increased, more CRC patients will face challenges regarding work ability and quality of life. Therefore, additional information on work ability and quality of life is necessary for improvements in patient support towards a return to work.

COST-EFFECTIVENESS

In order to maintain access to novel treatments and technologies in a social system such as used in the Netherlands, the available financial resources have to be divided. To achieve this, cost-effectiveness studies have gained much importance. The panel on cost-effectiveness in Health and Medicine, has previously made recommendations regarding the evaluation of costs versus effectiveness (defined as life year (LY) gained or quality of life year (QALY) gained).[29] A QALY is defined as a measure expressing live years gained as a result of treatment corrected for utilities.[29] Generic and non-disease specific questionnaires, such as the EQ-5D, are used to calculate health-related quality of life utilities.[26,30,56] Incremental costs are dependent of the perspective of the cost-effectiveness study. For instance, a hospital perspective will include predefined hospital costs, while disregarding any costs or cost reductions for society. The primary outcome in cost-effectiveness models is the incremental cost per QALY gained ratio or incremental cost per LY gained ratio (ICER).[29]

Additionally, to standardize the evaluation of clinical benefit of novel cancer treatments and to improve the use of healthcare resources, tools to quantify the clinical benefit of cancer treatment therapies have developed the American Society of Clinical Oncology (ASCO) Value in Cancer Task Force and the European Society for Medical Oncology (ESMO).[57,58] The ASCO value tool also includes a cost evaluation in addition to the clinical benefit evaluation.[57] In the Netherlands, the committee for evaluation of systemic cancer treatment of the Dutch Society for Medical Oncology (NVMO) also uses a value tool according to the 'PASKWILL' criteria in order to achieve national consistency regarding the adoption of novel treatments in medical oncology. These criteria evaluate efficacy (overall and progression free survival benefit), the ESMO Magnitude of Clinical Benefit scale, toxicity and QoL. Treatment costs are additionally reported upon, although not included in the final recommendation.[59]

Recently, a negative correlation between incremental cancer drug cost (monthly cost of the novel regimen compared to the control or standard treatment arm) and clinical benefit as assessed by the ASCO value tool has been demonstrated. In other words, higher drug costs were correlated with less clinical benefit.[60] This might seem a somewhat inappropriate analysis, however cost aspects in healthcare are surrounded

by much debate among policy makers, medical societies, healthcare insurances, pharmaceutical companies and patients. After approval of a novel medicinal product by the European Medicines Agency (EMA), an application at the "Zorginstituut Nederland" to attain reimbursement in the basic healthcare insurance is required in the Netherlands. In addition to clinical benefit outcomes, it is required to provide a cost-effectiveness evaluation for this application.[61] For approval of reimbursement, reference thresholds can be used, but are not strictly applied.[62] In addition, after approval reimbursement can be blocked for cost-reducing negotiations with the manufacturer (or license holder).

Despite measures to reduce prices of cancer drugs, no sufficiently effective and long-term solution is currently available. It has therefore been suggested to centrally (i.e. at the level of the European Union) adopt a novel pricing system in which a maximum cancer drug price is calculated. The proposed algorithm for price calculation would include research and development costs, clinical efficacy and a profit margin in order to support innovation.[63] It is yet unclear if and when such a novel pricing system would be approved of by central and local authorities.

Finally, CRC treatments may seem costly if only considering drug or intervention costs at the time of intervention, but may be cost-effective according to a reference threshold if this leads to improved outcomes (such as symptom management), reduced costs in the later course of the disease, improvements in personalized treatment choices based on clinical predictive markers regarding clinical efficacy or improvements in societal patient participation (such as work participation). This emphasizes the importance of well-designed clinical trials or patients cohorts to assess clinical benefit and cost-effectiveness evaluations to maintain a social healthcare system.

OUTLINE OF THIS THESIS

The research described in this thesis addresses cost-effectiveness, quality of life and work ability in patients with CRC.

Over the last decades costs for CRC treatment have increased, especially since the availability of targeted therapies.[9,10] Due to the increase of health-care costs, it is increasingly important for policy makers, payers (such as health-care insurances) and medical societies to consider cost-effectiveness of treatment choices. In **Chapter 2**, we address the complexity of cost-effectiveness studies and present our view on the standardized methodology for cost-effectiveness studies. This is further illustrated in **Chapter 2b**, in which we discuss the pitfall of result interpretation if cost-effectiveness outcomes are based on inappropriate data assumptions, such as survival outcomes.

In the Dutch CAIRO3 study, the efficacy of 6 cycles first-line capecitabine (an oral fluoropyrimidine), oxaliplatin, and bevacizumab (CAPOX-B) followed by either capecitabine and bevacizumab (CAP-B) maintenance or observation for mCRC patients was evaluated. A clinical benefit in terms of progression-free survival was demonstrated for patients treated with CAP-B maintenance, while no detriment in QoL due to CAP-B maintenance was shown.[19] In **Chapter 3**, cost-effectiveness of CAP-B maintenance therapy compared to observation based on the CAIRO3 study was evaluated. For this cost-effectiveness study, a state-transition model (or Markov model) was designed. This state-transition methodology was subsequently compared with an alternative methodology, discrete event simulation, to assess effects of a modelling strategy on model outcomes. Results of this methodological evaluation are described in **Chapter 4**.

Cost-effectiveness outcomes are generally expressed in incremental cost per quality adjusted life year (QALY).[29] For QALY calculations, utilities (a value to quantify quality of life) are required using generic quality of life questionnaires, such as the EQ-5D. Unfortunately, many clinical studies only include cancer-specific questionnaires, such as the EORTC QLQ-C30. In **Chapter 5**, we propose a mapping algorithm to convert EORTC QLQ-C30 outcomes towards EQ-5D-3L utilities.

In **Chapter 6**, we address work ability and quality of life for patients diagnosed with colon cancer and enrolled in the Prospective Dutch Colorectal Cancer Cohort (PLCRC, Prospectief Landelijk CRC cohort).[64] In this longitudinal observational cohort study, clinical data and patient reported outcome measurements are registered. Patients who completed questionnaires addressing quality of life and work ability at different time points in their disease were selected for this analysis.

A thesis summary and the general discussion are presented in **Chapter 7 and 8,** respectively.

REFERENCES

1. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and moratlity worldwide: Sources, methods and major patterns in GLOBOCAN2012. *Int J Cancer*. 2014;136:E359-386. doi:10.1002/ijc.29210
2. Brenner H, Kloor M, Pox CP. Colorectal Cancer. *Lancet*. 2014;383:1490-1502. doi:10.1016/S0140-6736(13)61649-9
3. Araghi M, Soerjomataram I, Bardot A, et al. Changes in colorectal cancer incidence in seven high-income countries: a population-based study. *Lancet Gastroenterol Hepatol*. 2019;4(7):511-518. doi:10.1016/S2468-1253(19)30147-5
4. Edwards BK, Noone A-M, Mariotto AB, et al. Annual Report to the Nation on the status of cancer, 1975-2010, featuring prevalence of comorbidity and impact on survival among persons with lung, colorectal, breast, or prostate cancer. *Cancer*. 2014;120(9):1290-1314. doi:10.1002/cncr.28509
5. van Steenbergen LN, Elferink MAG, Krijnen P, et al. Improved survival of colon cancer due to improved treatment and detection: A nationwide population-based study in The Netherlands 1989-2006. *Ann Oncol*. 2010;21(11):2206-2212. doi:10.1093/annonc/mdq227
6. IKNL. cijfers over kanker. Accessed September 10, 2019. https://www.cijfersoverkanker.nl/
7. van der Geest LGM, Lam-Boer J, Koopman M, Verhoef C, Elferink MAG, de Wilt JHW. Nationwide trends in incidence, treatment and survival of colorectal cancer patients with synchronous metastases. *Clin Exp Metastasis*. 2015;32(5):457-465. doi:10.1007/s10585-015-9719-0
8. Elferink MAG, de Jong KP, Klaase JM, Siemerink EJ, de Wilt JHW. Metachronous metastases from colorectal cancer: a population-based study in North-East Netherlands. *Int J Colorectal Dis*. 2015;30(2):205-212. doi:10.1007/s00384-014-2085-6
9. 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(12):1165-1174. doi:10.1016/S1470-2045(13)70442-X
10. Sullivan R, Peppercorn J, Sikora K, et al. Delivering affordable cancer care in high-income countries. *Lancet Oncol*. 2011;12(10):933-980. doi:10.1016/S1470-2045(11)70141-3
11. Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB. Colorectal cancer. *Lancet*. 2019;394(10207):1467-1480. doi:10.1016/S0140-6736(19)32319-0
12. Labianca R, Nordlinger B, Beretta GD, Mosconi S, Cervantes A, Arnold D. Early colon cancer : ESMO Clinical Practice Guidelines for diagnosis , treatment and follow-up. *Ann Oncol*. 2013;24(suppl. 5):v70-v77. doi:10.1093/annonc/mdt354
13. ESMO. eUpdate: Early Colon Cancer Treatment Recommendations. Published 2019. Accessed October 24, 2019. eupdate: Early Colon Cancer Treatment Recommendations
14. Adjuvante systemische therapie coloncarcinoom.
15. Grothey A, Sobrero AF, Shields AF, et al. Duration of adjuvant chemotherapy for stage III colon cancer. *N Engl J Med*. 2018;378(13):1177-1188. doi:10.1056/NEJMoa1713709
16. Chakrabarti S, Peterson CY, Sriram D, Mahipal A. Early stage colon cancer: Current treatment standards, evolving paradigms, and future directions. *World J Gastrointest Oncol*. 2020;12(8):808-832. doi:10.4251/wjgo.v12.i8.808
17. Punt CJA, Koopman M, Vermeulen L. From tumour heterogeneity to advances in precision treatment of colorectal cancer. *Nat Rev Clin Oncol*. 2017;14(4):235-246. doi:10.1038/nrclinonc.2016.171
18. Van Cutsem E, Cervantes A, Nordlinger B, Arnold D. Metastatic colorectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up†. *Ann Oncol*. 2014;25 Suppl 3:iii1-9. doi:10.1093/annonc/mdu260
19. Simkens LHJ, van Tinteren H, May A, et al. Maintenance treatment with capecitabine and bevacizumab in metastatic colorectal cancer (CAIRO3): a phase 3 randomised controlled trial of the Dutch Colorectal Cancer Group. *Lancet*. 2015;385(9980):1843-1852. doi:10.1016/S0140-6736(14)62004-3

20. Hegewisch-Becker S, Graeven U, Lerchenmüller CA, et al. Maintenance strategies after first-line oxaliplatin plus fluoropyrimidine plus bevacizumab for patients with metastatic colorectal cancer (AIO 0207): a randomised, non-inferiority, open-label, phase 3 trial. *Lancet Oncol*. 2015;16(13):1355-1369. doi:10.1016/S1470-2045(15)00042-X

21. Quidde J, Hegewisch-Becker S, Graeven U, et al. Quality of life assessment in patients with metastatic colorectal cancer receiving maintenance therapy after first-line induction treatment: A preplanned analysis of the phase III AIO KRK 0207 trial. *Ann Oncol*. 2016;27(12):2203-2210. doi:10.1093/annonc/mdw425

22. Cremolini C, Schirripa M, Antoniotti C, et al. First-line chemotherapy for mCRC-areviewandevidence-basedalgorithm. *Nat Rev Clin Oncol*. 2015;12(10):607-619. doi:10.1038/nrclinonc.2015.129

23. Mol L, Koopman M, van Gils CWM, Ottevanger PB, Punt CJA. Comparison of treatment outcome in metastatic colorectal cancer patients included in a clinical trial versus daily practice in The Netherlands. *Acta Oncol (Madr)*. 2013;52(5):950-955. doi:10.3109/0284186X.2013.777158

24. Hamers PAH, Elferink MAG, Stellato RK, et al. Informing metastatic colorectal cancer patients by quantifying multiple scenarios for survival time based on real-life data. *Int J Cancer*. Published online 2020. doi:10.1002/ijc.33200

25. Schandelmaier S, Conen K, von Elm E, et al. Planning and reporting of quality-of-life outcomes in cancer trials. *Ann Oncol*. 2015;26(9):1966-1973. doi:10.1093/annonc/mdv283

26. Williams A. EuroQol - A new facility for the measurement of health-related quality of life. *Health Policy (New York)*. 1990;16:199-208. doi:10.1016/0168-8510(90)90421-9

27. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organisation for Research and Treatment of Cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst*. 1993;85:365-376.

28. Torrance GW. Measurement of health state utilities for economic appraisal. *J Health Econ*. 1986;5:1-30. doi:10.1016/0167-6296(86)90020-2

29. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the Panel on Cost-Effectiveness in Health and Medicine. *JAMA*. 1996;276(15):1253-1258. doi:10.1097/00132586-199712000-00019

30. Lamers LM, Stalmeier PFM, McDonnell J, Krabbe PFM, van Busschbach JJ. [Measuring the quality of life in economic evaluations: the Dutch EQ-5D tariff]. *Ned Tijdschr Geneeskd*. 2005;149(28):1574-1578. Accessed October 30, 2014. http://www.ncbi.nlm.nih.gov/pubmed/16038162

31. Versteegh M, M. Vermeulen K, M. A. A. Evers S, de Wit GA, Prenger R, A. Stolk E. Dutch Tariff for the Five-Level Version of EQ-5D. *Value Heal*. 2016;19(4):343-352. doi:10.1016/j.jval.2016.01.003

32. Carter HE, Zannino D, John Simes R, et al. The cost effectiveness of bevacizumab when added to capecitabine, with or without mitomycin-C, in first line treatment of metastatic colorectal cancer: results from the Australasian phase III MAX study. *Eur J Cancer*. 2014;50(3):535-543. doi:10.1016/j.ejca.2013.09.028

33. Färkkilä N, Sintonen H, Saarto T, et al. Health-related quality of life in colorectal cancer. *Colorectal Dis*. 2013;15(5):e215-22. doi:10.1111/codi.12143

34. Ramsey SD, Ph D, Andersen MR, et al. Quality of Life in Survivors of Colorectal Carcinoma. Published online 2000:1294-1303.

35. Schuurhuizen CSEW, Braamse AMJ, Konings IRHM, et al. Does severe toxicity affect global quality of life in patients with metastatic colorectal cancer during palliative systemic treatment? A systematic review. *Ann Oncol*. 2017;28(3):478-486. doi:10.1093/annonc/mdw617

36. Verhaar S, Vissers PAJ, Maas H, Van De Poll-Franse L V., Van Erning FN, Mols F. Treatment-related differences in health related quality of life and disease specific symptoms among colon cancer survivors: Results from the population-based PROFILES registry. *Eur J Cancer*. 2015;51(10):1263-1273. doi:10.1016/j.ejca.2015.04.004

37. Sanoff HK, Goldberg RM, Pignone MP. A systematic review of the use of quality of life measures in colorectal cancer research with attention to outcomes in elderly patients. *Clin Colorectal Cancer*. 2007;6(10):700-709. doi:10.3816/CCC.2007.n.039

38. Arndt V, Merx H, Stegmaier C, Ziegler H, Brenner H. Restrictions in quality of life in colorectal cancer patients over three years after diagnosis: A population based study. *Eur J Cancer*. 2006;42(12):1848-1857. doi:10.1016/j.ejca.2006.01.059

39. Caravati-Jouvenceaux A, Launoy G, Klein D, et al. Health-Related Quality of Life Among Long-Term Survivors of Colorectal Cancer: A Population-Based Study. *Oncologist*. 2011;16(11):1626-1636. doi:10.1634/theoncologist.2011-0036

40. Bouvier AM, Jooste V, Bonnetain F, et al. Adjuvant treatments do not alter the quality of life in elderly patients with colorectal cancer: A population-based study. *Cancer*. 2008;113(4):879-886. doi:10.1002/cncr.23629

41. Wright P, Downing A, Morris EJA, et al. Identifying social distress: A cross-sectional survey of social outcomes 12 to 36 months after colorectal cancer diagnosis. *J Clin Oncol*. 2015;33(30):3423-3430. doi:10.1200/JCO.2014.60.6129

42. Giesinger JM, Kieffer JM, Fayers PM, et al. Replication and validation of higher order models demonstrated that a summary score for the EORTC QLQ-C30 is robust. *J Clin Epidemiol*. 2016;69:79-88. doi:10.1016/j.jclinepi.2015.08.007

43. Van Der Noordt M, IJzelenberg H, Droomers M, Proper KI. Health effects of employment: A systematic review of prospective studies. *Occup Environ Med*. 2014;71(10):730-736. doi:10.1136/oemed-2013-101891

44. Van Rijn RM, Carlier BE, Schuring M, Burdorf A. Work as treatment? the effectiveness of re-employment programmes for unemployed persons with severe mental health problems on health and quality of life: A systematic review and meta-analysis. *Occup Environ Med*. 2016;73(4):275-279. doi:10.1136/oemed-2015-103121

45. Duijts SFA, Kieffer JM, van Muijen P, van der Beek AJ. Sustained employability and health-related quality of life in cancer survivors up to four years after diagnosis. *Acta Oncol (Madr)*. 2017;56(2):174-182. doi:10.1080/0284186X.2016.1266083

46. Mehnert A. Employment and work-related issues in cancer survivors. *Crit Rev Oncol / Hematol*. 2011;77(2):109-130. doi:10.1016/j.critrevonc.2010.01.004

47. Syse A, Tretli S, Kravdal Ø. Cancer's impact on employment and earnings – a population-based study from Norway. *J Cancer Surviv*. 2008;2:149-158. doi:10.1007/s11764-008-0053-2

48. Sharp L, O'Leary E, O'Ceilleachair A, Skally M, Hanly P. Financial Impact of Colorectal Cancer and Its Consequences: Associations between Cancer-Related Financial Stress and Strain and Health-Related Quality of Life. *Dis Colon Rectum*. 2018;61(1):27-35. doi:10.1097/DCR.0000000000000923

49. Beesley VL, Vallance JK, Mihala G, Lynch BM, Gordon LG. Association between change in employment participation and quality of life in middle – aged colorectal cancer survivors compared with general population controls. *Psychooncology*. 2017;26:1354-1360. doi:10.1002/pon.4306

50. Hauglann BK, Benth JŠ, Fosså SD, Tveit KM, Dahl A. A controlled cohort study of sickness absence and disability pension in colorectal cancer survivors. *Acta Oncol (Madr)*. 2014;53(6):735-743. doi:10.3109/0284186X.2013.844354

51. Hauglann B, Benth JŠ, Fosså SD, Tveit KM, Dahl AA. A controlled cohort study of long-term income in colorectal cancer patients. *Support Care Cancer*. 2014;22:2821-2830. doi:10.1007/s00520-014-2258-4

52. Rottenberg Y, Ratzon NZ, Cohen M, Hubert A, Uziely B, de Boer AGEM. Unemployment risk at 2 and 4 years following colorectal cancer diagnosis: a population based study. *Eur J Cancer*. 2016;69:70-76. doi:10.1016/j.ejca.2016.09.025

53. Gordon LG, Beesley VL, Lynch BM, et al. The return to work experiences of middle-aged Australian workers diagnosed with

colorectal cancer: A matched cohort study. *BMC Public Health*. 2014;14(1):1-11. doi:10.1186/1471-2458-14-963

54. Bains M, Munir F, Yarker J, et al. The impact of colorectal cancer and self-efficacy beliefs on work ability and employment status : a longitudinal study. *Eur J Cancer Care*. 2012;21:634-641. doi:10.1111/j.1365-2354.2012.01335.x

55. Boer AGEM De, Bruinvels DJ, Tytgat KMAJ, Schoorlemmer A, Klinkenbijl JHG, Frings-Dresen MHW. Employment status and work-related problems of gastrointestinal cancer patients at diagnosis : a cross-sectional study. *BMJ Open*. 2011;2:e000190. doi:10.1136/bmjopen-2011-000190

56. Brooks R, De Charro F. EuroQol: The current state of play. *Health Policy (New York)*. 1996;37(1):53-72. doi:10.1016/0168-8510(96)00822-6

57. Schnipper LE, Davidson NE, Wollins DS, et al. American Society of Clinical Oncology Statement: A Conceptual Framework to Assess the Value of Cancer Treatment Options. *J Clin Oncol*. 2015;33(23):2563-2577. doi:10.1200/JCO.2015.61.6706

58. Cherny NI, Sullivan R, Dafni U, et al. A standardised, 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(8):1547-1573. doi:10.1093/annonc/mdv249

59. NVMO. Over de commissie BOM. Accessed September 17, 2019. https://www.nvmo.org/nvmo/commissie-bom/over-de-commissie-bom/

60. Del Paggio JC, Sullivan R, Schrag D, et al. Delivery of meaningful cancer care: a retrospective cohort study assessing cost and benefit with the ASCO and ESMO frameworks. *Lancet Oncol*. 2017;18(7):887-894. doi:10.1016/S1470-2045(17)30415-1

61. ZorginstituutNederland. Richtlijn voor economische evaluatie. Accessed September 17, 2019. https://www.zorginstituutnederland.nl/over-ons/werkwijzen-en-procedures/adviseren-over-en-verduidelijken-van-het-basispakket-aan-zorg/beoordeling-van-geneesmiddelen/richtlijnen-voor-economische-evaluatie

62. Zorginstituut Nederland. Ziektelast in de praktijk. Published online 2018. https://www.zorginstituutnederland.nl/binaries/zinl/documenten/rapport/2018/05/07/ziektelast-in-de-praktijk/Ziektelast+in+de+praktijk_definitief.pdf

63. Uyl-De Groot CA, Löwenberg B. Sustainability and affordability of cancer drugs: A novel pricing model. *Nat Rev Clin Oncol*. 2018;15(7):405-406. doi:10.1038/s41571-018-0027-x

64. Burbach JPM, Kurk SA, Coebergh van den Braak RRJ, et al. Prospective Dutch colorectal cancer cohort: an infrastructure for long-term observational, prognostic, predictive and (randomized) intervention research. *Acta Oncol (Madr)*. 2016;55(11):1273-1280. doi:10.1080/0284186X.2016.1189094

2

COST-EFFECTIVENESS IN COLORECTAL CANCER: CHALLENGES ON QUALITY AND COMPARABILITY

Mira D. Franken

Elisabeth M. van Rooijen

Carin A. Uyl-de Groot

Martijn G.H. van Oijen

Miriam Koopman

Colorectal Cancer, Future Medicine, 2016, 5 (1): 21-31

ABSTRACT

Costs in colorectal cancer treatment are rising, especially since the availability of expensive targeted drugs. Comparisons between cost-effectiveness evaluations are restricted to differences in applied methodology, e.g. differences in model assumptions and design, health care systems. Cost-effectiveness analyses should be performed and reported upon in a standardized manner within a disease area and in particular in oncology to facilitate better comparisons between treatments. Ideally, to evaluate the cost-effectiveness of a treatment in daily practice, models should be based on patient cohort studies or registries with appropriate prospective data collection from a societal perspective. Randomized clinical trials remain most suitable for comparison of treatment strategies and could estimate the budget impact of a novel treatment introduction.

PRACTICE POINTS

- Cost-effectiveness models differ greatly in design thereby constraining model comparisons.
- Clinical studies and cohort studies or patient registries should prospectively collect data in a standardized manner, ideally from a societal perspective, to facilitate cost-effectiveness evaluations.
- As the debate on cost-effectiveness becomes increasingly important, physicians should be able to adequately communicate on treatment benefits and costs with patients, health economists and policy makers to achieve the best evidence-based and affordable treatment for their patients.

INTRODUCTION

Over the last decades, colorectal cancer death has declined as a result of advances in the prevention, detection and treatment of colorectal cancer.[1] However, cancer care also has a significant impact on health care costs. In 2009, European health-care costs of colorectal cancer (stage I-IV disease) were estimated at €5.57 billion of which €4.04 billion was spent on in-patient care and €565 million was spent on drugs.[2] Over the past years, health care expenditures continued to rise. For instance, between 2003 and 2011, the cost of colorectal cancer treatment in the Netherlands has doubled to a total of €427 million. Drug costs –accounting for 5% of the total Dutch colorectal cancer expenditure - almost tripled to a total of €24.1 million (Figure 1).[3]

Especially, the budget spent on novel targeted agents, which have become a backbone in the treatment of metastatic colorectal cancer (mCRC) patients[4], is subject to debate. For instance, the addition of bevacizumab to fluoropyrimidine-containing first line chemotherapy has resulted in significant improvements of progression-free and in some studies overall survival.[5–12] Recently, multiple cost-effectiveness analyses (CEAs) and cost evaluations of bevacizumab containing regimens for the first line treatment of mCRC patients have been published. Bevacizumab-containing chemotherapy was not regarded cost-effective in multiple publications, including a NICE (United Kingdom's National Institute for Health and Clinical Excellence) appraisal.[12–17] In contrast, others did regard bevacizumab as a cost-effective first line treatment of mCRC if compared to other novel anti-cancer therapies.[18–23] Lange et al. have reported a systematic review on CEAs containing antibody regimens in the treatment of mCRC and concluded that bevacizumab containing first line treatment regimens do not seem to be cost effective.[24] However, despite guidelines for cost-effectiveness research [25,26], CEAs are difficult to compare due to important differences in model design, e.g. primary outcome, patient population, country of origin, health care systems, and model assumptions. As a result comparison of different model outcomes can be hampered and conclusions on cost-effectiveness of a treatment strategy are difficult to compare, even when systematically reviewed.[24,27,28]

Even so, in an attempt to temper further increase in health care costs, cost-effectiveness evaluations are becoming more important in reimbursement decisions. In order to preserve an evidence based and sustainable health-care system, physicians should at least consider the economic impact of a treatment choice. The necessity to educate physicians in the field of treatment value and costs for this purpose, was previously alluded to by the American Society of Clinical Oncology (ASCO) in their guidance statement on the cost of cancer.[29] This paper illustrates some challenges on quality and comparability of cost-effectiveness analyses for physicians involved in cancer care, for which multiple manuscripts reporting on the costs and effectiveness of bevacizumab containing regimens in the first line treatment of mCRC were chosen as an example to illustrate these challenges.

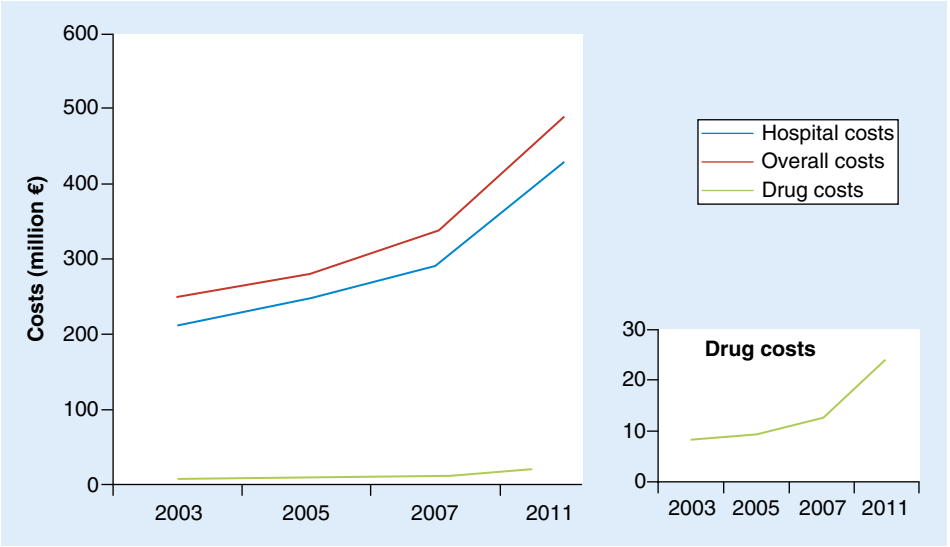


Figure 1. Colorectal cancer health-care costs in million euros in the Netherlands between 2003 and 2011. Overall costs (red) include hospital costs, but also costs such as drug costs, general practitioner costs, home care, costs of nursing houses. Hospital costs (blue) include costs of chemotherapy. Inlay shows drug costs (green) in million euros.

WHAT DO PHYSICIANS NEED TO KNOW REGARDING THE QUALITY OF A COST-EFFECTIVENESS MODEL?

First, one has to realize that similarly to randomized controlled trials (RCTs), CEAs compare two or more treatment strategies for which a cost-effectiveness model is designed. Cost-effectiveness models can be based on clinical trial datasets, published results of RCTs, data collected in prospective or retrospective patient cohorts, health care databases or could even be fully based on assumptions (Table 1, model population). Generally, CEAs are complex models that include many assumptions, based on expert opinion or literature, to compensate for data unavailable in clinical trials or patient cohort studies such as medical resource use (physician visits, number of imaging investigations, treatment initiated for adverse events, medical care outside the hospital, etc.). Moreover, cost-effectiveness models sometimes even compare treatment strategies based on indirect outcome comparison of multiple RCTs (Table 1).[20,21,30] Indirect comparison of time to progression and overall survival can introduce an additional bias in the cost-effectiveness model. Ideally, CEAs are based on RCTs, patient cohorts or registries designed to prospectively include a cost-effectiveness analysis in order to minimize the number of assumption made in the model. In the design of a clinical trial, a health-economist should be consulted beforehand when considering the inclusion of cost-effectiveness outcomes to assure the data-collection is appropriate for a CEA. Unfortunately, this is hardly ever the case.

Second, the cost perspective and included costs in the cost-effectiveness model is important. There are various options regarding the cost perspective: the societal perspective, third-party payer perspective, health care or hospital perspective and patient's perspective. The societal perspective considers all costs and health consequences related to the treatment and medical condition (medical costs, home nursing, but also work absenteeism, etc.) and can even include costs of medical conditions unrelated to the investigated treatment strategy (indirect medical costs).[25] In contrast, a third-party payer will only include health care costs remunerated by a health insurance or health care system; a health care payer or hospital perspective includes medical costs made within the establishment, such as drug costs, hospitalizations etc.; while a patient perspective includes out-of-pocket expenses of patients, such as medical bills or additional payments for uninsured care. Most cost-effectiveness models are developed with a health care payer or hospital perspective (Table 1). Even when the cost perspective is specified in the methods section of a CEA publication, it is worthwhile to consider which costs were included in the model, as different definitions and interpretations for perspectives are at hand: some models might only include drug costs, while others also include costs such as outpatient clinic treatment, adverse events, and hospitalizations, which obviously influence the total cost estimate. Even though, a health care payer or hospital perspective seems an attractive choice as it is easiest to define the costs to be included in the cost-effectiveness model and model outcomes can be used to estimate the budget impact for the health care payer, in our opinion a societal perspective should be the perspective of choice in order to evaluate treatment value in the broadest context possible. The best value for money for a hospital or health care payer, might not be the best value for money for society when including the economic impact of work absenteeism of patients and caring family, home nursing, general practitioner visits, etc. Our treatment proposition to patients should not only be the best evidence based treatment option, but we should concurrently also attempt to pursue the treatment option with the best value for society. Unfortunately, information on societal impact of treatment choices is often lacking and a health care payer or hospital perspective is most often chosen. Clearly, there is a necessity for clinical studies, patient cohort studies or registries to objectively document information necessary to enable cost-effectiveness evaluations with a societal perspective. Pursuing cost-effectiveness analyses from a societal perspective will require additional effort and funding for the supplementary data collection on work absenteeism, costs, etc. Even when a societal perspective is chosen, it is impossible to include every single cost experienced by patients (and their families), as cost-effectiveness models are a simplified reflection of a patient population.

Third, it is important to realize which patient population is included in the model. Not surprisingly, the choice of patient population influences both costs - e.g. treatment duration, adverse events - and outcome -in terms of time to progression, survival and

Table 1.

Table 1. Summary of cost–effectiveness publications on first-line bevacizumab treatment for metastatic colorectal cancer patients.

Study (year)	Country	Perspective	Treatment	Model population	Time horizon	Incremental cost	QALY gained	LY gained	Other effect outcome	Incr. cost/ QALY	Incr. Cost/LY	Other cost outcome	Ref.
Goldstein *et al.* (2015)	USA	Healthcare/ hospital	FU/LV/OX vs FU/LV/OX-B	NO1966 RCT	Life time	€46,475	0.10	0.14		€438,444	€336,776		[13]
Koeberle *et al.* (2015)	Switzerland	Healthcare/ hospital	B vs obs. following SD or better of (16–24weeks) standard induction therapy	SAKK41/06	Trial duration, median follow-up 36.7 months	€22,217[†]						B: €4,443/ month Obs: €1558/ month	[12]
Kovacevic *et al.* (2015)	Serbia	Healthcare/ hospital	CTx vs CTx antibody treatment (bevacizumab or cetuximab)	Retrospective (n = 62)	5-year	€16,975		0.53			€32,102		[17]
Carter *et al.* (2014)	Australia	Healthcare/ hospital	CAP vs CAP-B	MAX trial	Trial duration, truncated at 18 months	€15,871			QPFS: 0.154			Cost/QPFS: €103,061	[14]
Ruiz-Millo *et al.* (2014)	Spain	Healthcare/ hospital	IRI/LV/FU vs IRI/LV/FU-B	Retrospective (n = 69)	Duration of treatment (median 12 and 16 cycles)	€13,368							[15]
Ewara *et al.* (2014)	Canada	Third-party payor	FOLFIRI-B vs FOLFIRI-C and vs FOLFIRI-P	Ontario health administrative data and RCT, *KRAS* wild-type	100 months	–				FOLFIRI-B dominated			[18]
Shankaran *et al.* (2014)	USA	Healthcare/ hospital	CTx-B vs CTx and vs BSC in patients ≥65 years	SEER medicare	Life time	vs CTx: €30,296 vs BSC: €85,934		vs CTx: 0.42 vs BSC: 1.16			vs CTx: €71,284 vs BSC: €74,188		[19]

Costs are expressed in euro and were corrected for inflation and power parity [32] based on costs reported in the original publications. If data were not reported in euros, costs were converted to euros using conversion rates of 26 July 2015.

[†]Calculated based on data in publication.

[‡]FBC vs FBC bevacizumab.

[§]Incremental cost, LY gained and ICER of FOLFIRI-C over FOLFOX-B.

[¶]Results for CAPOX vs CAPOX-B presented only.

B: Bevacizumab; C: Cetuximab; CAP: Capecitabine; CAPOX: Capecitabine/oxaliplatin; CTx: Chemotherapy; FBC: FU-based chemotherapy; FOLFIRI: Leucovorin/irinotecan/fluorouracil; FOLFOX: Leucovorin/fluorouracil/oxaliplatin; FU: Fluorouracil; IFL: Irinotecan/fluorouracil/leucovorin: Incr.: Incremental; IRI: Irinotecan; LV: Leucovorin; LY: Life year; Obs.: Observation; P: Panitumumab; QALY: Quality-adjusted life year; QPFS: Quality of life-adjusted progression-free survival; RCT: Randomized controlled trial; SD: Stable disease.

Table 1. (continued)

Table 1. Summary of cost–effectiveness publications on first-line bevacizumab treatment for metastatic colorectal cancer patients (cont.).

Study (year)	Country	Perspective	Treatment	Model population	Time horizon	Incremental cost	QALY gained	LY gained	Other effect outcome	Incr. cost/ QALY	Incr. Cost/LY	Other cost outcome	Ref.
Lawrence *et al.* (2013)	Canada	Third-party payor	FU-based CTx(FBC) or FBC-B or FBC-C or FBC-P	Multiple RCTs, *KRAS* wild-type	Life time	€41,983‡	0.49			€85,030‡			[20]
Lee *et al.* (2012)	Korea	Healthcare/ hospital	FOLFIRI vs FOLFIRI-B	Indirect RCT comparison	8 years/ life time	€40,251		1.18			€34,199		[21]
Hedden *et al.* (2012)	Canada	Healthcare/ hospital	CTx prior to B registration and CTx-B	Real-world data extracted from the British Colombia Cancer Agency	12 years	€2566	0.06	0.325		€42,272	€10,568		[22]
Asseburg *et al.* (2011)	Germany	Third-party payor	FOLFIRI-C vs FOLFOX-B	Indirect RCT comparison, *KRAS* wild-type	10 years	€7,657§		0.5§		€15,190§			[30]
Shiroiwa (2010)	Japan	Patient	CAPOX vs CAPOX-B	Retrospective evaluation based on RCTs	–	–			QPFS: 0.11			Cost/QPFS ranged €609–16,224 dependent of age and income	[23]
Shiroiwa (2007)	Japan	Healthcare/ hospital	CTx vs CTx-B	Multiple RCTs	Life time	€500,662		0.86¶			€57,508¶		[31]
Tappenden *et al.* (2007)	UK	Third-party payor	IFL or 5–FU/ LV vs IFL-B or 5–FU/LV-B	Multiple RCTs	Life time						IFL ± B: €87,951 5FU/LV ± B: €123,741		[16]

Costs are expressed in euro and were corrected for inflation and power parity [32] based on costs reported in the original publications. If data were not reported in euros, costs were converted to euros using conversion rates of 26 July 2015.

†Calculated based on data in publication.

‡FBC vs FBC bevacizumab.

§Incremental cost, LY gained and ICER of FOLFIRI-C over FOLFOX-B.

¶Results for CAPOX vs CAPOX-B presented only.

B: Bevacizumab; C: Cetuximab; CAP: Capecitabine; CAPOX: Capecitabine/oxaliplatin; CTx: Chemotherapy; FBC: FU-based chemotherapy; FOLFIRI: Leucovorin/irinotecan/fluorouracil; FOLFOX: Leucovorin/fluorouracil/oxaliplatin; FU: Fluorouracil; IFL: Irinotecan/fluorouracil/leucovorin: Incr.: Incremental; IRI: Irinotecan; LV: Leucovorin; LY: Life year; Obs.: Observation; P: Panitumumab; QALY: Quality-adjusted life year; QPFS: Quality of life-adjusted progression-free survival; RCT: Randomized controlled trial; SD: Stable disease.

quality of life- and ultimately the incremental cost per effect ratio. Koeberle et al. chose to evaluate costs of bevacizumab maintenance for patients responding to standard first-line chemotherapy. As a consequence, costs related to the induction treatment for responders and non-responders were disregarded in this cost-evaluation (Table 1).[12] Different cost per effect results are to be expected if the CEA would also include the non-responders to induction chemotherapy treatment. The necessity of clearly identifying the patient population in cost-effectiveness evaluations is further illustrated with CEAs evaluating bevacizumab-containing regimens versus anti-EGFR containing regimens (cetuximab or panitumumab). Previously, KRAS mutation status was recognized as a negative predictive factor for response on anti-EGFR therapy: a progression free survival or overall survival benefit of the addition of anti-EGFR treatment to first line chemotherapy was only demonstrated in KRAS wild type patients.[31–34] Anti-EGFR treatment should therefore only be considered for patients with (K)RAS wild type tumors as activity of anti-EGFR regimens is confined to this patient sub-group.[4] For this reason, CEAs evaluating anti-EGFR regimens should clearly specify the patient population selected for the model as this influences the effect parameter of anti-EGFR containing regimens in the model. Ewara et al., Lawrence et al. and Asseburg et al. chose to only include KRAS wild type patients in their cost-effectiveness models. As a consequence, costs and effects related to mutation status testing and treatment for KRAS mutant type patients were disregarded in these analyses (Table 1).[18,20,30] With our increasing knowledge on patient's genetic characteristics, we expect an increase in patient sub-classification in the near future, enabling us to propose individualized and thus more effective treatments to patients. This will, however, result in an additional challenge in the interpretation of cost-effectiveness results both within the context of patient subgroups and the total mCRC population.

Fourth, the time horizon in a model also influences outcome. For example, a more expensive, but more efficacious first line treatment might reduce the total lifetime health care expenditures. Carter et al. evaluated the cost-effectiveness of first line capecitabine monotherapy versus capecitabine-bevacizumab therapy and truncated the time horizon at 18 months. Future treatment benefit and costs were thus not included in the primary outcome (incremental cost per quality of life adjusted progression free survival year gained (QPFS)). The incremental cost/QPFS was calculated at €103,061 for the addition of bevacizumab to capecitabine.[14] If hypothetically, the subsequent treatment of the first line capecitabine and bevacizumab treated patients is less costly compared to the subsequent treatment of patients treated with capecitabine monotherapy, the incremental treatment cost would become less. Naturally, the incremental cost/QPFS could also increase if the incremental costs were to increase over time. If an initial treatment benefit would result in reduced effect duration of subsequent treatment lines; this influences the magnitude of the overall effect parameter and ultimately the incremental cost per

effect ratio. Therefore, a lifetime horizon for cost per effect calculations is generally accepted as the most optimal time horizon for CEAs.

Finally, the most frequently used effectiveness parameter in cost-effectiveness models is quality adjusted life year gained (QALY), which is a measure reflecting the additional years lived corrected for quality of life. The primary outcome in cost-effectiveness models, as recommended by the panel on cost-effectiveness in health and medicine, is an incremental cost per QALY ratio (ICER).[26] For QALY calculations, quality of life must be measured with a generic quality of life questionnaire such as the EQ-5D questionnaire.[35] Based on this questionnaire, patient scores are transformed into health-related quality of life utility, on a scale of 1 - being best imaginable health - and 0 - reflecting worst imaginable health or death.[26,36] The calculation from patient scores to utilities is country specific. In other words, the quality of life results of a clinical study can translate to somewhat different utilities in the USA compared to the Netherlands as a different value is assigned to each of five quality of life domains: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Furthermore, many clinical trials only include a disease specific quality of life questionnaire and lack generic questionnaires, precluding the possibility of calculating study based utilities. Therefore, utilities used in models are often assumptions based on literature and might not reflect the actual quality of life difference between compared treatments. In addition, there is some debate on whether costs per QALY is the best method to evaluate cost-effectiveness.[37–39] Even though, the QALY was designed to allow cost-effectiveness outcome comparison between different treatments and diseases, there are some difficulties in the assignment of a value to a QALY that hamper its use in decision-making. Generic questionnaires such as the EQ-5D are fairly insensitive to disease specific changes in quality of life.[38,40] As can be seen in Table 1, a variety of other outcome measures are being used. A commonly calculated outcome parameter is incremental cost per life years (LYs) gained ratio, while others defined incremental cost per progression free survival year gained or mean monthly costs. Thus, life years gained might be a reasonable alternative as effect measure in oncology CEAs and could be added as a secondary outcome to QALYs to reflect the impact of quality of life on the cost-effectiveness model.

