Early Health Technology Assessment of a Prognostic Biomarker for the Stratification of Stage II and III Colorectal Cancer Patients

A Case Study of Histotyping in the Norwegian Perspective

Anna Ruth Kenseth
Supervised by Eline Aas

Thesis submitted as a part of the joint degree
European Master in Health Economics and Management

University of Oslo
November 2020

Early Health Technology Assessment of a Prognostic Biomarker for the Stratification of Stage II and III Colorectal Cancer Patients

A Case Study of Histotyping in the Norwegian perspective

http://www.duo.uio.no/

Print: Reprosentralen, Universitetet i Oslo

Abstract

Objective: To apply an early HTA framework to Histotyping plus the standard of care versus the standard of care alone for the treatment of stage II and III CRC patients. This framework was developed to model and understand the cost and effect implications of integrating Histotyping into the clinical routine in a structure that is later updateable with new clinical evidence.

Method: A partitioned survival model with a 10-year time horizon, 1-year cycle length and 4% discount rate applied to cost and health outcomes was created for an early HTA of Histotyping integrated into the standard of care compared to the standard of care alone. Histotyping is intended to inform adjuvant chemotherapy assignment by stratifying patients based on recurrence risk, impacting over- and undertreatment. KM curves were generated using QUASAR 2 validation data and formed the basis of our model inputs. Survival analysis was modeled with a flexible RCS method using life table data and stage IV recurrence rates. Effects were measured in LYs, QALYs gained and NMB of the integration of Histotyping. Costs included were for drug administration, adverse events, progressed disease maintenance and end of life care.

Results: In the stage II Histotyping arm versus the standard of care alone, expected life-years were 7.54 and 6.90 respectively, and QALYs accumulated were 6.52 and 5.93 QALYs, respectively. The Histotyping arm cost was 85,106 NOK lower than the standard of care. The NMB for Histotyping integration for stage II was 375,666 NOK. In the stage III Histotyping arm versus the standard of care alone, expected life-years were 6.74 and 6.47 respectively, and QALYs accumulated were 5.88 and 5.53 QALYs, respectively. The Histotyping arm cost was 91,932 NOK lower than the standard of care. The NMB for Histotyping integration for stage III was 267,139 NOK based on a threshold of 495,000 NOK.

Conclusion: Integrating Histotyping in the standard of care showed good potential to be a cost-effective integration under the conditions we defined in the early HTA framework also when considering the tradeoffs of increased progression and increased therapy intensity. Histotyping shows promise in its potential for meaningful patient implications when applied to adjuvant chemotherapy assignment.

Keywords: Stage II and III CRC, Histotyping, early HTA, cost-effectiveness

Acknowledgements

First and foremost, my utmost thanks go to my supervisor, Assistant Professor Eline Aas for sharing her knowledge and expertise with me for this thesis. If not for your incredibly helpful and valuable input, feedback and encouragement, this thesis may not have come to fruition.

Thank you Professor Håvard Danielsen and my wonderful colleagues at the Institute for Cancer Genetics and Informatics for accommodating my transition to part-time work while I pursued this master's degree over these past two years, including a study abroad, no less.

To the three professionals that granted meetings with me to elucidate the topics around my thesis and discuss the potential of this tool, thank you for your time. You provided invaluable input and helped me understand the intricacies of colorectal cancer, prognostic biomarkers, the market, and beyond.

To my friends, especially Sanne, Kristin, Jen and Sarah, I don't know how I got so lucky to be surrounded by so many intelligent and inspiring women. Thank you for your friendship and support that filled me up so that I could persevere.

I would be remiss if I did not give special thanks to Dominika, my fellow Eu-HEM classmate who became a dear friend through our study sessions, academic discussions, Hatta residence and more. Thanks for helping to make this time so much fun.

I am so grateful for my Baumann/Kenseth/Eriksson/Grodås family for their support and love and to my three brothers, Lars, Christian, and Eric, for always making me laugh and keeping my spirits up during this time. To those of you who took the time to read through this thesis for the NorwEnglish that has developed in my time in Norway, I'm sorry and thank you.

Finally, thank you to my dear husband Christian for being such a considerate and doting partner as I finished this thesis. You are such a source of light and strength through good times and bad and I'm grateful for whatever forces in the universe brought us together. Thanks for all of the walks you've taken our doggy Penny on while I was inside writing.

Anna Ruth Kenseth
November, 2020 Oslo

Abbreviations

AE	Adverse Event
AI	Artificial Intelligence
AIC	Akaike Information Criterion
ALL	Acute Lymphocytic Leukemia
BIA	Budget Impact Analysis
BIC	Bayesian Information Criterion
CBA	Cost Benefit Analysis
CEA	Cost-effectiveness analysis
CMA	Cost Minimization Analysis
CRC	Colorectal cancer
CT	Computerized Tomography
CUA	Cost Utility Analysis
DNA	Deoxyribonucleic acid
DRG	Diagnosis-Related Groups
GBP	Great British Pound
GP	General Practitioner
H&E	Hematoxylin and Eosin
HRQoL	Health-related Quality of Life
HTA	Health Technology Assessment
ICER	Incremental cost-effectiveness ratio
ICGI	Institute for Cancer Genetics and Informatics
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
IQR	Inter-quartile Range
KM	Kaplan–Meier plot
LY	Life years
MDT	Multidisciplinary Team
MRI	Magnetic Resonance Imaging
MSI	Microsatellite instability
MSI-H	Microsatellite instability-high
NIH	National Institutes of Health

NIPH	Norwegian Institute of Public Health
NOK	Norwegian krone
NPV	Negative Predictive Value
OP	Outpatient
OS	Overall Survival
PD	Progressed disease
PF	Progression free
PICO	Population, Intervention, Comparator, Outcome
PPV	Positive Predictive Value
PSA	Probabilistic Sensitivity Analysis
QALY	Quality-adjusted life-year
QI/II/III/IV	Quadrant 1/2/3/4
RCT	Randomized Controlled Trial
RCS	Restricted cubic splines
RHA	Regional Health Authorities
SoC	Standard of care
TNM	Tumor Node Metastasis
WHO	World Health Organization
WTP	Willingness to pay

Table of Contents

ABSTRACT III
ACKNOWLEDGEMENTS V
ABBREVIATIONS VII

TABLE OF CONTENTS ..... **IX**

LIST OF FIGURES XI
LIST OF TABLES XIII
1. INTRODUCTION 1
2. BACKGROUND 4
2.1 Colorectal Cancer *4*
2.2 Standard of Care *4*
2.3 The Age of Personalized Medicine *8*
Prognostication 8
Prognostic Tools in Practice: Oncologist Stakeholder Input 8
Biomarkers for Stratification in CRC: Surgeon Stakeholder Input 9
The Future of Precision Medicine: Biotech Analyst Stakeholder Input 11
2.4 Histotyping: A Case Study *11*
DoMore! 11
Histotyping 12
2.5 Norwegian Healthcare System *14*
2.6 Reimbursement in Norwegian Healthcare system *15*
3. THEORETICAL FRAMEWORK 17
3.1 Economic Evaluation and Health Technology Assessments *17*
3.2 Early HTA *17*
Methods of Economic Evaluation 20
The QALY 20
Cost Perspective 21
Time Horizon and Time Preference 21
The Decision Rule 22
Uncertainty 23
The CE Plane 24
3.3 Evaluating Prognostic Tools *25*
Sensitivity 25
Specificity 26
Using Models 28
Survival Analysis 30
Extrapolating Survival 32
Sensitivity Analysis 33
4. METHODOLOGY 35
4.1 Population *35*
4.2 Intervention and Comparator *35*
4.3 Perspective *38*
4.4 Outcome *38*
4.5 Time Horizon *38*
4.6 Model Structure *38*
4.7 Data Inputs *41*
Transition Probabilities 42
Adverse Events 49
Health Outcome 50
Costs 51
Modeling the Effect of Histotyping 55
Sensitivity Analysis 57
5. RESULTS 60
5.1 Early Cost-Utility Analysis *60*
Disaggregated Results 61

Scenario Analysis 62
5.2 Protocol *65*
6. DISCUSSION 66
6.1 Main Findings *66*
6.2 Limitations and Future Research *67*
6.3 Strengths *70*
7. CONCLUSION 71
REFERENCES 72
APPENDIX 77
Literature Search *77*
Survival Analysis *78*
Treatment Schedules *80*
DECLARATION IN LIEU OF OATH 81

List of Figures

Figure 1	Stage II CRC patient stratification
Figure 2	Stage III CRC patient stratification
Figure 3	A simplified depiction of the CRC patient pathway where Histotyping is likely to be integrated
Figure 4	Comparison of late and early-CEAs of medical tests
Figure 5	Cost-effectiveness plane as described by Briggs and Penn
Figure 6	Depiction of sensitivity and specificity in a Histotyping test
Figure 7	Illustrative example of how prevalence impacts positive and negative predictive values
Figure 8	The conceptual model of Histotyping integrated with the standard of care, reflecting the different pathways
Figure 9	Model structure for partitioned survival analysis health outcomes. Each health state depends on stage (II or III) and prognosis group (good, uncertain, poor).
Figure 10	Stage II survival extrapolated using RCS on all years of observation. Stage II CRC survival from a Norwegian population is included for comparison.
Figure 11	Stage III survival extrapolated using RCS on all years of observation. Stage III CRC survival from a Norwegian population is included for comparison.
Figure 12	Stage II survival extrapolated using RCS with three years of observation. Stage II CRC survival from a Norwegian population is included for comparison.
Figure 13	Stage III survival extrapolated using RCS with three years of observation. Stage III CRC survival from a Norwegian population is included for comparison.
Figure 14	Stage III survival extrapolated using RCS with five years of observation. Stage III CRC survival from a Norwegian population is included for comparison.
Figure 15	Stage II survival extrapolated hybrid of RCS with three years of observation and Norwegian life table data. Stage II CRC survival from a Norwegian population is included for comparison.

Figure 16 Stage III survival extrapolated hybrid of RCS with censored data from five years and adapted curve for uncertain prognosis. Stage III CRC survival from a Norwegian population is included for comparison.

Figure 17 EQ-5D utility evolution over time by treatment arms. Average and 95% CI

Appendix Figure 1 Smoothed hazard function for stage II poor prognosis

Appendix Figure 2 Smoothed hazard function for stage III poor prognosis

Appendix Figure 3 Survival was underestimated by all parametric functions when applied to the stage II poor prognosis KM

Appendix Figure 4 Survival was strongly underestimated by all parametric functions when applied to the stage III poor prognosis KM

List of Tables

Table 1	Stage II and III CRC classification rates and 5-year survival
Table 2	Explanation of adjuvant chemotherapy options in the Norwegian guidelines
Table 3	Stata OS survival output of spline models in Stage II good, uncertain and poor prognosis groups
Table 4	Stata OS survival output of spline models in Stage III good, uncertain and poor prognosis groups
Table 5	Incidence of adverse events for CAPOX 3-month and 6-month arms
Table 6	Incidence of adverse events for Capecitabine monotherapy 6-month arm in stage II CRC
Table 7	Norwegian stage II and III colorectal cancer adjuvant chemotherapy guidelines
Table 8	Chemotherapy assignment
Table 9	PF, PD and Death health state costs and sources
Table 10	Base case scenario versus the standard of care
Table 11	PSA parameter distribution assumptions
Table 12	PSA parameter standard error assumptions
Table 13	Early cost-effectiveness of Histotyping plus the standard of care vs. the standard of care alone in Stage II and III CRC patients with a time horizon of 10 years
Table 14	Disaggregated and undiscounted costs of Histotyping plus the standard of care vs. the standard of care alone in Stage II and III CRC patients in the base case scenario with a time horizon of 10 years
Table 15	Intensified chemotherapy regimen for stage III poor prognosis patients treated with Histotyping plus the standard of care vs. the standard of care alone.
Table 16	Recurrence rate increases for undertreatment scenarios in the good prognosis patients in both stage II and III treated with Histotyping plus the standard of care.
Appendix Table 1	Limited literature search strategy
Appendix Table 2	Adjuvant chemotherapy schedules by therapy type and duration

1. Introduction

As the third most common diagnosed malignancy worldwide and the second most common cause of cancer deaths, colorectal cancer (CRC) accounted for an estimated 1.85 million new cases and 881,000 deaths in 2018 (1) with older age groups (≥65) shown to contribute the greatest proportion of incident cases (2). Cancer incidence is on the rise globally, with estimates from the World Health Organization (WHO) projecting a 45% increase in cancer deaths between 2008 and 2030 (3). These figures hold true in Norway where more than 4,000 new CRC cases are diagnosed every year (4).

Most stage II and III CRC patients will survive five years or longer, however, survival will greatly depend on the ability to select appropriate treatment (5). While stage II low risk CRC is effectively treated with surgery alone, high risk stage II and all stage III cancer patients will be recommended adjuvant chemotherapy to reduce the risk that a cancer will recur (6).

Adjuvant chemotherapy has been shown to improve survival in some cases and is offered at the discretion of the oncologist based on a patient's risk profile, but without a strong prognostication method there is a risk of over- or under-treatment (7). With this risk comes the potential that patients will experience severe treatment-induced side effects, or recurrence due to undertreatment, which greatly impacts their health-related quality of life (HRQoL). It is crucial that CRC patients who could benefit from adjuvant chemotherapy are correctly identified and offered the appropriate treatment in order to lessen their risk of recurrence and reduce the treatment burden, not add to it.

A biomarker developed at the Institute for Cancer Genetics and Informatics (ICGI), DoMore-v1-CRC, shows promise for stage II and III CRC patient stratification and potential for

supplementing adjuvant chemotherapy treatment decision making in the clinical setting. Histotyping, as the method is called, is a novel tool and its success will depend on its real-world utility in a clinical trial or clinical setting. Unlike with other markets, the adoption of new products in healthcare depends on the ability to demonstrate cost-effectiveness (8) and what added value it can provide.

In the latest report from the Cancer Registry of Norway, there was an observed increase in colon cancer incidence by 1.3% in men, and 5.2% in women when the most recent 5-year period (2014-2018) was compared to the previous period (4). This is expected to increase in the coming decades due to an aging population, potentially imposing a substantial economic burden on the healthcare system due to treatment and care costs.