HOW TO INTERPRET COST-EFFECTIVENESS OUTCOMES: COST PER QALY AND COST PER LY?

In evaluating ICERs (using QALYs or LY gained), one should always consider the crude costs in addition to the incremental costs and effects (QALYs or LYs gained). For instance, Shankaran et al. presented CEA results of chemotherapy in combination with bevacizumab compared to best supportive care and chemotherapy in combination with bevacizumab compared to chemotherapy alone.[19] The ICERs (in costs/life year) were comparable for both analysis, €74,188 and €71,284 respectively (Table 1). However, the incremental

costs of bevacizumab containing chemotherapy compared to best supportive care were much higher than compared to chemotherapy, €85,934 and €30,296, respectively. The incremental LY gain of chemotherapy in combination with bevacizumab compared to best supportive care was also much higher in contrast to bevacizumab-containing chemotherapy compared to chemotherapy alone (1.16 and 0.42, respectively), ultimately yielding similar ICERs for both treatment options while both comparisons are quite dissimilar and given the current treatment standards a comparison to best supportive care might be regarded as obsolete (Figure 2). Another example to illustrate the importance of evaluating the incremental costs and life years gained in addition to the ICER, are the results of Asseburg et al. who compared cost-effectiveness of FOLFIRI-cetuximab compared to FOLFOX-bevacizumab, in a KRAS wild type patient population using an indirect comparison of RCTs for survival projection. Asseburg et al. found an incremental cost of €7,657 for FOLFIRI-cetuximab with an estimated incremental gain of 0.5 LYs, yielding an ICER of €15,190.[30] Recently, the randomized, clinical phase III FIRE study in which FOLFIRI-cetuximab was compared to FOLFIRI-bevacizumab as first line treatment in mCRC resulted in a median OS of 28.7 months and 25.0 months (HR 0.77, 95% CI 0.62-0.96) respectively, equivalent to an incremental gain of 0.31 LYs, approximately.[41] The assumption of an incremental gain of 0.5 LYs made by Asseburg et al. therefore seems to be overestimated and the ICER would increase if recalculated based on the survival

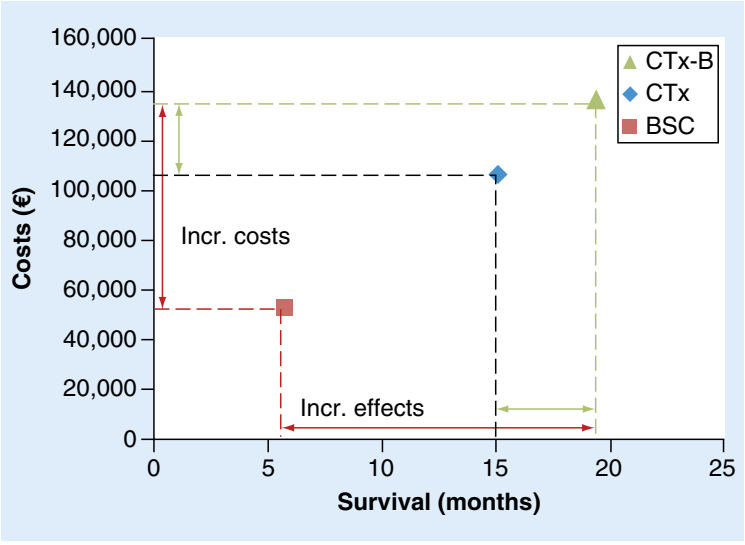


Figure 2. Mean treatment cost in euros (corrected for inflation and power parity) per median survival (months). The red arrow indicates the incremental effects of CTx-B versus best supportive care, while the green arrow indicates the incremental effects of CTx-B versus chemotherapy. Costs and effects were retrieved from Shankaran et al.[19]
BSC best supportive care, CTx chemotherapy, CTx-B chemotherapy bevacizumab, Incr. Incremental.

data of the FIRE-3 study. Even so, results would yield a viable ICER, while both treatments are regarded to be expensive and have an important impact on the health-care payer or hospital budget.

If compelled to compare different CEAs, one should realize that there are a number of issues regarding transferability of outcomes between countries for which a correction is needed: differences in health care resources, treatment standards, correction for inflation, etc.[42] Additionally, the so called purchasing power parity or adjustment on relative value of different currencies, i.e. the value of the same chemotherapy vial in different countries, should be accounted for, as differences in prosperity and costs between countries and continents exist. Even with important efforts to optimize comparisons between CEAs, costs may change significantly over time, as a drug might come of patent resulting in an important price reduction. After adjustment for inflation and purchasing power parity (Table 1), the most we can conclude is that the addition of bevacizumab to the first-line treatment of mCRC consistently resulted in an increase in cost for health-care payers and hospitals with an increase of effect, either QALY or LY.

CONCLUSION AND FUTURE PERSPECTIVE

We have shown that cost-effectiveness models are complex and might not generate reproducible results in the evaluation of treatments due to differences in modeling assumptions between research groups. We have addressed important areas wherein these differences due to applied methodology may occur, such as model assumptions, included costs, study population, time horizon of the model, and the primary outcome. To reduce the number of assumptions in a cost-effectiveness model, prospective data collection on medical resource use and societal impact (e.g. home care, loss of work) should be pursued to enable more accurate estimates on cost per QALY or life year gained. For cost evaluations of daily practice, cohort studies or registries are appropriate. However, this incurs the uncertainty whether the most optimal treatment strategies were proposed to patients during the course of their disease. Cost-effectiveness evaluations of different treatment strategies are most appropriately evaluated with RCT data.

Physicians must take up their role in an overarching discussion on cost of cancer care – not limited to cost of cancer drugs - together with patient advocates, industry, health care payers, health-economists and policy makers to ensure the sustainability of the current health care system. While, colorectal cancer treatment guidelines enable us to propose the most optimal treatment strategy for a patient, they do not provide us with information on the magnitude of clinical benefit nor cost-effectiveness. Recently, both the European Society for Medical Oncology (ESMO) and the American Society of Clinical Oncology (ASCO) proposed standards to assess the magnitude of clinical benefit in comparative studies to rank different strategies by clinical benefit.[43,44] The framework proposed by ASCO not only includes a clinical benefit ranking, but also presents drug

costs per month.[44] Standardized evaluation of treatment benefits with the inclusion of cost evaluation and preferably cost-effectiveness evaluations, could and should empower physicians to participate in the discussion on cost of cancer care. Moreover, standardized treatment evaluations could help physicians to engage a discussion with their patients on the most optimal treatment strategy.

In our opinion, comparisons of cost-effectiveness outcomes might improve with country specific standardization of cost-effectiveness models for specific disease areas, such as mCRC. Standardized cost-effectiveness models should be designed based on patient cohort studies or registries, in which data on medical resource use from a societal perspective and quality of life are prospectively collected to evaluate treatment strategies in daily practice. Additionally, the design of RCTs should include generic quality of life questionnaires, such as EQ-5D, and prospectively include cost-effectiveness as outcome parameter. When that will be the case, cost-effectiveness outcomes will become more reliable and comparable, and better estimates of budget impact of a novel treatment introduction will follow.

REFERENCES

1. Edwards BK, Noone A-M, Mariotto AB, et al. Annual Report to the Nation on the status of cancer, 1975-2010, featuring prevalence of comorbidity and impact on survival among persons with lung, colorectal, breast, or prostate cancer. *Cancer*. 2014;120(9):1290-1314. doi:10.1002/cncr.28509
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(12):1165-1174. doi:10.1016/S1470-2045(13)70442-X
3. Rijksinstituut voor Volksgezondheid en Milieu - ministerie van Volksgezondheid Welzijn en Sport. Cost of illness in the Netherlands. http://www.kostenvanziekten.nl/cijfers/. Published 2015. Accessed July 7, 2015.
4. Van Cutsem E, Cervantes A, Nordlinger B, Arnold D. Metastatic colorectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up†. *Ann Oncol*. 2014;25 Suppl 3:iii1-9. doi:10.1093/annonc/mdu260
5. Wagner A, Arnold D, Grothey A, Haerting J, Unverzagt S. Anti-angiogenic therapies for metastatic colorectal cancer. *Cochrane Collab*. 2009;(3). http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD005392.pub3/pdf/standard. Accessed January 30, 2015.
6. Hurwitz H, Fehrenbacher L, Novotny W, et al. Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. *NEJM*. 2004;350(23):2335-2342. doi:10.1056/NEJMoa032691
7. Saltz LB, Clarke S, Díaz-Rubio E, Scheithauer W, Figer A, Wong R, Koski S, Lichinitser M, Yang TS, Rivera F, Couture F, Sirzén F, Cassidy J. Bevacizumab in combination with oxaliplatin-based chemotherapy as first-line therapy in metastatic colorectal cancer: a randomized phase III study. *J Clin Oncol*. 2008;26(12):2013-2019. doi:10.1200/JCO.2007.14.9930
8. Cunningham D, Lang I, Marcuello E, et al. Bevacizumab plus capecitabine versus capecitabine alone in elderly patients with previously untreated metastatic colorectal cancer (AVEX): an open-label, randomised phase 3 trial. *Lancet Oncol*. 2013;14(11):1077-1085. doi:10.1016/S1470-2045(13)70154-2
9. Tebbutt NC, Wilson K, Gebski VJ, et al. Capecitabine, bevacizumab, and mitomycin in first-line treatment of metastatic colorectal cancer: Results of the Australasian Gastrointestinal Trials Group randomized phase III MAX study3643. *J Clin Oncol*. 2010;28(19):3191-3198. doi:10.1200/JCO.2009.27.7723
10. Cassidy J, Saltz LB, Giantonio BJ, Kabbinavar FF, Hurwitz HI, Rohr UP. Effect of bevacizumab in older patients with metastatic colorectal cancer: Pooled analysis of four randomized studies. *J Cancer Res Clin Oncol*. 2010;136(5):737-743.
11. Simkens LHJ, van Tinteren H, May A, et al. Maintenance treatment with capecitabine and bevacizumab in metastatic colorectal cancer (CAIRO3): a phase 3 randomised controlled trial of the Dutch Colorectal Cancer Group. *Lancet*. 2015;385(9980):1843-1852. doi:10.1016/S0140-6736(14)62004-3
12. Koeberle D, Betticher DC, von Moos R, et al. Bevacizumab continuation versus no continuation after first-line chemotherapy plus bevacizumab in patients with metastatic colorectal cancer: a randomized phase III non-inferiority trial (SAKK 41/06). *Ann Oncol*. 2015;(January):709-714. doi:10.1093/annonc/mdv011
13. Goldstein D a, Chen Q, Ayer T, et al. First- and second-line bevacizumab in addition to chemotherapy for metastatic colorectal cancer: a United States-based cost-effectiveness analysis. *J Clin Oncol*. 2015;33(10):1112-1118. doi:10.1200/JCO.2014.58.4904
14. Carter HE, Zannino D, John Simes R, et al. The cost effectiveness of bevacizumab when added to capecitabine, with or without mitomycin-C, in first line treatment of

metastatic colorectal cancer: results from the Australasian phase III MAX study. *Eur J Cancer*. 2014;50(3):535-543. doi:10.1016/j.ejca.2013.09.028

15. Ruiz-Millo O, Albert-Mari A, Sendra-Garcia A, Jimenez-Torres NV. Comparative cost-effectiveness of bevacizumab-irinotecan-fluorouracil versus irinotecan-fluorouracil in first-line metastatic colorectal cancer. *J Oncol Pharm Pract*. 2014;20(5):341-350. doi:10.1177/1078155213508437
16. Tappenden P, Jones R, Paisley S, Carroll C. The cost-effectiveness of bevacizumab in the first-line treatment of metastatic colorectal cancer in England and Wales. *Eur J Cancer*. 2007;43(17):2487-2494. doi:10.1016/j.ejca.2007.08.017
17. Kovacevic A, Dragojevic-Simic V, Tarabar D, et al. Five-year survival and costs of care in metastatic colorectal cancer: conventional versus monoclonal antibody-based treatment protocols. *Expert Rev Anticancer Ther*. 2015;15(8):1-8. doi:10.1586/14737140.2015.1059280
18. Ewara EM, Zaric GS, Welch S, Sarma S. Cost-effectiveness of first-line treatments for patients with KRAS wild-type metastatic colorectal cancer. *Curr Oncol*. 2014;21:541-550.
19. Shankaran V, Mummy D, Koepl L, et al. Survival and lifetime costs associated with first-line bevacizumab use in older patients with metastatic colorectal cancer. *Oncologist*. 2014;19(8):892-899. doi:10.1634/theoncologist.2013-0209
20. Lawrence D, Maschio M, Leahy KJ, Yunger S, Easaw JC, Weinstein MC. Economic analysis of bevacizumab, cetuximab, and panitumumab with fluoropyrimidine-based chemotherapy in the first-line treatment of KRAS wild-type metastatic colorectal cancer (mCRC). *J Med Econ*. 2013;16(12):1387-1398. doi:10.3111/13696998.2013.852097
21. Lee E-K, Revil C, Ngoh C a, et al. Clinical and cost effectiveness of bevacizumab + FOLFIRI combination versus FOLFIRI alone as first-line treatment of metastatic colorectal cancer in South Korea. *Clin Ther*. 2012;34(6):1408-1419. doi:10.1016/j.clinthera.2012.05.001
22. Hedden L, Kennecke H, Villa D, et al. Incremental cost-effectiveness of the pre- and post-bevacizumab eras of metastatic colorectal cancer therapy in British Columbia, Canada. *Eur J Cancer*. 2012;48(13):1969-1976. doi:10.1016/j.ejca.2012.01.012
23. Shiroiwa T, Fukuda T, Tsutani K. Out-of-pocket payment and cost-effectiveness of XELOX and XELOX plus bevacizumab therapy: from the perspective of metastatic colorectal cancer patients in Japan. *Int J Clin Oncol*. 2010;15(3):256-262. doi:10.1007/s10147-010-0045-x
24. Lange A, Prenzler A, Frank M, Kirstein M, Vogel A, von der Schulenburg JM. A systematic review of cost-effectiveness of monoclonal antibodies for metastatic colorectal cancer. *Eur J Cancer*. 2014;50(1):40-49. doi:10.1016/j.ejca.2013.08.008
25. Siegel JE, Torrance GW, Russell LB, Luce BR, Weinstein MC, Gold MR. Guidelines for pharmacoeconomic studies. Recommendations from the panel on cost effectiveness in health and medicine. Panel on cost Effectiveness in Health and Medicine. *Pharmacoeconomics*. 1997;11(2):159-168.
26. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the Panel on Cost-Effectiveness in Health and Medicine. *JAMA*. 1996;276(15):1253-1258. doi:10.1097/00132586-199712000-00019
27. Leung HWC, Chan ALF, Leung MSH, Lu C-L. Systematic review and quality assessment of cost-effectiveness analysis of pharmaceutical therapies for advanced colorectal cancer. *Ann Pharmacother*. 2013;47(4):506-518. doi:10.1345/aph.1R152
28. Shankaran V. Cost Considerations in the Treatment of Colorectal Cancer. *Curr Treat Options in Oncol*. 2015;16(8):354-365. doi:10.1007/s11864-015-0354-4
29. Meropol NJ, Schrag D, Smith TJ, et al. American Society of Clinical Oncology guidance statement: The cost of cancer care. *J Clin Oncol*. 2009;27(23):3868-3874. doi:10.1200/JCO.2009.23.1183

30. Asseburg C, Frank M, Köhne C-H, et al. Cost-effectiveness of targeted therapy with cetuximab in patients with K-ras wild-type colorectal cancer presenting with initially unresectable metastases limited to the liver in a German setting. *Clin Ther*. 2011;33(4):482-497. doi:10.1016/j.clinthera.2011.04.010

31. Douillard J-Y, Siena S, Cassidy J, et al. Randomized, phase III trial of panitumumab with infusional fluorouracil, leucovorin, and oxaliplatin (FOLFOX4) versus FOLFOX4 alone as first-line treatment in patients with previously untreated metastatic colorectal cancer: the PRIME study. *J Clin Oncol*. 2010;28(31):4697-4705. doi:10.1200/JCO.2009.27.4860

32. Van Cutsem E, Köhne C-H, Láng I, et al. Cetuximab plus irinotecan, fluorouracil, and leucovorin as first-line treatment for metastatic colorectal cancer: updated analysis of overall survival according to tumor KRAS and BRAF mutation status. *J Clin Oncol*. 2011;29(15):2011-2019. doi:10.1200/JCO.2010.33.5091

33. Bokemeyer C, Bondarenko I, Hartmann JT, et al. Efficacy according to biomarker status of cetuximab plus FOLFOX-4 as first-line treatment for metastatic colorectal cancer: the OPUS study. *Ann Oncol*. 2011;22(7):1535-1546. doi:10.1093/annonc/mdq632

34. Dahabreh IJ, Terasawa T, Castaldi PJ, Trikalinos TA. Systematic review: Anti-epidermal growth factor receptor treatment effect modification by KRAS mutations in advanced colorectal cancer. *Ann Intern Med*. 2011;154(1):37-49. doi:10.7326/0003-4819-154-1-201101040-00006

35. Williams A. EuroQol - A new facility for the measurement of health-related quality of life. *Health Policy (New York)*. 1990;16:199-208. doi:10.1016/0168-8510(90)90421-9

36. Torrance GW. Measurement of health state utilities for economic appraisal. *J Health Econ*. 1986;5:1-30. doi:10.1016/0167-6296(86)90020-2

37. Uyl-de Groot C a., de Vries EGE, Verweij J, Sullivan R. Dispelling the myths around cancer care delivery: It's not all about costs. *J Cancer Policy*. 2014;2(1):22-29. doi:10.1016/j.jcpo.2014.01.001

38. Miller JD, Foley KA, Russell MW. Current Challenges in Health Economic Modeling of Cancer Therapies : *Am Heal Drug Benefits*. 2014;7(3):153-162.

39. McGregor M, Caro JJ. QALYs: Are they helpful to decision makers? *Pharmacoeconomics*. 2006;24(10):947-952. doi:10.2165/00019053-200624100-00002

40. Garau M, Shah KK, Mason AR, Wang Q, Towse A, Drummond MF. Using QALYs in cancer: a review of the methodological limitations. *Pharmacoeconomics*. 2011;29(8):673-685. doi:10.2165/11588250-000000000-00000

41. Heinemann V, von Weikersthal LF, Decker T, et al. FOLFIRI plus cetuximab versus FOLFIRI plus bevacizumab as first-line treatment for patients with metastatic colorectal cancer (FIRE-3): a randomised, open-label, phase 3 trial. *Lancet Oncol*. 2014;15(10):1065-1075. doi:10.1016/S1470-2045(14)70330-4

42. Drummond, M., Mcguire A. *Economic Evaluation in Health Care: Merging Theory with Practice*. Oxford University press; 2001.

43. Cherny NI, Sullivan R, Dafni U, et al. A standardised, 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(8):1547-1573. doi:10.1093/annonc/mdv249

44. Schnipper LE, Davidson NE, Wollins DS, et al. American Society of Clinical Oncology Statement: A Conceptual Framework to Assess the Value of Cancer Treatment Options. *J Clin Oncol*. 2015;33(23):2563-2577. doi:10.1200/JCO.2015.61.6706

2b

LETTER TO THE EDITOR:

CETUXIMAB AS FIRST-LINE TREATMENT FOR METASTATIC COLORECTAL CANCER: *CAUTION WITH INTERPRETATION OF COST-EFFECTIVENESS RESULTS TOWARD MEDICAL DECISION MAKING*

Mira D. Franken

Miriam Koopman

Martijn G.H. van Oijen

American Journal of Clinical Oncology, 2016: 39 (2): 214

LETTER TO THE EDITOR

Shankaran et al. are to be complimented for their excellent and important work to address the cost-effectiveness of FOLFIRI plus cetuximab as first-line treatment of metastatic colorectal cancer (mCRC). The reported cost-effectiveness model was based on the randomized clinical phase-3 FIRE-3 study.[1, 2] An incremental cost per effect ratio (ICER) of $86,487 per life year gained (LY) for FOLFIRI plus cetuximab compared with FOLFIRI plus bevacizumab was presented. Interestingly, Shankaran et al. performed an additional analysis using CALGB/SWOG 80405 results (Table 4); the results of this additional analysis seem conflicting with the data presented by Shragh et al. in their ASCO 2015 presentation.[3]

The benefit of chemotherapy plus cetuximab compared to chemotherapy plus bevacizumab remains a matter of debate due to conflicting overall survival results of the FIRE-3 and CALGB/SWOG80405 study.[2, 4] Influence of the subsequent treatment regimens has been suggested as a possible explanation for the overall survival difference demonstrated in the FIRE-3 study.[5] Shankaran et al. have calculated cost-effectiveness in a scenario using the CALGB/SWOG 80405 outcomes, yielding an ICER of $121,501/LY. It is important to realize that for this incremental cost per effect ratio calculation, the FOLFIRI plus cetuximab strategy resulted in an increment cost of $37,191. This incremental cost is subsequently divided by the increment in effect (life years or quality adjusted life years gained) to calculate an ICER. According to Venook et al. both strategies yield similar survival results; we can only conclude that the cetuximab containing strategy results in more costs, while this strategy does not provide a clinically meaningful benefit (median survival difference of 0.9 months, p=0.34).[4] Calculating an ICER over treatment strategies that do not yield a difference in clinical outcomes, does neither seem meaningful nor appropriate.

The FIRE-3 study design did not include a cost-effectiveness evaluation, thus information on resource utilization and possible differences between treatment arms was not readily available. As a consequence, Shankaran et al. needed to make assumptions on medical resource use based on reported adverse events only. However, differences in medical resource utilization may exist between treatment arms, such as differences in diagnostic work-up, number and type of imaging used, paracentesis, etc. These differences are yet unaccounted for.

Finally, the difference in life years gained between FOLFIRI plus cetuximab or FOLFIRI plus bevacizumab used for the model calculations, seem large compared to the medians reported in both the FIRE-3 and the CALGB/SWOG 80405 studies. We recognize that cost-effectiveness outcomes are calculated based on the incremental difference between mean costs and *mean* effects. However, the difference in means between treatment arms should be within range of the difference between medians. In example, the difference between means of the FIRE-3 scenario (or base case) was 0.63 LY, while we calculate a difference in medians of 0.31 LY between study arms. A difference in effect of 0.31 LY would yield

an ICER of $175,912/LY instead of the reported $86,487. For the scenario calculation based on the CALGB/SWOG 80405 study, a mean difference of 0.31 LY was reported, while we calculated a median difference of 0.075 LY yielding an ICER of $495,880/LY. In our opinion, this difference is mainly caused by the fact that the (Kaplan-Meier) survival curves are skewed and this cannot be disregarded in the considerations and interpretation of cost-effectiveness outcomes comparing both treatment strategies.

Even though, this cost-effectiveness study was very well performed, reported results again illustrate that cost-effectiveness models are complex. We should be cautious in interpreting cost-effectiveness results towards clinical decision-making.

REFERENCES

1. Shankaran V, Ortendahl JD, Purdum AG, et al. Cost-Effectiveness of Cetuximab as First-line Treatment for Metastatic Colorectal Cancer in the United States. *Am J Clin Oncol* 2015; published online ahead of print.
2. Heinemann V, von Weikersthal LF, Decker T, et al. FOLFIRI plus cetuximab versus FOLFIRI plus bevacizumab as first-line treatment for patients with metastatic colorectal cancer (FIRE-3): a randomised, open-label, phase 3 trial. *Lancet Oncol* 2014;15: 1065–75.
3. Schrag D, Dueck AC, Naughton MJ, et al. Cost of chemotherapy for metastatic colorectal cancer with either bevacizumab or cetuximab: Economic analysis of CALGB/SWOG 80405. | 2015 ASCO Annual Meeting | Abstracts | Meeting Library. *J Clin Oncol* 2015;33: abstract 6504.
4. Venook AP, Niedzwiecki D, Lenz H-J, et al. CALGB/SWOG 80405: Phase III trial of irinotecan/5-FU/leucovorin (FOLFIRI) or oxaliplatin/5-FU/leucovorin (mFOLFOX6) with bevacizumab (BV) or cetuximab (CET) for patients (pts) with KRAS wild-type (wt) untreated metastatic adenocarcinoma of the colon or re. *J Clin Oncol* 2014; 32:suppl; abstr LBA3.
5. Modest DP, Stintzing S, von Weikersthal LF, et al. Impact of Subsequent Therapies on Outcome of the FIRE-3/AIO KRK0306 Trial: First-Line Therapy With FOLFIRI Plus Cetuximab or Bevacizumab in Patients With KRAS Wild-Type Tumors in Metastatic Colorectal Cancer. *J Clin Oncol* 2015.

3

COST-EFFECTIVENESS OF CAPECITABINE AND BEVACIZUMAB MAINTENANCE TREATMENT AFTER FIRST-LINE INDUCTION TREATMENT IN METASTATIC COLORECTAL CANCER

M.D. Franken

E.M. van Rooijen

A.M. May

H. Koffijberg

H. van Tinteren

L. Mol

A.J. ten Tije

G.J. Creemers

A.M.T. van der Velden

B.C. Tanis

C.A. Uyl-de Groot

C.J.A. Punt

M. Koopman*

M.G.H. van Oijen*

* These authors contributed equally to this work.

European Journal of Cancer, 2017, 75: 204-212

ABSTRACT

Aim

Capecitabine and bevacizumab (CAP-B) maintenance therapy has shown to be more effective compared to observation in metastatic colorectal cancer (mCRC) patients achieving stable disease or better after six cycles of first-line capecitabine, oxaliplatin, bevacizumab (CAPOX-B) treatment in terms of progression-free survival. We evaluated the cost-effectiveness of CAP-B maintenance treatment.

Methods

Decision analysis with Markov modelling to evaluate the cost-effectiveness of CAP-B maintenance compared to observation was performed based on CAIRO3 study results (n=558). An additional analysis was performed in patients with complete or partial response. The primary outcomes were the incremental cost-effectiveness ratio (ICER) defined as the additional cost per life year (LY) and quality-adjusted life years (QALY) gained, calculated from EQ-5D questionnaires and literature, and life years (LY) gained. Univariable sensitivity analysis was performed to assess the influence of input parameters on the ICER, and a probabilistic sensitivity analysis represents uncertainty in model parameters.

Results

CAP-B maintenance compared to observation resulted in 0.21 QALYs (0.18LYs) gained at a mean cost increase of €36,845, yielding an ICER of €175,452 per QALY (€204,694 per LY). Varying the difference in health-related quality of life (HRQoL) between CAP-B maintenance and observation influenced the ICER most. For patients achieving complete or partial response upon CAPOX-B induction treatment, an ICER of €149,300 per QALY was calculated.

Conclusion

CAP-B maintenance results in improved health outcomes measured in QALYs and LYs compared to observation, but also in a relevant increase in costs. Despite the fact that there is no consensus on cost-effectiveness thresholds in cancer treatment, CAP-B maintenance may not be considered cost-effective.

INTRODUCTION

Recently, the results of the phase 3 CAIRO3 study showed that metastatic colorectal cancer (mCRC) patients with stable disease or better after 6 cycles of treatment with capecitabine, oxaliplatin and bevacizumab (CAPOX-B) had a significant benefit from capecitabine and bevacizumab (CAP-B) maintenance treatment compared to observation.[1] In this trial, reintroduction of CAPOX-B treatment was planned in all patients who had progressive disease following either CAP-B maintenance or observation. A statistically significant improvement in the primary endpoint of second progression-free survival (PFS-2), defined as the time from randomization until progression of disease after CAPOX-B reintroduction, was shown for maintenance treatment versus observation, 11.7 months and 8.5 months, respectively (HR 0.67, 95% CI 0.56-0.81). Although the study was not designed to detect a difference in overall survival (OS), an absolute median OS benefit of 3.5 months was observed which was not statistically significant (HR 0.89, 95% CI 0.73-1.07). Median OS from the time of randomization was 21.6 months for patients receiving maintenance treatment and 18.1 months for observation.[1] A statistically significant OS benefit in favour of CAP-B maintenance treatment was demonstrated in patients achieving complete response (CR) or partial response (PR) during induction treatment (24.1 months and 18.8 months, respectively (log-rank p=0.0002)).[1] However, results for this subgroup analysis require further validation. Maintenance treatment did not impair quality of life (mean change in global quality of life 0.03, 95% CI -0.35-0.41).[1] Our findings are supported by the results of the AIO 0207 study, which had a comparable study design.[2]

Despite these results, economic concerns may hamper the implementation of CAP-B maintenance therapy in daily practice. Multiple cost-effectiveness analyses (CEAs) of bevacizumab containing first-line regimens for mCRC treatment have been published with different results: some analyses did[3–6], but others did not show that the addition of bevacizumab to chemotherapy was cost-effective.[7–12] This diversity in results arises due to differences in methodology applied for these cost-effectiveness studies, such as therapy of comparison and country of origin.[13] Additionally, as recently described, a cost-effectiveness study can be fully designed and calculated based upon assumptions, such as duration of bevacizumab treatment continuation, which might importantly influence cost and effect outcomes.[3,14]

Cost-effectiveness of CAP-B maintenance treatment has not been previously evaluated. Therefore, we evaluated the cost-effectiveness of CAP-B maintenance compared to the observational strategy following first-line CAPOX-B induction treatment for mCRC patients based on the CAIRO3 study.

METHODS

Patient population

Results of the CAIRO3 study (NCT00442637)[1] were used for this post-hoc cost-effectiveness model. The CAIRO3 study was a Dutch multicentre randomized clinical study in which mCRC patients (n=558) with stable disease (SD) or better after six cycles of CAPOX-B induction therapy were either assigned to observation or capecitabine (625mg/m^2 orally twice daily continuously) and bevacizumab (7.5mg/kg intravenously every 3 weeks) maintenance treatment. Upon progression, patients were scheduled to receive reintroduction of CAPOX-B therapy. Progression free survival 1 (PFS1), time to second progression (TT2PD) and overall survival (OS) were used for this cost-effectiveness model. PFS1 was defined as the interval between randomization and the date of first progression while under maintenance or observation, or until death, or discontinuation, or end of trial for patients who did not progress. The time to second progression on any treatment including reintroduction of CAPOX-B (TT2PD) was defined as the interval between randomization and the date of second progression (for those who had a first progression). In patients for whom no further treatment was registered beyond PFS1, time to second progression on any treatment was similar to overall survival (i.e., death as endpoint or censored if still alive). We defined overall survival (OS) as the interval between randomization and the date of death (or censored if still alive). Mean values from the CAIRO3 dataset were used for base case deterministic analysis.

Model structure

A Markov decision analytic model was designed (using TreeAge Pro Healthcare v.2014, *TreeAge Software, Williamstown, MA, U.S.A.*) incorporating both the observation and CAP-B maintenance treatment strategies and subsequent treatments for the management of mCRC. The Markov model (Figure 1) consisted of four health states throughout the treatment of mCRC (post-induction, reintroduction or other treatment, salvage and death) and included probabilities of adverse events with related costs. Health-related quality of life (HRQoL) and costs were attributed to each health state. A lifetime horizon, based on a median duration of follow-up of 48 months in the CAIRO3 study, with Markov transition cycles of three weeks was chosen, to reflect the original treatment cycle length in the CAIRO3 study.

Health state transitions and clinical probability estimates

Time–dependent transition probabilities (time to first and second progression, and survival) were derived directly from the CAIRO3 dataset on a patient-level in three week intervals using Life Tables in IBM SPSS Statistics software, version 23 (Armonk, NY, IBM Corp.). Although, the CAIRO3 study was not designed to analyse specific subgroups, such as patients achieving CR or PR upon CAPOX-B induction, we performed a separate

analysis on this patient subgroup as CAP-B maintenance resulted in significantly improved PFS2 and OS. As a result, transition probabilities for this subgroup analysis were based on a smaller patient population (n=366). Additionally, time-dependent transition probability tables were also derived for the subgroup of patients with CR/PR after CAPOX-B induction treatment, since this subgroup appeared to benefit most from maintenance treatment. Clinical probabilities for grade 3 and 4 adverse events were derived from the CAIRO3 dataset and a literature search in the PubMed database (Supplementary Table S1).[15–18] Assumptions on incidence and treatment of grade 3 and 4 adverse events were validated following expert opinion within the study team.

Health-related quality of life scores and outcomes

HRQoL scores were calculated using the EQ-5D (3L) questionnaires for 492 patients (n=250 and 242 for the observation and maintenance arm, respectively) for the post-induction health state. EQ-5D questionnaires were completed at baseline and every 9 weeks thereafter. For 96 patients, no EQ-5D questionnaires were available. HRQoL scores for the reintroduction or other treatment and salvage health states were derived from literature (Table 1).[4,5,7,9,19–26] HRQoL scores reflect quality of life in a range between 0 (worst imaginable health/death) to 1 (optimal health) and were calculated with the Dutch EQ-5D tariff.[27] Quality-adjusted life years (QALYs) were subsequently calculated as the product of utility and time.[28] QALY is a measure reflecting years lived adjusted for quality of life and is widely used to compare health outcomes between treatment strategies.[29]

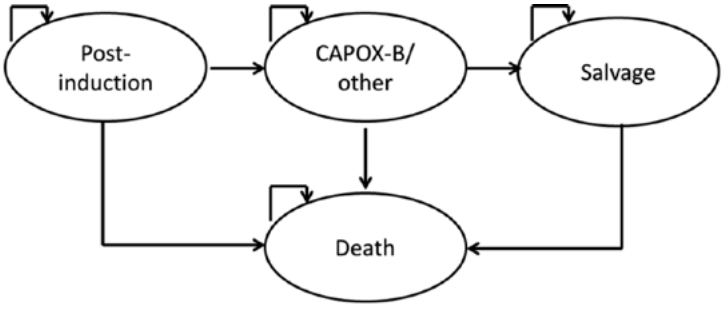


Figure 1. Markov model with 3 week cycle length. All patients start in the post-induction phase. After 3 weeks a patient can remain in this health state or transition to another health state: upon progression CAPOX-B reintroduction/ other treatment or death. Patients transitioned to CAPOX-B reintroduction/ other treatment could remain in this health state or transition to salvage treatment upon progression or death every Markov cycle. Patients in the salvage health state remained in this health state until death.

Resource use and cost-estimates

Resource use was determined based on the CAIRO3 clinical dataset and retrospective additional data collection for 146 patients (n=75 in the observation arm and n=71 in the maintenance arm) in six hospitals participating in the CAIRO3 study (2 academic centres, 2 teaching hospitals and 2 general hospitals), listed in the supplementary Table S1. Data collected in selected hospitals were regarded as a reflection of average daily practice. Only medical resource use data within the hospital participating to the CAIRO3 study were collected from the time of randomization until death or last date alive. Medical resource utilization of patients outside the participating centre was outside the scope of the retrospective data collection.

A hospital payer's perspective was chosen in this study, which include drug costs for the treatment of mCRC and associated adverse events, hospital and physician fees, and all work-up performed within the hospital. Cost calculations for medication were based on an average patient of 79 kg with a body surface area of 1.92 m^2 (derived from the CAIRO3 dataset). Medication costs included to this study are limited to inpatient medication for adverse events and medication prescribed by the medical oncologist. Treatment costs were derived from the Dutch Healthcare Institute, and medication costs were derived from the Healthcare Insurance Board.[30–33] The Dutch Healthcare institute provides standardized information on costs for i.e. physician fees, medical work-up, etc., averaged for hospital types (academic, teaching and general hospitals). All costs are presented in euros and indexed to 2014 using consumer price indices, if necessary, in the supplementary Table S2. Costs and health outcome effects were discounted at an annual rate of 4% and 1.5%, respectively, as per Dutch guidelines (base case).[30] In addition, both cost and health outcome effects were also discounted at an annual rate of 3% [34], and no discounting for costs and effects was also applied for comparison.

Outcomes and sensitivity analysis

Our primary outcomes were the incremental cost-effectiveness ratios (ICERs): the ratio of incremental costs per QALY and per life year (LY) gained. Univariable sensitivity analyses was used to demonstrate the influence of all model parameters, including a ratio reflecting the difference in the HRQoL scores between CAP-B maintenance and observation on the ICER (ranges tested as shown in Table 1). Additionally, the mean cost-effectiveness and confidence intervals were computed in a probabilistic sensitivity analysis using Monte Carlo simulation (n=10,000) on the costs (gamma distributions), rates (normal distributions), probabilities (beta distributions) and utilities (beta distributions) with a 25% standard deviation used per parameter (supplementary Table S2).

Cost-effectiveness outcomes were also calculated based on clinical outcomes of patients with PR or CR upon induction treatment with CAPOX-B, as this subgroup appeared to derive the most benefit of maintenance treatment in CAIRO3 study.[1]

Table 1. Model parameters –probabilities and HRQoL scores

	Base case	Range tested in sensitivity analysis	Source
Probabilities of 1st progression and 2nd progression	Life-time tables	-	CAIRO3 dataset [1]
Probabilities of death	Life-time tables	-	CAIRO3 dataset [1]
HRQoL score maintenance health state	0·84	0·74-0·95	Base case: CAIRO3 dataset and range: [4,5,7,9,19–22]
HRQoL score observation health state	0·83	0·74-0·95	Base case: CAIRO3 dataset and range: [4,5,7,9,19–22]
HRQoL score for the reintroduction or 2nd line health state	0·73	0·58-0·78	Base case and range: [5,9,19,20,23–25]
HRQoL score for the salvage health state	0·69	0·64-0·73	Base case and range: [20,22,23,25,26]

Table 2. Average chemotherapy costs* in euro's for a patient with mean weight 79 kg, mean BSA 1,92 m^2 per Markov cycles of 3 weeks– base case

Cost category	Cost (€)	Range tested in sensitivity analysis	Source
CAPOX-B induction treatment (for a total of 6 cycles)	16,473	-	33
CAP-B maintenance	2,165	1,624 – 2,706	33
CAPOX-B reintroduction	2,743	2,057-3,429	33
Other systemic treatment (defined as 1.5 cycles of FOLFIRI)	1,470	1,103-1,838	33
Salvage treatment†	1,338	1,004-1,673	33

* Includes costs for systemic treatment including prophylaxis and preparation of systemic agents by the (hospital) pharmacist.

† Based on additional data collected retrospectively for 146 patients.

RESULTS

HRQoL scores from the CAIRO3 dataset for the observation (mean 0.83) and CAP-B maintenance (mean 0.84) health states were not significantly different between treatment arms. HRQoL scores for reintroduction or other treatment, and salvage treatment were averaged from previously reported HRQoL scores and shown in Table 1. The mean systemic treatment costs (including mean costs for prophylaxis and handling of systemic agents by the hospital pharmacist) per 3-week cycle of CAP-B maintenance and salvage therapy were €2,165 and €1,338, respectively (Table 2). In the CAIRO3 study, a median of 11 treatment cycles were administered to patients in the CAP-B maintenance treatment arm eventuating in a systemic therapy cost of €23,815.

In the deterministic analysis of this cost-effectiveness model, CAP-B maintenance treatment resulted in mean incremental cost of €36,845 with a mean incremental benefit of 0.21 QALYs and 0.18 LYs gained. This yields an incremental cost per effectiveness ratio (ICER) of €175,452 per QALY (quality-adjusted life year) gained and €204,694 per life year gained (Table 3). For patients achieving complete response (CR) or partial response (PR) following six cycles of CAPOX-B induction treatment, CAP-B maintenance treatment resulted in a mean incremental cost of €46,283 with a mean incremental benefit of 0.31 QALYs and 0.31 LYs, yielding an ICER of €149,300 per QALY and per LY gained. No important differences were observed between survival curves based on the cost-effectiveness model and the CAIRO3 study (Figure 2).

Sensitivity analyses

Univariable sensitivity analyses for influence of model parameters showed that the model was most influenced by a larger difference in HRQoL between patients receiving CAP-B maintenance treatment and observation in the post-induction treatment phase. The ICER reduced to €94,454 per QALY if the HRQoL score of maintenance was 1 or equal to perfect health and the HRQoL score for observation remained 0.83. In contrast, observation dominated (i.e. higher HRQoL and lower costs) maintenance if the HRQoL for maintenance was 0.8 times the HRQoL score for observation. The second most influential parameter in the model was the HRQoL score during the post-induction phase (while the difference in HRQoL scores between CAP-B maintenance and observation remained constant) for both maintenance and observation (range: 0.74 to 0.95). This yielded an ICER range of €134,288 per QALY (for a utility of 0.95) to €231,172 per QALY (for a utility of 0.74). The third most influential parameter on the model was the cost of bevacizumab, yielding an ICER of €218,168 and €136,758 for a 25% bevacizumab cost increase or decrease, respectively. The ICER range (in € per QALY) of the ten most influential parameters are listed in Table 4.

Univariable sensitivity analyses were also performed on cost parameters only. The following three cost parameters were of most influence on the ICER, the cost of bevacizumab (see above for the ICER range), cost of salvage (ICER range €172,353-€182,574

per QALY) and cost of hospitalization (ICER range €174,207-€180,720 per QALY). The ICER range (in € per QALY) of the ten most influential cost parameters are listed in Table 5.

With a probabilistic sensitivity analysis for 10,000 samples (plotted in Figure 3) a mean incremental cost of €36,832 (95% CI €36,660 – 37,004) with a mean incremental benefit of 0.21 QALYs (95% CI 0.21 - 0.21), resulted in a mean ICER of €181,346 per QALY (95% CI €102,271 – €260,421).

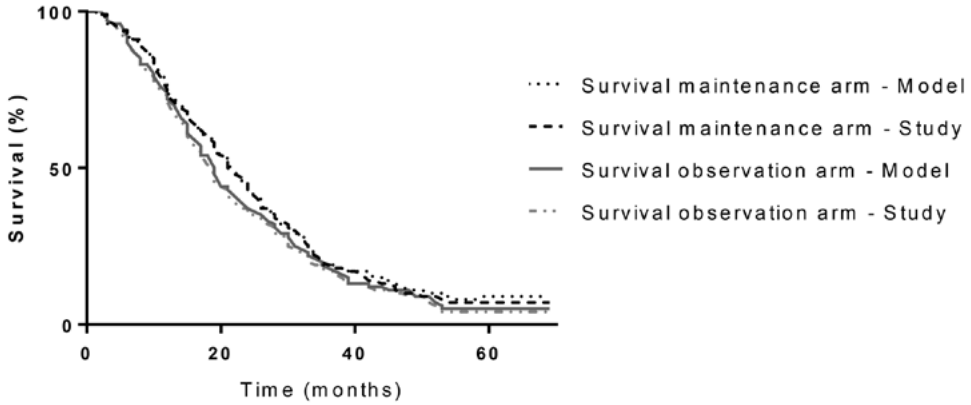


Figure 2. Survival curves for patients for the maintenance and observation arm of the CAIRO3 study and survival curves for the maintenance and observation arm generated from the cost-effectiveness model.