If we are able to accurately, and in an automated fashion, categorize patients by their likely outcomes and assign them to the appropriate treatment pathway, there is undoubtably much to be gained both by patients through greater access to a tool that can provide a reduced treatment burden and increased quality of life, and by the healthcare system through a reduced cost, and a broadened provision of care.

The purpose of this thesis is to develop and apply an early health technology assessment framework in order to study the impact of the integration of Histotyping with the standard of care for stage II and III CRC patients. A meeting conducted with a gastrointestinal surgeon strongly informed the model framework and provided clarity to complexities of the patient pathway in the standard of care. Meetings with an oncologist and a biotech analyst contextualized the current use of prognostic biomarkers in cancer and explained the market entry process and current need for predictive prognostic tools like Histotyping.

With this thesis, we objectively explore the costs and benefits of Histotyping. The literature that exists on the topic is from one source in a scientific discovery and validation study (9). This thesis is the first analysis of Histotyping's performance relating quantitatively and qualitatively how its properties as a tool might have implications for patient use. To our knowledge at the time of submission, this thesis is also the first instance in which Norwegian survival data has been juxtaposed with Histotyping survival data. Finally, we contribute a well-researched structure for an early cost-effectiveness analysis and a protocol including suggestions for next steps.

2. Background

2.1 Colorectal Cancer

Colon- and rectal cancer are often grouped together under the term colorectal cancer (CRC) due to similarities they share in mortality, treatment and anatomical proximity. A primary CRC diagnosis is generally determined by colonoscopy and other imaging such as MRI or CT scanning, and a pathologist's microscopic analysis provides insight into what's known as the Tumor, Node, Metastasis (TNM) classification (10). TNM classification is essential for a primary diagnosis, where "T" implies the presence of cancerous cells and the size of the primary tumor, "N" indicates if and how many of the nearby lymph nodes the cancer has infiltrated, and "M" signifies whether the cancer has metastasized, or spread to other organs (11). These pieces of evidence are the basis for what is commonly known as the stage, ranging in severity from localized stage I to metastasized stage IV cancer.

Today, TNM classification is determined through individual assessments by a pathologist, and inform the clinician about a cancer's likely outcome (prognosis), guiding the treatment decision. For cancer patients, the likelihood of a cancer's return following treatment is called the risk of recurrence, with increasing risk corresponding with increasing stage assignment (12).

2.2 Standard of Care

Surgery is the main line of treatment for CRC and can be curative, however about 16% of low risk stage II CRC patients and 50% of stage III CRC patients will experience a recurrence in the five years following surgery (13). Low risk and high risk classifications will guide the treatment decision, but a substantial proportion of CRC patients will fall in an

intermediate risk group (Table 1) for which there are no clear procedures in the Norwegian guidelines (5). As a result, the majority of these patients will be recommended adjuvant chemotherapy. Adjuvant chemotherapy was integrated in the early 1990's as part of the clinical routine for the treatment of CRC, with the intention to eradicate residual cancer cells with the potential to multiply uncontrollably, leading to cancer recurrence, metastasis and death (14, 15).

Table 1. Stage II and III CRC classification rates and 5-year survival (5)

Stage	T status	N status	M status	Classification (US SEER)	5-year survival (US patients ≥ 65)
IIA (II low risk)	T3	N0	M0	85%	84.4
IIB (II high risk)	T4	N0	M0	15%	55
IIIA (III low risk)	T1, T2	N1	M0	11%	85.1
IIIB (III intermediate risk)	T3, T4	N1	M0	57%	64.6
IIIC (III high risk)	T (any)	N2	M0	32%	45.5

For patients under the age of 75, adjuvant chemotherapy is a standard recommendation following surgical resection of stage III CRC, and is recommended for stage II CRC high risk patients where tumor-related risk factors are present (11). High risk criteria include pathological and histological characteristics such high grade tumor (T4 status), poor differentiation in the tissue morphology, and tumoral perforation or invasion into the nerves, blood vessels or lymphatic system; weighing risk factors against one another presents a challenge to decision makers (16).

Adjuvant chemotherapy has been shown to prevent recurrence, improve survival and ameliorate symptoms (17) which is of particular relevance for stages II and III where recurrence risk is increased but potential for survival remains good (6). Depending on the

patient and tumor characteristics, chemotherapy can be offered as a single agent (monotherapy) or as multiple agents (combination therapy). The type and dose of chemotherapy is selected at the discretion of the oncologist and tailored to what best suits the individual patient.

Chemotherapy is cytotoxic and as a result, can be harmful to a patient's healthy cells, with the risk of severe adverse events (AE). AEs can vary in grade, ranging from mild discomfort to life-threatening complications as a result of the agent(s) given. Common chemotherapy-induced AEs are hair loss, nausea, diarrhea, anemia, neutropenia, febrile neutropenia, and peripheral neuropathy.

Neutropenia and peripheral neuropathy are particularly common AEs experienced by patients treated with platinum-based chemotherapy, such as Oxaliplatin. Oxaliplatin is limited to particular doses due to its cytotoxic properties, but Oxaliplatin-containing regimens have a demonstrated superiority when compared to non-Oxaliplatin containing therapies for treating stage II and III CRC (6, 18). As seen in Table 2, Oxaliplatin is a component present in all combination therapies recommended in the Norwegian treatment guidelines.

Table 2. Explanation of adjuvant chemotherapy options in the Norwegian guidelines

Name(s)	Name Description	Dosage	Known AE
5/FU	monotherapy 5-FU or Capecitabine	Capecitabine: 2500 mg/m^2/day (14 out of 21 days)	anorexia, nausea, vomiting, and diarrhea (19)
XELOX (CAPOX)	combination therapy XEL/CAP = Capecitabine (Xeloda) OX = Oxaliplatin (Eloxatin)	Oxaliplatin (I.V.): 130 mg/m2 on day 1 Capecitabine (oral): 1,000 mg/m2 twice daily on days 1-14 every 3 weeks for a total of either 4 or 8 cycles	neurotoxicity, leucopenia, thrombocytopenia, nausea, diarrhea, hand-foot syndrome (20-22)
FOLFOX	combination therapy FOL = Folinic acid (Leucovorin) F=Fluorouracil (5-FU) Ox=Oxaliplatin (Eloxatin)	Oxaliplatin 85mg/m2 IV : 2 hrs, day 1 Leucovorin 400mg/m2 IV : 2 hrs, day 1 Fluorouracil 400mg/m2 IV day 1, then 1,200mg/m2/day × 2 days (total 2,400mg/m2 over 46–48 hours) IV continuous infusion. Repeat cycle every 2 weeks for 12 cycles	nausea, vomiting, skin reactions, diarrhea, stomatitis, visual problems, hair loss, neutropenia (22, 23)
FLOX	combination therapy F= Fluorouracil (5-FU) L=Folinic Acid (Leucovorin) OX=Oxaliplatin (Eloxatin)	FLOX: 5-FU 500 mg/m2 IV weekly plus leucovorin 500 mg/m2 IV weekly for 6-wk (days 1, 8, 15, 22, 29, and 36) of each 8-wk cycle plus oxaliplatin 85 mg/m2 IV administered on days 1, 15, and 29 of each 8-wk cycle for three cycles	nausea, vomiting, diarrhea, anemia, neutropenia, thrombocytopenia

Neutropenia is characterized by a steep decrease in a type of white blood cell, neutrophils, which makes patients more prone to serious and life-threatening infection, requiring immediate hospitalization. It is also a major factor in the overall cost for adults with cancer. A study of United States hospital costs in 2012 showed the cost for adults treated for cancer-related neutropenia was 2.3 billion USD, accounting for 8% of all cancer-related costs for adults (24).

A patient's risk of recurrence can be reduced by adjuvant chemotherapy but comes at a cost and a substantial toxicity profile. Therefore, it is crucial that CRC patients who could benefit from adjuvant therapy are correctly identified and offered the appropriate treatment in order to reduce their risk of recurrence and the treatment burden, not add to it.

2.3 The Age of Personalized Medicine

Prognostication

In the field of cancer medicine, a growing base of literature focuses on the identification of prognostic predictive tools, called prognostic biomarkers, to understand a patient's likely outcome and survival. These biomarkers can be particularly helpful in stratifying patients based on their risk of cancer recurrence, which can aid in selecting the appropriate treatment.

If patients can be identified on the basis of their risk of recurrence, they can receive appropriate treatment and avoid AEs from overtreatment, rather investing valuable time and health resources on effective treatment. For example, if a patient is stratified into a low risk of recurrence, he or she could be offered less intensive adjuvant chemotherapy, potentially reducing morbidity and mortality caused by the therapy itself. Similarly, if we can identify patients with a high risk of recurrence that would benefit from a more intensive regimen, then there is potential to save lives or improve survival by intensifying therapy.

Additionally, accurate risk stratification is associated with more efficient allocation and use of health care resources, which can improve the overall efficiency of health care coordination and provision (25).

Prognostic Tools in Practice: Oncologist Stakeholder Input

In some fields, prognostic biomarkers are a standard of treatment selection. For example, risk-stratifying biomarkers are successfully being used in pediatrics where the prevalence of rare tumor cases requires tools to supplement decision making, as we were informed through our meeting with an oncologist.

In the case of acute lymphoblastic leukemia (ALL), patients are risk-stratified on the basis of the genetics of their leukemia with the use of a molecular biomarker classifying patients as low, standard, high and very high risk. Patients with high and very high risk features receive more intensive treatment, while standard risk and low risk patients receive reduced treatment. The result of this is minimized long-term chemotherapy-induced toxicity as well as drastically improved 5-year survival, from 10% in the 1960's to 90% in 2000 (26). This example provides a promising outlook for other fields of oncology. If prognostic tools can be used to reliably stratify patients by risk, they can be applied to tailor treatment.

In another example of how prognostic biomarkers have driven prognostication and treatment choices, the WHO recently reclassified the staging of pediatric brain cancer from its purely histological basis. As more has been learned about tumor profiling, diffuse midline pediatric glioma tumors that harbor specific mutations are now classified as stage IV cancer regardless of their histological presentation under the microscope (27). This is an interesting case in which the classic clinical information which once informed the likely outcome and choice of treatment is now overruled by a more powerful piece of prognostic information in the form of a prognostic biomarker.

Biomarkers for Stratification in CRC: Surgeon Stakeholder Input

In recent years, research has been devoted to the study of molecular biomarkers for Microsatellite Instability (MSI), and genes *KRAS* and *BRAF* for their potential to stratify CRC patients by risk of recurrence.

KRAS and *BRAF* are the most frequently mutated cancer genes in CRC, with *KRAS* appearing in 15-37% of early stage CRC, and *BRAF* appearing in 10% of cases (28). Despite their

prevalence and years of targeting attempts for drug development, neither have yet to show any predictive prognostic value for CRC, nor are they recommended for the routine assessment of risk of recurrence in non-metastatic patients (29).

As we learned through meeting with a specialist gastrointestinal surgeon, *KRAS* and *BRAF* mutation testing are used today as predictive markers to refine the type of chemotherapy recommended. Requested in most cases by the oncologist in the palliative or recurrent setting, if a patient tests positive for either of these mutations, they will be resistant to treatment with cetuximab (30).

According to the surgeon, testing for MSI was integrated in the Norwegian clinical routine in the last few years and can identify areas within the DNA where the genes responsible for single nucleotide mismatch repair malfunction (31). Stage II CRC patients with a high frequency of MSI (MSI-H) have been shown to have improved survival compared to microsatellite stable patients (32). Treatment guidelines recommend the inclusion of MSI test results as part of the pathological report for patients with CRC as it is of consequence for treatment selection. These patients encompass a 10-15% subset of stage II CRC for whom the benefits of adjuvant chemotherapy are marginal and likely unnecessary given their low likelihood of recurrence (33).

These three biomarkers have emerged within the last 10 years and as such are in their early stages of use. Although they show potential for targeted treatment, future cost-effectiveness analyses will depend on their predictive power and the quality of information provided in clinical trials.

The Future of Precision Medicine: Biotech Analyst Stakeholder Input

In a meeting with a biotech analyst, we learned the market for prognostic tools and precision medicine has a strong and growing foothold. In the field of cancer medicine, we can expect the coming three to five years will show much better results for patients with individually tailored treatment choices. The main challenges for prospective new tools will be in demonstrating an added benefit to both the consumer and the healthcare system and will depend on the current knowledge base in order for confident adoption to the clinical routine. In this regard, a well-evidenced prognostic power of the biomarker will be crucial.

2.4 Histotyping: A Case Study

DoMore!

Robust cancer diagnostics are necessary in order to provide the optimal course of treatment to cancer patients. The Institute for Cancer Genetics and Informatics (ICGI) at The Norwegian Radium Hospital in Oslo is one group that has focused on this since its inception in 2004. Under the guidance of founder Professor Håvard E. G. Danielsen, ICGI initiated the DoMore! project in 2016 aimed at applying medical informatics to solving the societal burden of cancer. The DoMore! team is an interdisciplinary collaboration of international experts within digital imaging, processing, pathology, cell biology, surgery and oncology (34). This initiative is funded by the Research Council of Norway through its IKTPLUSS Lighthouse Project grant, which allotted 60 million Norwegian Kroner (NOK) over a 5-year period. DoMore! was selected in part based on its commitment to solving large societal challenges through the development of new technologies.

Histotyping

One product that has emerged from the project is called "Histotyping" and culminated in 2019 in a landmark artificial intelligence (AI) publication in the Lancet (9). Histotyping involves an automated analysis of high resolution scans of cancerous tissue that has been stained with routine DNA-specific stains hematoxylin and eosin (H&E). To develop this method, ICGI utilized convolutional neural networks trained on large retrospective clinical trial data with known patient outcomes, and trained deep learning models to identify and analyze the morphology of a tumor in a surgically resected tissue section. Histotyping estimates the probability of a good, poor or uncertain outcome (Figure 1 & 2), i.e. the likelihood of a recurrence, and could be employed as a supplementary diagnostic service to support clinicians in their adjuvant chemotherapy decision making for patients with stages II and III CRC.