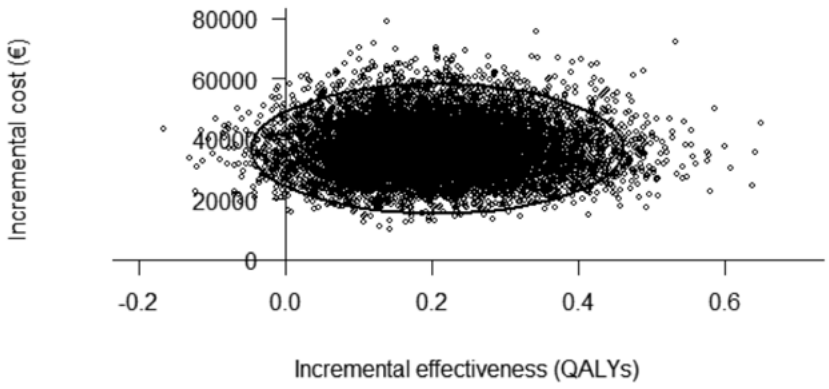


Figure 3. Probabilistic Sensitivity Analysis (n=10,000) for the incremental cost-effectiveness of CB maintenance versus observation. Each dot represents an ICER, =95% confidence incidence.

Table 3. Incremental Cost per Effectiveness Ratio (ICER in € per QALY or € per LY)

	Observation	CAP-B maintenance	Difference	ICER (€) Discount: 4% for costs, 1.5% for effect	ICER (€) Discount: 3% for costs and effects	ICER (€) Without discounting
Cost (€)	94,111	130,956	36,845			
QALY	1.42	1.62	0.21	175,452	186,815	177,413
LY	2.25	2.43	0.18	204,694	207,572	205,426

Table 4. ICER ranges (in € per QALY) of the ten most influential model parameters.*

		Low value (€)	High value (€)
1	Post-induction HRQoL ratio maintenance:observation	Dominated‡	491,270
2	HRQoL post-induction	134,288	231,172
3	Cost of bevacizumab	136,758	218,168
4	HRQoL re-induction	158,504	184,666
5	Cost salvage	172,353	182,574
6	Duration of hospitalization	172,849	182,078
7	Cost of hospitalization	174,207	180,720
8	Physician visit	174,513	180,414
9	Day-care hospital admission	174,632	180,294
10	Discount rate for costs	176,422	178,504

* A 25% percent range around the base case value was used for each model parameter.
‡ If the post-induction HRQoL of maintenance is 25% lower compared to the HRQoL of observation, then the observation treatment arm is regarded as superior in terms of both effectiveness and costs

Table 5. ICER ranges (in € per QALY) of the ten most influential cost parameters.*

		Low value (€)	High value (€)
1	Bevacizumab	136,758	218,168
2	Salvage	172,353	182,574
3	Hospitalization	174,207	180,720
4	Physician visit	174,513	180,414
5	Day-care hospital admission	174,632	180,294
6	Capecitabine	175,110	179,817
7	Oxaliplatin	175,653	179,273
8	Blood count & chemistry	176,022	179,008
9	Radiotherapy	176,088	178,841
10	CT-scan	176,422	178,505

* A 25% percent range around the base case value was used for each cost parameter.

DISCUSSION

In our analysis, patients with mCRC receiving CAP-B maintenance treatment after CAPOX-B induction treatment generate more costs at an improved effectiveness in terms of gained quality-adjusted life years (QALYs) or life years gained (LY) compared to observation. The CAP-B maintenance treatment strategy results in an incremental cost per additional QALY ratio (ICER) of €175,452 compared to the treatment strategy with observation. Patients achieving CR or PR after CAPOX-B induction treatment appear to have a somewhat more beneficial ICER of €149,300 per QALY. However, the predictive value of CR/PR after induction treatment for the outcome of maintenance treatment should be further validated. Currently, subgroup analyses of the CAIRO3 study are

being performed to identify patients that benefit most from CAP-B maintenance treatment strategy.

Previously, contradictory conclusions have been reported for cost-effectiveness studies regarding the addition of bevacizumab to first-line systemic treatment. However, none of these cost-effectiveness models evaluated bevacizumab-containing maintenance treatment. Koeberle et al., who reported about costs – but not cost-effectiveness – found average costs of €4,443/month for bevacizumab maintenance and €1,558/month for the observation strategy (cost converted to euros and corrected for inflation and power parity).[35] Unfortunately, lack of a reported ICER hampers proper comparison with our results. The previously reported ICERs for the addition of bevacizumab to systemic treatment (converted in euros and corrected for inflation and power parity) ranged from €42,272 to €438,444 per QALY and €10,567-336,776 per LY.[13] These different cost-effectiveness models all compared different first-line treatment strategies, yielding important differences in incremental costs and effects. Differences in costs originate for instance from differences in cost-perspectives (e.g. hospital perspective or third party payer perspectives), assumptions made (e.g. treatment duration), different health-care systems between countries and differences in medical practice.[13,14]

Even though the CAIRO3 study did not include a cost-effectiveness evaluation in the study design, we are able to develop a robust model incorporating data from both the original dataset of the CAIRO3 study as well as detailed retrospective data of a patient sub-set (n=146) of the CAIRO3 patient population. The probabilistic sensitivity analysis confirms the outcomes of the deterministic analysis.

We have shown that our model outcome is most influenced by the difference in HRQoL score between patients in the maintenance and observation arm in the post-induction health state, i.e. when different HRQoL scores are assigned to patients in the maintenance and observation arm. The calculated post-induction HRQoL scores in our model were directly derived from the CAIRO3 study and are not significantly different between treatment arms. Furthermore, we have previously shown that the EORTC QLQ-C30 quality of life results of the CAIRO3 study were not clinically relevantly different between treatment arms in the post-induction health state.[1] As a result, we do not expect the post-induction HRQoL scores to have an important influence on the interpretation of our results towards daily practice even though these HRQoL scores demonstrated to be the most influential parameter on model outcomes by univariable sensitivity analysis.

In our decision-analytic model, the cost of bevacizumab is the most influential cost parameter in the model, with a range of €136,758 to €218,168 per QALY gained. A reduction in the cost of bevacizumab will therefore yield a more favourable cost-effectiveness outcome. The second most influential cost parameter in the model was the cost of salvage treatment (ICER range €172,353-€182-574 per QALY). The influence of salvage treatment costs on model outcomes is limited as can be derived from the ICER

range. A standard cost for all patients achieving the salvage treatment health state was used in our model. Detailed data, including information on systemic treatment dosage and duration of treatment, collected retrospectively on salvage treatment were only available for a subset of the patients included in the CAIRO3 study (n=146). However, real life costs of subsequent treatment lines may differ between response subgroups and treatment arms.

There are some limitations to our study. Firstly, our results may not reflect the cost-effectiveness in the general mCRC population due to in- and exclusion criteria of the study. For instance, one of the inclusion criteria was stable disease or better after six cycles of CAPOX-B induction treatment. Thus, costs and effects for patients achieving no more than progressive disease upon CAPOX-B induction treatment were not included in the model. Although the general population is dissimilar to patients participating in clinical trials, we have previously demonstrated external validity of trial outcomes for patients not included in clinical trials, but fulfilling inclusion criteria.[36] Our outcomes may therefore, with caution, be externalized towards patients in general practice who meet the CAIRO3 study inclusion criteria. Secondly, the analysis was performed from a hospital perspective for which data on medical resource utilization was collected retrospectively for a subset of 146 patients. A societal perspective may provide a more accurate estimate of costs, a better insight of the impact of treatment choices on society and is to be preferred. Unfortunately, it was not feasible to perform this analysis from a societal perspective due to the retrospective nature of this cost-effectiveness study. Finally, a survival benefit or clinically relevant improvement in quality of life for CAP-B maintenance treatment have not been demonstrated previously.[1-2] Therefore, the clinical relevance of the QALYs or LYs gained in this cost-effectiveness study should be interpreted with caution.

In conclusion, our data show that CAP-B maintenance treatment in mCRC incurs a relevant increase in costs. Given the increased burden of costs of systemic cancer treatment for healthcare budgets, cost-effectiveness analyses are increasingly being included in the decision-making on the implementation of treatments with clinically meaningful benefits in daily practice. Clearly, evaluation of costs versus clinical benefit should be included in the design of treatment guidelines.[37] Differences between cost-effectiveness models (e.g. costs included in the model) and the lack of a general consensus on defined financial thresholds complicate model comparisons, which hamper the adoptability of cost-effectiveness results towards treatment guidelines.[13,38] Even so, CAP-B maintenance treatment may not be considered as cost-effective.

REFERENCES

1. Simkens LHJ, van Tinteren H, May A, et al. Maintenance treatment with capecitabine and bevacizumab in metastatic colorectal cancer (CAIRO3): a phase 3 randomised controlled trial of the Dutch Colorectal Cancer Group. *Lancet*. 2015;385(9980):1843-1852. doi:10.1016/S0140-6736(14)62004-3
2. Hegewisch-Becker S, Graeven U, Lerchenmüller CA, et al. Maintenance strategies after first-line oxaliplatin plus fluoropyrimidine plus bevacizumab for patients with metastatic colorectal cancer (AIO 0207): a randomised, non-inferiority, open-label, phase 3 trial. *Lancet Oncol*. 2015;16(13):1355-1369. doi:10.1016/S1470-2045(15)00042-X
3. Goldstein D a, Chen Q, Ayer T, et al. First- and second-line bevacizumab in addition to chemotherapy for metastatic colorectal cancer: a United States-based cost-effectiveness analysis. *J Clin Oncol*. 2015;33(10):1112-1118. doi:10.1200/JCO.2014.58.4904
4. Carter HE, Zannino D, John Simes R, et al. The cost effectiveness of bevacizumab when added to capecitabine, with or without mitomycin-C, in first line treatment of metastatic colorectal cancer: results from the Australasian phase III MAX study. *Eur J Cancer*. 2014;50(3):535-543. doi:10.1016/j.ejca.2013.09.028
5. Tappenden P, Jones R, Paisley S, Carroll C. The cost-effectiveness of bevacizumab in the first-line treatment of metastatic colorectal cancer in England and Wales. *Eur J Cancer*. 2007;43(17):2487-2494. doi:10.1016/j.ejca.2007.08.017
6. Kovacevic A, Dragojevic-Simic V, Tarabar D, et al. Five-year survival and costs of care in metastatic colorectal cancer: conventional versus monoclonal antibody-based treatment protocols. *Expert Rev Anticancer Ther*. 2015;15(8):1-8. doi:10.1586/14737140.2015.1059280
7. Ewara EM, Zaric GS, Welch S, Sarma S. Cost-effectiveness of first-line treatments for patients with KRAS wild-type metastatic colorectal cancer. *Curr Oncol*. 2014;21:541-550.
8. Shankaran V, Mummy D, Koepl L, et al. Survival and lifetime costs associated with first-line bevacizumab use in older patients with metastatic colorectal cancer. *Oncologist*. 2014;19(8):892-899. doi:10.1634/theoncologist.2013-0209
9. Lawrence D, Maschio M, Leahy KJ, Yunger S, Easaw JC, Weinstein MC. Economic analysis of bevacizumab, cetuximab, and panitumumab with fluoropyrimidine-based chemotherapy in the first-line treatment of KRAS wild-type metastatic colorectal cancer (mCRC). *J Med Econ*. 2013;16(12):1387-1398. doi:10.3111/13696998.2013.852097
10. Lee E-K, Revil C, Ngoh C a, et al. Clinical and cost effectiveness of bevacizumab + FOLFIRI combination versus FOLFIRI alone as first-line treatment of metastatic colorectal cancer in South Korea. *Clin Ther*. 2012;34(6):1408-1419. doi:10.1016/j.clinthera.2012.05.001
11. Hedden L, Kennecke H, Villa D, et al. Incremental cost-effectiveness of the pre- and post-bevacizumab eras of metastatic colorectal cancer therapy in British Columbia, Canada. *Eur J Cancer*. 2012;48(13):1969-1976. doi:10.1016/j.ejca.2012.01.012
12. Shiroiwa T, Fukuda T, Tsutani K. Out-of-pocket payment and cost-effectiveness of XELOX and XELOX plus bevacizumab therapy: from the perspective of metastatic colorectal cancer patients in Japan. *Int J Clin Oncol*. 2010;15(3):256-262. doi:10.1007/s10147-010-0045-x
13. Franken MD, van Rooijen EM, Uyl-de Groot CA, van Oijen MGH, Koopman M. Cost-effectiveness in colorectal cancer: challenges on quality and comparability. *Color Cancer*. 2016;5(1):21-31. doi:doi: 10.2217/crc.15.33
14. Saltz LB. Can money really be no object when cancer care is the subject? *J Clin Oncol*. 2015;33(10):1093-1094. doi:10.1200/JCO.2014.60.1401
15. Tol J, Koopman M, Cats A, et al. Chemotherapy, bevacizumab, and cetuximab in metastatic colorectal cancer. *N Engl J Med*. 2009;360(6):563-572. doi:10.1056/NEJMoa0808268
16. Diaz-Rubio E, Gomez-Espana a, Massuti B, et al. First-line XELOX plus bevacizumab

followed by XELOX plus bevacizumab or single-agent bevacizumab as maintenance therapy in patients with metastatic colorectal cancer: The phase III MACRO TTD study1477. *Oncologist*. 2012;17(1):15-25. doi:10.1634/theoncologist.2011-0249

17. Cassidy J, Clarke S, Díaz-Rubio E, et al. XELOX vs FOLFOX-4 as first-line therapy for metastatic colorectal cancer: NO16966 updated results. *Br J Cancer*. 2011;105(1):58-64. doi:10.1038/bjc.2011.201

18. Yalcin S, Uslu R, Dane F, et al. Bevacizumab + capecitabine as maintenance therapy after initial bevacizumab + XELOX treatment in previously untreated patients with metastatic colorectal cancer: phase III "Stop and Go" study results--a Turkish Oncology Group Trial. *Oncology*. 2013;85(6):328-335. doi:10.1159/000355914

19. Bennett L, Zhao Z, Barber B, et al. Health-related quality of life in patients with metastatic colorectal cancer treated with panitumumab in first- or second-line treatment. *Br J Cancer*. 2011;105(10):1495-1502. doi:10.1038/bjc.2011.409

20. Graham CN, Hechmati G, Hjelmgren J, et al. Cost-effectiveness analysis of panitumumab plus mFOLFOX6 compared with bevacizumab plus mFOLFOX6 for first-line treatment of patients with wild-type RAS metastatic colorectal cancer. *Eur J Cancer*. September 2014. doi:10.1016/j.ejca.2014.08.016

21. Ramsey SD, Ph D, Andersen MR, et al. Quality of Life in Survivors of Colorectal Carcinoma. 2000:1294-1303.

22. Adler A, Armstrong R, Aronson J, et al. *Cetuximab, Bevacizumab and Panitumumab for the Treatment of Metastatic Colorectal Cancer after First - Line Chemotherapy.*; 2012. http://www.nice.org.uk/guidance/ta242/chapter/appendix-a-appraisal-committee-members-and-nice-project-team.

23. Stein D, Joulain F, Naoshy S, et al. Assessing health-state utility values in patients with metastatic colorectal cancer: a utility study in the United Kingdom and the Netherlands. *Int J Colorectal Dis*. 2014;29(10):1203-1210. doi:10.1007/s00384-014-1980-1

24. Starling N, Tilden D, White J, Cunningham D. Cost-effectiveness analysis of cetuximab/irinotecan vs active/best supportive care for the treatment of metastatic colorectal cancer patients who have failed previous chemotherapy treatment. *Br J Cancer*. 2007;96(2):206-212. doi:10.1038/sj.bjc.6603561

25. Färkkilä N, Sintonen H, Saarto T, et al. Health-related quality of life in colorectal cancer. *Colorectal Dis*. 2013;15(5):e215-22. doi:10.1111/codi.12143

26. Odom D, Barber B, Bennett L, et al. Health-related quality of life and colorectal cancer-specific symptoms in patients with chemotherapy-refractory metastatic disease treated with panitumumab. *Int J Colorectal Dis*. 2011;26(2):173-181. doi:10.1007/s00384-010-1112-5

27. Lamers LM, Stalmeier PFM, McDonnell J, Krabbe PFM, van Busschbach JJ. [Measuring the quality of life in economic evaluations: the Dutch EQ-5D tariff]. *Ned Tijdschr Geneeskd*. 2005;149(28):1574-1578. http://www.ncbi.nlm.nih.gov/pubmed/16038162. Accessed October 30, 2014.

28. Siegel JE, Torrance GW, Russell LB, Luce BR, Weinstein MC, Gold MR. Guidelines for pharmacoeconomic studies. Recommendations from the panel on cost effectiveness in health and medicine. Panel on cost Effectiveness in Health and Medicine. *Pharmacoeconomics*. 1997;11(2):159-168.

29. National Institute for Health and Care Excellence. Measuring effectiveness and cost effectiveness: The QALY. http://www.nice.org.uk/newsroom/features/measuringeffectivenessandcosteffectivenesstheqaly.jsp. Published 2010.

30. Hakkaart-van Roijen L, Tan SS, Bouwmans CAM. *Handleiding Voor Kostenonderzoek Methoden En Standaard Kostprijzen Voor Economische Evaluaties*. College voor zorgverzekeringen; 2010. http://www.zorginstituutnederland.nl/binaries/content/documents/zinl-www/documenten/publicaties/overige-publicaties/1007-handleiding-voor-kostenonderzoek/1007-handleiding-voor-kostenonderzoek/Handleiding+voor+kostenonderzoek.pdf.

31. NZa Zorgproducten Tariefapplicatie. http://dbc-zorgproducten-tarieven.nza.nl/nzaZpTarief/ZoekfunctieDot.aspx.

32. *Uitkomstenonderzoek Dure Geneesmiddelen. Pilot Studies Bortezomib En Oxaliplatin: Resultaten En Methodologische Aanbevelingen.*; 2012. https://www.zorginstituutnederland.nl/binaries/content/documents/zinl-www/documenten/publicaties/rapporten-en-standpunten/2012/1207-uitkomstenonderzoek-dure-geneesmiddelen-pilot-studies-bortezomib-en-oxaliplatin-resultaten-en-methodologische-aanbevelingen/Uitkomstenonderzoek+dure+geneesmiddelen+(Pilot+studies+bortezomib+en+oxaliplatin-resultaten+en+methodologische+aanbevelingen).pdf. Accessed January 30, 2015.

33. Medicijnkosten - Zorginstituut Nederland. www.medicijnkosten.nl. Published 2014.

34. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the Panel on Cost-Effectiveness in Health and Medicine. *JAMA*. 1996;276(15):1253-1258. doi:10.1097/00132586-199712000-00019

35. Koeberle D, Betticher DC, von Moos R, et al. Bevacizumab continuation versus no continuation after first-line chemotherapy plus bevacizumab in patients with metastatic colorectal cancer: a randomized phase III non-inferiority trial (SAKK 41/06). *Ann Oncol*. 2015;(January):709-714. doi:10.1093/annonc/mdv011

36. Mol L, Koopman M, van Gils CWM, Ottevanger PB, Punt CJA. Comparison of treatment outcome in metastatic colorectal cancer patients included in a clinical trial versus daily practice in The Netherlands. *Acta Oncol (Madr)*. 2013;52(5):950-955. doi:10.3109/0284186X.2013.777158

37. Schnipper LE, Davidson NE, Wollins DS, et al. American Society of Clinical Oncology Statement: A Conceptual Framework to Assess the Value of Cancer Treatment Options. *J Clin Oncol*. 2015;33(23):2563-2577. doi:10.1200/JCO.2015.61.6706

38. Marseille E, Bruce L, Kazi DS, Kahn JG, Rosen S. WHO | Thresholds for the cost–effectiveness of interventions: alternative approaches. WHO.

SUPPLEMENTARY FILES

Table S1. Medical resource use (mean rates per 3 weeks for outpatient visits and hospitalizations) and, probabilities of central line placement, RAS mutation status analysis and grade 3 and 4 adverse events, subdivided in three health states: 1) post-induction (observation or maintenance), 2) reintroduction or other treatment and 3) salvage treatment, based on an additional retrospective data collection for 146 patients.

	Observation			CAP-B maintenance			
Health state phase	**1**	**2**	**3**	**1**	**2**	**3**	**Source**
Medical resource use:							
Day-care hospital admission	NA	1	1.20	1	1	1.25	CAIRO3 protocol, retrospective data collection
Outpatient visit	0.77	10.5	1.31	2.12	0.23	1.46	Retrospective data collection
Telephone consultation	0.18	0.37	1.45	0.29	0.83	0.51	Retrospective data collection
Blood count and chemistry	2.77	3.24	3.71	5.69	2.62	3.66	Retrospective data collection
X-ray	0.45	0.41	0.71	0.51	0.38	0.38	Retrospective data collection
CT-scan	0.88	0.77	0.82	1.30	0.47	0.67	Retrospective data collection
MRI	0.05	0.06	0.06	0.08	0.03	0.04	Retrospective data collection
PET-CT	0.00	0.01	0.01	0.02	0.02	0.07	Retrospective data collection
(Abdominal) ultrasound	0.18	0.13	0.10	0.20	0.08	0.08	Retrospective data collection
Skeletal scintigraphy	0.01	0.00	0.01	0.03	0.01	0.03	Retrospective data collection
ERCP	0.03	0.03	0.01	0.00	0.02	0.02	Retrospective data collection
MRCP	0.01	0.00	0.00	0.00	0.01	0.00	Retrospective data collection
Colonoscopy	0.02	0.02	0.01	0.07	0.02	0.03	Retrospective data collection
Hospitalization*	0.05	0.08	0.29	0.11	0.11	0.24	Retrospective data collection
Intensive care admission	0.00	0.00	0.00	0.00	0.00	0.00	Retrospective data collection
Paracentesis	0.00	0.01	0.07	0.01	0.02	0.05	Retrospective data collection
Blood transfusion	0.00	0.02	0.04	0.01	0.01	0.01	Retrospective data collection
Laparotomy	0.01	0.01	0.01	0.01	0.01	0.01	Retrospective data collection
Radiotherapy	0.00	0.04	0.18	0.01	0.00	0.14	Retrospective data collection
Antibiotic treatment**	0.01	0.00	0.06	0.01	0.01	0.06	Retrospective data collection
Pain medication	0.13	0.25	0.34	0.13	0.25	0.34	Retrospective data collection
Central line placement	0.04	0.09	0.19	0.10	0.07	0.08	Retrospective data collection
RAS mutation status analysis	NA	0.31	0.31	NA	0.31	0.31	Retrospective data collection, expert opinion
Adverse events:							
Hypertension	0.18	0.02	-	0.24	0.02	-	[1, 15–17]
Sensory neuropathy	0.05	0.18	-	0.10	0.18	-	[1, 15–17]
Diarrhoea	-	0.17	-	0.09	0.17	-	[15–18]
Hypersensitivity	-	0.15	-	-	0.15	-	[15]
Nausea and vomiting	-	0.11	-	0.03	0.11	-	[15–18]
Thromboembolic event	-	0.04	-	-	0.04	-	[15]
Bleeding	-	0.01	-	-	0.01	-	[15, 16]

Table S1. (continued)

	Observation			CAP-B maintenance			
Health state phase	**1**	**2**	**3**	**1**	**2**	**3**	**Source**
Gastro-intestinal perforation	-	0.03	-	0.03	0.03	-	[15–17]
Fatigue	-	-	0.25	-	-	0.25	Expert opinion

Abbreviations: CT – computed tomography, MRI - Magnetic resonance imaging, NA – not applicable, PET-CT - Positron emission tomography computed tomography, ERCP - Endoscopic retrograde cholangiopancreatography, MRCP – Magnetic resonance cholangiopancreatography.

* Additional imaging and blood count and chemistry performed during hospitalizations were added to the model (data on file).

** Antibiotic treatment during hospitalizations

Table S2. Model cost parameters in euro's per day or unit – base case

Units	**Unit cost (€)**	**Range tested in sensitivity analysis**	**Source**
Day-care hospital admission	184	138-230	[32]
Outpatient visit	121	91-151	[32]
Telephone consultation physician	14	11-18	[32]
Complete blood count and blood chemistry	58	44-73	[32]
RAS mutation status analysis	974	730-1,217	[31]
X-ray	46	35-58	[31]
CT	229	172-286	[32]
MRI	251	188-314	[31]
PET-CT	1,163	872-1,454	[31]
Abdominal ultrasound	88	66-110	[31]
Skeletal scintigraphy	282	212-353	[31]
ERCP	128	96-160	[31]
MRCP	79	59-99	[31]
(Emergency) Colonoscopy	94	71-118	[31]
Hospitalization	504	378-630	[30]
Intensive care admission	2,406	1,805-3,008	[30]
Central line placement	224	168-280	[31]
Paracentesis	209	157-261	[31]
Blood transfusion (incl. 2 packed cells and outpatient clinic visit)	627	470-784	[30, 32]
Laparotomy	6,392	4,794-7,990	[31]
Standard radiotherapy	907	680-1,134	[31]

Rates corrected for inflation and indexed for the year 2014, where applicable.

4

MATCHING THE MODEL WITH THE EVIDENCE: COMPARING DISCRETE EVENT SIMULATION AND STATE-TRANSITION MODELING FOR TIME-TO-EVENT PREDICTIONS IN A COST-EFFECTIVENESS ANALYSIS OF TREATMENT IN METASTATIC COLORECTAL CANCER PATIENTS

Mira D. Franken*

Koen Degeling*

Anne M. May

Martijn G.H. van Oijen

Miriam Koopman

Cornelis J.A. Punt

Maarten J. IJzerman

Hendrik Koffijberg

* These authors contributed equally to this work.

Cancer Epidemiology, 2018, 57: 60-67

ABSTRACT

Background

Individual patient data, e.g. from clinical trials, often need to be extrapolated or combined with additional evidence when assessing long-term impact in cost-effectiveness modeling studies. Different modeling methods can be used to represent the complex dynamics of clinical practice; the choice of which may impact cost-effectiveness outcomes. We compare the use of a previously designed cohort discrete-time state-transition model (DT-STM) with a discrete event simulation (DES) model.

Methods

The original DT-STM was replicated and a DES model developed using AnyLogic software. Models were populated using individual patient data of a phase III study in metastatic colorectal cancer patients, and compared based on their evidence structure, internal validity, and cost-effectiveness outcomes. The DT-STM used time-dependent transition probabilities, whereas the DES model was populated using parametric distributions.

Results

The estimated time-dependent transition probabilities for the DT-STM were irregular and more sensitive to single events due to the required small cycle length and limited number of event observations, whereas parametric distributions resulted in smooth time-to-event curves for the DES model. Although the DT-STM and DES model both yielded similar time-to-event curves, the DES model represented the trial data more accurately in terms of mean health-state durations. The incremental cost-effectiveness ratio (ICER) was €172,443 and €168,383 per Quality Adjusted Life Year gained for the DT-STM and DES model, respectively.

Conclusion

DES represents time-to-event data from clinical trials more naturally and accurately than DT-STM when few events are observed per time cycle. As a consequence, DES is expected to yield a more accurate ICER.

INTRODUCTION

Healthcare expenditures have increased importantly over the last decades, especially in oncology due to expensive novel targeted agents and personalized treatments based on molecular markers in order to provide patients with the best possible care [1, 2]. Cost-effectiveness analysis of such novel medical technologies is becoming increasingly relevant, as it may inform treatment, resource allocation, and research prioritization decisions. This is illustrated by the standardized approaches to value cancer treatment options in terms of efficacy and costs for clinicians [3, 4] and guidance for performing cost-effectiveness analysis alongside clinical trials.[5]

High quality individual patient data (IPD) on health outcomes, resource use, and care procedures, e.g. obtained from randomized controlled trials (RCTs), are the preferred source of evidence for cost-effectiveness analysis. However, single individual patient datasets do not always provide all (or the only) evidence required for estimating the (long-term) cost-effectiveness of medical technologies [6, 7], indicating the need for cost-effectiveness models to synthesize evidence from additional sources or to extrapolate beyond the time horizon of e.g. RCTs.[5, 8] Such cost-effectiveness models should adequately represent clinical practice and, therefore, reflect the true nature of the evidence used to define them, including evidence obtained from RCTs and other sources of IPD. In other words, *the model should match the evidence*.

The primary outcome of many clinical oncology studies is the time until an event of interest occurs, e.g. the patients' overall survival or progression-free survival from the moment of randomization, which are typically recorded continuously over time. However, the most frequently applied cost-effectiveness modeling method, i.e. discrete-time state-transition modeling (DT-STM) [9], uses transition probabilities over discrete time cycles with a fixed length to represent the progression of time. For example, in an DT-STM with time cycles of three weeks patients can only progress to another health state after this predefined and rigid time length, even though in daily practice patients may progress at any time instead of only at a multiple of three weeks. The length of these time cycles needs to be chosen so that the complex dynamics of clinical practice are appropriately represented [9]. For DT-STM to represent clinical practice better, shorter cycle lengths would be preferable [10]. Although half-cycle corrections may be applied to avoid bias and to better approximate clinical practice [11], this still insufficiently allows complex clinical dynamics if the cycle length is too long.[12]

Using shorter cycles lengths can be disadvantageous, mainly because of increase in number of cycles that needs to be simulated. Besides increasing the computational burden of the simulation [9, 12], the larger number of cycles makes it more challenging to represent the uncertainty in the transition probabilities, as the uncertainty in the numerous cycle-specific probabilities needs to be reflected while also maintaining the correlation between them. Furthermore, because the expected number of

observations within a cycle decreases with decreasing cycle length, the likelihood of substantial irregularities in transition probabilities between successive cycles is expected to increase. These irregularities are likely to impact the simulation outcomes and do not correspond to clinical practice, as the probability of an event is commonly expected to be similar between successive moments, i.e. the transition-curves follow a smooth pattern over time.

Discrete event simulation (DES) is an alternative modeling technique to which the challenges associated with discrete time cycles do not apply. Events can occur at any time in a DES model, because the time to these events are typically modeled using smooth time-to-event distributions, e.g. Gamma or Weibull distributions.[13] In DES, the behavior of a system is translated into an ordered sequence of well-defined events, which comprise specific changes in the system's state at a specific point in time.[13] DES is well suitable for modeling clinical processes, as it is able to incorporate patient-level characteristics and clinical histories, competing resources, and interactions between different actors, e.g. physicians and patients.[14] Although originating from the operations research field, DES is increasingly being used for cost-effectiveness modeling.[15]

Several studies have compared the use of DT-STM and DES for cost-effectiveness analyses of medical technologies. Using the same model structure and evidence, quantitative outcomes such as the incremental cost-effectiveness ratio (ICER), are unlikely to be substantially different between these modeling methods.[16, 17] However, substantial differences in outcomes may occur, if the use of DES results in a more appropriate representation of clinical practice compared to DT- STM, for example by including patient characteristics or considering resource constraints.[18] Especially in the scenario in which insufficient observations are available for the chosen cycle length, and irregularities in the cycle-specific transition probabilities are substantial when using DT-STM, the use of DES might be preferable.

The objective of this study is to compare the evidence structure and outcomes of a recently published cost-effectiveness DT-STM [19] with those of a newly developed DES model. The comparison will be performed based on the dataset of the randomized clinical phase 3 CAIRO3 study, in which maintenance treatment with capecitabine and bevacizumab (CAP-B) or observation in metastatic colorectal cancer patients after six induction cycles of capecitabine, oxaliplatin, and bevacizumab (CAPOX-B) was evaluated.[20] The results of this study should facilitate a better understanding of the potential impact of selecting a modeling method for cost-effectiveness modeling studies informed by IPD.

METHODS

Maintenance treatment in metastatic colorectal cancer

The CAIRO3 study (NCT00442637) is a randomized clinical phase III study, which was carried out by the Dutch Colorectal Cancer Group (DCCG) in 64 hospitals in

the Netherlands. A total of 558 metastatic colorectal patients with stable disease or better after six cycles of CAPOX-B induction therapy were randomized to either receive CAP-B maintenance treatment or observation until progression, which is referred to as the post-induction stage. CAPOX-B treatment was to be re-introduced upon progression on either maintenance or observation, and continued until second progression (primary end-point), which is referred to as the reintroduction stage. Although second progression was the primary end-point of the CAIRO3 study, the cost-effectiveness analysis of the CAIRO3 study also considered additional lines of treatment after second progression [19], which is referred to as the salvage therapy stage. Study results have been previously published.[20]

State-transition model

A cohort DT-STM, i.e. Markov model, was originally developed for the cost-effectiveness analysis of the CAIRO3 study and included four health states: post-induction, reintroduction, salvage therapy, and death (Figure 1a). The model was defined using cohort level cycle-specific transition probabilities, which were estimated from the CAIRO3 trial using Life Tables in IBM SPSS Statistics software, version 23, IBM Corp. (Armonk, NY, USA). This indicates that the probability of moving from one state to another depended only on the time passed since the start of the simulation, e.g. time from randomization until first progression. Half-cycle correction was applied and 100 cycles of three weeks were simulated in total. The DT-STM was built using TreeAge Pro Healthcare v.2014*, TreeAge Software (Williamstown, MA, USA),* and is described in detail elsewhere.[19]

To facilitate an adequate comparison between the two modeling methods, the DT-STM was first replicated in AnyLogic multi-method simulation software, v.7.3, The AnyLogic Company (Chicago, IL, USA), the environment also used for developing the DES model. This replicated DT-STM was then compared to the original DT-STM to assess potential variation in outcomes due to the use of different software environments. In total, 100 events were generated at intervals of three weeks, corresponding to the setup in the original DT-STM. Following each event, the occupation of the health states was recorded and used to calculate health and economic outcomes at the corresponding point in time. The model was validated by structured "walk-throughs", comparing (intermediate) results with calculations by hand, extreme value analysis, trace analysis, and cross validation with the original DT-STM during model development, and sensitivity analysis using the final model.[21, 22]

Discrete Event Simulation model

The DES model was defined on patient-level using AnyLogic software and according to the ISPOR-SMDM Modeling Good Research Practice Task Force guidelines (14). The model was defined to have the same health states as the DT-STM (Figure 1b). Although DES

allows for constrained resources to be accounted for, resource use was not considered in the DT-STM and, consequently, also not in the DES model.

Both Weibull and Gamma parametric distributions [13] were estimated from the CAIRO3 trial data with maximum likelihood estimation (MLE), or methods of moments estimation (MME) where MLE was not successful, using the *fitdist* function of the *fitdistrplus* [23] package in R Statistical Software.[24] Estimated parametric distributions were compared graphically based on density plots, Q-Q plots, and P-P plots, and numerically based on the Akaike information criterion and Bayesian information criterion. Since performance was similar without meaningful differences, and all Weibull distributions could be estimated with MLE, whereas MME was required for some Gamma distributions, Weibull distributions were assumed for all health state-specific time-to-event parameters in the DES model. Transitions between health states, i.e. events, were based on patient-level processing times, which were randomly drawn from Weibull distributions. Competing risks were handled by stratifying state-specific time-to-event distributions according to the two competing events that were considered, i.e. progression and death, and selecting the event to occur based on the respective observed probabilities of progression and survival.[25] To illustrate this, for a patient entering the reintroduction state a randomly drawn value compared to the chance of progression determined whether the patient would survive and progress to the salvage therapy state. Next, the time to the selected event, i.e. progression or death, was randomly drawn from the corresponding Weibull distribution.

A total of 10,000 patients were simulated per treatment strategy in the DES model, resulting in relative standard errors for the mean costs and effects of approximately 0.5%. No fixed runtime was assumed, so the simulation terminated when all patients had left the model, i.e. reached the death state. Patient-level outcomes were calculated using the time spent in each health state and summarized to enable comparison of the two treatment strategies. The DES model was validated by structured "walk-throughs", comparing (intermediate) results with calculations by hand, extreme value analysis, trace analysis, and cross validation with both DT-STMs during model development, and by sensitivity analysis.[21, 22]

Model Comparison

First, the original DT-STM and the replicated DT-STM were compared based on the cost-effectiveness outcomes of the CAIRO3 case study, to assess potential variation in outcomes due to differences in software environments. For this analysis, the incremental cost-effectiveness ratio (ICER) expressed in incremental costs per Quality Adjusted Life Year (QALY) gained served as the primary outcome. Costs and effects were discounted at discount rates of 4% and 1.5% per year, respectively, according to Dutch pharmacoeconomic guidelines.[26] Probabilistic sensitivity analysis (PSA) was performed

using Monte Carlo simulation with 10,000 samples to assess the effect of the uncertainty surrounding the input parameters on the primary outcome measure.[27] Since the original cost-effectiveness analysis did not account for uncertainty in the correlated cycle-specific transition probabilities [19], uncertainty in the correlated distribution parameters used to represented the time-to-event evidence in the DES model, was also not considered to maintain comparability between both models. Parameter values used to populate both

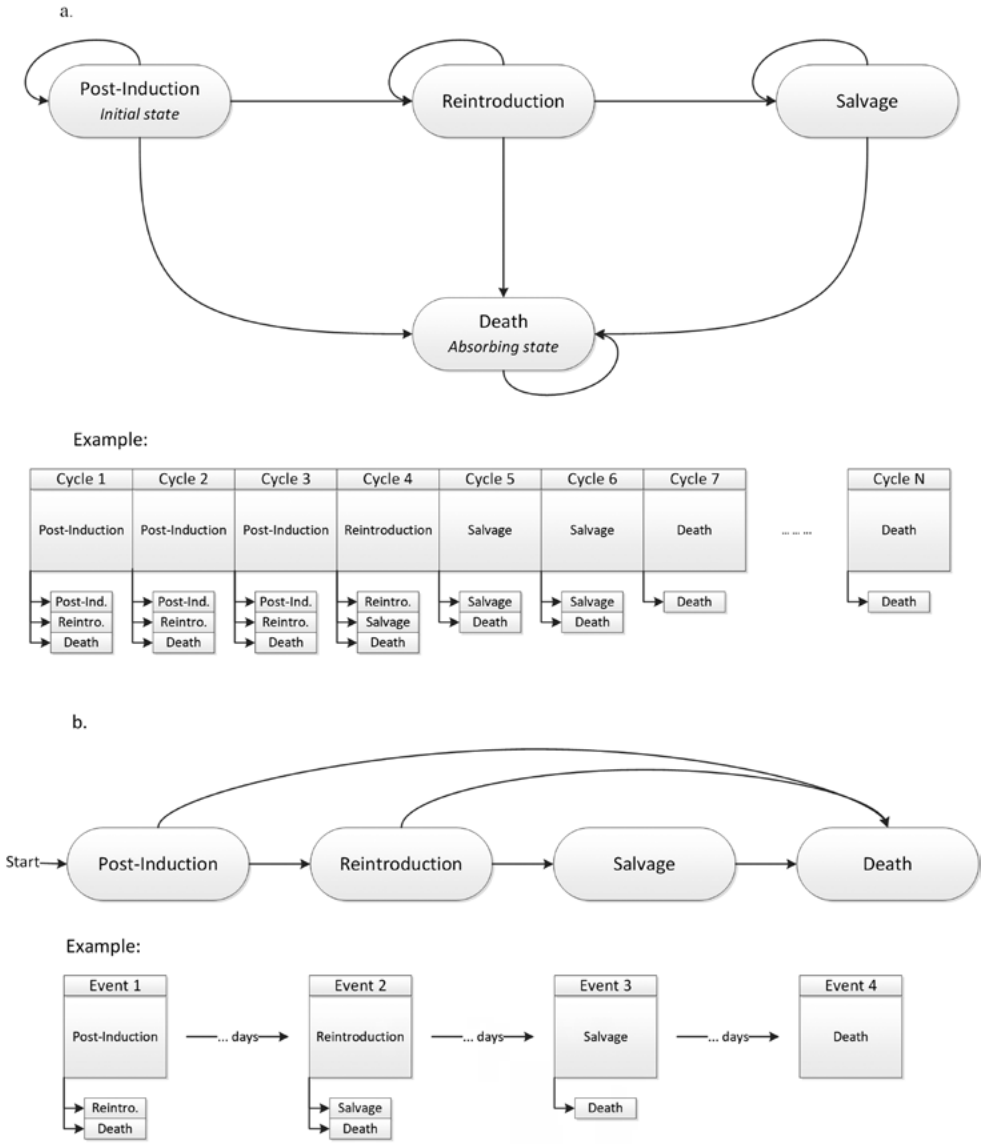


Figure 1. Graphical representation of the health states defined in the Discrete-Time State-Transition (a) and the Discrete Event Simulation (b) model.

models, including their distributions, are listed in the publication of the original CAIRO3 cost-effectiveness analysis [19], as well as in Supplementary Materials 1.

Subsequently, the replicated DT-STM and the DES model were qualitatively and quantitatively compared based on the case study, to assess potential differences between the two modeling methods. The models were qualitatively compared based on the evidence structure. Thereafter, modeling methods were quantitatively compared based on cost-effectiveness outcomes and simulation outcomes, i.e. the simulated health-state durations. All results were graphically represented, using Kaplan-Meier curves for the simulation outcomes and incremental cost-effectiveness planes and cost-effectiveness acceptability curves (CEAC) for the cost-effectiveness outcomes.

RESULTS

The replicated DT-STM developed for this study yielded comparable cost-effectiveness outcomes as the original DT-STM developed in a different software environment. The results for the original DT-STM have been previously published elsewhere and are not presented here for the sake of readability.[19] The replicated DT-STM will be referred to just as "DT-STM" in the subsequent part of this manuscript.

Simulation of health state-transitions

Health state-transitions in the DES model yield smooth time-to-event curves defined using Weibull distributions estimated based on the CAIRO3 data. In contrast, the time-dependent probabilities used for health state-transitions in the DT-STM become irregular (non-smooth) when only few events are observed for some transitions. The irregularities in these transition probabilities are caused by a decreasing number of patients retained in a health state over time, causing large variations in the observed subsequent probability of a health state-transition. An example of this is presented in Figure 2, which depicts the difference between the DT-STM and DES model in health state-transitions from the post-induction state to the reintroduction state for the maintenance treatment strategy.