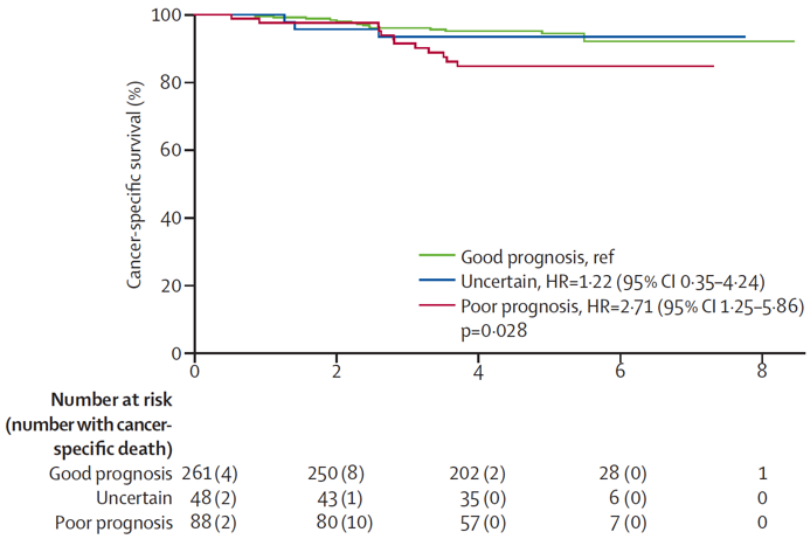


Figure 1. Stage II CRC patient stratification

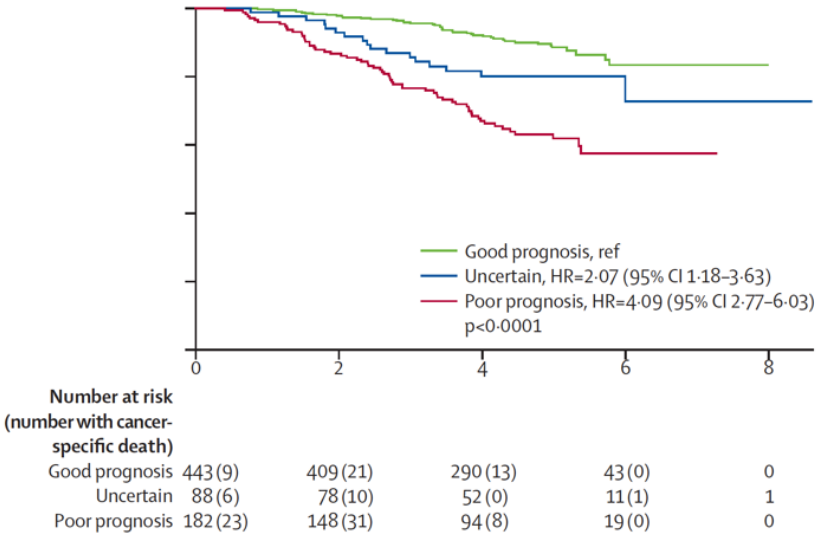


Figure 2. Stage III CRC patients stratification

Norwegian treatment guidelines provide suggestions for adjuvant chemotherapy, which are updated annually based on the latest leading publications. Selection of the type and duration of chemotherapy for individual patients is at the discretion of the oncologist based on patient characteristics following a discussion with a multidisciplinary team (MDT) as seen in Figure 3 (35). Opinions remain split on the decision making around appropriate adjuvant chemotherapy assignment, and research remains devoted to the refinement of duration and intensity of therapy choices.

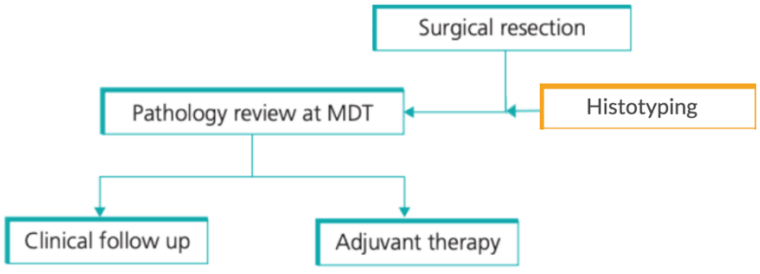


Figure 3. A simplified depiction of the CRC patient pathway where Histotyping is likely to be integrated

Histotyping is intended for use by clinicians as a software or offered as a service to stratify CRC patients for the prevision of appropriate treatment. The potential benefits of this are two-fold, both in identifying patients for whom the absolute benefit of adjuvant chemotherapy is low (a good prognosis assignment) as well as those that could benefit from prolonged or more intensive chemotherapy based on their poor prognosis assignment. This is interesting from a health economics perspective, as there is potential for greater accessibility based on the nature of this automated test, at low cost for improved outcomes, and at a lessened treatment burden. Histotyping is a novel biomarker with its real-world clinical utility remaining to be demonstrated, and commercialization will be contingent upon what added value it can first provide in the Norwegian healthcare setting.

2.5 Norwegian Healthcare System

In Norway, universal access to high quality medical care is the ultimate goal. Healthcare follows a two-tiered Nordic welfare state model primarily funded publicly through taxes, but allowing the opportunity for patients to seek additional services privately and at their own expense. Minimizing inequality and maximizing health are central philosophies to this system, underscored by a principle of patient autonomy, which facilitates active patient participation in choosing care providers and determining care decisions (36).

Healthcare prioritization is a process that takes natural effect regardless of the choice to prioritize or not. While the UK maintains a willingness to pay (WTP) threshold of 20,000 – 30,000 GBP (37) and the United States slightly higher at 50,000 USD (38). Norway has three foundational principles for priority setting: expected benefit, cost-effectiveness, and illness severity (39). In 2015, a new set of guidelines were published with explicit differential WTP thresholds, based on a starting value of 275,000 NOK that is weighted by severity (40).

In Norway, the resources allocated to healthcare are limited. Funding decisions will be made actively or passively and the consequences for these decisions are widespread (41). Patients have the right to choose their healthcare provider for all non-emergent care. For those living in more rural settings without direct access to specialized services and the larger hospitals, inequalities in access are inevitable. In this way, ethics and equal access to care is a point of concern and emphasizes the importance of health interventions to be scalable and therefore easily accessed. As Nolte and McKee note, health services have a growing role in health promotion and in narrowing health inequalities resulting from social disadvantage (42).

2.6 Reimbursement in Norwegian Healthcare system

The procedure for access to the market in Norway is straightforward for pharmaceuticals, with clear guidelines for requirements surrounding quality, safety and medical efficacy. The entire process of economic evaluation and market authorizations is carried out by the Norwegian Medicines Agency, a subordinate agency of the Ministry of Health and Care Services.

When new diagnostic and prognostic tools and technologies attempt to gain market access, the procedure is different and falls under the purview of multiple players. Upon the identification of a new technology, The Norwegian Directorate of Health forwards proposals for HTAs to The Norwegian Institute of Public Health (NIPH) to ascertain its relevance. If relevant, the Medical Directors from each of the four Regional Health Authorities (RHA) delegate tasks for HTA, which will be performed by the NIPH and evaluated by the RHA for implementation and monitoring. Extensive documentation is required from the applicant including cost-effectiveness information. For this reason, early collection of these data for

evidence-gathering and case-building purposes is essential to have in place once market access is sought.

3. Theoretical Framework

3.1 Economic Evaluation and Health Technology Assessments

Norwegian healthcare resources are limited, so there is pressure to be fiscally sustainable. Wherever healthcare spending is concerned, there is a goal of efficiency where costs are contained and state-of-the-art treatment, responsiveness and choice are achieved (43). For this reason, economic evaluation serves an important role in the consideration and ranking of health interventions selected for reimbursement. Serving as an analytic and advisory tool, economic evaluation aims to provide an assurance of value for resources used and enables transparency in decision making.

3.2 Early HTA

Where new technologies are concerned, health technology assessments (HTA) are used to compare two or more alternatives in terms of cost and benefit. As precision medicine is a growing field and biomarker discovery is much sought after, cost-effectiveness is likely to become more central to the early phases of development in the form of "early HTA". Early HTA employs all of the traditional methods used in standard economic evaluation within HTA, but without the availability of clinical effect evidence. Given the time and resources devoted to product research and development, there are strong arguments in favor of early application of current models in order to explore potential benefits earlier in the process (44). One strong point in favor of early HTA is it facilitates us to make an initial assessment of a new technology to define its potential for use and cost-effectiveness, as well as whether there is likely to be a return on the investments made.

As opposed to standard HTA, which is a post-estimation performed when the efficacy is well assessed and evaluated, an early HTA defines the conditions under which a tool would be

potentially cost-effective. Furthermore, the framework for early HTA is iterative and adjustable once new information or potential areas of use become apparent. Early HTA is characterized by certain hallmarks as has been described in several publications, including one by Buisman et al., which clearly depicts the differences between early and late HTA in the form of a CEA (Figure 4) (44). Other references are cited in the steps detailed below.

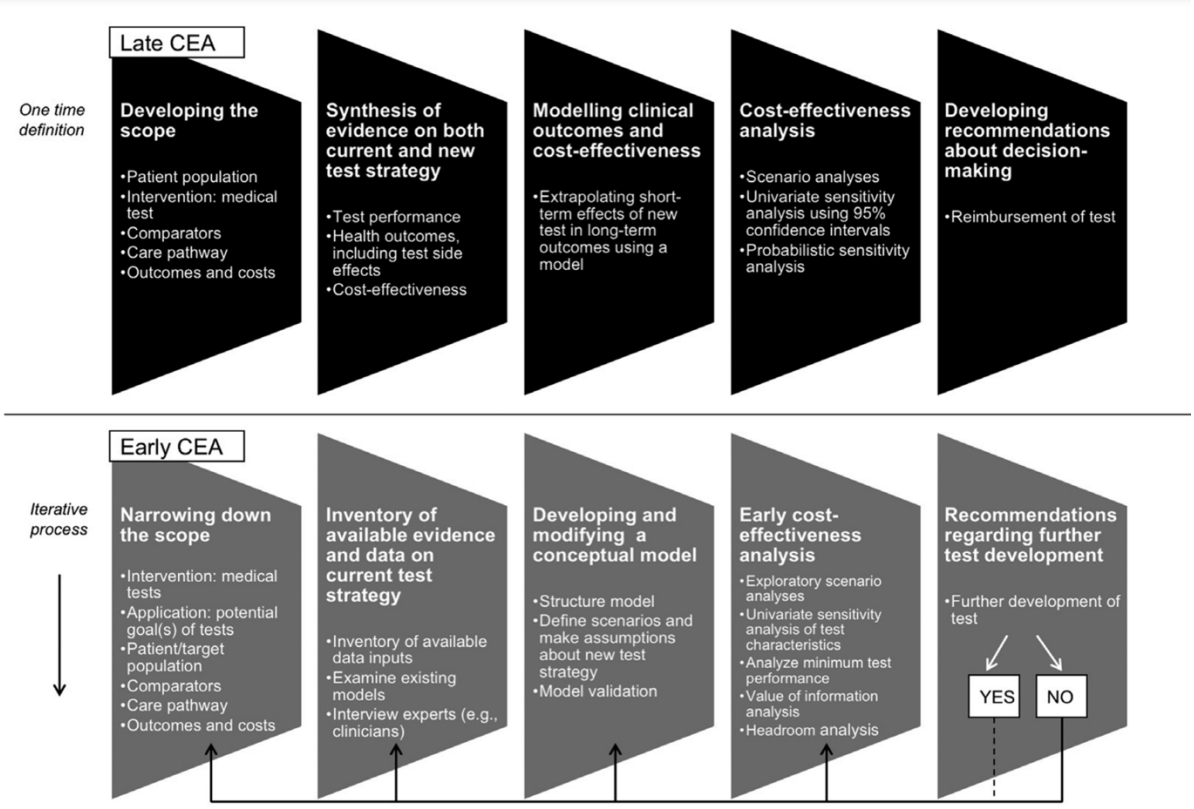


Figure 4. Comparison of late and early-CEAs of medical tests (44)

A first step in the process is narrowing down the scope of where the technology is likely to be implemented, commonly using the Population, Intervention, Comparator and Outcomes (PICO) method (45).

Second, we make an inventory of available evidence and data on our intervention. The data available for cost and health outcomes tend to be limited and expert opinion is typically used to strengthen assumptions (44). There is a lot of value in expert opinion, as it enables us to

handle gaps in evidence on probabilities such as those used between health states, or effect sizes such as likely prognostic performance (46). In addition, considering stakeholder input and preference can allow for estimation of relative value for the new technology, or tool of interest.

Third, an integral part of early HTA is drafting a conceptual model to scope out the various possible scenarios of population characteristics, disease prevalence, test performance, and cost among other aspects (47).

Fourth, based on the information gathered, a cost-effectiveness analysis can be performed in order to determine a benchmark of performance requirements for cost-effectiveness. Another aspect of this is to include a headroom analysis, which assesses the maximum price-setting possibility where the intervention is still cost-effective and reimbursable (8).

The final step in an early HTA framework is the development of recommendations based on the results and inputs from the previous steps. This is an essential step if an early HTA is being performed to determine whether further investments should be made (44). Quality assurance of early HTA, and economic evaluations in general, involves complying with the latest recommended guidelines according to the International Society for Pharmacoeconomics and Outcomes Research (ISPOR). ISPOR offers a wealth of material on standards and practices for health economics and outcomes research such as the best practices for conceptualizing a model and defining the objectives, scope and policy context (48), estimating parameters and analyzing uncertainty (49), and transparency and validation (50).

Methods of Economic Evaluation

Economic evaluation can take many forms depending on the outcome of interest. Types include: Budget Impact Analysis (BIA), Cost Minimization Analysis (CMA), Cost Benefit Analysis (CBA), Cost-Effectiveness Analysis (CEA) and Cost Utility Analysis (CUA).

Prior to adopting a new technology or product, a BIA can be used to determine the impact of implementation (51). When costs are the sole desired outcome measure, a CMA provides a comparison of the costs alone, with all other outcomes being equal. Cost benefit analyses (CBA) are commonly used in economic analyses outside the field of health where costs and effects are both expressed monetarily.

In healthcare, the most common analyses are the CEA and CUA. With a CEA, costs are expressed in monetary terms and effects in natural units (i.e. life years gained). A CUA sets costs in monetary terms and effects in quality adjusted life years (QALYs) (52).

The QALY

Health is multi-dimensional and requires many considerations and QALYs are an appropriate measure of benefits where treatment has the potential to not only prolong life but to impact a patient's quality of life.