The Kaplan-Meier curves for the health state-durations simulated in the DT-STM and DES model, compared to the CAIRO3 data, demonstrate that both modeling methods represent the clinical data well (Figure 3). However, when the mean time-to-transition presented in the descriptive statistics below the figure are considered, the DES model seems to represent the trial data more accurately. In example, the mean health-state duration of the post-induction state for the observation strategy was 175.7 days, 207.5 days and 173.4 days for the trial data, DT-STM and the DES model, respectively.

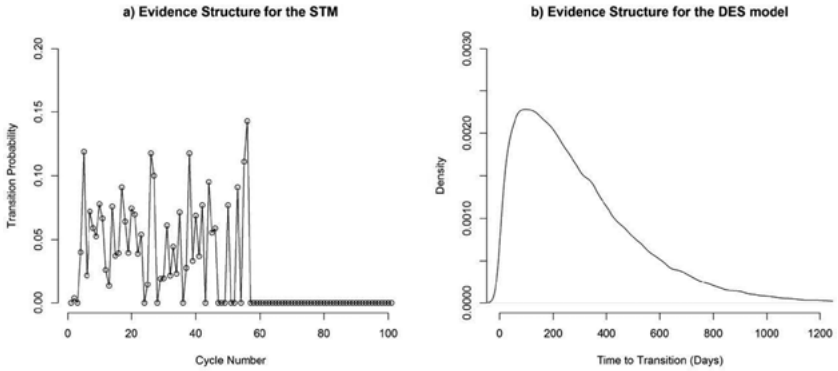
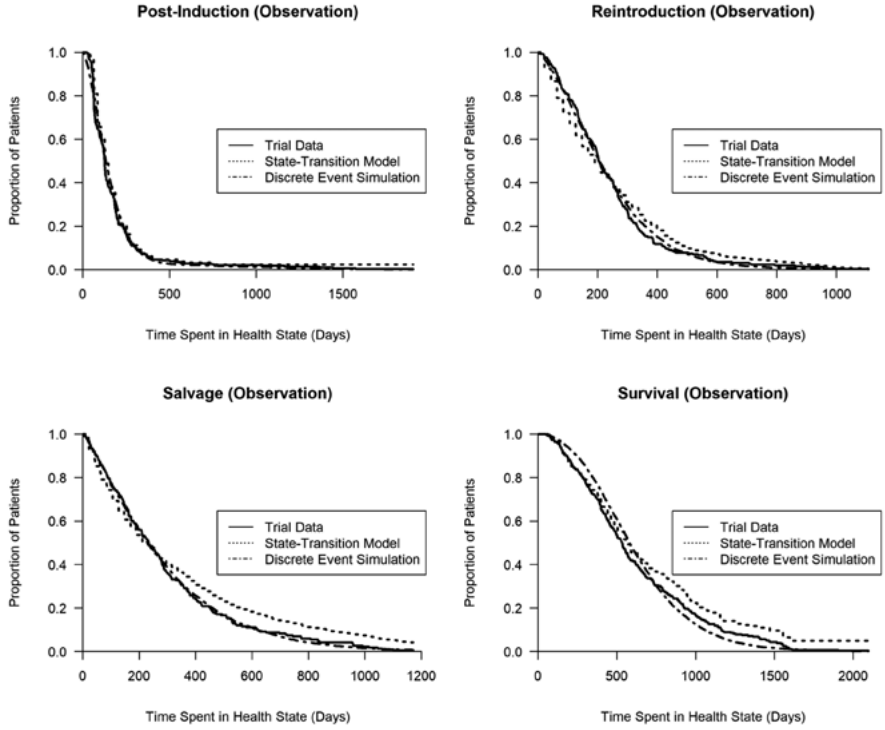
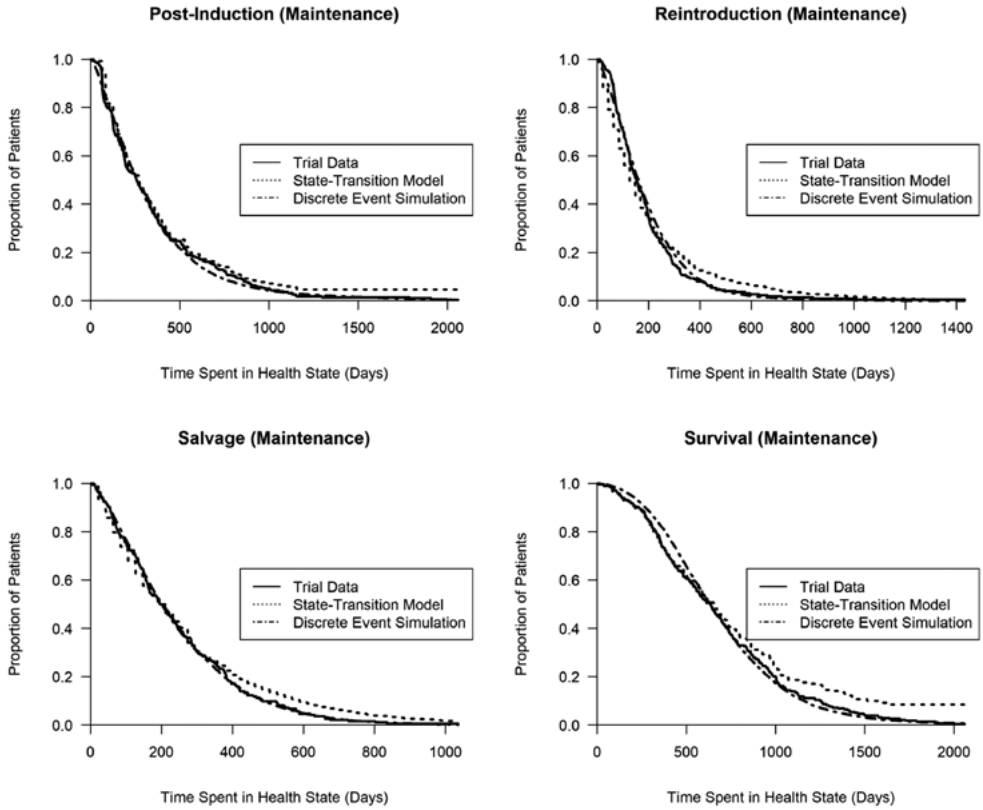


Figure 2. a) probability curve for the time to transition of the post-introduction to the reintroduction health state for the maintenance strategy per cycle (with a 3-week duration) for the DT-STM, and b) probability density curve for the time to transition of the post-introduction to the reintroduction health state for the maintenance strategy in the DES model.



	Post-Induction			Re-Induction			Salvage Therapy			Survival		
	Trial	STM	DES	Trial	STM	DES	Trial	STM	DES	Trial	STM	DES
2.5%	31.9	42.0	14.5	29.7	21.0	18.6	17.1	21.0	15.4	101.4	84.0	124.2
Media	124.0	126.0	132.3	203.0	189.0	205.8	233.0	231.0	228.5	532.0	567.0	580.0
Mean	175.7	207.5	173.4	238.8	253.4	238.8	292.3	349.9	291.8	619.4	707.0	626.6
97.5%	717.9	756.0	529.8	711.0	903.0	643.2	978.6	1323.0	913.5	1556.9	2100.0	1384.2
Size	279	10000	10000	254	9196	9698	216	7616	7599	279	10000	10000

Figure 3a. Kaplan-Meier curves for the time-spent in the health states for the observation strategy.



	Post-Induction			Re-Induction			Salvage Therapy			Survival		
	Trial	STM	DES	Trial	STM	DES	Trial	STM	DES	Trial	STM	DES
2.5%	48.9	63.0	20.5	26	21.0	13.0	17.5	21.0	16.2	92.9	84.0	133.8
Median	257.0	273.0	259.1	153.0	126.0	159.7	198.0	210.0	200.2	635.0	651.0	633.1
Mean	354.4	410.9	349.4	192.0	207.2	194.2	242.6	266.6	242.5	680.2	766.2	688.0
97.5%	1131.9	2100.0	1246.0	608	890.4	569.5	662.4	924.0	692.6	1639.9	2100.0	1568.3
Size (n)	279	10000	10000	225	8385	8971	194	6810	6776	279	10000	10000

Figure 3b. Kaplan-Meier curves for the time-spent in the health states for the maintenance strategy.

Cost-effectiveness analysis

The cost-effectiveness outcomes obtained from the DT-STM and the DES model are presented in the incremental cost-effectiveness planes of Figure 4. The incremental effectiveness estimates, including their 95% confidence intervals (CI), for CAP-B maintenance therapy compared to the observation strategy are 0.21 (CI: 0.015; 0.430) and 0.18 (CI: 0.006; 0.374) QALYs, and the incremental costs are €35,536 (CI: 19,945; 54,629) and €30,053, (CI: 17,047; 46,132) for the DT-STM and DES model, respectively. The mean ICERs are €172,443 and €168,383 per QALY gained for the DT-STM and DES model, respectively.

The PSA for both models only demonstrated a small difference in the amount of uncertainty surrounding the mean ICER point-estimates (Figure 4). This is illustrated

by the magnitude of the 95%-confidence ellipses surrounding these estimates, being slightly smaller for the DES model. However, as both mean ICER point-estimates and corresponding confidence ellipses are located rather similarly compared to the willingness-to-pay (WTP) threshold, the CEACs for both models are similar (Figure 5).

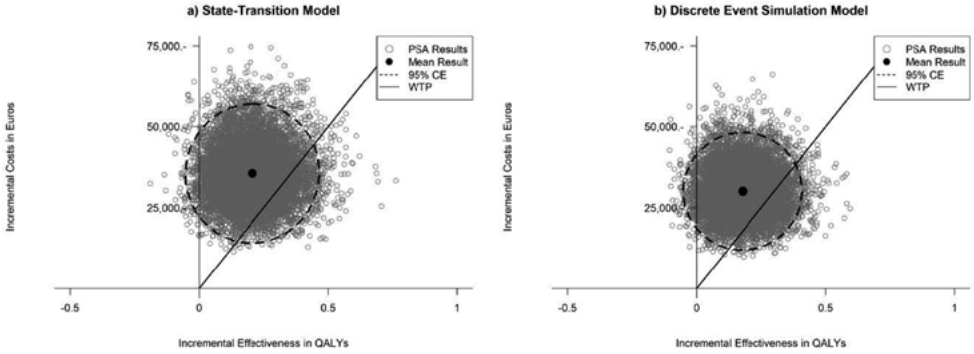


Figure 4. Incremental Cost-Effectiveness Planes comparing the maintenance treatment strategy with the observation strategy at a Willingness to Pay (WTP) of €100,000.- per Quality Adjusted Life-Year (QALY) gained, for a) the discrete-time state-transition model, and b) the discrete event simulation model. CE = Confidence Ellipse.

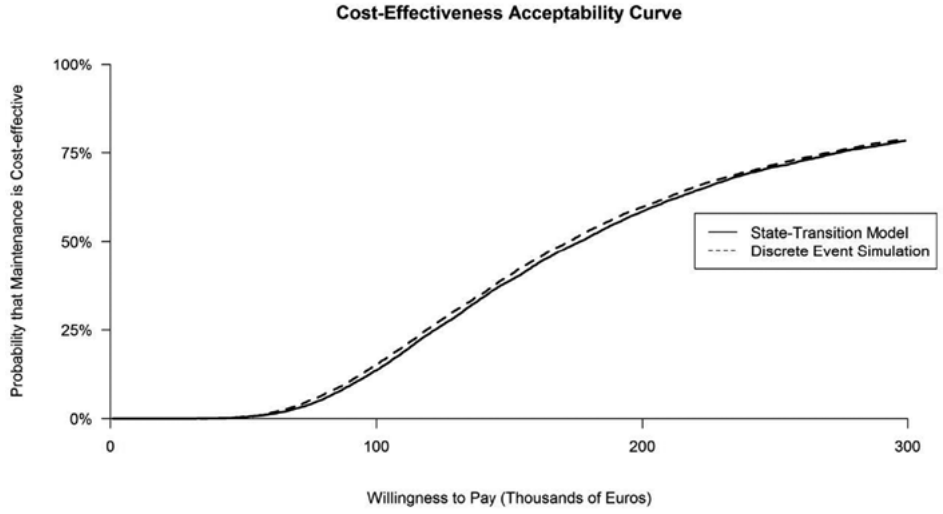


Figure 5. Cost-Effectiveness Acceptability Curves representing the probability that the maintenance treatment strategy is cost-effective compared to the observation strategy for a range of Willingness to Pay threshold values.

DISCUSSION

Smooth health state-transition curves served as input for the DES model, presenting the data in an informative manner. Conversely, it is more complicated to interpret the time-dependent health state-transitions probabilities used as input for the DT-STM. We have shown that these probabilities are irregular over time, due to scarce observations in many of the time cycles. Therefore, the health state-transition curves used in the DES model were much more representative of a "natural" patient flow through health states over time. Additionally, the Kaplan-Meier curves per health state simulated from the DES model matched the original study Kaplan-Meier curves slightly better, especially with regard to the mean time to transition from one health state to another, e.g. from randomization to the start of therapy reintroduction. The increasing difference between the trial data and STM over time, suggests a cumulative effect over successive health states, which may be amplified by a combination of irregularities in transition probabilities and their time-dependency.

Cost-effectiveness outcomes were comparable for the DT-STM and the DES model (ICER €172,443 and €168,383, respectively). The rather small differences observed, can be explained by the disparities in simulated mean time to transitions between both models. Furthermore, the magnitude of uncertainty surrounding the mean ICER point-estimate was smaller for the DES model. The observed difference in the uncertainty might be caused by the irregularities in the health state-transition probabilities in the DT-STM, consequently causing more extreme effects compared to the smooth health-state transition curves of the DES model. Results of this study did not alter the previously published conclusion that CAP-B maintenance may not be regarded as cost-effective.[19]

These results confirm that cost-effectiveness outcomes are not expected to be substantially different between DT-STM and DES models, if both models are based on the same evidence.[16, 17] It is, however, imaginable that ICER outcomes closer to a country's willingness to pay threshold might incur different conclusions on cost-effectiveness depending on the choice of modeling method. This was previously demonstrated by Jahn et al comparing a DES model and a DT-STM evaluating decision tools for adjuvant chemotherapy treatment in breast cancer.[28]

Even though the DES methodology may initially seem more complex for novices, its model structure and evidence structure more closely match transitions and events as observed in clinical trials, compared to that of DT-STM. Once familiar with the DES methodology, the parametric distributions used to describe time-to-event data are straightforward to estimate and interpret. Furthermore, these parametric distributions enable uncertainty in their parameter estimates to be included in the PSA more easily than the (correlated) uncertainty that is present in every individual time-dependent transition probability.[29] However, by discretizing parametric time-to-event distributions

into transition probabilities that can be used to populate a DT-STM, uncertainty in these transition probabilities can be represented. Additionally, by discretizing a parametric distribution rather than directly estimating transition probabilities from individual patient data, issues with regard to irregularities in these time-dependent transition probabilities may also be addressed. Furthermore, extrapolation beyond the time horizon of RCTs, although challenging, can be performed by fitting these parametric distributions.[30] Although parametric distributions can be used to address these general and DT-STM related challenges, doing so can be considered suboptimal due to the required discretization, whereas these parametric distributions can be incorporated directly in DES. In this respect, issues regarding appropriately reflecting uncertainty surrounding health economic outcomes, scarce events, and extrapolation may more easily be addressed using DES methodology. Regardless of these advantages to DES, DT-STM typically is computationally simpler, can be implemented using spreadsheets, and requires limited (programming) skills to do so, whereas implementation of DES is mainly limited to specialized simulation and statistical software.[15, 17, 31] Hence, regarding external review of models, DT-STM currently has an advantage, while experience with DES in health economics is developing.[15, 31]

DES seems the preferable modeling method compared to DT-STM for the evaluation of individual patient time-to-event data, which is also supported by the health economic modeling literature.[15, 31] In particular when time cycle size needs to be very small to adequately reflect dynamic treatment and monitoring processes, leading to irregularities in the estimated time-dependent transition probabilities due to a lack of observed events. However, DT-STM is still the most commonly used modeling method in cost-effectiveness modeling, for which different reasons can be identified. Firstly, as mentioned before, DES might initially be thought of as a more complex methodology requiring more evidence. This study demonstrated that DES models do not necessarily require more evidence or are more complex. Secondly, comprehensive guidance is available on how to use a (cohort) DT-STM for the evaluation of healthcare interventions [9], whereas the available guidance on the use of DES is less specific.[14] Researchers and clinicians with interest in health economics alike, however, should be aware of the potential advantages of DES compared to DT-STM, especially with regard to cost-effectiveness analyses informed by patient-level time-to-event data obtained from e.g. obtained from clinical trials.

This study compared a cohort state-transition model qualitatively and quantitatively based on an extensive health economic evaluation informed by patient-level time-to-event data obtained from the CAIRO3 study. Both models were developed in the same software environment and analyzed according to health economic good practices guidelines, optimizing the validity of our results. However, the generalizability of these single case study results is limited, though the results found are in line with literature.[16, 17, 28] Furthermore, the full potential of DES was not utilized, since no patient-

level characteristics were incorporated and, deliberately, parameter uncertainty in the time-to-event distributions parameter estimates themselves were not considered. The inclusion of patient-level characteristics in DES models undoubtedly allow even better representation of clinical practice. Finally, Weibull distributions were assumed for representing health state durations in the DES model, which may potentially influence health economic outcomes. To assess the impact of this design choice, simulations were additionally performed with Gamma distributions instead of Weibull distributions, which did not result in meaningfully different results.

In conclusion, the results show that the DT-STM and DES model did not yield substantially different outcomes if they are developed based on the same health states and evidence. Which modeling method should be applied, depends on the complexity of the clinical process to be modeled, the available evidence, and the modelers' experience. In our opinion, DES is the preferable modeling method in the scenario that patient-level time-to-event data is available, e.g. from clinical studies, as its model structure and evidence structure represent the dynamics of daily clinical practice more naturally.

REFERENCES

1. Meropol NJ, Schrag D, Smith TJ et al. American Society of Clinical Oncology Guidance Statement: The Cost of Cancer Care. J Clin Oncol. 2009;27(23):3868-74.
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(12):1165-74.
3. Cherny NI, Sullivan R, Dafni U et al. A standardised, 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(8):1547-73.
4. Schnipper LE, Davidson NE, Wollins DS et al. American Society of Clinical Oncology Statement: A Conceptual Framework to Assess the Value of Cancer Treatment Options. J Clin Oncol. 2015;33(23):2563-77.
5. Ramsey S, Willke R, Briggs A et al. Good Research Practices for Cost-Effectiveness Analysis Alongside Clinical Trials: The ISPOR RCT-CEA Task Force Report. Value Health. 2005;8(5):521-33.
6. Sculpher MJ, Claxton K, Drummond M, McCabe C. Whither trial-based economic evaluation for health care decision making? Health Econ. 2006;15(7):677-87.
7. Saramago P, Manca A, Sutton AJ. Deriving Input Parameters for Cost-Effectiveness Modeling: Taxonomy of Data Types and Approaches to Their Statistical Synthesis. Value Health. 2012;15(5):639-49.
8. Buxton MJ, Drummond MF, Van Hout BA et al. Modelling in Ecomomic Evaluation: An Unavoidable Fact of Life. Health Econ. 1997;6(3):217-27.
9. Siebert U, Alagoz O, Bayoumi AM et al. State-Transition Modeling: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. Value Health. 2012;15(6):812-20.
10. Sonnenberg FA, Beck JR. Markov Models in Medical Decision Making. Med Decis Making. 1993;13(4):322-38.
11. Elbasha EH, Chhatwal J. Theoretical Foundations and Practical Applications of Within-Cycle Correction Methods. Med Decis Making. 2016;36(1):115-31.
12. Chhatwal J, Jayasuriya S, Elbasha EH. Changing Cycle Lengths in State-Transition Models. Med Decis Making. 2016;36(8):952-64.
13. Law AM. Simulation Modeling and Analysis. 4th ed. Singapore: McGraw-Hill Higher Education; 2007.
14. Karnon J, Stahl J, Brennan A, Caro JJ, Mar J, Möller J. Modeling using Discrete Event Simulation: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-4. Value Health. 2012;15(6):821-7.
15. Karnon J, Haji Ali Afzali H. When to use discrete event simulation (DES) for the economic evaluation of health technologies? A review and critique of the costs and benefits of DES. Pharmacoeconomics. 2014;32(6):547-58.
16. Karnon J. Alternative decision modelling techniques for the evaluation of health care technologies: Markov processes versus discrete event simulation. Health Econ. 2003;12(10):837-48.
17. Simpson KN, Strassburger A, Jones WJ, Dietz B, Rajagopalan R. Comparison of Markov Model and Discrete-Event Simulation Techniques for HIV. Pharmacoeconomics. 2009;27(2):159-65.
18. Standfield LB, Comans TA, Scuffham PA. An empirical comparison of Markov cohort modeling and discrete event simulation in a capacity-constrained health care setting. Eur J Health Econ. 2017;18(1):33-47.
19. Franken MD, van Rooijen EM, May AM et al. Cost-effectiveness of capecitabine and bevacizumab maintenance treatment after first-line induction treatment in metastatic colorectal cancer. Eur J Cancer. 2017;75:204-12.
20. Simkens LHJ, van Tinteren H, May A et al. Maintenance treatment with capecitabine and bevacizumab in metastatic colorectal cancer (CAIRO3): a phase 3 randomised controlled trial of the Dutch Colorectal Cancer Group. Lancet. 2015;385(9980):1843-52.

21. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB. Model Transparency and Validation: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. Value Health. 2012;15(6):843-50.

22. Vemer P, Corro Ramos I, van Voorn GAK, Al MJ, Feenstra TL. AdViSHE: A Validation-Assessment Tool of Health-Economic Models for Decision Makers and Model Users. Pharmacoeconomics. 2016;34:349-61.

23. Delignette-Muller ML, Dutang C. fitdistrplus: An R Package for Fitting Distributions. J Stat Softw. 2015;64(4):1-34.

24. R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/.

25. Barton P, Jobanputra P, Wilson J, Bryan S, Burls A. The use of modelling to evaluate new drugs for patients with a chronic condition: the case of antibodies against tumour necrosis factor in rheumatoid arthritis. Health Technol Assess. 2004;8(11):104.

26. National Health Care Institute. Richtlijn voor het uitvoeren van economische evaluaties in de gezondheidszorg. 2016. https://www.zorginstituutnederland.nl/over-ons/publicaties/publicatie/2016/02/29/richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg. Accessed 17 July 2017.

27. Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, Paltiel AD. Model Parameter Estimation and Uncertainty Analysis: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group–6. Med Decis Making. 2012;32(5):722-32.

28. Jahn B, Rochau U, Kurzthaler C et al. Lessons Learned from a Cross-Model Validation between a Discrete Event Simulation Model and a Cohort State-Transition Model for Personalized Breast Cancer Treatment. Med Decis Making. 2016;36(3):375-90.

29. Degeling K, IJzerman MJ, Koopman M, Koffijberg H. Accounting for parameter uncertainty in the definition of parametric distributions used to describe individual patient variation in health economic models. BMC Med Res Methodol. 2017;17(1):170.

30. Williams C, Lewsey JD, Mackay DF, Briggs AH. Estimation of Survival Probabilities for Use in Cost-effectiveness Analyses: A Comparison of a Multi-state Modeling Survival Analysis Approach with Partitioned Survival and Markov Decision-Analytic Modeling. Med Decis Making. 2017;37(4):427-39.

31. Caro JJ, Moller J. Advantages and disadvantages of discrete-event simulation for health economic analyses. Expert Rev Pharmacoecon Outcomes Res. 2016;16(3):327-9.

SUPPLEMENTARY MATERIAL

Parameter Values used for the Base-case Analysis

Parameter Name	Description	Parameter Value
c_Bevacizumab	Costs of Bevacizumab	2062,35
c_bleeding	Costs of Bleeding	42,62
c_bloodtransfusion	Costs of Bloodtransfusion	627,09
c_central_line	Costs of Central Line	224,03
c_coloscopy	Costs of Coloscopy	94,42
c_CTscan	Costs of CT Scan	229,00
c_diarrhea	Costs of Diarrhea	16,72
c_echo	Costs of Echo	87,66
c_embolic_event	Probability of Embolic Event	2,52
c_ERCP	Costs of ERCP	127,87
c_fatigue	Costs of Fatigue	19,95
c_GIperforation	Costs of GI Perforation	6392,00
c_home_administration	Costs of Home Administration	504,00
c_hospitalisation	Costs of Hospitalisation	504,00
c_hospitalisation_armA_PI	Costs of Hospitalisation	364,91
c_hospitalisation_armA_Re	Costs of Hospitalisation	378,54
c_hospitalisation_armA_Sal	Costs of Hospitalisation	690,01
c_hospitalisation_armB_PI	Costs of Hospitalisation	425,26
c_hospitalisation_armB_Re	Costs of Hospitalisation	360,02
c_hospitalisation_armB_Sal	Costs of Hospitalisation	612,58
c_hypersensitivity	Costs of Hypersensitivity	1,58
c_hypertension_medication	Costs of Hypertension Medication	0,46
c_IC_hospitalisation	Costs of IC Hospitalisation	2406,00
c_induction_B	Costs of Bevacizumab Induction	12374,10
c_induction_CAP	Costs of Capecitabine Induction	710,16
c_induction_OX	Costs of Oxaliplatin Induction	3363,12
c_KRAS_testing	Costs of KRAS Testing	973,82
c_laboratory	Costs of Laboratory	58,00
c_laparotomy	Costs of Laparotomy	6392,00
c_maintenance_CAP	Costs of Capecitabine Maintenance	102,48
c_MOvisit	Costs of MO Visit	184,00
c_MRCP	Costs of MRCP	79,00
c_MRI	Costs of MRI	251,02
c_nausea	Costs of Nausea	0,30
c_neuropathymed	Costs of Neuropathymedication	2,52
c_painmedication	Costs of Painmedication	8,14
c_paracentesis	Costs of Paracentesis	209,09
c_PETCT	Costs of PETCT Scan	1162,93
c_phone_consultation	Costs of Phone Consultation	14,00
c_physician_visit	Costs of Physician Visit	121,00
c_post_induction_maintenance	Accumulated Costs in Post-Induction	3416,66
c_post_induction_obs	Accumulated Costs in Post-Induction	624,23

Parameter Values used for the Base-case Analysis, (continued)

Parameter Name	Description	Parameter Value
c_prohylaxis_CAPOXB	Costs of Prohylaxis CAPOXB	4,34
c_prohylaxis_Sal	Costs of Prohylaxis Salvage	4,99
c_prophylaxis_induction	Costs of Prophylaxis Induction	26,04
c_radiotherapy	Costs of Radiotherapy	907,21
c_reinduction_CAP	Costs of Capecitabine	115,36
c_reinduction_FOL	Costs of Fluorouracil and Leucovorin	362,70
c_reinduction_FU	Costs of Fluorouracil	142,01
c_reinduction_heparine	Costs of Heparine	2,55
c_reinduction_IRI	Costs of Irinotecan	956,31
c_reinduction_maintenance	Accumulated Costs in Reintroduction	3098,06
c_reinduction_obs	Accumulated Costs in Reintroduction	3419,59
c_reinduction_OX	Costs of Oxaliplatin	560,52
c_reinduction_prophylaxis_FOLFIRI	Costs of Prophylaxis FOLFIRI	6,51
c_salvage	Costs of Salvage	1332,63
c_salvage_maintenance	Accumulated Costs in Salvage	2790,85
c_salvage_obs	Accumulated Costs in Salvage	2881,68
c_skeletscintigrafie	Costs of Skeletscintigrafie	282,03
c_Xray	Costs of Xray	46,37
dis_maintenance	Disutility for Maintenance Therapy	1,012
hospitalization_duration	Duration of Hospitalization	3
p_bleeding	Probability of Bleeding	0,0006
p_bloodtransfusion_armA_PI	Rate of Bloodtransfusion	0,0000
p_bloodtransfusion_armA_Re	Rate of Bloodtransfusion	0,0210
p_bloodtransfusion_armA_Sal	Rate of Bloodtransfusion	0,0391
p_bloodtransfusion_armB_PI	Rate of Bloodtransfusion	0,0147
p_bloodtransfusion_armB_Re	Rate of Bloodtransfusion	0,0094
p_bloodtransfusion_armB_Sal	Rate of Bloodtransfusion	0,0104
p_central_line_armA_PI	Probability of Central Line	0,0400
p_central_line_armA_Re	Probability of Central Line	0,0933
p_central_line_armA_Sal	Probability of Central Line	0,1867
p_central_line_armB_PI	Probability of Central Line	0,0986
p_central_line_armB_Re	Probability of Central Line	0,0704
p_central_line_armB_Sal	Probability of Central Line	0,0845
p_coloscopy_armA_PI	Rate of Colonoscopy	0,0043
p_coloscopy_armA_Re	Rate of Colonoscopy	0,0030
p_coloscopy_armA_Sal	Rate of Colonoscopy	0,0000
p_coloscopy_armB_PI	Rate of Colonoscopy	0,0163
p_coloscopy_armB_Re	Rate of Colonoscopy	0,0046
p_coloscopy_armB_Sal	Rate of Colonoscopy	0,0143
p_CTscan_armA_PI	Rate of CT Scan	0,2640
p_CTscan_armA_Re	Rate of CT Scan	0,3398
p_CTscan_armA_Sal	Rate of CT Scan	0,4466
p_CTscan_armB_PI	Rate of CT Scan	0,5774
p_CTscan_armB_Re	Rate of CT Scan	0,2388

Parameter Values used for the Base-case Analysis, (continued)

Parameter Name	Description	Parameter Value
p_CTscan_armB_Sal	Rate of CT Scan	0,2717
p_diarrhea_armB_PI	Probability of Diarrhea	0,0900
p_diarrhea_Re	Probability of Diarrhea	0,1700
p_echo_armA_PI	Rate of Echo	0,0150
p_echo_armA_Re	Rate of Echo	0,0271
p_echo_armA_Sal	Rate of Echo	0,0319
p_echo_armB_PI	Rate of Echo	0,0245
p_echo_armB_Re	Rate of Echo	0,0230
p_echo_armB_Sal	Rate of Echo	0,0429
p_embolic_event	Probability of Embolic Event	0,0400
p_ERCP_armA_PI	Rate of ERCP	0,0000
p_ERCP_armA_Re	Rate of ERCP	0,0000
p_ERCP_armA_Sal	Rate of ERCP	0,0000
p_ERCP_armB_PI	Rate of ERCP	0,0000
p_ERCP_armB_Re	Rate of ERCP	0,0000
p_ERCP_armB_Sal	Rate of ERCP	0,0143
p_fatigue	Probability of Fatigue	0,2500
p_GIperforation_armB_PI	Probability of GI Perforation	0,0017
p_GIperforation_Re	Probability of GI Perforation	0,0017
p_hypersensitivity	Probability of Hypersensitivity	0,1500
p_hypertension_armA_PI	Probability of Hypertension	0,1800
p_hypertension_armB_PI	Probability of Hypertension	0,2400
p_hypertension_Re	Probability of Hypertension	0,0210
p_IC_hospitalisation_armA_PI	Rate of Hospitalisation	0,0017
p_IC_hospitalisation_armA_Re	Rate of Hospitalisation	0,0039
p_IC_hospitalisation_armA_Sal	Rate of Hospitalisation	0,0000
p_IC_hospitalisation_armB_PI	Rate of Hospitalisation	0,0018
p_IC_hospitalisation_armB_Re	Rate of Hospitalisation	0,0000
p_IC_hospitalisation_armB_Sal	Rate of Hospitalisation	0,0000
p_KRAS_testing_Sal	Probability of KRAS Testing	0,3080
p_laboratory_armA_PI	Rate of Laboratory	0,6380
p_laboratory_armA_Re	Rate of Laboratory	1,1489
p_laboratory_armA_Sal	Rate of Laboratory	1,6057
p_laboratory_armB_PI	Rate of Laboratory	2,0347
p_laboratory_armB_Re	Rate of Laboratory	0,7210
p_laboratory_armB_Sal	Rate of Laboratory	1,6873
p_laparotomy_armA_PI	Probability of Laparotomy	0,0066
p_laparotomy_armB_PI	Probability of Laparotomy	0,0105
p_laparotomy_Re	Probability of Laparotomy	0,0064
p_laparotomy_Sal	Probability of Laparotomy	0,0060
p_MOvisit_armA_Sal	Rate of MOVisit	1,2035
p_MOvisit_armB_Sal	Rate of MOVisit	1,2456
p_MRCP_armA_PI	Rate of MRCP	0,0000
p_MRCP_armA_Re	Rate of MRCP	0,0000

Parameter Values used for the Base-case Analysis, (continued)

Parameter Name	Description	Parameter Value
p_MRCP_armA_Sal	Rate of MRCP	0,0000
p_MRCP_armB_PI	Rate of MRCP	0,0000
p_MRCP_armB_Re	Rate of MRCP	0,0000
p_MRCP_armB_Sal	Rate of MRCP	0,0000
p_MRI_armA_PI	Rate of MRI	0,0086
p_MRI_armA_Re	Rate of MRI	0,0120
p_MRI_armA_Sal	Rate of MRI	0,0232
p_MRI_armB_PI	Rate of MRI	0,0109
p_MRI_armB_Re	Rate of MRI	0,0046
p_MRI_armB_Sal	Rate of MRI	0,0000
p_nausea_vomiting_armB_PI	Probability of Nausea Vomiting	0,0300
p_nausea_vomiting_Re	Probability of Nausea Vomiting	0,1100
p_neuropathy_armA_PI	Probability of Neuropathy	0,0500
p_neuropathy_armB_PI	Probability of Neuropathy	0,1000
p_neuropathy_Re	Probability of Neuropathy	0,1800
p_painmedication_armA_PI	Probability of Painmedication	0,1330
p_painmedication_armB_PI	Probability of Painmedication	0,2390
p_painmedication_Re	Probability of Painmedication	0,2530
p_painmedication_Sal	Probability of Painmedication	0,3360
p_paracentesis_armA_PI	Rate of Paracentesis	0,0033
p_paracentesis_armA_Re	Rate of Paracentesis	0,0078
p_paracentesis_armA_Sal	Rate of Paracentesis	0,0674
p_paracentesis_armB_PI	Rate of Paracentesis	0,0053
p_paracentesis_armB_Re	Rate of Paracentesis	0,0170
p_paracentesis_armB_Sal	Rate of Paracentesis	0,0495
p_PETCT_armA_PI	Rate of PETCT	0,0000
p_PETCT_armA_Re	Rate of PETCT	0,0000
p_PETCT_armA_Sal	Rate of PETCT	0,0000
p_PETCT_armB_PI	Rate of PETCT	0,0082
p_PETCT_armB_Re	Rate of PETCT	0,0046
p_PETCT_armB_Sal	Rate of PETCT	0,0572
p_phone_consultation_armA_PI	Rate of Phone Consultation	0,1769
p_phone_consultation_armA_Re	Rate of Phone Consultation	0,3725
p_phone_consultation_armA_Sal	Rate of Phone Consultation	0,4507
p_phone_consultation_armB_PI	Rate of Phone Consultation	0,2931
p_phone_consultation_armB_Re	Rate of Phone Consultation	0,8307
p_phone_consultation_armB_Sal	Rate of Phone Consultation	0,5128
p_physician_visit_armA_PI	Rate of Physician Visit	0,7672
p_physician_visit_armA_Re	Rate of Physician Visit	1,0531
p_physician_visit_armA_Sal	Rate of Physician Visit	1,3125
p_physician_visit_armB_PI	Rate of Physician Visit	2,1228
p_physician_visit_armB_Re	Rate of Physician Visit	0,2332
p_physician_visit_armB_Sal	Rate of Physician Visit	1,4559
p_radiotherapy_armA_PI	Probability of Radiotherapy	0,0033

Parameter Values used for the Base-case Analysis, (continued)

Parameter Name	Description	Parameter Value
p_radiotherapy_armA_Re	Probability of Radiotherapy	0,0395
p_radiotherapy_armA_Sal	Probability of Radiotherapy	0,1827
p_radiotherapy_armB_PI	Probability of Radiotherapy	0,0053
p_radiotherapy_armB_Re	Probability of Radiotherapy	0,0049
p_radiotherapy_armB_Sal	Probability of Radiotherapy	0,1365
p_skeletscintigrafie_armA_PI	Rate of Skeletscintigrafie	0,0021
p_skeletscintigrafie_armA_Re	Rate of Skeletscintigrafie	0,0030
p_skeletscintigrafie_armA_Sal	Rate of Skeletscintigrafie	0,0000
p_skeletscintigrafie_armB_PI	Rate of Skeletscintigrafie	0,0109
p_skeletscintigrafie_armB_Re	Rate of Skeletscintigrafie	0,0046
p_skeletscintigrafie_armB_Sal	Rate of Skeletscintigrafie	0,0143
p_Xray_armA_PI	Rate of Xray	0,0601
p_Xray_armA_Re	Rate of Xray	0,0451
p_Xray_armA_Sal	Rate of Xray	0,2871
p_Xray_armB_PI	Rate of Xray	0,0518
p_Xray_armB_Re	Rate of Xray	0,0276
p_Xray_armB_Sal	Rate of Xray	0,0858
p_KRAS_testing_Re	Probability of KRAS Testing	0,3080
u_post_induction_maintenance	Utility during Post-Induction	0,048
u_post_induction_obs	Utility during Post-Induction	0,048
u_reinduction	Utility during Reintroduction	0,042
u_salvage	Utility during Salvage	0,040

5

EVALUATION OF THE PERFORMANCE OF ALGORITHMS MAPPING EORTC QLQ-C30 ONTO THE EQ-5D INDEX IN A METASTATIC COLORECTAL CANCER COST-EFFECTIVENESS MODEL

Mira D. Franken

Anne de Hond

Koen Degeling

Cornelis J.A. Punt

Miriam Koopman

Carin A. Uyl-de Groot

Matthijs Versteegh*

Martijn G.H. van Oijen*

* These authors contributed equally to this work.

Health and Quality of Life outcomes, 2020, 18: 240

ABSTRACT

Background

Cost-effectiveness models require quality of life utilities calculated from generic preference-based questionnaires, such as EQ-5D. We evaluated the performance of available algorithms for QLQ-C30 conversion into EQ-5D-3L based utilities in a metastatic colorectal cancer (mCRC) patient population and subsequently developed a mCRC specific algorithm. Influence of mapping on cost-effectiveness was evaluated.

Methods

Three available algorithms were compared with observed utilities from the CAIRO3 study. Six models were developed using 5-fold cross-validation: predicting EQ-5D-3L tariffs from QLQ-C30 functional scale scores, continuous QLQ-C30 scores or dummy levels with a random effects model (RE), a most likely probability method on EQ-5D-3L functional scale scores, a beta regression model on QLQ-C30 functional scale scores and a separate equations subgroup approach on QLQ-C30 functional scale scores. Performance was assessed, and algorithms were tested on incomplete QLQ-C30 questionnaires. Influence of utility mapping on incremental cost/QALY gained (ICER) was evaluated in an existing Dutch mCRC cost-effectiveness model.

Results

The available algorithms yielded mean utilities of 1: 0.87±sd:0.14,2: 0.81±0.15 (both Dutch tariff) and 3: 0.81±sd:0.19. Algorithm 1 and 3 were significantly different from the mean observed utility (0.83± 0.17 with Dutch tariff, 0.80±0.20 with U.K. tariff). All new models yielded predicted utilities drawing close to observed utilities; differences were not statistically significant. The existing algorithms resulted in an ICER difference of €10,140 less and €1,765 more compared to the observed EQ-5D-3L based ICER (€168,048). The preferred newly developed algorithm was €5,094 higher than the observed EQ-5D-3L based ICER. Disparity was explained by minimal differences in incremental QALYs between models.

Conclusion

Available mapping algorithms sufficiently accurately predict utilities. With the commonly used statistical methods, we did not succeed in developing an improved mapping algorithm. Importantly, cost-effectiveness outcomes in this study were comparable to the original model outcomes between different mapping algorithms. Therefore, mapping can be an adequate solution for cost-effectiveness studies using either a previously designed and validated algorithm or an algorithm developed in this study.

BACKGROUND

Measurement of health-related quality of life (HRQoL) with generic questionnaires (e.g. EQ-5D-3L) and disease specific questionnaires (e.g. EORTC QLQ-C30) are of great interest to clinicians and researchers, especially in the context of cost-effectiveness research. In oncology, cost-effectiveness research becomes more important rapidly, as it provides information for decision-makers in establishing the content of the basic benefit package of a health insurance in some countries. Cost-effectiveness outcomes are more often reported in addition to clinical outcome parameters, and the incremental cost per quality adjusted life year (QALY) is generally chosen as primary outcome in cost-effectiveness models.[1] To calculate the total QALYs gained due to treatment, both length and quality of life have to be established. Quality of life can be measured through a generic preference-based quality of life questionnaire such as the commonly used EQ-5D-3L questionnaire, which is requested by some reimbursement authorities [2]. Based on this questionnaire, patient scores are transformed into health-related quality of life utilities, on a scale of 1 - being full health- to 0 - reflecting death (and even negative values reflecting health states worse than death), which can be combined with the duration (survival) of a patient to calculate the QALY [1,3].

In industry sponsored oncology studies, both the EORTC QLQ-C30 and the EQ-5D questionnaires are often used to capture clinically meaningful changes in quality of life and enable health-economic evaluations [2,4]. However, the lack of generic preference-based questionnaires in for instance academic clinical studies or clinical registries hamper the calculation of health-related quality of life utilities for cost-effectiveness research. To overcome this issue, researchers often revert to the translation of disease specific quality of life outcomes (such as those captured by QLQ-C30 in oncology) into utilities (such as captured by EQ-5D-3L) using so called 'mapping algorithms' for their cost-effectiveness models. Mapping algorithms are regression models developed and tested in specific patient population datasets, which make them 'sample dependent'. Consequently, Doble *et al.* [5] demonstrated that in oncology only two out of 10 eligible mapping algorithms, performed sufficiently well in the estimation of utilities (Versteegh *et al.* using a Dutch tariff for EQ-5D-3L, developed in a multiple myeloma and non-Hodgkin lymphoma dataset, and Longworth *et al.* for EQ-5D-3L, developed in a multiple myeloma and breast cancer dataset) [5–7]. As shown by Doble *et al.*, QLQ-C30 outcomes between development and validation datasets demonstrated clinically relevant differences on multiple QLQ-C30 dimensions, although congruence of QLQ-C30 outcomes between datasets was not predictive for mapping algorithm performance [5]. Even so, disease related effects could influence the outcomes of mapping algorithms and it has been previously advised to use a mapping algorithm with similar clinical characteristics compared to the sample on which the mapping is to be applied [8]. More recently, Marriott *et al.* proposed a mapping algorithm developed with a metastatic colorectal cancer (mCRC) patient dataset using an

U.K. tariff for EQ-5D-3L [9]. Even so, we question whether the currently available mapping algorithms, which were not all developed with mCRC datasets and an mCRC disease specific algorithm based on a U.K. tariff, are sufficiently suitable to translate QLQ-C30 outcomes to Dutch EQ-5D-3L based utilities for mCRC patients.

Our first objective was to evaluate the accuracy of available mapping algorithms for conversion of QLQ-C30 outcomes to EQ-5D-3L utilities in a population of mCRC patients. Our second objective was to design an mCRC specific mapping algorithm using a Dutch tariff for the conversion of QLQ-C30 outcomes to EQ-5D-3L based utilities. Finally, we evaluated the influence of utility mapping on the incremental cost per QALY gained (ICER) in an existing mCRC cost-effectiveness model [10].

METHODS

Patient population

The CAIRO3 study is a randomized phase 3 study (NCT00442637) sponsored by the Dutch Colorectal Cancer Group (DCCG), in which mCRC patients with stable disease or better (n=558) following six cycles of initial therapy with capecitabine, oxaliplatin and bevacizumab (CAPOX-B). Patients were either randomized to the observation strategy or capecitabine (625 mg/m^2 orally twice daily continuously) and bevacizumab (7.5mg/kg intravenously every 3 weeks) (CB) maintenance treatment [11]. Patients completed both the disease specific QLQ-C30 version 3.0 and generic EQ-5D-3L questionnaires every 9 weeks simultaneously [2,4]. Only patients participating in the completion of QLQ-C30 and EQ-5D questionnaires were selected and all time points were pooled for this study. Descriptive statistics were used for baseline characteristics.

Questionnaires

The EORTC QLQ-C30 questionnaire version 3.0 comprises 30 questions evaluating quality of life in five functional scales (physical, role, cognitive, emotional and social functioning), three symptom scales (fatigue, pain, nausea and vomiting), global health status and single items for the assessment of symptoms commonly reported by cancer patients (dyspnea, appetite loss, insomnia, constipation, diarrhea and financial difficulties) [4]. QLQ-C30 outcomes were calculated using the EORTC QLQ-C30 scoring manual. After linear transformation and calculation of raw score for the questions ranging not at all (0) to very much (4) for functional and symptom scale scores and very poor (0) to excellent (7) for global health, scale scores range 0 to 100. For functional scales and global health, a high score represents a higher level of functioning, while for the symptoms scales a low outcome represents less symptomatology [12].

The EQ-5D-3L contains 5 questions each addressing a different domain: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each of these domains has 3 levels [2]. An EQ-5D-3L based utility is derived from an EQ-5D questionnaire using

a country specific value set, i.e. tariff. EQ-5D-3L outcomes in this study were transformed to Dutch and U.K. tariff EQ-5D-3L -based utilities [13,14].

Evaluation of existing algorithms

The algorithms by Versteegh *et al*. and Longworth *et al*. were initially selected as these performed best in the analysis by Doble and Lorgelly, and is appropriate to the Dutch setting as both can predict Dutch tariff EQ-5D-3L utilities [5,6]. The mapping algorithm by Marriott *et al*. was additionally selected as this algorithm was developed in an mCRC patient dataset appropriate to a U.K. setting [8]. All three mapping algorithms were used for prediction of an EQ-5D-3L based utility using concurrently collected EORTC QLQ-C30 outcomes. As the algorithm by Versteegh *et al*. was based on version 2 of the QLQ-C30 questionnaire, while version 3 was used in the CAIRO3 trial, QLQ-C30 question 1 through 5 were converted into a binary response to fit the mapping algorithm. All algorithms were developed for non-patient level modelling purposes and the performance analysis is therefore focused on their sample means. Some individual level performance characteristics were also used for the mapping algorithms, albeit the well documented suboptimal performance of these algorithms on the individual level in the lower utility ranges. The algorithms were compared to the observed EQ-5D-3L based utilities using the root mean square error (RMSE), mean absolute error (MAE), t-test and Spearman correlation. The data was formatted in STATA. All analyses were performed using R.

Mapping algorithm design

Methodology according to the MAPS statement was used for developing the mapping algorithm [15]. The mCRC specific mapping algorithms that were developed with commonly used statistical methods and evaluated used 5-fold cross-validation.

Each fold provided a test set in which the trained model, which was developed based on the other 4 folds, could be tested, resulting in 5 estimates for each performance measure.

First, the EQ-5D-3L based utility was regressed on the QLQ-C30 functional and symptom scale scores using a random effects model (RE) with a random intercept: model 1. In a second RE model (model 2), the QLQ-C30 questions were treated as continuous variables and in a third model as dummy variables (model 3). Dummy variables essentially are a redefinition of the four QLQ-C30 answer categories (categories: 1 (no problem at all) to 4 (very much a problem)) and seven categories (categories: 1 (very poor) to 7 (excellent)) for the last two QLQ-C30 questions. For each QLQ-C30 question dummies for outcome categories were regressed on utility prediction. All abovementioned RE models assume a continuous and normal distribution for EQ-5D utilities. Although this assumption is hardly realistic considering the well-studied skewed distribution of utilities, it is by far the most popular form of mapping in the literature and generally performs quite well compared to more complex models [16].

Model 4 is a two-step model, also known as a response mapping model. The advantage of a response mapping model is that it is independent of tariff calculations and it can therefore compute any country utility score for which tariffs are available. First, in model 4, ordered logit regression was used to predict the EQ-5D-3L domain score. An ordered logit model was chosen to preserve the ordering of the categories in the dependent variable.* For this method, input variables were the QLQ-C30 functional scale scores. Secondly, a utility was calculated using the most likely probability method. With the most likely probability method, the probabilities of the EQ-5D-3L response levels (no problem, some problems and severe problems) per EQ-5D domain (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) were predicted based on the QLQ-C30 functional scale scores. The following formulas were used for this:

$$\boldsymbol{Prob1_{level_1}} = \frac{\mathbf{1}}{\mathbf{1}+e^{EQ5D}}$$

$$\boldsymbol{Prob2_{level2}} = \frac{\mathbf{1}}{\mathbf{1}+e^{EQ5D-\kappa}} - \frac{\mathbf{1}}{\mathbf{1}+e^{EQ5D}}$$

$$\boldsymbol{Prob3_{level_3}} = \mathbf{1} - \boldsymbol{Prob1_{level1}} - \boldsymbol{Prob2_{level2}}$$

Where level stands for the EQ-5D-3L response level, EQ5D stands for the latent EQ-5D functional or symptom scale score regressed on the QLQ dimensions, stands for the estimated threshold between different response levels. These predicted probabilities were subsequently scored with the EQ-5D scoring system [17].

Model 5 used beta regression to restrict the EQ-5D-3L utilities to the 0,1 interval. The advantage of this method is that it cannot lead to unrealistic utility predictions exceeding 1. However, it will not be able to produce negative utilities. In the current analyses, the number of individuals with negative utilities was so small (0.2 %) that this is unlikely to notably affect the results. Moreover, it cannot model values of exactly 1 or 0, so these values were rescaled prior to the mapping. All utilities were first transformed to disutilities. All values ≥1 (which were utilities of 0 or less than 0) were selected to be approximated so that the disutilities would return a value <1 and thus included in the beta regression. To do so, a standardized value was subtracted from the disutility. All values of exactly 0 (which were utilities of 1) were selected to be adapted so that the disutilities

* A multinomial logit model was also developed; however the ordered logit model outperformed the multinomial logit model. Hence, we only report on the ordered logit model in this manuscript.

would return values >0. The standardized transformation applied was: (disutility*(N-1)+0.5)/N. Nevertheless, the beta distribution is in theory a better approximation of the EQ-5D utility distribution compared to the normal distribution underlying OLS regression, at least in samples with very few health state observations worse than dead. This regression was also conducted on the QLQ-C30 functional scale scores.

The final model (model 6) consisted of a separate equations subgroup approach. In the first step, probabilities are calculated on the basis of a multinomial logistic regression for having a EQ-5D-3L utility score lower than 0.6 (related to scoring 'extreme problems' on any EQ-5D-3L dimension [18], higher than 0.6 but lower than 1 and equal to 1. In the next step, RE models are trained on individuals with utility scores lower than 0.6 and higher than 0.6 separately. Finally, the predicted utilities of these two sub-models and of having a 1 are combined with the probabilities from the first step. The advantage of this approach is that it relaxes the assumption of a continuous linear relation between EQ-5D utilities and QLQ-C30 functional and symptom scale scores. Poor health states often adhere to a different (approximate) linear relation with the EQ-5D utilities compared to higher scores, often leading to the overvaluing of low health states in the literature [18].

All models were developed using a backward selection procedure, where non-significant coefficients based on the QLQ-C30 items were removed one-by-one (cut-off value p=0.05) until all coefficients were at or below the cut-off value. Except for model 4 and 6 (in part), backward selection was performed to minimize the mapping algorithm length without compromising the model performance, which has previously been done by others [6,7]. In a second step, non-logical coefficients were removed. Non-logical coefficients were defined as coefficients that carried an incongruous sign, for example a coefficient for nausea leading to a better utility when one would expect a reduction in the assigned utility. Random effects with cluster robust standard errors were introduced to correct for multiple responses from one patient for all OLS models (models 1, 2, 3, and 6 in part). The beta, ordered logit and multinomial logit regressions (models 4, 5 and 6 in part) used normal standard errors as there were no cluster robust standard errors available for these methods.

Validation of the developed mapping algorithms

After development of the six mapping algorithms using each of the five training data sets consecutively, the algorithms were tested in the corresponding folds. Performance of the algorithms was reported as mean predicted utility, the root mean squared error (RMSE) and mean absolute error (MAE). The RMSE will give a better insight into the performance of the mapping algorithm alongside MAE, as it is more sensitive to outliers and hence helps identify the mapping algorithm with the least extreme deviations between predicted and observed values. The resulting algorithms were analyzed for logical consistency using scatter plots comparing observed and predicted utilities, i.e.

worse outcomes of the observed EQ-5D-3L based utility also lead to worse outcomes in the predicted utilities with the six methods described above. Lastly, Spearman correlation coefficients and t-tests were used to illustrate the performance of the various algorithms. The model of preference was selected based on best fit: smallest value for RMSE, MAE and highest value for the Spearman correlation.

Performance of the mapping algorithms based on QLQ-C30 functional scale scores, developed with OLS, response mapping, beta regression and the separate equations model, were tested on incomplete QLQ-C30 questionnaires. Quality of life functional scale scores (e.g. physical functioning) can be calculated with a minimal completion of half of the questions included in the QLQ-C30 questionnaires [12]. Incomplete questionnaires, for which functional scale scores calculations remained possible and with a concurrently collected EQ-5D-3L, were selected to test mapping algorithm performance with those algorithms based on functional scale scores. No imputations were performed on QLQ-C30 questionnaires. Results were compared with concurrently collected EQ-5D-3L questionnaires. Outcomes were compared with observed utilities as previously described.

Algorithm influence on cost-effectiveness model outcomes

The influence of the mapping algorithms on the primary outcome, the incremental cost per QALY gained (ICER), was evaluated using a Dutch cost-effectiveness model comparing CB maintenance and observation following 6 cycles of first line CAPOX-B for patients with mCRC. For this purpose, a discrete event simulation model, developed in AnyLogic (multi-method simulation software, v.8.2.3, The AnyLogic Company (Chicago, IL, USA) was used for the current analysis [19]. ICERs comparing CB maintenance and observation were calculated for 1) observed EQ-5D-3L based utilities as was done in the original study, 2) utilities obtained with the mapping algorithm developed by Versteegh *et al*.[6] (mapping algorithm for a Dutch tariff conversion), 3) utilities obtained with the mapping algorithm developed by Longworth *et al*. using a Dutch tariff and 4) utilities obtained with the preferred mapping algorithm developed in this study (model 1). The mapping algorithm developed by Marriott *et al*.[9] uses a U.K. tariff conversion and was therefore not included. Only concurrently collected EQ-5D and QLQ-C30 observations during either maintenance treatment and observation, defined as the first health-state, were used in this analysis. Utilities in subsequent health-states (re-introduction of therapy, salvage therapy, death) were derived from literature as these could not be derived from the CAIRO3 study [10].

A total of 10,000 hypothetical patients per treatment strategy were simulated for a patient-level outcome calculation. Subsequently, a probabilistic analysis was performed to calculate the ICERs with a 95% confidence interval based on 10,000 samples. To reflect parameter uncertainty in the probabilistic analysis, distributions for the utilities were defined according to the method of moments using the mean and a standard

error for each of the utilities derived from the selected mapping algorithms in line with the original cost-effectiveness evaluation of the CAIRO3 study. With the exception of the uncertainty around utilities only, distributions for the other parameters, such as costs, health-state transitions, were defined as in the original cost-effectiveness evaluation of the CAIRO3 study [10].

RESULTS

From a total of 2440 observations, 1905 concurrently collected, complete QLQ-C30 and EQ-5D-3L questionnaires were included in this analysis. The concurrent observations were obtained from 473 patients enrolled in the CAIRO3 study (238 patients in the observation arm and 235 patients in the maintenance treatment arm). In table 1, characteristics of the QLQ-C30 and EQ-5D data set are presented. The distribution of EQ-5D based utilities can be viewed in Additional File 1. Incomplete QLQ-C30 or EQ-5D-3L questionnaires were excluded for mapping algorithm development. For the purpose of the mCRC specific mapping algorithm design, we randomly divided the data in 5 folds (n=381 each).

Performance of existing mapping algorithms on an mCRC dataset

The mean observed utility based on completed EQ-5D-3L questionnaires of the mCRC dataset included in this analysis was 0.834±sd: 0.171 (Dutch tariff) and 0.803±sd: 0.197 (U.K. tariff). The algorithm by Versteegh *et al*. resulted in a mean utility of 0.866±0.135 with a Spearman correlation of 0.76 ($p<0.01$) (Table 2). The algorithm by Longworth *et al*. resulted in a mean utility of 0.835±0.127 and 0.810±0.152, with a Spearman correlation of 0.77 and 0.79, for the Dutch tariff and the U.K. tariff respectively. The algorithm by Longworth for Dutch tariff performed very well and was not significantly different compared to observed utilities ($p=0.687$). The algorithm by Marriott *et al*. (U.K. tariff) resulted in a mean utility of 0.813±sd:0.185 with a Spearman correlation of 0.75 ($p<0.01$) (Table 2).

Design and validation of a new mapping algorithm on a mCRC dataset

Algorithm coefficients for the RE based algorithms are presented in Table 3 (model 1), 4 (model 2) and 5 (model 3). These algorithms concern the RE model with QLQ-C30 functional scale scores (model 1), RE model with QLQ-C30 question outcomes as continuous variable (model 2) and RE model with the QLQ-C30 questions as dummy variables (model 3). The ordered logit regressions for prediction of the EQ-5D-3L based utility (model 4) can be viewed in the Additional file 2: Tables 1-3. The beta regression (model 5) output can be found in Table 6 and the separate equations subgroup approach model (model 6) in Additional file 2 Tables 4-6.

Observed and mean predicted utility resulting from the six developed mapping algorithms are presented in Table 7. The mean observed utility was 0.834±0.171,

Table 1. Patient characteristics for concurrently collected EQ-5D and QLQ-C30 questionnaires

		Complete N = 1905
Age (years)		64 (8.4)
Male gender (%)		69
EQ-5D-3L*	N	1905
	Mobility 1/2/3 (%)	57.9/41.8/0.3
	Self-cae 1/2/3 (%)	93.4/6.1/0.4
	Usual activities 1/2/3 (%)	57.5/38.5/3.9
	Pain/discomfort 1/2/3 (%)	60.2/38.4/1.4
	Depression/anxiety 1/2/3 (%)	77.2/21.8/1
	EQ-5D utility, mean (SD)	0.834 (0.171)
	EQ-5D range	-0.134 to 1
QLQ-C30 v.3.0	Questionnaires, N	1905
	Physical functioning, mean (SD)	82.681 (17.195)
	Role functioning, mean (SD)	76.947 (24.218)
	Emotional functioning, mean (SD)	85.744 (15.829)
	Cognitive functioning, mean (SD)	89.221 (15.294)
	Social functioning, mean (SD)	86.177 (18.718)
	Global health, mean (SD)	74.711 (17.464)
	Fatigue, mean (SD)	24.205 (20.059)
	Nausea/vomiting, mean (SD)	4.234 (11.286)
	Pain, mean (SD)	13.508 (20.705)
	Dyspnea, mean (SD)	10.866 (19.061)
	Insomnia, mean (SD)	15.083 (22.297)
	Appetite, mean (SD)	9.729 (19.651)
	Constipation, mean (SD)	6.824 (15.917)
	Diarrhea, mean (SD)	10.569 (19.363)
	Financial difficulties, mean (SD)	6.229 (15.978)
Concurrent EQ-5D and incomplete QLQ-C30 with retainment of functional scale scores, N**		120

* Percentages at level 1, 2 and 3 represent no problems at all, some problems and extreme problems, respectively.
** Patient characteristics for concurrently collected incomplete QLQ-C30 questionnaires available in Online Resource file 3.

while the mean predicted utilities for model 1 to 6 were nearly identical, 0.832±0.134, 0.832±0.134, 0.833±0.133, 0.830±0.145, 0.838±0.156 and 0.834±0.138, respectively. A utility prediction drawing close to the observed utility was achieved in all models. Differences between observed and predicted utilities were non-significant. The lowest RMSE and MAE was achieved by model 1 (RMSE 0.098, MAE 0.072) and model 4 (RMSE 0.098, MAE 0.072). Note that comparable to the Longworth algorithm, model 4 is an algorithm for EQ-5D response prediction and is thus independent of country tariff. For the purpose of comparison between model performance, a Dutch tariff was applied to the Longworth algorithm and model 4. Mapping algorithms based on functional scale scores are more forgiving towards incomplete questionnaires, as quality of life functional scale scores (e.g.

Table 2. Utility, observed and predicted, for all patients with complete questionnaires (n = 1905)

	Mean utility	SD	Min.	Max.	RMSE	MAE	Spearman correlation	p-value
Observed utility (Dutch tariff)	0.834	0.171	-0.134	1	-	-	-	-
Observed utility (U.K. tariff)	0.803	0.197	-0.239	1	-	-	-	-
Predicted utility (Versteegh [6])	0.866	0.135	-0.298	0.978	0.113	0.080	0.76	<0.001*
Predicted utility (Longworth[7] (Dutch tariff))	0.835	0.127	-0.088	0.959	0.106	0.078	0.77	0.687*
Predicted utility (Longworth[7] (U.K. tariff))	0.810	0.152	-0.307	0.955	0.114	0.085	0.79	0.026**
Predicted utility (Marriott [9])	0.813	0.185	-0.159	1.061	0.122	0.089	0.75	0.001**

*p-value tested against Dutch tariff ** p-value tested against U.K. tariff; p-values result from a t-test.

Table 3. Regression results for model 1: EQ-5D-3L based utility values on QLQ-C30 domain scores

Variable	Coefficient (SD)	t-value	p-value	95% CI
Constant	0.2993 (0.027)	10.940	<0.001	[0.246, 0.353]
Physical functioning	0.0021 (0.000)	7.949	<0.001	[0.002, 0.003]
Role functioning	0.0011 (0.000)	5.738	<0.001	[0.001, 0.001]
Emotional functioning	0.0025 (0.000)	10.901	<0.001	[0.002, 0.003]
Cognitive functioning	0.0005 (0.000)	2.279	0.023	[0.000, 0.001]
Social functioning	0.0006 (0.000)	2.814	0.005	[0.000, 0.001]
Symptom scale: Pain	-0.0023 (0.000)	-13.519	<0.001	[-0.003, -0.002]
Symptom scale: Insomnia	-0.0005 (0.000)	-3.166	0.002	[-0.001, 0.000]
Symptom scale nausea and vomiting was removed as non-logical coefficient				

p-values result from a t-test.

physical functioning) can be calculated with a minimal completion of half of the questions included in the QLQ-C30 questionnaires. Performance of all newly developed mapping algorithms using QLQ-C30 functional scale scores (model 1, 4, 5 and 6), were additionally tested in incomplete QLQ-C30 questionnaires for which functional scale scores could still be calculated for which EQ-5D outcomes were concurrently available (n=120). Patient characteristics of incomplete questionnaires are presented in Additional file 3. The mean observed utility in 120 incomplete QLQ-C30 questionnaires was 0.760±0232. The best predicted mean utilities were 0.767±0.177, 0.756±0.222, 0.764±0.222, for model 1, model 4 and model 5 respectively (Table 8). The lowest RMSE an MAE were achieved for model 1, which was chosen as preferred model. The algorithm based on the QLQ-C30 functional scale scores (preferred model) was regarded effective based on correlation between observed and mapped utilities (Figure 1).

Figures depicting the error of predicted utilities compared to the observed utilities for each algorithm are available in the Additional file 4: Figures 2 and 3. As is well documented

Table 4. Regression results for model 2: EQ-5D-3L based utility values QLQ-C30 questions as continuous variables

Variable	Coefficient (SD)	t-value	p-value	95% CI
Constant	1.340 (0.015)	86.755	<0.001	[1.310, 1.370]
QLQ3	-0.031 (0.006)	-5.553	<0.001	[-0.042, -0.020]
QLQ5	-0.077 (0.011)	-7.056	<0.001	[-0.098, -0.055]
QLQ6	-0.048 (0.005)	-9.660	<0.001	[-0.057, -0.038]
QLQ9	-0.053 (0.006)	-9.305	<0.001	[-0.064, -0.042]
QLQ11	-0.018 (0.005)	-3.686	<0.001	[-0.027, -0.008]
QLQ19	-0.021 (0.007)	-3.150	0.002	[-0.033, -0.008]
QLQ22	-0.021 (0.005)	-4.126	<0.001	[-0.031, -0.011]
QLQ23	-0.025 (0.006)	-4.010	<0.001	[-0.038, -0.013]
QLQ24	-0.040 (0.007)	-6.113	<0.001	[-0.053, -0.027]
QLQ26	-0.026 (0.006)	-4.546	<0.001	[-0.037, -0.015]
QLQ28	-0.012 (0.006)	-1.913	0.056	[-0.025, 0.000]
QLQ15 (vomiting) was removed as non-logical coefficient				

p-values result from a t-test.

in the literature [18], all mapping algorithms show overestimation of lower utilities and underestimation of high utilities.

Algorithm influence on ICERs in a mCRC cost-effectiveness model

The influence of the mapping algorithms on the ICER, was tested in an existing Dutch cost-effectiveness model comparing two different treatment strategies (CB maintenance versus observation following 6 cycles of first line CAPOX-B) in an mCRC patient population. For the first health state in this cost-effectiveness model, utilities were estimated using a total of 1654 observations (709 observations for 223 patients in the observation arm and 945 observations for 225 patients in the maintenance arm), utilities of subsequent health states (first progression and theirafter) were derived from literature as was done in the original cost-effectiveness study. The ICERs presented in Table 9 were obtained with 1) observed EQ-5D-3L based utilities, 2) utilities obtained with the mapping algorithm developed by Versteegh *et al.*, 3) utilities obtained with the mapping algorithm developed by Longworth *et al* using a Dutch tariff and 4) utilities obtained with the preferred model 1. The calculated ICER based on observed utilities in this analysis was €168,048/QALY. Previously developped mapping algorithm by Versteegh *et al*. compared to the observed EQ-5D-3L based utility lead to a negative ICER difference in the point estimate of €10,140 per QALY gained, while a positive difference of €5,094 and €1,765 was shown for the preferred algorithm (model 1) and the Longworth algorithm, respectively (Figure 2).

Table 5. Regression results for model 3: EQ-5D-3L based utilities on QLQ-C30 questions as dummy variables

Variable	Coefficient (SD)	t-value	p-value	95% CI
Constant	0.966 (0.006)	158.046	<0.001	[0.954, 0.978]
QLQ1_quite a bit	-0.020 (0.009)	-2.142	0.032	[-0.038, -0.002]
QLQ2_a little	-0.014 (0.006)	-2.324	0.020	[-0.026, -0.002]
QLQ3_a little	-0.028 (0.008)	-3.670	<0.001	[-0.043, -0.013]
QLQ3_quite a bit	-0.046 (0.014)	-3.231	0.001	[-0.074, -0.018]
QLQ5_ a little	-0.065 (0.013)	-4.903	<0.001	[-0.091, -0.039]
QLQ5_ quite a bit	-0.225 (0.037)	-6.097	<0.001	[-0.297, -0.153]
QLQ6_ a little	-0.047 (0.007)	-6.978	<0.001	[-0.06, -0.034]
QLQ6_ quite a bit	-0.076 (0.011)	-6.780	<0.001	[-0.098, -0.054]
QLQ6_ very much	-0.259 (0.020)	-13.215	<0.001	[-0.297, -0.22]
QLQ9_ a little	-0.068 (0.007)	-10.468	<0.001	[-0.081, -0.055]
QLQ9_quite a bit	-0.099 (0.012)	-8.327	<0.001	[-0.123, -0.076]
QLQ9_very much	-0.168 (0.023)	-7.383	<0.001	[-0.213, -0.123]
QLQ18_very much	-0.058 (0.020)	-2.933	0.003	[-0.097, -0.019]
QLQ19_quite a bit	-0.062 (0.014)	-4.475	<0.001	[-0.089, -0.035]
QLQ22_a little	-0.027 (0.006)	-4.251	<0.001	[-0.039, -0.014]
QLQ22_quite a bit	-0.046 (0.012)	-3.733	<0.001	[-0.07, -0.022]
QLQ23_quite a bit	-0.060 (0.018)	-3.294	0.001	[-0.096, -0.024]
QLQ23_ very much	-0.211 (0.049)	-4.338	<0.001	[-0.306, -0.115]
QLQ24_a little	-0.045 (0.007)	-6.117	<0.001	[-0.059, -0.03]
QLQ24_quite a bit	-0.108 (0.020)	-5.264	<0.001	[-0.148, -0.068]
QLQ24_very much	-0.213 (0.035)	-6.039	<0.001	[-0.283, -0.144]
QLQ26_a little	-0.033 (0.007)	-4.874	<0.001	[-0.047, -0.02]
QLQ26_quite a bit	-0.068 (0.016)	-4.385	<0.001	[-0.099, -0.038]
QLQ28_ very much	-0.056 (0.022)	-2.603	0.009	[-0.098, -0.014]

p-values result from a t-test.

Table 6. Beta regression results for model 5: EQ-5D-3L based disutility values on QLQ-C30 domain scores

Variable	Coefficient (SD)	t-value	p-value	95% CI
Constant	2.081 (0.210)	9.898	<0.001	[1.669,2.493]
Global health	-0.004 (0.002)	-1.892	0.058	[-0.007,0]
Physical functioning	-0.018 (0.002)	-8.269	<0.001	[-0.022,-0.014]
Role functioning	-0.010 (0.002)	-6.125	<0.001	[-0.013,-0.007]
Emotional functioning	-0.015 (0.002)	-8.479	<0.001	[-0.019,-0.012]
Cognitive functioning	-0.005 (0.002)	-3.063	0.002	[-0.009,-0.002]
Symptom scale: pain	0.014 (0.001)	9.939	<0.001	[0.011,0.017]
Symptom scale: insomnia	0.005 (0.001)	4.119	<0.001	[0.002,0.007]
Symptom scale: financial	0.007 (0.001)	4.448	<0.001	[0.004,0.009]
Symptom scale nausea and vomiting was removed as non-logical coefficient				

p-values result from a t-test.

Table 7. Mean, standard deviation, minimum and maximum of utility values, RMSE and MAE for the predicted utilities (p-values result from a t-test)

	Observed utility NL	Model 1 RE functional scale scores*	Model 2 RE: continuous	Model 3 RE: dummy variables	Model 4 Ordered logit	Model 5 Beta regression	Model 6 Separate equations
w	0.834	0.832	0.832	0.833	0.830	0.838	0.834
St.Dev	0.171	0.134	0.134	0.133	0.145	0.156	0.138
Min	-0.134	0.069	-0.055	0.057	-0.206	0.034	0.178
Max	1	0.982	0.969	0.966	0.975	0.959	0.990
RMSE	-	0.098	0.098	0.103	0.098	0.106	0.100
MAE	-	0.072	0.075	0.077	0.072	0.081	0.070
Spearman correlation	-	0.781	0.780	0.779	0.786	0.774	0.787
p-value	-	0.564	0.501	0.615	0.125	0.108	0.943

p-values result from a t-test; * Preferred model.

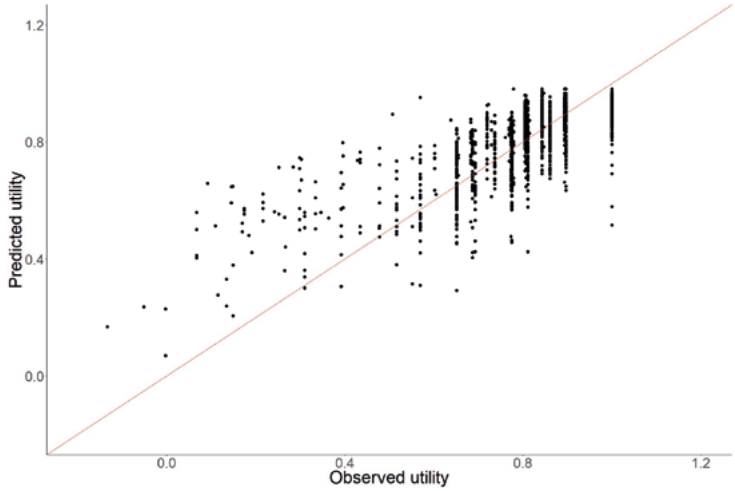


Figure 1. Correlation of observed versus predicted utility for model 1. Observed utility values were based on the EQ-5D-3L questionnaire and regressed on the QLQ-C30 functional and symptom scale scores.

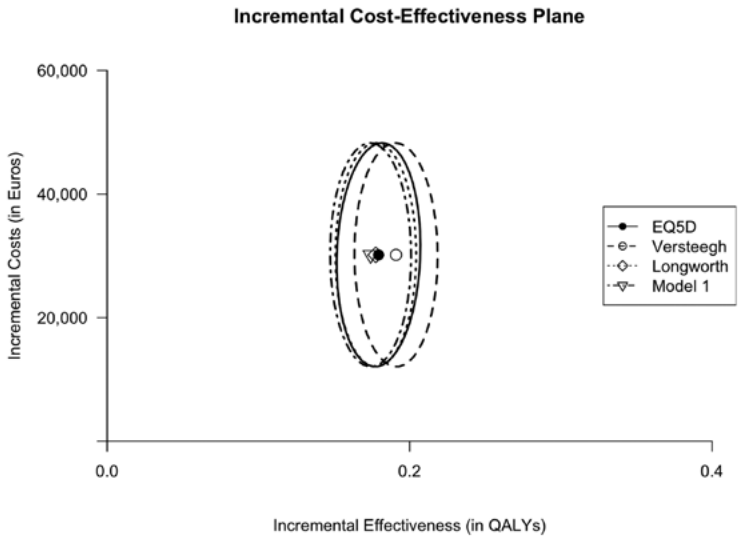


Figure 2. Incremental cost-effectiveness plans for observed and predicted utilities. Incremental cost-effectiveness planes comparing the effect of using observed EQ-5D-3L utility, the mapping algorithm by Versteegh et al., the mapping algorithm by Longworth et al (based on Dutch tariff). and predicted utility based on the preferred model (model 1 on OLS algorithm on QLQ-C30 functional scale scores). Ellipses represent the 95% confidence interval.

DISCUSSION

We have shown that the previously developed algorithm by Versteegh *et al*. and Marriott *et al*. for conversion of the disease-specific questionnaire EORTC QLQ-C30 into EQ-5D-3L based utilities resulted in a statistically significant difference between predicted and observed utilities. Still, the existing algorithms performed well as the mean predicted utilities drew close to the mean observed utilities (mean differences between the observed and respectively the mapped utilities by Versteegh *et al*., Longworth *et al*. and Marriott *et al*. were 0.03, 0.001 and 0.01 for the Dutch tariff EQ-5D utilities). No significant difference between, observed and predicted utilities were seen with the algorithm developed by Longworth *et al*. Even though the predicted utilities calculated with the algorithms by Versteegh *et al*. and Marriott *et al*. were significantly different, the outcome differences were not considered clinically meaningful. Previously, the minimal clinically relevant difference in utility for cancer patients was found to range 0.08-0.16, although this difference might vary per patient population [20,21]. Moreover, patients with different cancers types and stages of disease experience different symptoms and may thus respond differently on the QLQ-C30 functional scale scores [8]. In contrast, as was previously shown by Doble *et al*. disease severity is more likely to drive EQ-5D estimation based on QLQ-C30, and less by the cancer type [5]. Moreover, several studies developed condition-specific instruments, such as the EORTC QLU-C10D to derive health-related quality of life utilities, which might be more sensitive to disease-specific effects and in theory be preferred over EQ-5D. However, one can question whether these condition-specific instruments outperform EQ-5D [22–24]. Finally, with the emergence of novel treatment strategies in cancer treatment, such as immunotherapy, one could hypothesize a different value of QLQ-C30 functional scale or symptom scores, which could affect mapping outcomes.

Nevertheless, we pursued a better fitting algorithm for the mCRC patient population. All developed models demonstrated improved utility prediction ability with non-significant differences between observed and predicted utilities, although we acknowledge that the performance of the models developed in this study are not tested in a truly external dataset (as the models taken from the literature). Importantly, with the commonly used statistical methods to develop mapping algorithm, we did not succeed in the development of a better performing mapping algorithm. In case a mapping algorithm would be selected from our study, we would suggest the use of the RE model based on QLQ-C30 functional scale scores (model 1). This model provided the benefit of utility prediction for incomplete QLQ-C30 questionnaires (for which functional scale scores could be calculated), while retaining a good performance if tested on incomplete QLQ-C30 questionnaires. QLQ-C30 outcome conversion into EQ-5D-3L based utilities (Dutch tariff) could therefore be performed with the following algorithm, developed on functional scale scores *(model 1)*:

*$EQ\text{-}5D_{utility}$ = 0.2993 + 0.0021 * physical functioning score + 0.0011 * role functioning score + 0.0025 * emotional functioning score + 0.0005 * cognitive functioning score + 0.0006 * social functioning score + pain score * -0.0023 + insomnia score * -0.0005.*

The main purpose of mapping algorithms is to convert disease specific quality of life data into utilities for the purpose of cost-effectiveness research, if utilities cannot directly be derived from the dataset. We investigated the influence of a mapping algorithm on a cost-effectiveness model evaluating CB maintenance treatment compared to observation in mCRC patients. We demonstrated that the use of mappings results in comparable outcomes when used in a cost-effectiveness model. The newly developed algorithm slightly underperformed compared to the previously developed algorithm by Longworth *et al*. (ICER differences between in CEA using observed utilities and mapping: €1,765/QALY gained for the Longworth *et al*. mapping and €5,094 /QALY gained for the preferred model 1 in this study). An ICER difference of -€10,140/QALY gained was seen if compared to the Versteegh *et al*. mapping. Disparities were explained by small differences in incremental QALY estimation between treatment arms. The algorithm by Versteegh *et al*. and Longworth *et al*. slightly overestimated the utilities in both study arms; while the preferred model algorithm (model 1) overestimated the utilities in the observation arm and underestimated the utilities in the CB maintenance arm. Nevertheless, the Longworth algorithm outperformed our preferred model algorithm in this cost-effectiveness model. In a model with more pronounced utility differences, the impact of the chosen mapping algorithm might be different due to case mix effects. The good performance of the Longworth algorithm in this study is remarkable, as this algorithm had not been developed on colon cancer patients, and was estimated on an entirely different sample. Hence, its good performance, especially relative to the within-sample validation of the algorithm we developed, shows the usefulness of this flexible algorithm. Its performance raises the question if similarity of symptoms and severity of symptoms between the development sample and the application sample might not be of greater importance than type of cancer or tumor. While this study seems to suggest that indeed tumor type is less relevant, such a statement must be made with caution: many mapping algorithms, including the one by Versteegh *et al*., use only a selection of items of the QLQ-C30. As a consequence, out of sample prediction in patients with other cancer types with specific symptoms not captured by the included items might be complicated.

A strength of this study was the use of multiple statistical methods which enabled us to evaluate and select the best-performing algorithm, while also considering convenience in use. Furthermore, the analyses were conducted on a large population of patients, with a total of 1905 completed questionnaires. As previously mentioned, the algorithm by Versteegh *et al*. and the algorithm by Longworth *et al*. were not developed or validated in mCRC patient populations [6,7]. Only, the algorithm by Marriott *et al*. was developed and tested in an mCRC patient population using a U.K. tariff for EQ-5D-3L [9]. Patients with

different cancers types and stages of disease experience different symptoms and might thus respond differently on the QLQ-C30 domains functional scale scores. Thus, the most applicable algorithm in terms of cancer type and disease stage, should be applied for utility prediction, although it has previously been shown to be more dependent of disease severity than cancer type [5]. Of note, another colorectal cancer specific mapping algorithm estimating EQ-5D-5L values using a U.K. tariff was previously developed [25,26]. However, this mapping algorithm could not be tested and validated with the EQ-5D-3L values in our dataset, as this would require an additional mapping of EQ-5D-3L to EQ-5D-5L and we consequently would not been able to separate performance of the mapping algorithm due to differences in utilities. Currently, the EQ-5D-5L questionnaire is increasingly being adopted in clinical trials as it is regarded more sensitive to health effects and reduce ceiling effects [27]. Further research on mapping of QLQ-C30 outcomes towards EQ-5D-5L is therefore necessary.

The mapping algorithm was developed using a single sample, in which completed questionnaires were assigned to one of five folds that functioned as hold-out sample, which may be regarded as limitation of this study. Inevitably, the training and test datasets therefore contain comparable patients, who completed the quality of life questionnaires under similar circumstance. Preferably, validation of the developed algorithms should have occurred in another sample containing mCRC patient data on both the QLQ-C30 and the EQ-5D-3L questionnaires. Another limitation to this study, is the use of different time-points. The regression algorithms accounted for the panel data structure where possible through the use of random effects models. However, it has previously been shown that colorectal cancer patients continue to report high quality of life during the course of their disease [28–31]. Nonetheless, significant and clinically relevant changes in quality of life occur in the palliative stage of the disease, especially in the last few months of life a decline in quality of life has been demonstrated [32]. Therefore, it may be hypothesized that this could also apply for different time-points within a trial during which different dimensions of health are affected. The models developed in this study, are especially sensitive to this issue.

CONCLUSION

We have developed a QLQ-C30 to EQ-5D-3L mapping algorithm on a mCRC patient population with predicted utilities drawing close to the observed utilities. However, the mapping algorithm did not outperform existing mapping algorithms, especially compared with the response mapping algorithm by Longworth *et al*. Moreover, external validation of our preferred mapping algorithm remains desirable. The choice of mapping algorithm might only have a small impact on the predicted utility and cost-effectiveness, as was illustrated in the case study. Nonetheless, for studies only including disease-specific quality of life questionnaires, our results show that mapping is an adequate solution to

obtain utility estimates for use in cost-effectiveness analysis for mCRC patients, using either our newly developed mapping algorithm or one of the existing algorithms used in this study.