The QALY is a value function of quality ranging from the lowest value, 0 (death) and the greatest, 1 (perfect health) representing the overall quality of life in a particular health state (53). The formula for deriving the QALY is depicted in Equation 1. A year of life spent in perfect health will be equal to 1 QALY.

$$\underbrace{\text{QALY}}_{\text{Quality Adjusted Life Years}} = \underbrace{V}_{\text{Value function of quality (range 0-1)}} * \underbrace{Y}_{\text{\# Life years spent with this quality}} \quad [1]$$

The input for the value function, V, is patient-reported HRQoL data and the Norwegian guidelines require the use of a questionnaire called EQ-5D (54). The EQ-5D is a generic instrument, meaning it is not disease-specific and survey questions pertain to health status and ability to perform certain tasks (54). Data is often accessed directly from clinical trials, or otherwise from the literature. The advantage of the EQ-5D is that it allows for cross-comparability between health technology assessments.

Cost Perspective

The most important determinant of the type of costs that are included is the perspective through which the early HTA is conducted. Perspective defines the breadth of stakeholders and agents included within an analysis and helps with cost identification. While a patient and family perspective will cover expenses such as productivity costs, judicial costs, out of pocket costs, travel costs and informal care time costs, a healthcare perspective limits costs to fixed and variable costs pertaining only to primary and secondary medical care, and a societal perspective includes a range of costs including healthcare and other sectors, and patient and family costs. Healthcare is unlike the competitive market, where costs are transparent, so it can often be challenging to identify unit costs. Where the cost is unknown, list prices and publicly available Diagnosis Related Groups (DRGs) are used as proxies.

Time Horizon and Time Preference

Most interventions in healthcare require a long-term perspective for both costs and effects. A phenomenon of social time preference describes that costs later in time will weigh

less and effects later in time will have less value (55). As a result, it is common to integrate a process of discounting in order to bring future costs to their present value for an accurate comparison of interventions. Furthermore, the choice of whether to include marginal costs as opposed to average absolute costs can provide information about how changes in costs will be reflected in outcomes, which is valuable information when considering fiscal efficiency (56).

The Decision Rule

When the net benefits of an intervention outweigh the costs of the investment, it will have a positive impact on social welfare. It is not just the ratio of net costs and effects, but the ratio of incremental costs over incremental effects that allows broad cross-comparability across different interventions. The formula for incremental cost-effectiveness ratio (ICER) is described in Equation 2.

$$ICER = \frac{Cost_{new\ technology} - Cost_{standard\ of\ care}}{Effect_{new\ technology} - Effect_{standard\ of\ care}} = \frac{\Delta C}{\Delta E} \qquad [2]$$

In a typical decision rule where interventions are mutually exclusive, the ICER that is most favorable, and below a given threshold will be reimbursed. In early HTA, there will likely be several ICERs resulting from the various assumptions made and scenarios analyzed. With Histotyping, through stratification and adjuvant chemotherapy assignment, the aim is to maximize the health benefits and minimize the treatment burden, reflected in the effects, while minimizing the costs.

A negative ICER can result from either incrementally negative effects or incrementally negative costs. Generally, negative ICERs should not be reported as they can be difficult to

interpret. In order to avoid the ambiguity of a negative ICER, it is preferable to report results in terms of net monetary benefit (NMB) using Equation 3 and as Norway has well-defined thresholds, it is easy to do so.

$$NMB = (\underbrace{\lambda}_{\text{WTP threshold}} \times \Delta E) - \Delta C \qquad [3]$$

The WTP threshold (λ) is calculated based on an estimation of absolute shortfall, where total potential health in expected QALYs for the age of intervention is reduced by the number of average QALYs remaining for patients receiving standard care at the age of the intervention, as in Equation 4. This criterion identifies the amount of future health a patient is expected to lose as a result of their disease, reflecting the disease severity and informing the WTP threshold (40).

$$Absolute\ Shortfall = \underbrace{QALYs_A}_{\substack{\text{Accumulated lifetime QALYs in the general population at age "A"}}} - \underbrace{Prognosis_A}_{\substack{\text{Accumulated lifetime QALYs for a patient with the disease of interest under standard care at age "A"}}} \qquad [4]$$

Uncertainty

Any assessment of cost-effectiveness will be inherently uncertain, either in terms of the parameters used or the structure of the model itself. Parametric uncertainty is common, as the values used in early HTA are estimates often based on a sample population gathered from randomized controlled trials or observational data. Structural uncertainty, however, is more problematic and stems from incorrect assumptions made or the selection of an inappropriate

model type. Economic modeling allows for characterization of the uncertainty surrounding probability, utility and cost inputs, likely having a significant impact on the ICER.

The CE Plane

The cost-effectiveness plane (CE plane) illustrates how potential incremental changes in costs and effects relate to a WTP threshold (57). As seen in Figure 4, the x-axis indicates the potential incremental effect, and the y-axis indicates the potential incremental cost. The four quadrants of the CE plane are defined by these properties. Values that fall in quadrant IV (QIV) and quadrant II (QII) are characterized relative to the standard of care as never and always having the potential for cost-effectiveness, respectively. The threshold rotates around the axis intercept (C) and denotes a maximum acceptable ICER for all points falling below it for a given value of λ (52). With a decreasing threshold value, the maximum acceptable ICER rotates clockwise, and with an increasing value, rotates counter-clockwise.

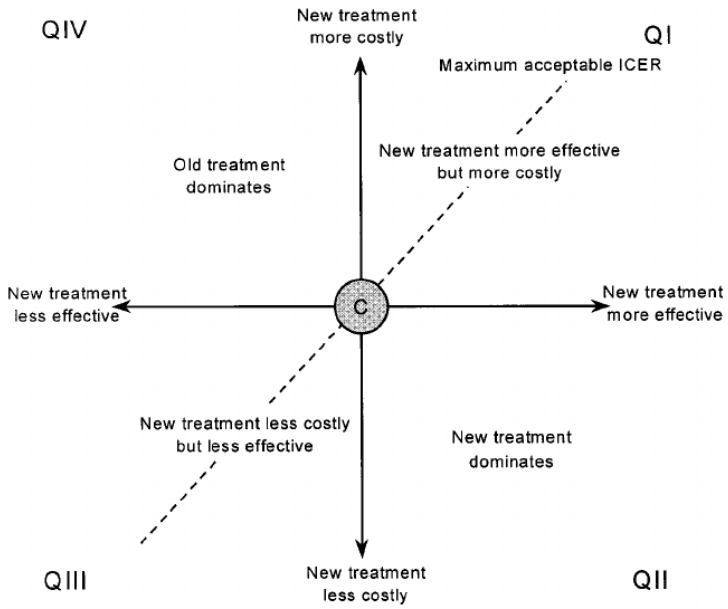


Figure 5. Cost-effectiveness plane as described by Briggs and Penn (58)

The CE plane is commonly referred to for decision making (52), however for early HTAs, results will provide an explorative analysis where potential cost-effectiveness can be examined.

3.3 Evaluating Prognostic Tools

Where diagnostic tools are intended to identify the likelihood that a patient is experiencing a specific disease, prognostic tools identify the risk of future outcomes in patients with a specific diagnosis.