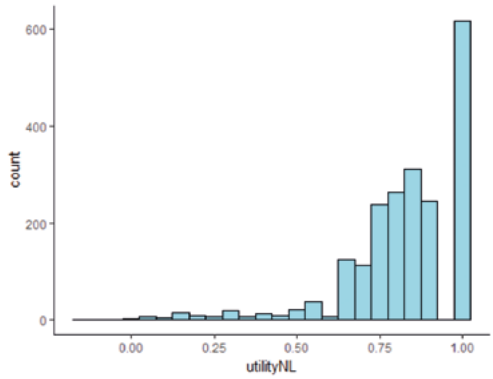
REFERENCES

1. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the Panel on Cost-Effectiveness in Health and Medicine. *JAMA*. 1996;276(15):1253-1258. doi:10.1097/00132586-199712000-00019
2. Williams A. EuroQol - A new facility for the measurement of health-related quality of life. *Health Policy (New York)*. 1990;16:199-208. doi:10.1016/0168-8510(90)90421-9
3. Torrance GW. Measurement of health state utilities for economic appraisal. *J Health Econ*. 1986;5:1-30. doi:10.1016/0167-6296(86)90020-2
4. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organisation for Research and Treatment of Cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst*. 1993;85:365-376.
5. Doble B, Lorgelly P. Mapping the EORTC QLQ-C30 onto the EQ-5D-3L: assessing the external validity of existing mapping algorithms. *Qual Life Res*. 2016;25(4):891-911. doi:10.1007/s11136-015-1116-2
6. Versteegh MM, Leunis A, Luime JJ, Boggild M, Uyl-de Groot C a, Stolk E a. Mapping QLQ-C30, HAQ, and MSIS-29 on EQ-5D. *Med Decis Making*. 2012;32(4):554-568. doi:10.1177/0272989X11427761
7. Longworth L, Yang Y, Young T, et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. *Health Technol Assess*. 2014;18(9):1-224. doi:10.3310/hta18090
8. Longworth L, Rowen D. Mapping to obtain EQ-5D utility values for use in nice health technology assessments. *Value Heal*. 2013;16(1):202-210. doi:10.1016/j.jval.2012.10.010
9. Marriott E-R, van Hazel G, Gibbs P, Hatswell AJ. Mapping EORTC-QLQ-C30 to EQ-5D-3L in patients with colorectal cancer. *J Med Econ*. 2017;20(2):193-199. doi:10.1080/13696998.2016.1241788
10. Franken M., van Rooijen E., May A., et al. Cost-effectiveness of capecitabine and bevacizumab maintenance treatment after first-line induction treatment in metastatic colorectal cancer. *Eur J Cancer*. 2017;75:204-212. doi:10.1016/j.ejca.2017.01.019
11. Simkens LHJ, van Tinteren H, May A, et al. Maintenance treatment with capecitabine and bevacizumab in metastatic colorectal cancer (CAIRO3): a phase 3 randomised controlled trial of the Dutch Colorectal Cancer Group. *Lancet*. 2015;385(9980):1843-1852. doi:10.1016/S0140-6736(14)62004-3
12. EORTC. EORTC QLQ-C30 Scoring Manual The EORTC QLQ-C30 Introduction. *EORTC QLQ-C30 Scoring Man*. 2001;30:1-67. doi:D/2001/6136/001
13. Lamers LM, Stalmeier PFM, McDonnell J, Krabbe PFM, van Busschbach JJ. [Measuring the quality of life in economic evaluations: the Dutch EQ-5D tariff]. *Ned Tijdschr Geneeskd*. 2005;149(28):1574-1578. Accessed October 30, 2014. http://www.ncbi.nlm.nih.gov/pubmed/16038162
14. Dolan P. Modeling valuations for EuroQol health states. *Med Care*. 1997;35(11):1095-1108.
15. Petrou S, Rivero-Arias O, Dakin H, et al. Preferred reporting items for studies mapping onto preference-based outcome measures: the MAPS statement. *Qual Life Res*. 2016;25(2):275-281. doi:10.1007/s11136-015-1082-8
16. Crott R. Direct Mapping of the QLQ-C30 to EQ-5D Preferences: A Comparison of Regression Methods. *PharmacoEconomics - Open*. 2018;2(2):165-177. doi:10.1007/s41669-017-0049-9
17. Le QA, Doctor JN. Probabilistic mapping of descriptive health status responses onto health state utilities using Bayesian networks: an empirical analysis converting SF-12 into EQ-5D utility index in a national US sample. *Med Care*. 2011;49(5):451-460. doi:10.1097/MLR.0b013e318207e9a8
18. Versteegh MM, Rowen D, Brazier JE, Stolk EA. Mapping onto Eq-5 D for patients in poor health. *Health Qual Life Outcomes*. 2010;8:141. doi:10.1186/1477-7525-8-141
19. Degeling K, Franken MD, May AM, et al. Matching the model with the evidence:

comparing discrete event simulation and state-transition modeling for time-to-event predictions in a cost-effectiveness analysis of treatment in metastatic colorectal cancer patients. *Cancer Epidemiol*. 2018;57:60-67. doi:10.1016/j.canep.2018.09.008

20. Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. *Health Qual Life Outcomes*. 2007;5:2-9. doi:10.1186/1477-7525-5-70
21. Revicki DA, Cella D, Hays RD, Sloan JA, Lenderking WR, Aaronson NK. Responsiveness and minimal important differences for patient reported outcomes. *Health Qual Life Outcomes*. 2006;4:1-5. doi:10.1186/1477-7525-4-70
22. King MT, Costa DSJ, Aaronson NK, et al. QLU-C10D: a health state classification system for a multi-attribute utility measure based on the EORTC QLQ-C30. *Qual Life Res*. 2016;25(3):625-636. doi:10.1007/s11136-015-1217-y
23. King MT, Viney R, Simon Pickard A, et al. Australian Utility Weights for the EORTC QLU-C10D, a Multi-Attribute Utility Instrument Derived from the Cancer-Specific Quality of Life Questionnaire, EORTC QLQ-C30. *Pharmacoeconomics*. 2018;36(2):225-238. doi:10.1007/s40273-017-0582-5
24. Versteegh MM, Leunis A, Uyl-De Groot CA, Stolk EA. Condition-specific preference-based measures: Benefit or burden? *Value Heal*. 2012;15(3):504-513. doi:10.1016/j.jval.2011.12.003
25. Ameri H, Yousefi M, Yaseri M, Nahvijou A, Arab M, Akbari Sari A. Mapping the cancer-specific QLQ-C30 onto the generic EQ-5D-5L and SF-6D in colorectal cancer patients. *Expert Rev Pharmacoeconomics Outcomes Res*. 2019;19(1):89-96. doi: doi.org/10.1080/14737167.2018.1517046
26. Ameri H, Yousefi M, Yaseri M, Nahvijou A, Arab M, Akbari Sari A. Mapping EORTC-QLQ-C30 and QLQ-CR29 onto EQ-5D-5L in Colorectal Cancer Patients. *J Gastrointest Cancer*. 2020;51(1):196-203. doi:10.1007/s12029-019-00229-6
27. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20(10):1727-1736. doi:10.1007/s11136-011-9903-x
28. Arndt V, Merx H, Stegmaier C, Ziegler H, Brenner H. Restrictions in quality of life in colorectal cancer patients over three years after diagnosis: A population based study. *Eur J Cancer*. 2006;42(12):1848-1857. doi:10.1016/j.ejca.2006.01.059
29. Caravati-Jouvenceaux A, Launoy G, Klein D, et al. Health-Related Quality of Life Among Long-Term Survivors of Colorectal Cancer: A Population-Based Study. *Oncologist*. 2011;16(11):1626-1636. doi:10.1634/theoncologist.2011-0036
30. Bouvier AM, Jooste V, Bonnetain F, et al. Adjuvant treatments do not alter the quality of life in elderly patients with colorectal cancer: A population-based study. *Cancer*. 2008;113(4):879-886. doi:10.1002/cncr.23629
31. Verhaar S, Vissers PAJ, Maas H, Van De Poll-Franse L V., Van Erning FN, Mols F. Treatment-related differences in health related quality of life and disease specific symptoms among colon cancer survivors: Results from the population-based PROFILES registry. *Eur J Cancer*. 2015;51(10):1263-1273. doi:10.1016/j.ejca.2015.04.004
32. Raijmakers NJH, Zijlstra M, van Roij J, Husson O, Oerlemans S, van de Poll-Franse L V. Health-related quality of life among cancer patients in their last year of life: results from the PROFILES registry. *Support Care Cancer*. 2018;26(10):3397-3404. doi:10.1007/s00520-018-4181-6

ADDITIONAL FILES



Additional file 1. Histogram of EQ-5D-3L based utilities of 1905 observations

Additional file 2

Table 1. Ordered logit regression (model 4) results for QLQ-C30 domain scores on EQ-5D-3L domain.

Mobility (MO)	**Coefficient**	**SD**	**t-value**	**p-value***	**95% CI**
Kappa	8.568	0.628	13.647	<0.001	[7.338,9.799]
Constant	9.741	0.953	10.221	<0.001	[7.873,11.609]
Global health	-0.007	0.005	-1.23	0.219	[-0.017,0.004]
Physical functioning	-0.093	0.007	-12.5	<0.001	[-0.108,-0.078]
Role functioning	-0.012	0.005	-2.734	0.006	[-0.021,-0.004]
Emotional functioning	0.011	0.005	1.983	0.047	[0,0.021]
Cognitive functioning	-0.011	0.005	-2.213	0.027	[-0.022,-0.001]
Social functioning	-0.005	0.005	-1.083	0.279	[-0.015,0.004]
Symptom scale: fatigue	-0.023	0.006	-4.174	<0.001	[-0.034,-0.012]
Symptom scale: nausea	-0.020	0.007	-2.963	0.003	[-0.033,-0.007]
Symptom scale: pain	0.014	0.004	3.393	0.001	[0.006,0.022]
Symptom scale: dyspnoea	0.002	0.004	0.624	0.532	[-0.005,0.01]
Symptom scale: insomnia	-0.000	0.003	-0.035	0.972	[-0.007,0.006]
Symptom scale: appetite loss	-0.001	0.004	-0.204	0.838	[-0.009,0.007]
Symptom scale: constipation	-0.001	0.005	-0.197	0.843	[-0.01,0.008]
Symptom scale: diarrhoea	0.005	0.003	1.553	0.120	[-0.001,0.012]
Symptom scale: financial diff.	0.017	0.004	3.996	<0.001	[0.009,0.025]
Self-care (SC)					
Kappa	3.582	0.466	7.687	<0.001	[2.669,4.495]
Constant	2.426	1.196	2.029	0.042	[0.082,4.77]
Global health	0.004	0.008	0.437	0.662	[-0.013,0.02]
Physical functioning	-0.056	0.010	-5.887	<0.001	[-0.075,-0.037]
Role functioning	-0.014	0.007	-1.995	0.046	[-0.028,0]
Emotional functioning	0.000	0.008	0.023	0.982	[-0.015,0.015]
Cognitive functioning	-0.010	0.007	-1.304	0.192	[-0.024,0.005]
Social functioning	0.004	0.007	0.604	0.546	[-0.01,0.019]
Symptom scale: fatigue	-0.008	0.008	-0.935	0.350	[-0.024,0.009]
Symptom scale: nausea	-0.002	0.009	-0.196	0.845	[-0.02,0.016]
Symptom scale: pain	0.011	0.005	2.046	0.041	[0,0.022]
Symptom scale: dyspnoea	-0.006	0.006	-1.005	0.315	[-0.017,0.005]
Symptom scale: insomnia	0.008	0.005	1.641	0.101	[-0.002,0.017]
Symptom scale: appetite loss	-0.007	0.006	-1.114	0.265	[-0.018,0.005]
Symptom scale: constipation	-0.010	0.006	-1.575	0.115	[-0.023,0.002]
Symptom scale: diarrhoea	0.010	0.005	2.020	0.043	[0,0.021]
Symptom scale: financial diff.	0.015	0.005	2.761	0.006	[0.004,0.026]

p-values result from a t-test.

Table 2. Ordered logit regression (model 4) results for QLQ-C30 domain scores on EQ-5D-3L domain.

Daily activities (DA)	**Coefficient**	**SD**	**t-value**	**p-value***	**95% CI**
Kappa	6.652	0.333	19.977	<0.001	[6,7.305]
Constant	9.823	0.957	10.267	<0.001	[7.948,11.698]
Global health	-0.014	0.006	-2.427	0.015	[-0.025,-0.003]
Physical functioning	-0.032	0.007	-4.613	<0.001	[-0.045,-0.018]
Role functioning	-0.062	0.005	-12.095	<0.001	[-0.072,-0.052]
Emotional functioning	0.006	0.006	0.999	0.318	[-0.005,0.017]
Cognitive functioning	-0.018	0.005	-3.361	0.001	[-0.029,-0.008]
Social functioning	-0.015	0.005	-3.039	0.002	[-0.025,-0.005]
Symptom scale: fatigue	0.021	0.006	3.654	<0.001	[0.01,0.033]
Symptom scale: nausea	-0.006	0.007	-0.918	0.359	[-0.019,0.007]
Symptom scale: pain	0.008	0.004	2.007	0.045	[0,0.017]
Symptom scale: dyspnoea	0.001	0.004	0.186	0.852	[-0.007,0.008]
Symptom scale: insomnia	0.003	0.003	0.972	0.331	[-0.003,0.01]
Symptom scale: appetite loss	0.003	0.004	0.752	0.452	[-0.005,0.011]
Symptom scale: constipation	0.002	0.005	0.459	0.646	[-0.007,0.011]
Symptom scale: diarrhoea	0.003	0.004	0.838	0.402	[-0.004,0.011]
Symptom scale: financial diff.	0.009	0.004	2.181	0.030	[0.001,0.018]
Pain and discomfort (PA)					
Kappa	8.559	0.506	16.929	<0.001	[7.568,9.55]
Constant	0.262	0.869	0.302	0.763	[-1.441,1.964]
Global health	-0.009	0.006	-1.5	0.134	[-0.02,0.003]
Physical functioning	-0.009	0.007	-1.346	0.178	[-0.023,0.004]
Role functioning	0.005	0.005	1.102	0.270	[-0.004,0.015]
Emotional functioning	-0.004	0.006	-0.682	0.495	[-0.015,0.007]
Cognitive functioning	-0.007	0.005	-1.287	0.198	[-0.018,0.004]
Social functioning	-0.002	0.005	-0.354	0.723	[-0.012,0.009]
Symptom scale: fatigue	0.003	0.006	0.567	0.571	[-0.008,0.015]
Symptom scale: nausea	-0.023	0.007	-3.182	0.001	[-0.037,-0.009]
Symptom scale: pain	0.109	0.006	17.809	<0.001	[0.097,0.121]
Symptom scale: dyspnoea	0.007	0.004	1.657	0.098	[-0.001,0.015]
Symptom scale: insomnia	0.005	0.003	1.509	0.131	[-0.002,0.012]
Symptom scale: appetite loss	-0.009	0.004	-2.114	0.034	[-0.018,-0.001]
Symptom scale: constipation	0.014	0.005	3.045	0.002	[0.005,0.024]
Symptom scale: diarrhoea	0.005	0.004	1.306	0.191	[-0.003,0.013]
Symptom scale: financial diff.	-0.000	0.005	-0.008	0.994	[-0.009,0.009]

p-values result from a t-test.

Table 3. Ordered logit regression (model 4) results for QLQ-C30 domain scores on EQ-5D-3L domain.

Anxiety and depression (AD)	**Coefficient**	**SD**	**t-value**	**p-value***	**95% CI**
Kappa	5.685	0.390	14.568	<0.001	[4.92,6.45]
Constant	11.065	1.033	10.715	<0.001	[9.041,13.089]
Global health	-0.007	0.006	-1.057	0.291	[-0.019,0.006]
Physical functioning	-0.005	0.007	-0.742	0.458	[-0.02,0.009]
Role functioning	-0.009	0.005	-1.79	0.073	[-0.019,0.001]
Emotional functioning	-0.123	0.007	-16.542	<0.001	[-0.138,-0.109]
Cognitive functioning	-0.002	0.005	-0.355	0.722	[-0.013,0.009]
Social functioning	-0.003	0.005	-0.556	0.578	[-0.013,0.007]
Symptom scale: fatigue	-0.005	0.006	-0.813	0.416	[-0.018,0.007]
Symptom scale: nausea	0.005	0.007	0.745	0.457	[-0.008,0.019]
Symptom scale: pain	-0.011	0.004	-2.437	0.015	[-0.02,-0.002]
Symptom scale: dyspnoea	-0.009	0.004	-2.074	0.038	[-0.018,0]
Symptom scale: insomnia	0.011	0.004	3.179	0.001	[0.004,0.018]
Symptom scale: appetite loss	-0.000	0.004	-0.024	0.981	[-0.009,0.008]
Symptom scale: constipation	-0.008	0.005	-1.714	0.087	[-0.018,0.001]
Symptom scale: diarrhoea	-0.005	0.004	-1.095	0.274	[-0.013,0.004]
Symptom scale: financial diff.	0.007	0.004	1.656	0.098	[-0.001,0.016

p-values result from a t-test.

Utilities can be calculated by applying the composing equations for each EQ-5D domain :

In example for EQ-5D domain mobility:

MO = 9.740831 -0.00654**global health score* -0.09308**physical functioning score* -0.01237**role functioning score* +0.010595**emotional functioning score* -0.01149**cognitive functioning score* -0.00532**social functioning score* -0.02323**fatigue score* -0.01976* *nausea score* +0.014166**pain score* +0.002403**dyspnoea score* -0.00011**insomnia score* -0.00083**appetite score* -0.00088**constipation score* +0.005426**diarrhoea score* +0.016817**financial difficulties score*

probmo1 = 1/(1 + exp(MO))

probmo2 = 1/(1 + exp(MO - kappa)) – 1/(1+exp(MO))

probmo3= 1-Probmo1-Promo2

Where prob stands for the predicted probability of the EQ-5D-3L response level (1, 2 or 3). Thus probmo1 stands for the probability of a level 1 response for the EQ-5D mobility domain. Probmo2 and probmo3 for level 2 and level 3 responses on the EQ-5D mobility domain respectively . Each probability for each EQ-5D domain can be used for EQ-5D tariff calculations, in example for the Dutch tariff:

Estimated EQ-5D=1-(probmo2*0.036)-(probmo3*0.161)-(probsc2*0.082)-(probsc3*0.152)-(probda2*0.032)-(probda3*0.057)-(probpa2*0.086)-(probpa3*0.329)-(probad2*-0.124)-(probad3*0.325)-(1-ProbPerfect)*0.071-ProbN3*0.234

ProbPerfect= probmo1*probsc1*probda1*probpa1*probad1

ProbN3= 1-(1-probmo3)*(1-probsc3)*(1-probda3)*(1-probpa3)*(1-probad3)

Table 4. Separate equations subgroup approach (model 6) results for QLQ-C30 domain scores on EQ-5D-3L utility of i) < 0.6, ii) 0.6 and < 1 and iii) 1.

Variable	Coefficient	SD	t-value	p-value*	95% CI
Category 1: utility < 0.6			(Base outcome)		
Category 2: utility 0.6 and < 1					
Constant	-4.355	1.305	-3.338	0.001	[-6.913,-1.798]
Global health	0.011	0.010	1.152	0.249	[-0.008,0.03]
Physical functioning	0.044	0.010	4.229	<0.001	[0.024,0.064]
Role functioning	0.022	0.008	2.684	0.007	[0.006,0.038]
Emotional functioning	0.031	0.008	3.771	<0.001	[0.015,0.048]
Cognitive functioning	0.008	0.008	1.015	0.310	[-0.007,0.023]
Social functioning	-0.002	0.008	-0.283	0.777	[-0.017,0.013]
Symptom scale: fatigue	-0.013	0.009	-1.364	0.173	[-0.031,0.005]
Symptom scale: nausea	-0.006	0.009	-0.656	0.512	[-0.023,0.011]
Symptom scale: pain	-0.023	0.006	-3.999	<0.001	[-0.034,-0.012]
Symptom scale: dyspnoea	0.006	0.006	1.016	0.310	[-0.006,0.018]
Symptom scale: insomnia	-0.006	0.005	-1.181	0.238	[-0.016,0.004]
Symptom scale: appetite loss	0.004	0.006	0.641	0.521	[-0.008,0.016]
Symptom scale: constipation	0.008	0.007	1.212	0.225	[-0.005,0.021]
Symptom scale: diarrhoea	0.000	0.006	-0.054	0.957	[-0.012,0.011]
Symptom scale: financial diff.	-0.012	0.007	-1.870	0.062	[-0.025,0.001]
Category 3: utility = 1					
Constant	-18.864	1.994	-9.461	<0.001	[-22.772,-14.956]
Global health	0.024	0.012	2.046	0.041	[0.001,0.048]
Physical functioning	0.116	0.015	7.926	<0.001	[0.087,0.145]
Role functioning	0.046	0.010	4.470	<0.001	[0.026,0.066]
Emotional functioning	0.055	0.012	4.748	<0.001	[0.032,0.077]
Cognitive functioning	0.028	0.011	2.579	0.010	[0.007,0.05]
Social functioning	0.009	0.011	0.890	0.373	[-0.011,0.03]
Symptom scale: fatigue	-0.021	0.011	-1.866	0.062	[-0.044,0.001]
Symptom scale: nausea	0.022	0.014	1.628	0.103	[-0.005,0.049]
Symptom scale: pain	-0.107	0.012	-9.208	<0.001	[-0.13,-0.085]
Symptom scale: dyspnoea	0.006	0.008	0.744	0.457	[-0.01,0.022]
Symptom scale: insomnia	-0.024	0.007	-3.459	0.001	[-0.038,-0.011]
Symptom scale: appetite loss	0.012	0.009	1.360	0.174	[-0.005,0.029]
Symptom scale: constipation	-0.001	0.010	-0.095	0.924	[-0.02,0.018]
Symptom scale: diarrhoea	-0.001	0.008	-0.149	0.882	[-0.016,0.014]
Symptom scale: financial diff.	-0.037	0.010	-3.890	<0.001	[-0.056,-0.018]

p-values result from a t-test.

Table 5. Regression results (model 6) for EQ-5D-3L based utility values < 0.6 on QLQ-C30 domain scores.

Variable	Coefficient	SD	t-value	p-value*	95% CI
Constant	0.313	0.054	5.826	<0.001	[0.208,0.418]
Emotional functioning	0.002	0.001	3.158	0.002	[0.001,0.003]
Symptom scale: pain	-0.001	0.001	-2.621	0.010	[-0.002,0]

p-values result from a t-test.

Table 6. Regression results (model 6) for EQ-5D-3L based utility values 0.6 and < 1 on QLQ-C30 domain scores.

Variable	Coefficient	SD	t-value	p-value*	95% CI
Constant	0.570	0.017	33.378	<0.001	[0.536,0.603]
Physical functioning	0.001	0.000	3.983	<0.001	[0,0.001]
Role functioning	0.000	0.000	2.373	0.018	[0,0.001]
Emotional functioning	0.002	0.000	13.325	<0.001	[0.002,0.002]
Symptom scale: pain	-0.001	0.000	-8.869	<0.001	[-0.001,-0.001]
Symptom scale: insomnia	0.000	0.000	-1.902	0.057	[0,0]

p-values result from a t-test.

Additional file 3

Additional file 3. Patient characteristics for concurrently collected EQ-5D and partially incomplete QLQ-C30 questionnaires for which functional scale scores could still be calculated.

		N = 120
Age (years)		66 (8.3)
Male gender (%)		60
EQ-5D-3L*	N	120
	Mobility 1/2/3 (%)	40.8/55/4.2
	Self-cae 1/2/3 (%)	83.3/13.3/3.3
	Usual activities 1/2/3 (%)	45.8/47.5/6.7
	Pain/discomfort 1/2/3 (%)	46.7/48.3/5
	Depression/anxiety 1/2/3 (%)	72.5/26.7/0.8
	EQ-5D utility, mean (SD)	0.76 (0.232)
	EQ-5D range	-0.086 to 1
QLQ-C30 v.3.0	Questionnaires, N	120
	Physical functioning, mean (SD)	74.801 (22.661)
	Role functioning, mean (SD)	69.861 (28.113)
	Emotional functioning, mean (SD)	80.949 (18.591)
	Cognitive functioning, mean (SD)	86.111 (18.629)
	Social functioning, mean (SD)	85.139 (20.245)
	Global health, mean (SD)	68.75 (20.247)
	Fatigue, mean (SD)	30.972 (22.942)
	Nausea/vomiting, mean (SD)	6.389 (13.351)
	Pain, mean (SD)	25.556 (28.903)
	Dyspnea, mean (SD)	16.111 (25.559)
	Insomnia, mean (SD)	19.444 (26.137)
	Appetite, mean (SD)	14.167 (22.314)
	Constipation, mean (SD)	5.278 (12.961)
	Diarrhea, mean (SD)	11.667 (22.723)
	Financial difficulties, mean (SD)	4.167 (13.363)

* Percentages at level 1, 2 and 3 represent no problems at all, some problems and extreme problems, respectively.

Additional file 4

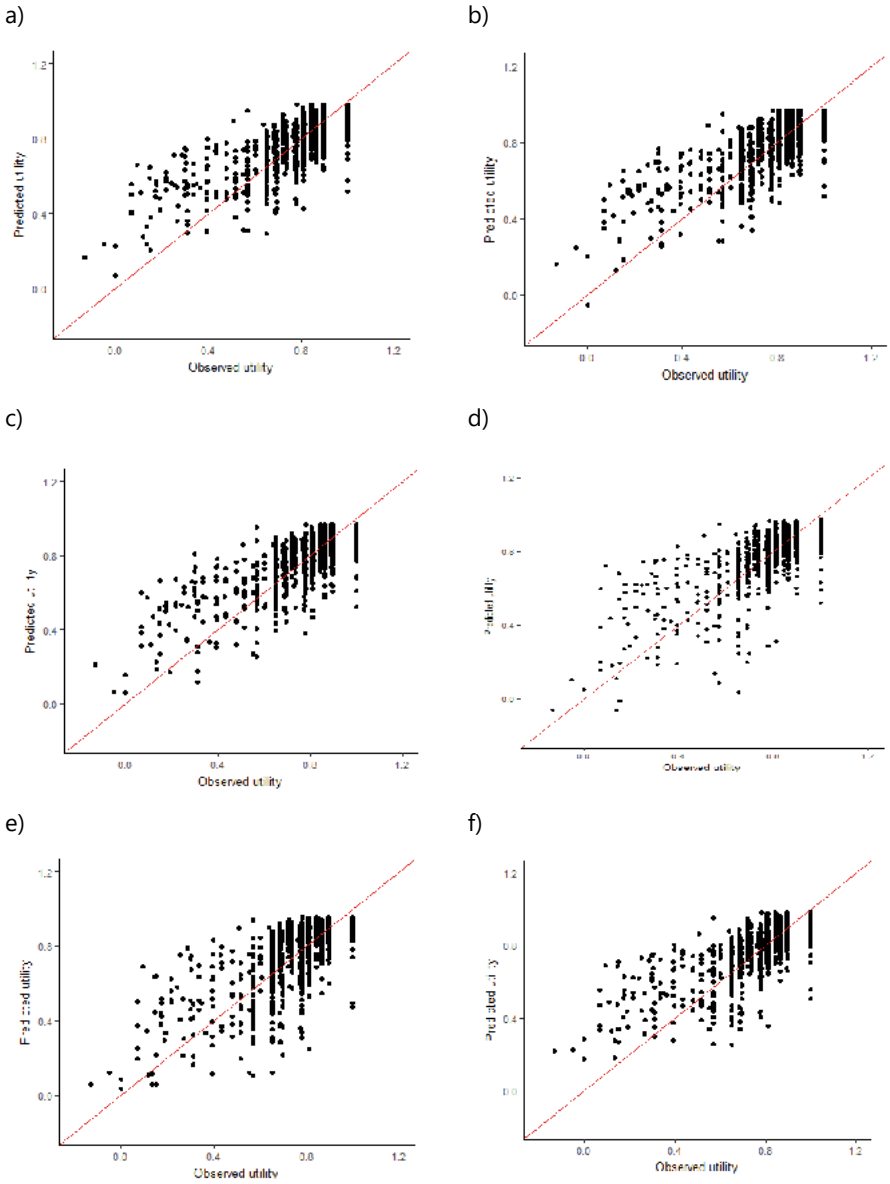


Figure 2. Predicted EQ-5D-3L utility versus the observed utility for a) the RE model with QLQ-C30 domain scores (preferred model 1); b) the RE model with continuous QLQ-C30 questions (model 2); c) the RE model with QLQ-C30 dummy questions (model 3); d) the ordered logit model on the EQ-5D-3L domains (model 4); e) beta regerssion (model 5) and; f) the separate equations subgroup approach (model 6).

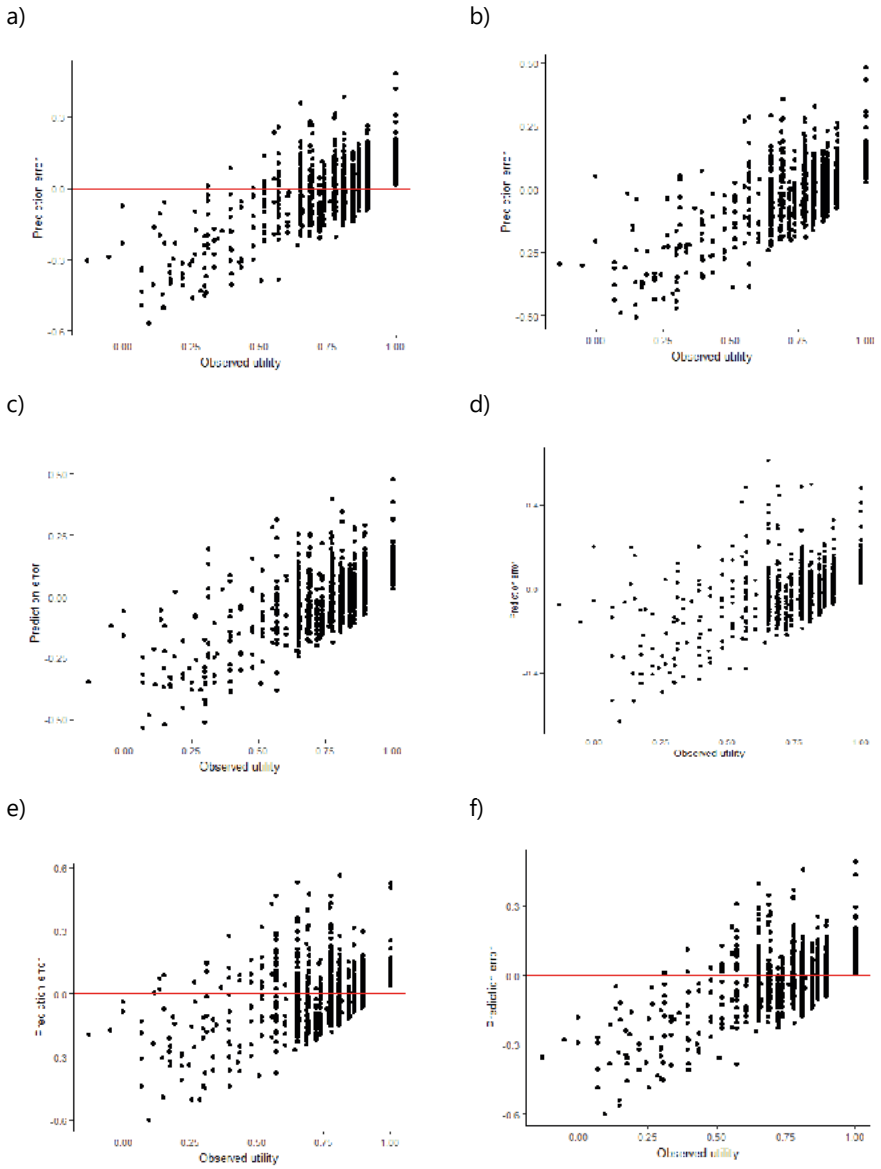


Figure 3. Prediction error (observed – predicted EQ-5D-3L uility) for a) the RE model with QLQ-C30 domain scores (preferred model 1); b) the RE model with continuous QLQ-C30 questions (model 2); c) the RE model with QLQ-C30 dummy questions (model 3); d) the ordered logit model on the EQ-5D-3L domains (model 4); e) beta regerssion(model 5) and; f) the separate equations subgroup approach (model 6).

6

WORK ABILITY IN PATIENTS WITH STAGE I-IV COLON CANCER, RESULTS OF THE DUTCH PROSPECTIVE COLORECTAL CANCER COHORT

Mira D. Franken

Geraldine R. Vink

Wilhelmina M.U. van Grevenstein

Helena M. Verkooijen

Cornelis J.A. Punt

Anne M. May

On behalf of the PLCRC study group *

* PLCRC study group listed after the references.

Submitted

ABSTRACT

Background

Colon cancer affects a patient's ability to work. Many colon cancer patients are employed at the time of diagnosis.

Objective

We evaluated work ability during the first two years after colon cancer diagnosis.

Design

This study is a national prospective study, the Prospective Dutch ColoRectal Cancer cohort, including clinical data and patient reported outcomes.

Settings

Data were collected in 59 medical centres in the Netherlands.

Patients

Patients <67 years, with stage I-IV colon cancer, who completed work ability index questionnaires, were selected.

Main outcome measures

Work ability was assessed at baseline, 3, 6, 12, 18 and 24 months. The work ability index (ranging from 0-49) was evaluated using linear mixed models. Outcomes were matched to population controls without cancer.

Results

Of 390 patients, 84% had payed employment. Work ability of stage I-IV patients was significantly lower at time of diagnosis compared to matched population controls (31±8.2 and 41±5.6, respectively). Patients with stage I-III disease receiving surgery only regained work ability index scores comparable to matched population controls at 18 months. Patients receiving adjuvant systemic treatment initially demonstrated a decrease in work ability with improvements from 6 months onwards and normalisation at 24 months. Stage IV patients did not demonstrate improvements in work ability outcomes over time. Work ability scores were negatively influenced by the administration of systemic treatment and ≥1 comorbidities.

Limitations

Only patients with patient reported outcomes and work at baseline were included in this analysis. Also, questionnaire response rates decreased over time.

Conclusion

Work ability in colon cancer patients is decreased for a prolonged time. Recovery depends on disease stage, type of treatment and comorbidity. Patients with stage I-III disease treated with curative surgery alone are first to regain work ability, followed by patients who receive adjuvant chemotherapy. Stage IV disease patients do not regain work ability.

INTRODUCTION

Work absenteeism or reduced work ability has been associated with reduced quality of life in cancer patients.[1,2] Cancer does not only reduce employability, but also results in income reduction shortly after cancer diagnosis and long term income reduction compared to persons without cancer.[3] More detrimental effects on employment are observed in advanced stages of disease.[3] There are important variations in return to work rates after cancer diagnosis and treatment, ranging from 24% up to 94%.[2] Return to work is influenced by multiple factors, such as work-related factors (i.e. employer agreements, counselling), demographic factors (i.e. younger age, higher levels of education) and cancer related factors (i.e. absence of surgery, continuity of care, less physical symptoms).[2] Cognitive work demands seem to reduce the negative impact of cancer on employment.[4] Furthermore, the ability to work is important for mental wellbeing and quality of life.[5,6]

A limited number of studies on work ability, sick leave or return to work in patients with colorectal cancer (CRC) have previously been reported, with the majority including only small numbers of patients and mainly focussing on patients with early stages of disease (i.e. stages I-III) who receive treatment with curative intent. CRC patients have an increased risk for unemployment or work decrease following treatment [7–9], and a reduction in annual labour income for men and women of 6% and 22%, respectively, has been shown[10]. Advanced disease stage, chemotherapy or a combination of treatments, extensive surgery, multiple comorbidities, sick leave in the year prior to diagnosis or unemployment, lower level of education, female sex and higher age have previously all been associated with increased risk of receiving a disability pension or reduced work ability.[8,11–14] Importantly, CRC patients who continue or return to work, report a better quality of life compared with those discontinuing their work.[7]

As the retirement age rises in many Western countries and with the observed increase in CRC in patients <50 years [15,16], the number of CRC patients who face challenges regarding work ability will increase. More research is warranted, to improve patient support regarding work ability. Also, as the local treatment strategies for colon cancer and rectal cancer are dissimilar, with major differences in surgical approaches and, use of radiotherapy or chemo(radio)therapy for rectal cancer which may affect the quality of life, both patient populations should be evaluated separately. Here, we perform a longitudinal evaluation on work ability for colon cancer patients (stage I-IV) in comparison to a working population without cancer.

MATERIALS & METHODS

The Prospective Dutch ColoRectal Cancer (PLCRC) cohort is an observational study in which clinical data, biospecimens and patient reported outcome measurements (PROMs) are collected in 59 participating clinical centres for CRC patients with all disease stages.[17] The PLCRC study was approved by the Medical Research Ethics

Committee of the University Medical Centre Utrecht (the Netherlands) (clinicaltrials.gov: NCT02070146). Informed consent for clinical data collection is mandatory, while consent for biospecimens and PROMs is optional. For the PLCRC cohort, the clinical information is retrieved from the Netherlands Cancer Registry (NCR), which collects this information for all Dutch cancer patients upon diagnosis until death. PROMs are collected within the Patient Reported Outcomes Following Initial treatment and Long term Evaluation of Survivorship (PROFILES) registry.[18]

Patients with stage I-IV colon cancer enrolled in PLCRC within 3 months of diagnosis between 2013 until 2019 and aged <67 years were included in this analysis. The age limit was set at the Dutch basic government pension age. Patients who did not consent for completion of PROMs were excluded. Demographic and work ability data were used in the current analyses. Patients were included if a baseline work ability questionnaire and at least one subsequent work ability questionnaire at time points 3, 6, 12, 18 and 24 months, were completed.

Work ability index

Work ability was assessed by calculating the work ability index (WAI) score. The WAI score is based on an algorithm covering 7 domains: work ability in relation to work demands (physical, cognitive demand or a combination of both), current work ability compared to best ever, number of comorbidities, estimated work impairment, sick leave in the previous 12 months, estimated work ability in 2 years' time and vitality.[19,20] The WAI score ranges from 7 (poorest work ability) to a maximum of 49 (best work ability). In addition, WAI scores can be categorized as poor (7-27), moderate (28-36), good (37-43) and excellent (44-49).[20]

Patients were matched to a general population without cancer in a 1:5 ratio on the following parameters: age, sex, educational level and work demands. Only exact matches were accepted, with the exception of age, where a 2-years range was allowed. General population scores, collected in 2018, were retrieved from "Stichting Blik op Werk" (www.blikopwerk.nl), which is an independent national institute specialized in labour participation and data collection on work ability.

Statistical analysis

Descriptive statistics were used for baseline characteristics and stratified by tumour stage.

To estimate mean WAI changes from diagnosis over time (3, 6, 12, 18 and 24 months), linear mixed models were used. Model fit was assessed using the AIC (Akaike Information Criterion). A heterogeneous autoregressive covariance structure of the first order (heterogeneous AR1) was chosen and using a restricted maximum likelihood function. The model included a random intercept. The mean WAI scores over time were estimated for each disease stage (I-IV) and for treatment choice (surgery only or surgery with

adjuvant systemic treatment) for patients with a curative treatment intent (stages I-III). The following fixed factors were included in a separate linear mixed model to estimate the association with the WAI score: age (continuous), sex, educational level, work demands, marital status, number of co-morbidities, disease stage and ((neo-)adjuvant) systemic therapy, and surgery (the latter only for the model including all disease stages).[8,11,12,14] Longitudinal missing data were accepted within this statistical model since linear mixed effect models allow for missing values.[21,22]

A Kruskal-Wallis test was used for comparison of WAI scores between disease stages at baseline and for comparison of cases versus matched population controls, while a Chi-square test was used to compare for WAI categories and sick leave.

All statistical analyses and case-control matching were performed in IBM SPSS Statistics for Windows, Version 25.0 (Armonk, NY: IBM Corp.).

RESULTS

A total of 390 colon cancer patients enrolled in the PLCRC study were selected for analysis in this study (Figure 1). Baseline patient characteristics are presented in Table 1. The mean age of the colon cancer patients was 58±7 years and 58% was male. Two or more comorbidities were reported by 73% of patients. The majority of patients with stage I, II and III disease underwent a primary tumour resection, which occurred in 97%, 100% and

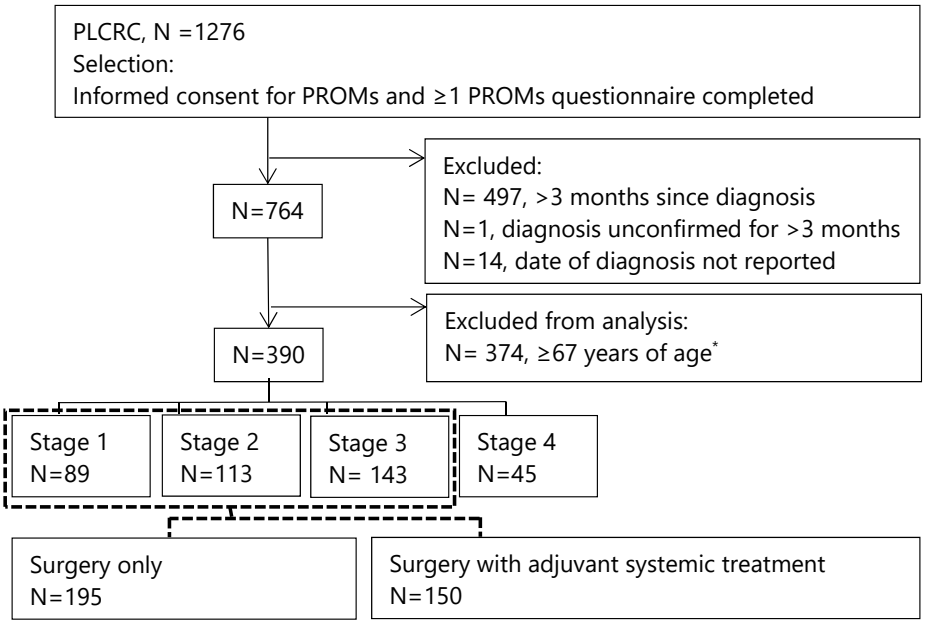


Figure 1. Flowchart of patient selection from the Prospective Dutch Colorectal Cancer cohort
*The Dutch basic government pension age is 67 years.

Table 1. Baseline characteristics of patients with colon cancer enrolled in the study within 3 months of diagnosis and <67 years.

	Disease stage				
	Stage I N=89 (%)	**Stage II N=113 (%)**	**Stage III N=143 (%)**	**Stage IV N=45 (%)**	**Overall N=390 (%)**
Mean age (y± SD) at baseline	59±5	58±6	57±7	55±7	58±7
Sex (male)	56 (63)	70 (62)	77 (54)	24 (53)	227 (58)
Number of comorbidities [1]					
0	8 (9)	9 (5)	11 (8)	2 (5)	30 (7)
1	17 (18)	21 (18)	25 (18)	10 (26)	73 (19)
≥2	63 (72)	79 (72)	103 (74)	31 (72)	276 (73)
Therapy					
Local excision primary tumour	11 (12)	1 (1)	3 (2)	2 (5)	17 (4)
Surgery	86 (97)	113 (100)	143 (100)	27 (60)	369 (95)
Systemic therapy	0 (0)	19 (17)	131 (92)	27 (60)	177 (45)
Median follow-up (months± SE)	15±1.2	16±1.2	12±0.9	16±1.4	14±0.6
Disease progression	0 (0)	3 (3)	7 (5)	16 (36)	26 (7)
Survival 2 years after inclusion[2]	85 (99)	106 (100)	134 (98)	32 (73)	357 (96)
Marital status, n (%) [3]					
Married/partner	46 (55)	60 (57)	67 (48)	22 (52)	195 (53)
Married/partner & children	19 (23)	28 (26)	44 (31)	17 (40)	108 (29)
Children	4 (5)	4 (4)	5 (4)	1 (2)	14 (4)
Alone	13 (16)	13 (12)	22 (16)	2 (5)	50 (13)
Educational level [4]					
Low	58 (65)	66 (59)	77 (54)	21 (47)	222 (57)
High	31 (35)	45 (41)	65 (46)	24 (53)	165 (43)
Employed at baseline [5]	65 (74)	79 (72)	98 (71)	33 (77)	275 (73)
Employed <60 years	33 (80)	50 (89)	72 (84)	23 (82)	178 (84)
Employer type					
Employed	53 (82)	66 (83)	83 (86)	28 (85)	230 (84)
Freelance/temporary worker	11 (17)	11 (14)	14 (14)	5 (15)	41 (15)
Company size [6]					
1-10 employees	14 (23)	14 (18)	12 (13)	3 (9)	43 (17)
10-50 employees	12 (19)	9 (12)	13 (15)	4 (13)	38 (15)
50-100 employees	4 (6)	6 (8)	4 (4)	7 (22)	21 (8)
≥100 employees	32 (52)	47 (62)	60 (67)	18 (56)	157 (61)

[1] Self-reported comorbidities, as reported in the WAI questionnaire, were not available for all patients (missing for 11 patients);

[2] Vital status not registered for 17 patients at the time of data analysis;

[3] Marital status not registered for 19 patients;

[4] Low educational level: secondary vocational education or lower. High educational level: Bachelor degree or higher. Educational level not registered for 3 patients;

[5] Employment status unavailable for 11 patients, of which 3 patients were aged <60;

[6] Company size was unavailable for 16 patients

100% of cases, respectively. Systemic adjuvant chemotherapy was administered in 17% of patients with stage II colon cancer and 92% in patients with stage III disease. Patients with stage IV colon cancer underwent surgery in 60% of cases and 60% received systemic therapy. The 2-year survival rate was 99%, 100%, 98% and 73% for patients with stage I, II, III and IV, respectively. Marital status and educational level were comparable between different disease stages.

Employment at baseline

Seventy-three percent of patients (n=275) reported to have work income at the time of study enrolment, which approximates the average of 74-78% in the Dutch population aged 16-65 years between 2015-2018.[23] Most patients received a paid income from an employer (84%) and 61% indicated to be employed in a company of over 100 employees.

Patients with work income at baseline were younger compared to patients without work income (age 57±7 and 61±6 years, respectively), and more often male (63% and 48%, respectively). No differences were observed in number of comorbidities, educational level and marital status.

Work Ability Index score

At baseline, 93% (n=255) of patients with paid employment completed the Work Ability Index (WAI) questionnaires. The WAI questionnaire completion rate was 79% at 3 months, 76% at 6 months, 67% at 12 months, 43% at 18 months and 22% at 24 months for this ongoing study. Matching resulted in 1350 matches, for 5 cases no matches were found.

At baseline, the mean WAI score of all patients - irrespective of disease stage - was significantly lower compared to matched population controls (mean WAI score 31±8.2 and 41±5.6 respectively, $p<0.01$). Baseline WAI scores between stage I to IV did not differ ($p=0.15$).The mean WAI score significantly improved over time, except for patients with stage IV disease (Figure 2A). The most pronounced decreases in WAI scores were seen for patients with stage III and IV disease. Compared to baseline, the WAI score at 24 months was significantly higher for patients with a mean difference for stage I to III of 6.6 (95%CI 3.3-9.9), 7.8 (4.8-10.8), 9.6 (6.3-12.8), respectively (Table 2A). No significant changes were found for stage IV, 1.6 (-4.9-8.0) (Table 2A). At 24 months all disease stages, except for stage IV disease, regained a work ability comparable to matched population controls. The mean WAI scores were significantly lower in female sex ($p=0.01$), patients receiving systemic therapy ($p<0.01$), a treatment without surgery ($p<0.01$), and presence of ≥1 comorbidities ($p<0.01$) (Table 3).

As the intensity of treatment in patients with curative disease differs, the mean WAI score over time was also stratified by treatment schedule (surgery vs. surgery with systemic treatment). At baseline, the WAI scores were significantly higher in patients receiving surgery only compared to patients receiving surgery with adjuvant systemic

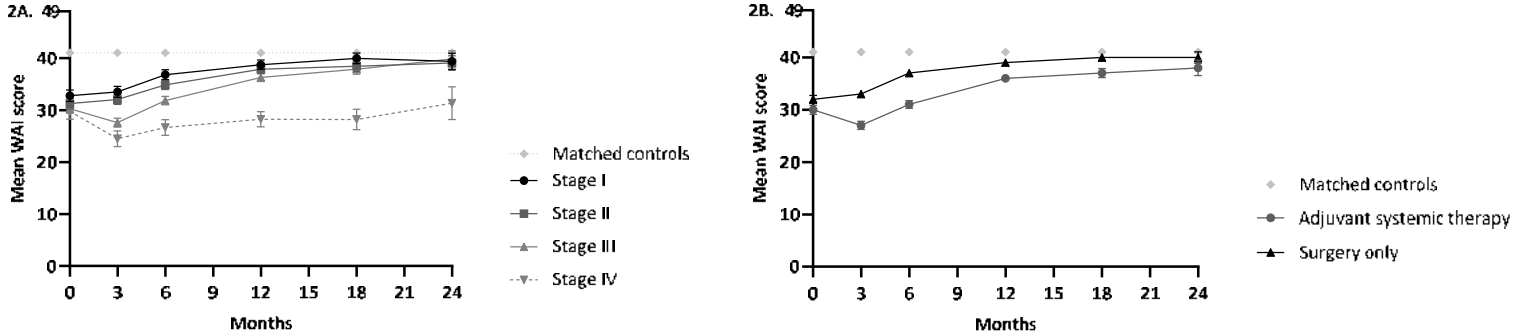


Figure 2. Modelled WAI over time displayed by disease stage (figure 2A) and displayed for patients with a curative intent (disease stage I-III) stratified by treatment strategy, either surgery only or surgery with adjuvant systemic treatment (figure 2B).

Table 2A. Mean changes (95% confidence intervals) on WAI-score in the first year after colon cancer diagnosis over time estimated by the mixed model. Table 2A displays the outcomes by disease stage (stage I-IV), while Table 2B displays the outcomes stratified by treatment strategy, either surgery with or without adjuvant systemic treatment, for patients with a curative intent (stage I-III).

	Stage I	Stage II	Stage III	Stage IV
Baseline - 3 months	0.7 (1.3-2.8)	0.8 (-1.0-2.6)	-2.7 (-4.3; -1.1)	-5.3 (-8.1;-2.4)
3 months - 6 months	3.3 (1.5-5.0)	2.8 (1.2-4.5)	4.3 (2.8-5.8)	2.2 (-0.7-5.0)
6 months - 12 months	2.0 (0.4-3.5)	3.0 (1.6-4.5)	4.5 (3.2-5.8)	1.6 (-1.0-4.2)
12 months -18 months	1.2 (-0.5-2.9)	0.6 (-1.1-2.2)	1.6 (-0.2-3.3)	-0.06 (-3.6-3.5)
18 months – 24 months	-0.6 (-3.5 – 2.3)	0.6 (-2.0-3.2)	2.0 (-1.0-5.0)	3.14 (-3.30-9.6)
Baseline – 24 months	6.6 (3.3-9.9)	7.8 (4.8-10.8)	9.6 (6.3 - 12.8)	1.6 (-4.9-8.0)

Table 2B.

	Adjuvant systemic therapy	Local treatment or surgery only
Baseline - 3 months	-2.8 (-4.4; -1.3)	1.0 (-0.4-2.3)
3 months - 6 months	4.2 (2.8 -5.6)	3.1 (1.9-4.3)
6 months – 12 months	5.0 (3.8-6.2)	2.1 (1.1-3.1)
12 months – 18 months	1.4 (-0.1-2.8)	0.9 (-0.2-2.0)
18 months – 24 months	1.0 (-1.8-3.8)	0.5 (-1.4-2.5)
Baseline – 24 months	8.7 (5.6-11.8)	7.6 (5.3-9.9)

therapy (32±8.0 and 29±8.0, respectively ($p<0.01$). The WAI score for patients receiving adjuvant systemic treatment decreased at 3 months and improvements in WAI scores were observed from 6 months onwards. The WAI scores remained lower for patients receiving adjuvant systemic treatment up to 24 months, although the mean WAI score drew close to matched population controls at 24 months (Figure 2B, Table 2B). The mean WAI score for patients receiving surgery only increased over time and returned towards matched population controls levels at 18 months (Figure 2B, Table 2B). Statistically significant differences over time and between treatment strategies were seen from baseline up to 18 months (Table 4). Adjuvant systemic therapy ($p<0.01$), a secondary vocational educational level or lower ($p=0.02$) and one or more comorbidities ($p<0.01$) (Table 4 and 5) were negatively associated with lower WAI scores.

Sick leave

Patients with stage I-III disease who received surgery reported significantly less sick leave days up to 24 months compared to patients treated with surgery and adjuvant systemic treatment. Sick leave of 100 days or more was reported less often for patients receiving surgery versus patients receiving surgery with adjuvant systemic treatment: 2% vs. 6% at baseline ($p=0.01$), 26% vs. 48% at 3 months ($p=0.01$), 27% vs. 71% at 6 months ($p<0.01$),

Table 3. Estimated association of patient factors with WAI scores for patients with stage I-IV colon cancer.

Parameter	Estimate	SE	p-value	95% Confidence Interval	
Intercept	31.7	5.1	<0.01	21.8	41.7
Disease stage					
Stage I	3.3	3.3	0.31	-3.2	9.9
Stage II	3.5	3.2	0.28	-2.9	9.9
Stage III	6.9	3.4	0.04	0.2	13.6
Stage IV	Ref				
Sex					
Male	2.1	0.7	0.01	0.6	3.5
Female	Ref				
Age					
years	-0.0	0.1	0.54	-0.2	0.1
Educational level*					
Low	-1.2	0.7	0.09	-2.7	0.2
High	Ref				
Living situation					
Alone	3.7	3.2	0.26	-2.7	10.0
Married/partner	4.4	3.1	0.16	-1.8	10.5
With children	7.6	3.5	0.03	0.64	14.54
Married/partner with children	4.53	3.07	0.14	-1.53	10.58
Others	Ref				
Systemic therapy					
No systemic therapy	4.4	1.3	<0.01	1.9	6.9
Systemic therapy	Ref				
Surgery					
No surgery	-6.9	2.1	<0.01	-10.9	-2.8
Surgery	Ref				
Work demands					
Cognitive work demand	1.3	0.6	0.03	0.1	2.5
Physical work demand	-0.4	0.8	0.64	-2.0	1.2
Combination of both	Ref				
Comorbidities					
≥ 4 comorbidities	-8.1	0.7	<0.01	-9.5	-6.8
2 or 3 comorbidities	-5.8	0.5	<0.01	-6.8	-4.8
1 comorbidity	-3.2	0.5	<0.01	-4.1	-2.2
No comorbidities	Ref				

* Low educational level: secondary vocational education or lower. High educational level: Bachelor degree or higher.

17% vs. 49% at 12 months (p<0.01), 4% vs. 21% at 18 months (p< 0.01) and 2% vs. 12% at 24 months (p=0.2) (Figure 3A and 3B). In matched-controls, sick leave of 100 days or more was only reported in 4% of cases. Patients with stage IV disease reported similar sick leave days compared to patients with curative disease receiving systemic therapy (p=0.12) (Figure 3C).

Table 4. Difference in mean estimated WAI between treatments, surgery only compared to surgery with adjuvant systemic treatment (stage I-III disease).

Parameter	Estimate	SE	p-value	95% Confidence Interval	
Time					
Baseline	2.9	1.1	0.01	0.6	5.2
3 months	6.7	1.0	<0.01	4.6	8.8
6 months	5.6	1.0	<0.01	3.7	7.5
12 months	2.7	0.9	<0.01	0.9	4.9
18 months	2.2	1.1	0.05	0.1	4.4
24 months	1.8	1.8	0.33	-1.9	5.4

Table 5. Estimated association of patient factors on mean WAI scores for disease stage I-III.

Parameter	Estimate	SE	p-value	95% Confidence Interval	
Intercept	39.0	4.3	<0.01	30.5	47.5
Time					
Baseline	-7.1	1.5	<0.01	-10.1	-4.2
3 months	-10.8	1.4	<0.01	-13.7	-8.0
6 months	-7.3	1.4	<0.01	-10.1	-4.5
12 months	-2.6	1.4	0.07	-5.3	0.2
18 months	-1.5	1.3	0.28	-4.2	1.2
24 months	Ref				
Sex					
Male	1.3	0.8	0.08	-0.1	2.8
Female	Ref				
Age					
years	-0.0	0.1	0.70	-0.1	0.1
Educational level*					
Low	-1.8	0.7	0.02	-3.2	-0.3
High	Ref				
Living situation					
Alone	2.8	3.1	0.36	-3.3	9.0
Married/partner	4.0	3.0	0.18	-1.9	10.0
With children	7.7	3.5	0.03	0.9	14.5
Married/partner with children	4.1	3.0	0.17	-1.8	10.0
Others	Ref				
Work demands					
Cognitive work demand	1.1	0.6	0.09	-0.2	2.3
Physical work demand	-0.4	0.8	0.62	-2.0	1.2
Combination of both	Ref				
Comorbidities					
≥ 4 comorbidities	-7.8	0.7	<0.01	-9.21	-6.42
2 or 3 comorbidities	-5.6	0.5	<0.01	-6.56	-4.65
1 comorbidity	-3.2	0.5	<0.01	-4.14	-2.33
No comorbidities	Ref				

* Low educational level: secondary vocational education or lower. High educational level: Bachelor degree or higher.

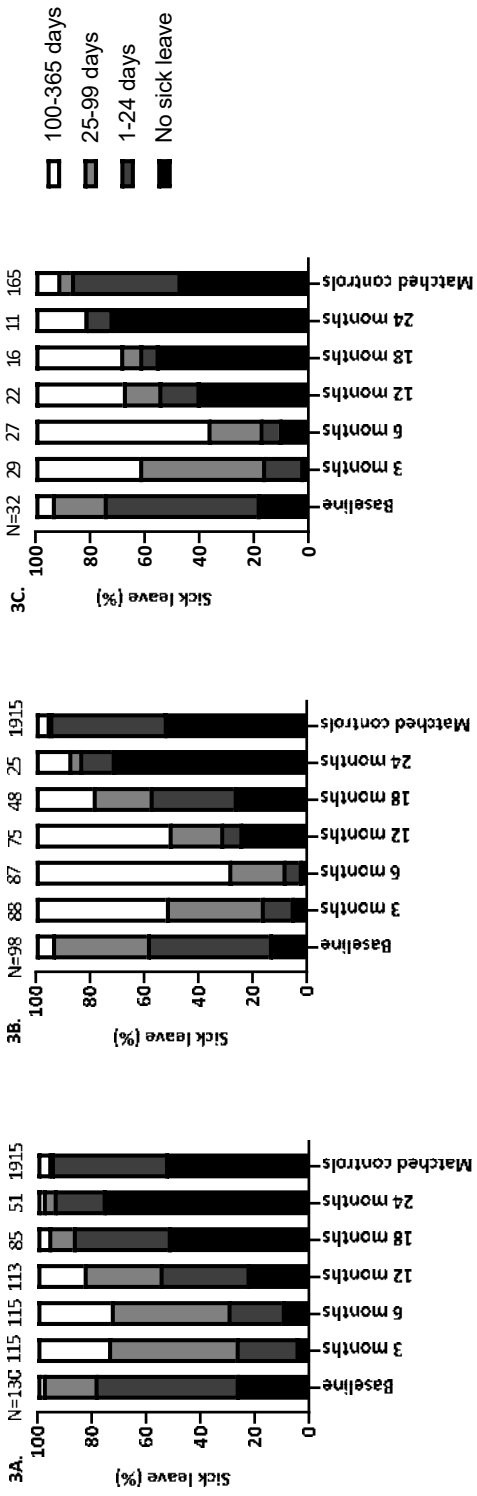


Figure 3. Sick leave score categories for patients with a curative intent (stage I-III) either receiving surgery only (figure 4A), surgery and adjuvant systemic treatment (figure 3B) and for patients with stage IV disease (figure 3C).

DISCUSSION

We demonstrated that at the time of colon cancer diagnosis, work ability is significantly reduced compared to matched population controls without cancer (mean WAI score 31±8.2 and 41±5.6 respectively, p<0.01). Stage I disease patients demonstrated the highest WAI scores over time, regaining scores towards matched population controls after one year. Patients with stage I-III disease, who were treated with surgery but did not receive adjuvant systemic treatment, regained work ability comparable to matched population controls at 18 months. WAI scores initially declined for patients with stage II-III disease receiving adjuvant systemic treatment. From 6 months onwards, work ability gradually improved. Nevertheless, work ability for patients receiving adjuvant systemic treatment remained lower compared to patients without adjuvant systemic treatment for a prolonged period of time. Work ability comparable to matched population controls was just reached at 24 months. In contrast, patients with stage IV disease did not demonstrate improvements in WAI scores over time.

In our study, systemic treatment and having ≥1 comorbidities were associated to lower WAI scores for all stages of disease. Moreover, a treatment without surgery and female sex were associated with lower WAI scores, which did not remain statistically significant if stage IV was excluded. In patients treated with curative intent (stage I-III), a lower educational level was associated with lower WAI scores. Previously, disease stage, sick leave prior to the diagnosis, unemployment and children aged under 18 were also identified as risk factors to receive a disability pension.[8,14] Systemic treatment as risk factor for poorer work ability outcomes has also previously been demonstrated by others.[11–14] In this study, disease stage was a significant risk factor only if systemic treatment was not included in the mixed model. Thus, disease stage might not be a risk factor in itself, but depends on the addition of systemic treatment to surgery, which is more likely to occur at higher disease stages. We lack information on sick leave or unemployment prior to the diagnosis in this study, and could therefore not assess the influence of these parameters on work ability. Living with children was also associated with higher WAI scores in this study, but only when living without spouse.

Recently, Couwenberg *et al*. have reported results for patients with rectal cancer receiving a treatment with curative intent. Compared to our study, patients reported higher baseline WAI scores and reported a more important decrease in WAI scores during treatment. This might be a result of more extensive surgery in rectal cancer compared to colon cancer. A similar time course in recovery was seen with WAI scores returning to matched population control levels 18 months after an initial decline up to 6 months after treatment initiation.[13]

There are some limitations to our study. Firstly, since not all PLCRC participants consented to complete PROMs. We cannot exclude that this selection influences our results. It has previously been shown that non-participants to PROMs have lower survival

rates and lower estimated health-related quality of life[24]. Also, reduced quality of life has been associated with reduced work ability in cancer patients.[1] This may have resulted in an overestimation of work ability in our study. Secondly, the response rates to PROMs in this study diminished over time, which is partly explained by the median follow-up time of 14 months. Thus, not all patients have completed all follow-up assessments at the time of analysis. Nevertheless, the most pronounced longitudinal improvements on work ability might still have been captured. Also, it has been shown previously that sick leave for CRC patients diminished to a nadir at 12 months and then remains constant or improves.[8,12,13] Nonetheless, longer follow-up is necessary to further evaluate work ability and sick leave for patients without disease recurrence. Finally, during the course of the current inclusion period, the advice on the duration of systemic combination treatment with capecitabine (a fluoropyrimidine) and oxaliplatin for the majority of patients with high risk stage II and stage III colon cancer has been reduced from 6 to 3 months,[25,26] which is likely to result in a swifter regain of work ability.

CONCLUSION

In conclusion, colon cancer patients experience prolonged negative effects on their self-estimated work ability, which was most pronounced for patients receiving systemic treatment. Our results may contribute to improvements in counselling patients on work ability and sick leave during colon cancer treatment.

ACKNOWLEDGEMENTS

The authors would like to thank Peter Langendijk from "Stichting Blik op Werk" for providing WAI questionnaire outcomes for the purposes of population control matching in this analysis.

REFERENCES

1. Duijts SFA, Kieffer JM, van Muijen P, van der Beek AJ. Sustained employability and health-related quality of life in cancer survivors up to four years after diagnosis. *Acta Oncol (Madr)*. 2017;56(2):174-182. doi:10.1080/0284186X.2016.1266083
2. Mehnert A. Employment and work-related issues in cancer survivors. *Crit Rev Oncol / Hematol*. 2011;77(2):109-130. doi:10.1016/j.critrevonc.2010.01.004
3. Syse A, Tretli S, Kravdal Ø. Cancer's impact on employment and earnings – a population-based study from Norway. *J Cancer Surviv*. 2008;2:149-158. doi:10.1007/s11764-008-0053-2
4. Heinesen E, Imai S, Maruyama S. Employment, job skills and occupational mobility of cancer survivors. *J Health Econ*. 2018;58. doi:10.1016/j.jhealeco.2018.01.006
5. Van Der Noordt M, IJzelenberg H, Droomers M, Proper KI. Health effects of employment: A systematic review of prospective studies. *Occup Environ Med*. 2014;71(10):730-736. doi:10.1136/oemed-2013-101891
6. Van Rijn RM, Carlier BE, Schuring M, Burdorf A. Work as treatment? the effectiveness of re-employment programmes for unemployed persons with severe mental health problems on health and quality of life: A systematic review and meta-analysis. *Occup Environ Med*. 2016;73(4):275-279. doi:10.1136/oemed-2015-103121
7. Beesley VL, Vallance JK, Mihala G, Lynch BM, Gordon LG. Association between change in employment participation and quality of life in middle — aged colorectal cancer survivors compared with general population controls. *Psychooncology*. 2017;26:1354-1360. doi:10.1002/pon.4306
8. Hauglann BK, Benth JŠ, Fosså SD, Tveit KM, Dahl A. A controlled cohort study of sickness absence and disability pension in colorectal cancer survivors. *Acta Oncol (Madr)*. 2014;53(6):735-743. doi:10.3109/0284186X.2013.844354
9. Rottenberg Y, Ratzon NZ, Cohen M, Hubert A, Uziely B, de Boer AGEM. Unemployment risk at 2 and 4 years following colorectal cancer diagnosis: a population based study. *Eur J Cancer*. 2016;69:70-76. doi:10.1016/j.ejca.2016.09.025
10. Hauglann B, Benth JŠ, Fosså SD, Tveit KM, Dahl AA. A controlled cohort study of long-term income in colorectal cancer patients. *Support Care Cancer*. 2014;22:2821-2830. doi:10.1007/s00520-014-2258-4
11. Bains M, Munir F, Yarker J, et al. The impact of colorectal cancer and self-efficacy beliefs on work ability and employment status : a longitudinal study. *Eur J Cancer Care*. 2012;21:634-641. doi:10.1111/j.1365-2354.2012.01335.x
12. de Boer AGEM, Verbeek JHAM, Spelten ER, et al. Work ability and return-to-work in cancer patients. *Br J Cancer*. 2008;98(8):1342-1347. doi:10.1038/sj.bjc.6604302
13. Couwenberg AM, Intven MPW, Gregorowitsch ML, Haaring C, Van Grevenstein W, Marieke Verkooijen H. Patient-Reported Work Ability during the First Two Years after Rectal Cancer Diagnosis. *Dis Colon Rectum*. 2020;63(5):578-587. doi:10.1097/DCR.0000000000001601
14. Den Bakker CM, Anema JR, Zaman ACGNM, et al. Prognostic factors for return to work and work disability among colorectal cancer survivors; A systematic review. *PLoS One*. 2018;13(8):1-18. doi:10.1371/journal.pone.0200720
15. Araghi M, Soerjomataram I, Bardot A, et al. Changes in colorectal cancer incidence in seven high-income countries: a population-based study. *Lancet Gastroenterol Hepatol*. 2019;4(7):511-518. doi:10.1016/S2468-1253(19)30147-5
16. Siegel RL, Miller KD, Fedewa SA, et al. Colorectal cancer statistics, 2017. *CA Cancer J Clin*. 2017;67(3):177-193. doi:10.3322/caac.21395
17. Burbach JPM, Kurk SA, Coebergh van den Braak RRJ, et al. Prospective Dutch colorectal cancer cohort: an infrastructure for long-term observational, prognostic, predictive and (randomized) intervention research. *Acta Oncol (Madr)*. 2016;55(11):1273-1280. doi:10.1080/0284186X.2016.1189094
18. Van De Poll-Franse L V., Horevoorts N, Eenbergen M Van, et al. The Patient Reported Outcomes Following Initial treatment

and Long term Evaluation of Survivorship registry: Scope, rationale and design of an infrastructure for the study of physical and psychosocial outcomes in cancer survivorship cohorts. *Eur J Cancer*. 2011;47(14):2188-2194. doi:10.1016/j.ejca.2011.04.034

19. Tuomi K, Ilmarinen J, Jhakola A, Katajarinne L, Tulkki A. Work Ability Index. In: *Helsinki: Finnish Institute of Occupational Health*. 2nd revise. ; 1998.
20. Blik op werk. *Work Ability Index (WAI), Werkwijzer.*; 2008.
21. Cnaan A, Laird NM, Slasor P. Mixed Models: Using the General Linear Mixed Model to Analyse Unbalanced Repeated Measures and Longitudinal Data. *Tutorials Biostat*. 2005;2:127-158. doi:10.1002/0470023724.ch1c(i)
22. Bonnetain F, Fiteni F, Efficace F, Anota A. Statistical challenges in the analysis of health-related quality of life in cancer clinical trials. *J Clin Oncol*. 2016;34(16):1953-1956. doi:10.1200/JCO.2014.56.7974
23. CBS. Statline. Published 2019. Accessed November 14, 2019. https://opendata.cbs.nl/statline/#/CBS/nl/dataset/82309NED/table?dl=2A99E
24. de Rooij BH, Ezendam NPM, Mols F, et al. Cancer survivors not participating in observational patient-reported outcome studies have a lower survival compared to participants: the population-based PROFILES registry. *Qual Life Res*. 2018;27(12):3313-3324. doi:10.1007/s11136-018-1979-0
25. ESMO. eUpdate: Early Colon Cancer Treatment Recommendations. Published 2019. Accessed October 24, 2019. eupdate: Early Colon Cancer Treatment Recommendations
26. Grothey A, Sobrero AF, Shields AF, et al. Duration of adjuvant chemotherapy for stage III colon cancer. *N Engl J Med*. 2018;378(13):1177-1188.

* PLCRC STUDY GROUP

Geerard L. Beets, Department of Surgery, Nederlands Kanker Instituut - Antoni van Leeuwenhoek Hospital, Amsterdam; *Eric J.Th. Belt*, Department of Surgery, Albert Schweitzer Hospital, Dordrecht; *Maaike Berbée*, Department of Radiotherapy, Maastro Clinic, Maastricht; *Frederique H. Beverdam*, Department of Surgery, Franciscus Gasthuis & Vlietland Hospital, Schiedam; *Ruud Blankenburgh*, Department of Medical Oncology, Saxenburgh Hospital, Hardenberg; *Peter Paul L.O. Coene*, Department of Surgery, Maasstad Hospital, Rotterdam; *Jan Willem B. de Groot*, Department of Medical Oncology, Isala Oncology Center, Zwolle; *Ignace H.J.T. de Hingh*, Department of Surgery, Catharina Hospital, Eindhoven; *Aad I. de Vos*, Department of Medical Oncology, Van Weel-Bethesda Hospital, Dirksland; *Johannes H.W. de Wilt*, Department of Surgery, Radboud University Medical Center, Nijmegen; *Jan Willem T. Dekker*, Department of Surgery, Reinier de Graaf Hospital, Delft; *Frans L.G. Erdkamp*, Department of Medical Oncology, Zuyderland Hospital, Heerlen; *Annebeth W. Haringhuizen*, Department of Medical Oncology, Ziekenhuis Gelderse Vallei, Ede; *Helgi H. Helgason*, Department of Medical Oncology, Haaglanden Medical Center, Den Haag; *Mathijs P. Hendriks*, Department of Medical Oncology, Northwest Clinics, Alkmaar; *Ronald Hoekstra*, Department of Medical Oncology, Ziekenhuisgroep Twente, Hengelo; *Jan N.M. Ijzermans*, Department of Surgery, Erasmus MC, University Medical Center Rotterdam, Rotterdam; *Jan Jansen*, Department of Surgery, Admiraal de Ruyter Hospital, Goes; *Frank W.H. Kloppenberg*, Department of Surgery, Treant Hospital, Emmen; *Maartje Los*, Department of Medical Oncology, St. Antonius Hospital, Nieuwegein; *Martijn R. Meijerink*, Department of Radiation and Nuclear medicine, Amsterdam University Medical Centre, Amsterdam; *Leonie J.M. Mekenkamp*, Department of Medical Oncology, Medisch Spectrum Twente, Enschede; *Peter Nieboer*, Department of Medical Oncology, Wilhelmina Hospital, Assen; *Koen C.M.J. Peeters*, Department of Surgery, Leiden University Medical Center, University of Leiden, Leiden; *Natascha A.J.B. Peters*, Department of Medical Oncology, Sint Jans Hospital, Weert; *Johannes F.M. Pruijt*, Department of Medical Oncology, Jeroen Bosch Hospital, Den Bosch; *Patricia Quarles van Ufford-Mannesse*, Department of Medical Oncology, Haga Hospital, Den Haag; *Ron C. Rietbroek*, Department of Medical Oncology, Rode Kruis Hospital, Beverwijk; *Anandi H.W. Schiphorst*, Department of Surgery, Diakonessenhuis Hospital, Utrecht; *Mark P.S. Sie*, Department of Medical Oncology, ZorgSaam Hospital, Terneuzen; *Lieke H.J. Simkens*, Department of Medical Oncology, Maxima Medical Center, Veldhoven; *Dirkje W. Sommeijer*, Department of Medical Oncology, Flevo Hospital, Almere **AND** Department of Medical Oncology, Amsterdam University Medical Center, University of Amsterdam, Amsterdam; *Leontine E.A. Spierings*, Department of Medical Oncology, Alrijne Hospital, Leiderdorp; *Hein B.A.C. Stockmann*, Department of Surgery, Spaarne Hospital, Haarlem; Koen Talsma, Department of Surgery, Deventer Hospital, Deventer; *Albert J. ten Tije*, Department of Medical Oncology,

Amphia ziekenhuis, Breda; *Frederiek Terheggen*, Department of Medical Oncology, Bravis Hospital, Roosendaal; *Manuel L.R. Tjin-A-Ton*, Department of Medical Oncology, Rivierenland Hospital, Tiel; *Liselot B.J. Valkenburg-van Iersel*, Department of Medical Oncology, Maastricht University Medical Center, Maastricht; *Hester van Cruijsen*, Department of Medical Oncology, Antonius Hospital, Sneek; *Ankie M.T. van der Velden*, Department of Medical Oncology, Tergooi Hospital, Hilversum; *Joyce M. van Dodewaard-de Jong*, Department of Medical Oncology, Meander Medical Center, Amersfoort; *Anja U.G. van Lent*, Department of Gastroenterology and Hepatology, OLVG Hospital, Amsterdam; *Theo van Voorthuizen*, Department of Medical Oncology, Rijnstate Hospital, Arnhem; *Maarten Vermaas*, Department of Surgery, Ijsselland Hospital, Capelle aan den Ijssel; *Wouter J. Vles*, Department of Surgery, Ikazia Hospital, Rotterdam; *F.Jeroen Vogelaar*, Department of Surgery, Viecuri Hospital, Venlo; *David D.E. Zimmerman*, Department of Surgery, Elisabeth-TweeSteden Hospital, Tilburg. Al

7

SUMMARY

Colorectal cancer (CRC) is the third most common cancer type worldwide. Approximately 20-25% of patients have (synchronous) metastatic CRC (mCRC) at the time of diagnosis and about 20% of patients with non-metastatic disease, will eventually develop (metachronous) metastatic disease. Treatment decisions depend on disease stage. Curative treatment for local disease and locally advanced disease (stage I-III) can be achieved with surgery. The addition of adjuvant chemotherapy in high-risk stage II and all stage III disease further improves survival rates. However, about 50% of patients will not benefit from adjuvant chemotherapy as these patients will not have disease recurrence after surgery only. This underscores the necessity for biomarkers to identify patients who may benefit of adjuvant treatment. For mCRC, the development of novel treatment strategies has resulted in improvements in progression-free survival and overall survival, at least for patients participating in clinical trials or who meet the original inclusion criteria of the trials in which the safety and efficacy of these drugs was evaluated. Despite the advances achieved in the treatment of CRC, healthcare expenses overall and expenses on targeted therapy continue to rise. Our main goal in the treatment of CRC remains to provide patients with the most efficacious treatment while retaining quality of life. However, the impact of treatment choices on healthcare resources is increasingly a matter of consideration, while the interpretation of cost-effectiveness outcomes is complex.

Chapter 2 and Chapter 2B illustrate the challenges in the interpretation of cost-effectiveness studies using mCRC as an example. Despite the reporting guidelines on cost-effectiveness, one should be aware that results are influenced by the assumptions used in a cost-effectiveness model. Cost-effectiveness models do not only include clinical outcomes (time to progression, survival, adverse events rates), but also many assumptions such as costs of adverse events treatments, duration of subsequent treatment, costs related to subsequent treatments, etc. These assumptions can therefore have important influence on model outcomes. Nevertheless, randomized controlled trials remain the most suitable for evaluation of cost-effectiveness for institutional budget impact estimations due thorough registration of events. However, a more appropriate evaluation of cost-effectiveness is based on patient cohort studies including real-world data on efficacy in the CRC patient population, quality of life and costs in order to assess a more realistic impact of choices for the society and improve healthcare resource allocations.

In Chapter 3, the cost-effectiveness of capecitabine-bevacizumab (CAP-B) maintenance compared to an observation strategy in mCRC patients achieving stable disease or better following six cycles of first-line capecitabine, oxaliplatin and bevacizumab (CAPOX-B) is presented. Model assumptions were based on the clinical outcomes of the randomized controlled phase 3 CAIRO3 study (n=558 patients), additional retrospective data collection on resource utilisation (n=146 patients from 6 different hospitals participating in the CAIRO3 study), literature and expert opinion. Cost-effectiveness was evaluated

using a deterministic state-transition model (STM) or Markov model. We have shown that CAP-B maintenance results in a mean gain of 0.21 quality of life adjusted life years (QALYs) at a mean incremental cost of €36,845, yielding a mean incremental-cost effectiveness ratio (ICER) of €175,452 per QALY gained. Patients achieving partial response or complete response have a more favourable mean ICER of €149,300 per QALY gained; mostly resulting from a larger gain in mean QALYs (0.31). Even though (inter)national cost-effectiveness thresholds are not strictly adhered to; we conclude that this treatment strategy is not cost-effective.

In Chapter 4, we compare the STM described in Chapter 3 with a Discrete Event Simulation (DES) model using the same model assumptions as reported in Chapter 3. This was done as STMs incur a rigidity towards time until clinical events, such as time to progression, and have a higher risk of irregularities in state transition probabilities. We demonstrated that indeed time-dependent transition probabilities in the STM were irregular and more sensitive to single events, while the DES model generated event transitions more naturally and slightly more accurately. The models resulted in a mean QALY gain of 0.21 and 0.18 for the STM and DES model, respectively. As a result of different time-to-event predictions, small differences in incremental costs were also seen. Similar mean ICERs were calculated for both the STM and DES model, €172,443 and €168,383 per QALY gained for the DT-STM and DES model, respectively. Nevertheless, DES models are expected to yield a more accurate ICER estimation due to the more natural distribution of time to events.

An often-reported primary outcome of cost-effectiveness models is cost per QALY gained. Generic questionnaires, such as EQ-5D, are used to calculate health-related utilities, which are subsequently used to calculate QALYs gained (or lost) in a cost-effectiveness model. However, disease specific questionnaires such as the EORTC QLQ-C30, are commonly preferred in oncology studies instead of generic questionnaires as disease specific questionnaires are deemed more sensitive to disease related changes in quality of life. In absence of generic questionnaire, a mapping algorithm suitable for translating QLQ-C30 outcomes towards EQ-5D-3L utilities can be used to enable cost-effectiveness analyses. In Chapter 5, we evaluate the performance of three existing mapping algorithms and six newly developed mapping algorithms to convert QLQ-C30 functional scores toward EQ-5D-3L utilities. Furthermore, we evaluated the difference on ICER point estimates using observed EQ-5D-3L utilities and mapped utilities using the mCRC cost-effectiveness model reported in Chapter 4. Two of the three selected algorithms yielded mean utilities which were significantly different from the mean utility (0.83± 0.17 with Dutch tariff) observed in the randomized, controlled phase 3 CAIRO3 study. All newly developed models yielded predicted utilities drawing close to observed utilities; differences were not statistically significant. In a cost-effectiveness model, comparable outcomes were found when using mean observed utilities and mean mapped

utilities. Mapping resulted in an ICER point estimate of €5,094 more per QALY gained compared to the observed utility ICER point (€168,048). The existing algorithms resulted in an ICER difference of €10,140 less and €1,765 more compared to the observed utility based ICER. Disparity was explained by minimal differences in incremental QALYs between model calculations with the mapped and observed utilities. Thus, mapping of QLQ-C30 onto EQ-5D-3L based utilities, with either existing or the newly developed algorithms, is an appropriate alternative for cost-effectiveness analyses if used for patients with comparable disease severity and not per se similar cancer type.

In Chapter 6, we evaluate the work ability of patients with colon cancer in the first two years after diagnosis. Results of 390 patients aged <67 years with stage I-IV disease and included in the Prospective Dutch Colorectal Cancer Cohort (PLCRC: prospectief Landelijk CRC cohort), and who consented in the completion of patient reported outcome questionnaires, were analysed. The work ability index (WAI) score in patients with stage I-IV colon cancer was reduced upon diagnosis compared to matched population-controls without cancer (mean work ability score: 31±SD 8 and 41±SD 6). By using a mixed linear model, we found that for patients with curative surgery only (stage I-III), WAI scores returned towards the mean score observed in matched population-controls at 18 months. Patients treated with adjuvant systemic treatment (stage I-III) regained WAI scores comparable to matched case-controls after 2 years. No improvements in WAI scores were observed in patients with stage IV disease. WAI scores were negatively influenced by the administration of systemic treatment and ≥1 comorbidities. The WAI scores were significantly lower in females, although this did not remain statistically significant if patients with stage IV disease were excluded. Lower educational levels (secondary vocational educational or lower) were associated with lower WAI scores only in patients with stage I-III disease. After 24 months, sick leave of 100 days or more, was reported in 12% and 2% of patients with stage I-III disease receiving surgery with additional systemic treatment reported or receiving surgery only, respectively (p=0.2). In matched population-controls, sick leave of 100 days or more was reported in 4% of cases. Thus, colon cancer patients experience prolonged negative effects on their self-estimated work ability. These results may contribute to improvements in patient counselling regarding work ability during the course of the treatment and for prolonged time thereafter.

8

GENERAL DISCUSSION

Despite the advances achieved in the treatment of CRC, healthcare costs continue to rise. This is only partly accountable to expensive targeted drugs, about 56% of cancer costs is a resultant of hospital inpatient care.[1,2] Previously, no differences in cancer death rates were seen between European countries with higher or lower healthcare expenditures.[3] More recently even a negative correlation between incremental costs spent on novel cancer therapies and benefit was demonstrated, suggesting that more expensive drugs did not result in clinical benefit.[4] These findings fuel the debate on increasing healthcare costs. Important steps towards the standardization of clinical benefit evaluations have been undertaken by the American Society of Clinical Oncology (ASCO) and the European Society for Medical Oncology (ESMO) as both organisations developed clinical benefit evaluation tools.[5,6] However, both value tools lack a standardized cost-effectiveness evaluation, although the ASCO Clinical benefit value tool does report on drug costs per months. Following approval of a novel medicinal product by the European Medicines Agency (EMA), "Zorginstituut Nederland" evaluates clinical benefit along with the cost-effectiveness prior to reimbursement in attempt to slow down the increase in healthcare expenses.[7] The committee for evaluation of systemic cancer treatment of the Dutch Society for Medical Oncology (NVMO) use the 'PASKWILL' criteria to evaluate clinical benefit and report on treatment costs, although costs are not included in the final recommendation.[8] An additional challenge in optimizing healthcare resource allocation, is the lack of international consensus on cost-effectiveness thresholds. The defined reference thresholds[9] in the Netherlands are not adhered to, resulting in a lack of incentive to reduce prices by pharmaceutical companies.

In this thesis, we have shown that there are important limitations regarding the interpretation of cost-effectiveness studies towards daily practice. Differences can arise at many different levels, such as choice of cost-effectiveness model methodology, dataset (i.e. randomized controlled trial, observational cohort, etc.), number of health-states modelled (i.e. progressive disease, survival, death), cost perspectives (i.e. institutional, societal) and assumptions based on expert opinion or literature for unavailable information in the dataset to be analysed.[10] Hereafter, we elaborate more on specific challenges regarding cost-effectiveness addressed in this thesis.

Cost-effectiveness evaluations should be tailored to the patient population of interest and not to a "one size fits all" strategy to optimize decisions on resource allocation. However, there is still a need for better predictive factors in CRC to guide treatment decisions and a uniform treatment approach is still mostly adopted. The impact of a uniform approach on cost-effectiveness can be illustrated by the cost-effectiveness study based on the randomized clinical phase 3 CAIRO3 study in mCRC patients, presented in this thesis. Patients with stable disease or better after six cycles of capecitabine, oxaliplatin and bevacizumab were randomized to either receive capecitabine-bevacizumab (CAP-B) maintenance or observation. Patients receiving CAP-B maintenance

demonstrated a significant improvement in the primary outcome, second progression-free survival (PFS2), 11.6 months versus 8.6 months (95% CI 0.53-0.77) for the observation strategy.[11] Despite this clinical benefit, we have shown that capecitabine-bevazicumab maintenance was not regarded cost-effective in a state-transition model (STM). Later, updated study results and subgroup analyses based on mutational status demonstrated the most pronounced clinical benefit in the *RAS/BRAF* wild-type subgroup, both PFS2 (13.3 and 9.0 months for respectively CAP-B maintenance and observation, 95% CI 0.39-0.84) and overall survival (25.7 and 19.0 months for respectively CAP-B maintenance and observation, 95% CI 0.46-1.00).[11] Hypothetically, it may be possible that CAP-B maintenance treatment would have been deemed cost-effective for the *RAS/BRAF* wild-type patient subgroup due to the larger clinical benefit, although treatment duration with resultant costs was most likely also increased. Others have previously demonstrated that anti-EGFR containing systemic therapy resulted in additional costs compared to bevacizumab containing systemic therapy in *KRAS* wild-type patients[12–14] and therefore bevacizumab containing therapy might be preferred. This example illustrates that in the future predictive markers will not only provide in the most optimal therapy choices for CRC patients, but also contribute in improved allocation of healthcare resources. Additionally, bevacizumab will soon be out of patent, which will reduce costs. A new analysis to determine whether the strategy with biosimilar VEGF antibodies is deemed cost-effective will be necessary for mCRC patients and/or defined subgroups, such as *RAS/BRAF* wild-type patients.

A frequently applied method for cost-effectiveness studies is state-transition modelling (STM).[15] A challenge to approximate clinical effects of treatment strategy is the duration of a health-state, as transition from one health-state to the other are defined by the duration of a fixed health state cycle length.[15] In contrast, discrete event simulation (DES) models allow clinical events to occur at any time instead of fixed cycle length.[16] We have demonstrated that time to events are predicted more smoothly in a DES model based compared to a previously developed STM based on the CAIRO3 study, although this did not result in large differences in cost-effectiveness outcomes. These results confirm, what has previously also been shown by others: if an STM and DES model use the same clinical evidence and assumptions, this does not result in relevant differences in cost-effectiveness outcomes.[17,18] However, for reasons of comparison, parameter uncertainty in time to event distributions for health-state transitions was not introduced in the DES model.[19] The impact of parameter uncertainty such as time to events, can importantly influence the uncertainty around the mean point estimate (95% confidence interval ellipses) of the ICER. This uncertainty should therefore be accounted for in the probabilistic sensitivity analysis of a model. This is especially important for patient cohorts with small sample sizes. Disregarding parameter uncertainty might result in an overall underestimation of the uncertainty around the mean ICER point estimate or cost-effectiveness. This itself could incur inappropriate allocation of healthcare resources.[20]

Therefore, DES modelling seems preferable over STM, when clinical patient-level data is available, in order to represent the course of a treatment strategy with subsequent events more naturally, including parameter uncertainty.

Incremental costs per quality of life adjusted life years (QALY) gained, is a frequently used primary outcome in cost-effectiveness studies.[21] Generic quality of life questionnaires, such as the EQ-5D, are used to calculate health-related quality of life utilities and subsequently enable clinical benefit and quality of life to be expressed in QALYs.[22] Preference values from the general populations are obtained and benchmarked to country specific outcomes to make the EQ-5D suitable for economic evaluations.[23,24] It has been argued previously that generic quality of life questionnaires, such as the EQ-5D are insensitive to disease specific aspects. This is why clinicians prefer disease specific questionnaires, such as the QLQ-C30. When the QLQ-C30 was benchmarked to preference values suitable for economic evaluations, it has been demonstrated that the EQ-5D-3L and the QLQ-C30 did not result in equal differences in utilities.[25] Therefore, if generic EQ-5D outcomes are lacking, one could choose to estimate the EQ-5D based utility by mapping the QLQ-C30 questionnaires. We have demonstrated that mapping is an adequate solution to enable cost-effectiveness analyses, although this also incurs additional uncertainty in the primary outcome (ICER in terms of incremental costs per QALY). Of note, the EQ-5D-5L has been introduced since 2009. The five level EQ-5D form is more discriminative in the detection of mild and severe health state decrements compared to the EQ-5D-3L.[26]

In the cost-effectiveness study based on the CAIRO3 study, surprisingly high utilities of 0.84 for capecitabine-bevacizumab maintenance and 0.83 for observation were observed.[27] Although, these findings were in line with previously reported utilities[28,29], it is possible that utilities are overestimated and thus introduce additional uncertainty in cost-effectiveness outcomes. Patients who discontinue the completion of patient reported outcomes (PROMs) have worse outcomes on multiple QLQ-C30 functional scale scores compared to study participants without drop-out.[30] Attrition in longitudinal studies and randomized clinical trials has been associated with increased symptom burden. This is important as attrition was seen in up to 25-30% of patients.[31–33] Non-participants to PROMs in an observational cohort also show lower survival rates and lower estimated health-related quality of life.[34] Moreover, it has previously been shown that the introduction incremental costs per QALYs to cost-effectiveness analyses compared to incremental costs per life years gained, did not substantially alter the estimated cost-effectiveness of an intervention in most cases.[35] Thus, with the additional uncertainty around utility estimates, it might be more correct to report outcomes with ICERs based on incremental costs per LYs gained. Decrements in quality of life as a result of adverse treatment effects are accounted for in cost-effectiveness models, by the inclusion of probabilities of adverse effects and related costs without the possible introduction of utility overestimation.

Finally, to retain a sustainable healthcare system, it is important to value cost-effectiveness from a societal perspective. A CRC diagnosis and treatment strategies result in prolonged reduced work ability. If treatment strategies would result in improved return to work rates and/or reduced time in return to work, an intervention may be more cost-effective from a societal point of view than if only viewed from an institutional or third party payer's perspective. Also, a societal perspective in cost-effectiveness analyses seems preferable for more effective allocation of healthcare resources.[36] A major caveat is the availability of information on societal health effects and costs resulting from an intervention as clinical studies are generally not designed to include such information.

FUTURE PERSPECTIVES AND CONCLUSION

Clinical studies will increasingly focus on molecular subgroups in CRC patients with the availability of predictive markers for treatment efficacy. This could incur important increases in costs related to clinical studies and difficulties in patient recruitment. To enable further development of targeted treatment modalities in CRC, novel study designs are necessary. In the Netherlands, the Prospective Dutch Colorectal Cancer Cohort (PLCRC) was developed to gather clinical data, biomaterial and PROMs of CRC patients (all disease stages) in a longitudinal manner, from primary diagnosis until death. In addition, the PLCRC serves as an infrastructure to conduct simultaneous (randomized controlled) studies.[37] The PLCRC study not only provides clinical and PROM outcomes in an unselected CRC population (unlike clinical studies), but also enables an integrated solution for clinical studies in molecular subgroups using a Trials within Cohort (TwiCs) design, allowing for randomization within the cohort. Ideally, the PLCRC should also harbour a foundation of cost-effectiveness evaluations from a societal perspective in the future, although some challenges are present. For instance, work ability is not assessed in patients with unpaid work, such as volunteer work, and no information on for instance return to work, time to return to work and income reduction as a result of the CRC diagnosis, is available. This complicates the estimation of cost-effectiveness for patient tailored treatment choices from a societal perspective. Still, to date, this is the most optimal study design to evaluate treatment efficacy in an unselected CRC patient population and which can thus improve healthcare resource allocations in the future. Although not addressed in this thesis, improving the efficiency of cancer care, both at institutional and scientific society level, can also contribute to more sustainable cancer care.[38]

Concluding, our main goal in cancer treatment clearly is clinical benefit while retaining the best achievable quality of life. With the emergence of more predictive markers, personalized treatment choices will deliver more efficacious treatments for CRC patients, which as a result are likely to be more cost-effective. In addition, policy makers should pursue more boldly the development of (European) legislation to contain healthcare costs and this should not be restricted to drug costs only. Finally, scientific societies

representing oncologists and oncologists themselves, should become less reluctant in the participation on the overarching discussion on cost of cancer care, as we can all contribute to improvements in healthcare resource allocations.

REFERENCES

1. 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(12):1165-1174. doi:10.1016/S1470-2045(13)70442-X
2. Sullivan R, Peppercorn J, Sikora K, et al. Delivering affordable cancer care in high-income countries. *Lancet Oncol*. 2011;12(10):933-980. doi:10.1016/S1470-2045(11)70141-3
3. Uyl-de Groot C a., de Vries EGE, Verweij J, Sullivan R. Dispelling the myths around cancer care delivery: It's not all about costs. *J Cancer Policy*. 2014;2(1):22-29. doi:10.1016/j.jcpo.2014.01.001
4. Del Paggio JC, Sullivan R, Schrag D, et al. Delivery of meaningful cancer care: a retrospective cohort study assessing cost and benefit with the ASCO and ESMO frameworks. *Lancet Oncol*. 2017;18(7):887-894. doi:10.1016/S1470-2045(17)30415-1
5. Schnipper LE, Davidson NE, Wollins DS, et al. American Society of Clinical Oncology Statement: A Conceptual Framework to Assess the Value of Cancer Treatment Options. *J Clin Oncol*. 2015;33(23):2563-2577. doi:10.1200/JCO.2015.61.6706
6. Cherny NI, Sullivan R, Dafni U, et al. A standardised, 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(8):1547-1573. doi:10.1093/annonc/mdv249
7. ZorginstituutNederland. Richtlijn voor economische evaluatie. Accessed September 17, 2019. https://www.zorginstituutnederland.nl/over-ons/werkwijzen-en-procedures/adviseren-over-en-verduidelijken-van-het-basispakket-aan-zorg/beoordeling-van-geneesmiddelen/richtlijnen-voor-economische-evaluatie
8. NVMO. Over de commissie BOM. Accessed September 17, 2019. https://www.nvmo.org/nvmo/commissie-bom/over-de-commissie-bom/
9. Zorginstituut Nederland. Ziektelast in de praktijk. Published online 2018. https://www.zorginstituutnederland.nl/binaries/zinl/documenten/rapport/2018/05/07/ziektelast-in-de-praktijk/Ziektelast+in+de+praktijk_definitief.pdf
10. Franken MD, van Rooijen EM, Uyl-de Groot CA, van Oijen MGH, Koopman M. Cost-effectiveness in colorectal cancer: challenges on quality and comparability. *Color Cancer*. 2016;5(1):21-31. doi:doi: 10.2217/crc.15.33
11. Goey KKH, Elias SG, van Tinteren H, et al. Maintenance treatment with capecitabine and bevacizumab versus observation in metastatic colorectal cancer: Updated results and molecular subgroup analyses of the phase 3 CAIRO3 study. *Ann Oncol*. 2017;28(9):2128-2134. doi:10.1093/annonc/mdx322
12. Ewara EM, Zaric GS, Welch S, Sarma S. Cost-effectiveness of first-line treatments for patients with KRAS wild-type metastatic colorectal cancer. *Curr Oncol*. 2014;21:541-550.
13. Lawrence D, Maschio M, Leahy KJ, Yunger S, Easaw JC, Weinstein MC. Economic analysis of bevacizumab, cetuximab, and panitumumab with fluoropyrimidine-based chemotherapy in the first-line treatment of KRAS wild-type metastatic colorectal cancer (mCRC). *J Med Econ*. 2013;16(12):1387-1398. doi:10.3111/13696998.2013.852097
14. Huxley N, Crathorne L, Varley-Campbell J, et al. The clinical effectiveness and cost-effectiveness of cetuximab (review of technology appraisal no. 176) and panitumumab (partial review of technology appraisal no. 240) for previously untreated metastatic colorectal cancer: A systematic review and economi. *Health Technol Assess (Rockv)*. 2017;21(38):V-241. doi:10.3310/hta21380
15. Siebert U, Alagoz O, Bayoumi AM, et al. State-transition modeling: A report of the ISPOR-SMDM modeling good research practices task force-3. *Value Heal*. 2012;15(6):812-820. doi:10.1016/j.jval.2012.06.014
16. Karnon J, Stahl J, Brennan A, Caro JJ, Mar J, Möller J. Modeling using discrete event

simulation: A report of the ISPOR-SMDM modeling good research practices task force-4. *Value Heal*. 2012;15(6):821-827. doi:10.1016/j.jval.2012.04.013

17. Karnon J. Alternative decision modelling techniques for the evaluation of health care technologies: Markov processes versus discrete event simulation. *Health Econ*. 2003;12(10):837-848. doi:10.1002/hec.770
18. Simpson KN, Strassburger A, Jones WJ, Dietz B, Rajagopalan R. Comparison of Markov model and discrete-event simulation techniques for HIV. *Pharmacoeconomics*. 2009;27(2):159-165. doi:10.2165/00019053-200927020-00006
19. Degeling K, Franken MD, May AM, et al. Matching the model with the evidence: comparing discrete event simulation and state-transition modeling for time-to-event predictions in a cost-effectiveness analysis of treatment in metastatic colorectal cancer patients. *Cancer Epidemiol*. 2018;57:60-67. doi:10.1016/j.canep.2018.09.008
20. Degeling K, IJzerman MJ, Koopman M, Koffijberg H. Accounting for parameter uncertainty in the definition of parametric distributions used to describe individual patient variation in health economic models. *BMC Med Res Methodol*. 2017;17(1):170. doi:10.1186/s12874-017-0437-y
21. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the Panel on Cost-Effectiveness in Health and Medicine. *JAMA*. 1996;276(15):1253-1258. doi:10.1097/00132586-199712000-00019
22. Williams A. EuroQol - A new facility for the measurement of health-related quality of life. *Health Policy (New York)*. 1990;16:199-208. doi:10.1016/0168-8510(90)90421-9
23. Versteegh M, M. Vermeulen K, M. A. A. Evers S, de Wit GA, Prenger R, A. Stolk E. Dutch Tariff for the Five-Level Version of EQ-5D. *Value Heal*. 2016;19(4):343-352. doi:10.1016/j.jval.2016.01.003
24. Lamers LM, Stalmeier PFM, McDonnell J, Krabbe PFM, van Busschbach JJ. [Measuring the quality of life in economic evaluations: the Dutch EQ-5D tariff]. *Ned Tijdschr Geneeskd*. 2005;149(28):1574-1578. Accessed October 30, 2014. http://www.ncbi.nlm.nih.gov/pubmed/16038162
25. Versteegh MM, Leunis A, Uyl-De Groot CA, Stolk EA. Condition-specific preference-based measures: Benefit or burden? *Value Heal*. 2012;15(3):504-513. doi:10.1016/j.jval.2011.12.003
26. Janssen MF, Birnie E, Bonsel GJ. Quantification of the level descriptors for the standard EQ-5D three-level system and a five-level version according to two methods. *Qual Life Res*. 2008;17(3):463-473. doi:10.1007/s11136-008-9318-5
27. Franken M., van Rooijen E., May A., et al. Cost-effectiveness of capecitabine and bevacizumab maintenance treatment after first-line induction treatment in metastatic colorectal cancer. *Eur J Cancer*. 2017;75:204-212. doi:10.1016/j.ejca.2017.01.019
28. Carter HE, Zannino D, John Simes R, et al. The cost effectiveness of bevacizumab when added to capecitabine, with or without mitomycin-C, in first line treatment of metastatic colorectal cancer: results from the Australasian phase III MAX study. *Eur J Cancer*. 2014;50(3):535-543. doi:10.1016/j.ejca.2013.09.028
29. Färkkilä N, Sintonen H, Saarto T, et al. Health-related quality of life in colorectal cancer. *Colorectal Dis*. 2013;15(5):e215-22. doi:10.1111/codi.12143
30. Ramsey I, de Rooij BH, Mols F, et al. Cancer survivors who fully participate in the PROFILES registry have better health-related quality of life than those who drop out. *J Cancer Surviv*. Published online 2019. doi:10.1007/s11764-019-00793-7
31. Perez-Cruz PE, Shamieh O, Paiva CE, et al. Factors Associated With Attrition in a Multicenter Longitudinal Observational Study of Patients With Advanced Cancer. *J Pain Symptom Manage*. 2018;55(3):938-945. doi:10.1016/j.jpainsymman.2017.11.009
32. Hui D, Glitza I, Chisholm G, Yennu S, Bruera E. Attrition rates, reasons, and predictive factors in supportive care and palliative oncology clinical trials. *Cancer*. 2013;119(5):1098-1105. doi:10.1002/cncr.27854

33. Roick J, Danker H, Kersting A, et al. Factors associated with non-participation and dropout among cancer patients in a cluster-randomised controlled trial. *Eur J Cancer Care (Engl)*. 2018;27(1):1-9. doi:10.1111/ecc.12645

34. de Rooij BH, Ezendam NPM, Mols F, et al. Cancer survivors not participating in observational patient-reported outcome studies have a lower survival compared to participants: the population-based PROFILES registry. *Qual Life Res*. 2018;27(12):3313-3324. doi:10.1007/s11136-018-1979-0

35. Chapman RH, Berger M, Weinstein MC, Weeks JC, Goldie S, Neumann PJ. When does quality-adjusting life-years matter in cost-effectiveness analysis? *Health Econ*. 2004;13(5):429-436. doi:10.1002/hec.853

36. Jönsson B. Editorial: Ten arguments for a societal perspective in the economic evaluation of medical innovations. *Eur J Heal Econ*. 2009;10(4):357-359. doi:10.1007/s10198-009-0173-2

37. Burbach JPM, Kurk SA, Coebergh van den Braak RRJ, et al. Prospective Dutch colorectal cancer cohort: an infrastructure for long-term observational, prognostic, predictive and (randomized) intervention research. *Acta Oncol (Madr)*. 2016;55(11):1273-1280. doi:10.1080/0284186X.2016.1189094

38. Wait S, Han D, Muthu V, et al. Towards sustainable cancer care: Reducing inefficiencies, improving outcomes—A policy report from the All.Can initiative. *J Cancer Policy*. 2017;13(May):47-64. doi:10.1016/j.jcpo.2017.05.004

APPENDICES

SAMENVATTING

Dikkedarm- en endeldarmkanker (colorectaal carcinoom, CRC) vormen wereldwijd tezamen de derde meest voorkomende kankersoort. Ongeveer 20-25% van de patiënten heeft uitgezaaide (gemetastaseerd) CRC ten tijde van de diagnose. Bij zo'n 20% van de patiënten, die ten tijde van diagnose geen uitzaaiingen had, zullen alsnog uitzaaiingen gevonden worden. Beslissingen rond de keuze van behandeling zijn afhankelijk van het stadium van de ziekte. Lokale ziekte en lokaal gevorderde ziekte (stadium I-III) kunnen met een operatie genezend behandeld worden. Bij hoog-risico stadium II en stadium III ziekte stijgen de overlevingskansen door toevoeging van aanvullende, oftewel adjuvante, chemotherapie. Ongeveer 50% van deze patiënten zal genezen zijn met alleen een operatie en zal geen meerwaarde ondervinden van de adjuvante chemotherapie, terwijl en ongeveer 20% van de patiënten juist zal genezen door de adjuvante chemotherapie. Er is daarom een dringende behoefte aan biomarkers om te kunnen bepalen welke patiënten nu baat hebben bij de adjuvante chemotherapie. Voor gemetastaseerd CRC, hebben nieuwe behandelmogelijkheden geleid tot een verbetering van de progressievrije overleving en overleving in het algemeen. Dit geldt althans voor patiënten die deelnamen aan klinisch onderzoek, waarin de werkzaamheid en de veiligheid van een behandeling werd onderzocht, of voor patiënten die zouden hebben voldaan aan de deelname criteria, maar niet meededen in het onderzoek. Ondanks alle vooruitgang die in de laatste decennia is geboekt bij de behandeling van CRC, blijven de gezondheidszorgkosten in het algemeen en de uitgaven voor de behandeling van kanker wel stijgen. Daarentegen, blijft het belangrijkste doel bij de behandeling van CRC, patiënten de meest effectieve behandeling met behoud van kwaliteit van leven te bieden. De invloed van deze behandelingen op het zorgbudget worden steeds groter en kunnen daarmee in de toekomst een bedreiging vormen voor het behoud van een evenwichtig zorgsysteem.

Hoofdstuk 2 en Hoofdstuk 2B illustreert de uitdagingen ten aanzien van het beoordelen van kosteneffectiviteitsstudies, hierbij is gemetastaseerd CRC als voorbeeld genomen. Ondanks het bestaan van richtlijnen om kosteneffectiviteit op gestandaardiseerde wijze te onderzoeken, is het belangrijk om in acht te nemen dat er aannames gedaan worden om kosteneffectiviteit in een model te onderzoeken. Kosteneffectiviteitsmodellen bevatten niet alleen klinische informatie (zoals tijd tot progressie, overleving, bijwerkingen), maar ook velerlei aannames zoals bijvoorbeeld de kosten gemoeid met de behandeling van bijwerkingen, de duur en kosten van vervolgbehandeling bij terugkeer van ziekte, etc. Om deze redenen, kunnen aannames in een belangrijke invloed hebben op de uitkomsten van kosteneffectiviteitsstudies. Gerandomiseerde studies blijven daarom het meest geschikt om de kosteneffectiviteit van een behandeling in het ziekenhuis in te schatten. Echter, de meest wenselijke afweging t.a.v. kosteneffectiviteit zou juist uit moeten gaan van cohortonderzoek met gegevens over de werkzaamheid van een behandeling, kwaliteit van leven en kosten (inclusief kosten gemaakt buiten het ziekenhuis), om zodoende

een inschatting te maken van kosteneffectiviteit voor de maatschappij en daarmee de mogelijkheid om het zorgbudget doelmatiger te kunnen inzetten.

In Hoofdstuk 3, wordt de kosteneffectiviteit van een onderhoudsbehandeling met capecitabine-bevacizumab (CAP-B) vergeleken met observatie (ofwel een periode zonder actieve behandeling) bij patiënten met gemetastaseerd CRC, die stabiele ziekte of beter bereikten na 6 kuren capecitabine, oxaliplatin en bevacizumab (CAPOX-B). Aannames in het kosteneffectiviteit model werden gedaan op basis van de resultaten uit het gerandomiseerde klinische fase 3 onderzoek, de CAIRO-3 studie waaraan 558 patiënten deelnamen. Daarnaast, werd gebruik gemaakt van aanvullende gegevens nog verkregen na het afronden van het gerandomiseerde onderzoek (in 6 verschillende ziekenhuizen voor 146 patiënten die aan de CAIRO-3 studie deelnamen), gegevens uit verschillende publicaties en advies van deskundigen op het vakgebied van CRC. Middels een deterministisch "state-transition model" (STM), ook wel Markov-model, werd de kosteneffectiviteit van de behandeling geëvalueerd. We toonden aan dat een onderhoudsbehandeling met CAP-B een gemiddelde winst van 0,21 voor kwaliteit van leven gecorrigeerde levensjaren (quality adjusted life years (QALY's)) werd bereikt tegen een gemiddelde incrementele kostprijs van €36.845. Dit levert een gemiddelde incrementele kosteneffectiviteitsratio (ICER) van €175.452 per QALY op. Bij patiënten met een partiële of complete respons op de behandeling, werd een gunstigere gemiddelde ICER berekend van €149.300 per QALY, grotendeels door een hogere gemiddelde QALY (0.31). We concluderen dat deze behandelstrategie niet kosteneffectief is, hoewel tot op heden de (inter)nationale maximale afgesproken waarde t.a.v. kosteneffectiviteit niet strikt wordt aangehouden.

In Hoofdstuk 4 vergelijken we het STM zoals beschreven in Hoofdstuk 3 met een "Discrete Event Simulation" (DES) model, waarbij voor beide modellen dezelfde modelaannames zijn gebruikt. Dit werd gedaan omdat STM's werken met cycli van een vaste tijdsduur, bijv. ten aanzien van tijd tot progressie, waardoor meer onregelmatigheden kunnen optreden in kansen om in het model te verschuiven van de ene ziekte toestand (bijv. stabiele ziekte) naar de andere (bijv. ziekte progressie). We hebben aangetoond dat tijdsafhankelijke kansen om van de ene ziekte toestand naar de andere in een STM inderdaad onregelmatig waren en gevoeliger voor gebeurtenissen, indien deze weinig voorkomen. De overgangen van de ene naar de andere ziektetoestand in het DES-model waren vloeiender en iets nauwkeuriger. Uit het STM- en het DES-model werd een gemiddelde QALY-winst van respectievelijk 0,21 en 0,18 berekend. Door de verschillen in tijd tot een gebeurtenis, werden tussen de beide modellen ook kleine verschillen gevonden in de incrementele kosten. Voor beide modellen, werden uiteindelijk vergelijkbare gemiddelde ICER's berekend (€172.443 per QALY voor het STM model en €168.383 per QALY voor het DES model). Voor de meest nauwkeurige ICER schatting, leveren DES-modellen betere resultaten door een meer natuurlijk verloop van tijd tot gebeurtenissen in het model.

Één van de meest gebruikte uitkomstmaten bij studies naar kosteneffectiviteit is de kosten per gewonnen QALY. Generieke vragenlijsten naar kwaliteit van leven, zoals de EQ-5D vragenlijst, worden gebruikt om gezondheid gerelateerde utiliteiten te berekenen. Deze utiliteiten worden vervolgens gebruikt om in een kosteneffectiviteit model, QALYs te berekenen. In klinische studies binnen het vakgebied van de oncologie, verdienen ziekte specifieke vragenlijsten, zoals de EORTC QLQ-C30, gewoonlijk de voorkeur aangezien deze vragenlijsten de ziekte gerelateerde veranderingen in kwaliteit van leven beter kunnen vastleggen. Bij gebrek aan EQ-5D vragenlijsten in een klinische studie, kan een algoritme gebruikt worden om resultaten van een QLQ-C30 vragenlijst om te rekenen naar utiliteiten om zo een kosteneffectiviteitsonderzoek toch mogelijk te maken. In Hoofdstuk 5, evalueerden we drie bestaande algoritmen en ontwikkelden we 6 nieuwe algoritmen om QLQ-C30 scores om te rekenen naar EQ-5D-3L gerelateerde utiliteiten. Daarnaast, evalueerden wij het gebruik daarvan in het kosteneffectiviteit model zoals beschreven in Hoofdstuk 4. Resultaten van twee van de drie onderzochte, bestaande algoritmen leverden significant verschillende utiliteiten t.o.v. de waargenomen utiliteiten in het gerandomiseerde, fase 3 CAIRO3 onderzoek (0,83± 0,17 o.b.v. Nederlandse referentie waarden). De door ons ontwikkelde nieuwe algoritmen resulteerden in utiliteiten zonder statisch significant verschil t.o.v. van de referentie waarde. In het kosteneffectiviteit model werden vergelijkbare resultaten verkregen bij EQ-5D-3L utiliteiten en utiliteiten o.b.v. een algoritme. Utiliteiten verkregen o.b.v. het nieuw ontwikkelde algoritme resulteerde in een ICER-puntschatting van €5.094 per QALY meer t.a.v. van de ICER-puntschatting o.b.v. EQ-5D-3L utiliteiten (€168.048). De bestaande algoritmen leverden een ICER-puntschatting van €10.140 per QALY minder en €1.765 per QALY meer in vergelijking met de ICER-puntschatting o.b.v. EQ-5D-3L utiliteiten. Het verschil in uitkomsten kan verklaard vanuit de minimale verschillen in incrementele QALY's bij modelberekeningen o.b.v. EQ-5D-3L utiliteiten en utiliteiten verkregen middels een algoritme. Het gebruiken van een algoritme om EQ-5D-3L utiliteiten te berekenen vanaf QLQ-C30 uitkomsten, is dus een geschikt alternatief om een kosteneffectiviteit analyse mogelijk te maken, mits dit wordt toegepast op een patiënten populatie met een vergelijkbare ziekte ernst.

In Hoofdstuk 6, evalueren we het werkvermogen van patiënten met CRC in de eerste 2 jaar na diagnose. De uitkomsten van 390 patiënten (jonger dan 67 jaar met stadium I-IV ziekte, participerend aan het Prospectief Landelijk CRC cohort (PLCRC), en deelnamen aan uitkomstvragenlijsten) werden onderzocht. Direct na diagnose, was de score voor werkvermogen (ook wel work ability index (WAI) score) bij patiënten met stadium I-IV CRC duidelijk verlaagd ten opzichte van een controle groep zonder kanker (respectievelijk 31±SD 8 en 41±SD 6). Bij patiënten behandeld met alleen een operatie (stadium I-III), herstelde het werkvermogen na 18 maanden naar een niveau vergelijkbaar met de controle groep. Voor patiënten die een operatie en adjuvante systeemtherapie ondergingen, duurde het herstel langer en werd dit bereikt na 2 jaar. Voor patiënten met

stadium IV ziekte werd er geen herstel van werkvermogen gezien. De uitkomsten op werkvermogen scores waren lager voor patiënten die systeem therapie kregen en één of meer bijkomende aandoeningen had. Het gemiddelde werkvermogen was significant lager bij vrouwen, hoewel dit niet statistisch significant bleef wanneer patiënten met stadium IV ziekte uitgesloten werden van de analyse. Middelbaar beroepsonderwijs of lager was geassocieerd met een lager werkvermogen bij patiënten met stadium I-III ziekte. Na 2 jaar, werd ziekteverlof van 100 dagen of meer gevonden in 12% van de patiënten die een operatie en adjuvante behandeling ondergingen en 2% van de patiënten die alleen curatieve operatie ondergingen (p=0,2). Ter vergelijking, dit is ongeveer 4% bij de populatie controle groep. Patiënten met CRC ervaren dus langdurig negatieve effecten op het werkvermogen. De uitkomsten van dit onderzoek kunnen bijdragen aan verbeteringen in de begeleiding van patiënten met betrekking tot het werkvermogen tijdens de behandeling, maar ook langere tijd daarna.

Dit proefschrift illustreert de complexiteit ten aanzien van kosteneffectiviteit onderzoek. De behandelingen voor CRC zullen in toekomst steeds meer gepersonaliseerd worden dankzij bijvoorbeeld de eerder genoemde biomarkers. Om de effectiviteit van gepersonaliseerde CRC behandelingen te kunnen onderzoeken, zullen klinische studies anders vormgegeven moeten worden. Een voorbeeld hiervan is het Prospectief Landelijk CRC cohort (PLCRC). Het PLCRC cohort biedt eveneens de mogelijkheid om kosteneffectiviteit van een behandeling te onderzoeken. Desondanks, zal het een uitdaging blijven om kosteneffectiviteit vanuit een maatschappelijk perspectief te evalueren aangezien onmogelijk alle kosten (buiten het ziekenhuis) vastgelegd kunnen worden in cohort onderzoek. Anderzijds, zouden politici wellicht moediger beleid na kunnen streven t.a.v. kosten in de gezondheidszorg, en zich daarbij niet beperken tot kosten voor geneesmiddelen alleen. Van oncologen en beroepsverenigingen binnen de oncologie, zou nog meer participatie mogelijk zijn t.a.v. de maatschappelijke discussie over kosten voor de behandeling van kanker. Tot slot, blijft de dialoog tussen de industrie, zorgverzekeraars, gezondheidseconomen, beleidsmakers, medische beroepsverenigingen en patiëntenverenigingen over doelmatigheid in de zorg noodzakelijk om een duurzame toekomst van ons zorgstelstel te waarborgen.

REVIEW COMMITTEE

Prof. dr. H.M. Verkooijen

Professor of Evaluation of Image-Guided treatment at the University Medical Center Utrecht, the Netherlands. Leader of multiple international research consortia, leader of the clinical research office within the Division of Imaging and Cancer at the University Medical Center Utrecht. Strategic Advisor for 'NTvG Gezonde Zorg', a campaign to ensure sustainability and affordability of health care in the Netherlands.

Prof. dr. E. van der Wall

Professor in internal medicine at the University Medical Center Utrecht, the Netherlands. Visiting professor at the Sidney Kimmel Cancer Institute at Johns Hopkins Medical Institutions (JHMI) in Baltimore, USA and member of the Medical Advisory Board of the Dutch Cancer Center Network, the KWF Advisory Board and member of the Scientific Advisory Board Princess Maxima Center/ University Medical Center Utrecht.

Prof. dr. N.J. de Wit

Professor in General Practice University Medical Center Utrecht, the Netherlands. Division chair ad interim of the Julius Center of Health Sciences and Primary Care at the University Medical Center Utrecht. Visiting professor at the University of Durham (U.K.), fellow of the Royal College of Physicians, Edinburg (U.K.). Member of the advisory board of the Quality Institute of the Dutch Ministry of Health and member of the scientific advisory board of the Dutch Cancer Foundation.

Prof. dr. C.A. Uyl-de Groot

Director and professor at the institute of Medical Technology Assessment at the Erasmus University, Rotterdam, the Netherlands. Professor of health technology assessment at the VU medical center in Amsterdam. Member of the Oncological Aids Review Committee and member of the Health Council.

Dr. G.W.J. Frederix

Associate professor Health Technology Assessment the Julius Center of Health Sciences and Primary Care at the University Medical Center Utrecht, the Netherlands.

LIST OF PUBLICATIONS

Published

Franken M.D., de Hond A., Degeling K., Punt C.J.A., Koopman M., Uyl-de Groot C.A., Versteegh M.*, van Oijen M.G.H.* Evaluation of the performance of algorithms mapping EORTC QLQ-C30 onto the EQ-5D index in a metastatic colorectal cancer cost-effectiveness model. Health and Quality of Life Outcomes, 2020, 18:240

Degeling K., Koffijberg H., **Franken M.D.**, Koopman M., IJzerman M.J. Comparing Strategies for Modelling Competing Risks in Discrete-Event Simulations: A Simulation Study and Illustration in Colorectal Cancer. Medical Decision Making. 2019, 39(1):57-73.

Franken M.D.*, Degeling K.*, May A.M., van Oijen M.G.H., Koopman M., Punt C.J.A., IJzerman M.J., Koffijberg H. Matching the model with the evidence: comparing discrete event simulation and state-transition modeling for time-to-event predictions in a cost-effectiveness analysis of treatment in metastatic colorectal cancer patients. Cancer Epidemiology, 2018, 57: 60-67.
* Authors contributed equally to this work.

Franken M.D., van Rooijen E.M., May A.M., Koffijberg H, van Tinteren H., Mol L., Ten Tije A.J., Creemers G.J., van der Velden A.M.T., Tanis B.C., Uyl-de Groot C.A., Punt C.J.A., Koopman M., van Oijen M.G.H. Cost-effectiveness of capecitabine and bevacizumab maintenance treatment after first-line induction treatment in metastatic colorectal cancer. European Journal of Cancer, 2017, 75: 204-212.

Franken M.D., Koopman M., van Oijen M.G.H. Cetuximab as first-line treatment for metastatic colorectal cancer: caution with interpretation of cost-effectiveness results toward medical decision making (*letter to the editor*). American Journal of Clinical Oncology, 2016, 39 (2): 214.

Franken M.D., van Rooijen E.M., Uyl-de Groot C.A., van Oijen M.G.H., Koopman M. Cost-effectiveness in colorectal cancer: challenges on quality and comparability. Colorectal Cancer – Future medicine, 2016, 5(1): 21-31.

Franken M.D., Horsting M.W.B., Meulenbelt J., van Klei W.A., de Lange D.W. The etiology and outcome of non-traumatic coma in the emergency department: a systematic literature review. BMC Anesthesiology, 2015, 15 (65).

Conditionally accepted for publication

Franken M.D., Vink G.R., van Grevenstein W.M.U., Verkooijen H.M., Punt C.J.A., Koopman M., May A.M., on behalf of the PLCRC study group. Work ability in patients with stage I-IV colon cancer, results of the Dutch Prospective ColoRectal Cancer cohort.

ACKNOWLEDGEMENTS (DANKWOORD)

Non, rien de rien... non, je ne regrette rien... - Edith Piaf

Na een omweg in mijn carrière, ben ik nu beland op het punt dat ik met trots mijn promotie mag afronden en op weg ben naar een carrière in de medische oncologie. Ik ben blij met de route die ik tot nu toe bewandeld heb en heb me op deze weg door velen gesteund gevoeld. Dank! De slotsom is: deze route was het allemaal waard. Je ne regrette rien! De volgende mensen, wil ik in het bijzonder bedanken.

Allereerst, wil ik alle **patiënten** bedanken die deelnamen aan de CAIRO3 studie en PLCRC. Zonder de belangeloze deelname van patiënten aan klinische studies, is een verbetering van behandeling met behoud van kwaliteit van leven onmogelijk. Zij dragen hiermee ook in belangrijke mate bij aan onderzoek naar de doelmatige inzet van behandelingen.

Prof. dr. M. Koopman, beste Miriam, dank voor het vertrouwen in onze samenwerking van de afgelopen jaren! Het is onmogelijk niet enthousiast te raken over CRC onderzoek als jij aan het woord bent. Je bent voor mij een rolmodel als het gaat om persoonlijk leiderschap.

Prof. dr. C.J.A. Punt, beste Kees, dank voor je betrokkenheid. De snelheid waarmee je met veel detail op versies van manuscripten wist te reageren, vond ik telkens weer indrukwekkend en altijd waardevol. Het heeft me buitengewoon geholpen om tot betere versies van mijn manuscripten te komen.

Dr. M.G.H. van Oijen, beste Martijn, je bent de grondlegger voor het onderwerp van mijn proefschrift en hebt me met veel geduld geholpen om de basis principes van kosteneffectiviteitsonderzoek onder de knie te krijgen. Dank voor je enthousiasme en betrokkenheid!

Leden van de beoordelingscommissie, **prof. Lenny Verkooijen**, **prof. Elsken van der Wall**, **prof. Niek de Wit** en **dr. Geert Frederix**: hartelijk dank voor jullie bereidheid en tijd om mijn proefschrift kritisch door te lezen en te beoordelen. Beste **Elsken**, door onze ontmoeting in 2014, heb je een belangrijke rol gespeeld in het verloop van mijn carrière. Dankzij jou ben ik over onderzoek in gesprek gegaan met Miriam Koopman en je gaf me het advies direct na het afronden van mijn SUMMA opleiding te solliciteren voor de opleiding interne geneeskunde. Dat advies heb ik ter harte genomen en daar ben ik je nog altijd dankbaar voor!

Alle coauteurs, hartelijk dank voor jullie samenwerking en bijdragen aan mijn manuscripten. In het bijzonder wil ik hierbij nog **prof. Anne May**, **dr. Koen Degeling**, **dr. Matthijs Versteegh** en **dr. Geraldine Vink** bedanken. Beste **Anne**, ik wil je bedanken voor het meedenken in al mijn projecten. Ik heb me altijd zeer gesteund gevoeld door jou. Beste **Koen**, de souplesse en het enthousiasme waarmee jij ingewikkelde modelleringstechnieken op het gebied van kosteneffectiviteit weet uit te leggen is bewonderenswaardig. Beste **Matthijs**, zonder jouw steun en adviezen t.a.v. het matching algoritme was dit stuk nooit tot een einde gekomen. Beste **Geraldine**, ik heb grote bewondering hoe razend snel je te werk gaat. Je betrokkenheid en kritische blik t.a.v. mijn onderzoek heb ik zeer gewaardeerd.

Principle investigators, research verpleegkundigen van de ziekenhuizen die deelnamen aan de CAIRO3 studie of deelnemen aan PLCRC, dank voor jullie bijdragen! Zeker als het gaat om het mogelijk maken van de retrospectieve verzameling van klinische gegevens voor de kosteneffectiviteitsanalyse o.b.v. de CAIRO3 studie.

Mijn promotietraject werd een stuk leuker door mijn kamergenoten en mede-PhDers: **Kaitlyn**, **Sophie**, **Karlijn**, **Patricia**, **Dave**, **Suzanna**, **Emerens**, **Medard**. Ik heb jullie in wisselende samenstellingen mogen meemaken en ik heb mooie tijden met jullie meegemaakt. **Kaitlyn,** ik ben blij dat wij samen het PhD-minion groepje hebben mogen starten. Naar de opera in Wenen blijft een prachtige herinnering! **Sophie**, **Jeroen** en **Patricia**, dank voor jullie statistiek hulp als ik weer in staat van ontbinding was t.a.v. mijn analyses. **Dave**, dank voor de "reality check", maar inmiddels weten we allebei dat ook "echte dokters" interesse kunnen hebben in kosteneffectiviteit.

Lieve familie en vrienden, dank voor jullie belangstelling en steun in dit traject. **Ester**, **Melanie**, we zijn al zo lang vriendinnen – ik vind het fantastisch dat we dit samen kunnen meemaken. **Anne-Joy**, sinds het laatste deel van mijn SUMMA opleiding hebben onze carrières elkaar gekruist en dat gaat met onze differentiatie keuze ook niet meer veranderen. Ik ben blij dat je mijn paranimf bent! **John** en **Stella**, dank voor jullie betrokkenheid, kaartjes, appjes en mailtjes.

Lieve **pap** en **mam**, dank voor jullie nimmer aflatende steun in alles wat ik doe. Jullie staan altijd voor me (ons) klaar! Dankzij de logeerpartijtjes voor Florian en Alice was er af en toe wat meer ruimte om aan mijn onderzoek en opleiding te werken. Pap, wat ben ik trots dat je de voorzijde van mijn boekje hebt gemaakt!

Lieve **Joep,** je bent mijn rots in de branding. Door alle ruimte die je me geeft, ben ik in staat geweest om nog arts te worden en zelfs te kunnen promoveren. Zonder jouw steun was ik hiertoe niet in staat geweest, ik hou van jou!

Last but not least, lieve **Sylvie**, lieve zus, jij bent degene die het vonkje om alsnog geneeskunde te gaan studeren, zuurstof heeft gegeven. Ik bewonder je in alle opzichten. Er is dan ook geen andere plek denkbaar voor jou, op de dag van mijn promotie, dan naast mij als paranimf om deze dag samen te vieren! Ik ben je voor altijd dankbaar!

CURRICULUM VITAE

Mira Desirée Franken was born in Delft, the Netherlands, on November 1st 1979. She spent her youth in the south-east of France with her parents and sister, Sylvie.

After relocating to the Netherlands, she went to the Huygens lyceum in Voorburg, the Netherlands. In 1997, she moved to Wageningen to study biotechnology. She performed a master thesis on multidrug resistance proteins at the Dutch Cancer Institute (NKI). Later, she also performed research on T-helper cell activation using virus like particles in a murine model at Otago University in New Zealand. She graduated as MSc in biotechnology in 2003. During her study, she was also a board member of the study association, Codon, where she was responsible for field trips to biotech-companies.

After her graduation she moved to the U.K. for 1 year, to work for a clinical research organisation, where she coordinated phase I-III studies in several countries. When she returned to the Netherlands, she started working at Janssen-Cilag to coordinate clinical phase IV studies and was later promoted to medical advisor.

Mira always had the wish to become a medical doctor. In 2011, Mira pursued her dream and was admitted to the Selective Utrecht Medical Master (SUMMA). She graduated as a Medical Doctor in 2015. During her SUMMA study, she started a PhD at the Medical Oncology department with prof. Miriam Koopman and prof. Kees Punt as promotors, and Martijn van Oijen as co-promotor. Meanwhile, she was also admitted as resident internal medicine and started her residency in hospital 'de Gelderse Vallei' in Ede. She is currently working as a resident internal medicine and fellow medical oncology since at the University Medical Centre (UMC) Utrecht.

Her future ambition is to participate in the discussion on sustainable cancer care in addition to patient care.