As introduced in section 2.3, prognostic biomarkers in cancer can serve an important role in guiding treatment decisions by identifying patients with particular genetic mutations or as with Histotyping, stratifying patients by their risk of recurrence. The importance of accurate test performance should not be underestimated, namely the sensitivity and specificity (Figure 6). Looking at these concepts through the lens of Histotyping, we can consider patients with stage II and III CRC as the pools of patients being stratified for treatment selection, and for whom the trade-offs are clear.

Sensitivity

If a patient with stage III CRC has a poor outcome, sensitivity is the likelihood of getting a positive Histotyping test result. In other words, sensitivity demonstrates the ability of a test to identify true positives. If sensitivity is high, there is a good degree of certainty that most or all stage III CRC patients with a positive test could benefit from a more intensive regimen of chemotherapy, potentially resulting in improved survival. If sensitivity is low, there is a risk of capturing false positives and the implication would be patients with lower risk stage III CRC being inappropriately selected for more intensive chemotherapy. If this happens, we risk

overtreating a patient that might have otherwise remained healthy or benefited from a reduced regimen.

As we covered, the consequences of overtreatment include AEs from the toxicity of chemotherapy, or in the worst case, chemotherapy can induce a recurrence or premature death. A tool with high sensitivity is necessary to reduce the burden of overtreatment among stage III CRC patients with a negative test, for whom intensive chemotherapy is not necessary.

Specificity

If we consider a stage III CRC patient with a good outcome, specificity is the likelihood of getting a negative test, a true negative. When specificity is high in this case, stage III CRC patients would be appropriately classified in the good prognosis group, which would likely receive a reduced duration or intensity of chemotherapy. If specificity is low, there is a greater chance that a test will capture false negatives in its classification, with the implication for stage III CRC patients being that patients likely to have a poor outcome will be inappropriately given a reduced dose or duration of chemotherapy than necessary.

For these patients, a low specificity will result in undertreatment potentially leading to a lost opportunity for improved survival and quality of life.

Ideally, a prognostic tool in clinical use should be both sensitive and specific in order to minimize the inevitable trade-off between over- and undertreatment and the serious consequences that may entail.

		Condition/disease		
		Poor Outcome (+ patients)	Good Outcome (- patients)	
Test Outcome	Poor Prognosis (+ test)	true positive	false positive	Positive Predictive Value: probability that those assigned to poor prognosis actually have a poor prognosis (greater likelihood of recurrence)
	Good Prognosis (- test)	false negative	true negative	Negative Predictive Value: probability that those assigned to good prognosis actually have a good prognosis (lesser likelihood of recurrence)
		Sensitivity: probability of correctly identifying those with poor prognosis	Specificity: probability of correctly identifying those with good prognosis	

Figure 6. Depiction of sensitivity and specificity in a Histotyping test

While sensitivity and specificity are perhaps the most important measures, the usefulness of the test will also depend on the positive predictive value (PPV) and the negative predictive value (NPV). While the PPV indicates the ability of a test to identify the true positives, those assigned to a positive test result (i.e. poor prognosis) that actually have a poor prognosis, an NPV does the opposite, identifying the true negatives (i.e. those with a good prognosis).

The properties of a test must be sufficiently good in order to avoid overtreatment and undertreatment. In the example shown in Figure 7, the values for sensitivity and specificity are uncertain and if put to use, 95 patients with good outcomes would be overtreated and 5 patients with poor outcomes would be undertreated. If the test properties were better, we would reduce the treatment burden for these 95 good outcome patients and increase treatment intensity for the 5 patients with poor outcome.

One component that is crucial to test performance is the prevalence of patients with good or poor prognosis. To borrow an example from stage II CRC, recall from Table 1 that the number of patients with a very poor prognosis are few. As illustrated in Figure 7, if prevalence within the diseased population is only 5%, the PPV will also be low, 32% in this

example. If this were a screening program, the results would not be especially informative, as in most cases it would not be useful to test an entire sample only to identify 32%. However, since Histotyping is a prognostic biomarker, there is a lot of value in the NPV as it can identify with a high degree of precision those with a good outcome, for whom we would like to reduce the treatment burden.

		Condition/disease			
		Poor Outcome (+ patients)	Good Outcome (- patients)	Total	
Test Outcome	Poor Prognosis (+ test)	45	95	140	Positive Predictive Value: $\frac{45}{140} = 32\%$
	Good Prognosis (- test)	5	855	860	Negative Predictive Value: $\frac{855}{860} = 99\%$
	Total	50	950	1000	
		Sensitivity: $\frac{45}{50} = 90\%$	Specificity: $\frac{855}{950} = 90\%$		
		Prevalence: $\frac{50}{1000} = 5\%$			

Figure 7. Illustrative example of how prevalence impacts positive and negative predictive values

Using Models

Modeling enables us to work within a structure in which we can simulate a disease's outcomes while calculating different scenarios based on changing input parameters. It is necessary to combine evidence on efficacy from underlying sources with evidence on country-specific resource use, unit costs and utilities, and to extrapolate over a longer time horizon, i.e. beyond the period captured in the original study. Furthermore, modeling allows for characterization of the uncertainty surrounding probability, utility and cost inputs, likely having a significant impact on the ICER.

Modeling can be applied in several different forms, such as: a decision-tree, a cohort model (e.g. Markov model), or a partitioned survival model, to name a few.

A decision tree is organized from left to right around different pathways, where each branching is signified by square decision nodes or circular chance nodes indicating points in the pathway where a decision is made, for example a type of treatment, or where different outcomes are assumed based on their likelihood of occurrence. Probabilities allotted to each branch are conditional on the events that transpired in previous branches to the left. Decision trees allow for pathway probabilities and costs to be calculated given a particular set of events and can also be useful in mapping out the testing and treatment pathways. Decision trees can be useful in clearly mapping out the potential pathways, however they are not recommended for complex treatment pathways as they can easily become bulky and difficult to interpret.

A Markov model is organized around defined health states, which are defined to represent the events in the disease process being modeled. Transitions are assumed to occur between the health states and transition probabilities relate to the likelihood of a population at risk of moving between the states at each cycle. A cycle length defines the period of transition, but Markov models reflect the continuous risk of experiencing an event over a longer period of time. One downside to the Markov model is that it does not take patient history into account, with each cycle independent of the cycles that occur previously (59).

A partitioned survival model is similar in structure to a Markov model, but makes use of a theoretical cohort as they progress through pre-defined health states over time as opposed to transition probabilities. The health states are determined by independently modeled survival

curves. The movements between health states in a partitioned survival model are dependent on overall survival and progression free survival and are frequently used in decision analytic modeling due to these two outcomes being widely available in the literature and easily understood by stakeholders (60).

Modeling is a useful tool in that it makes an explicit definition of the relevant patient group, clinical events, patient outcome, and costs, among other components. This conceptual framework shows which data and information are lacking and makes it possible to examine the impact of the input uncertainty on outcome. Furthermore, it is relatively fast, simple and cheap to perform when compared to empirical research.

Modeling has its limitations, however, being prone to bias both in the structuring of the model itself, the selection of health states, and the model input. For these reasons, it is imperative to follow good research practices as set forth by ISPOR (48). Firstly, consult with clinical experts with familiarity of the disease and the ability to qualify the relevant health states for costs and effects. Also, a clear understanding of the decision problem and objectives and a pre-defined scope including the target population, perspective and structure are important. As a principle, modeling allows the creation of a simplified version of a real world scenario, which in reality will be much more complex and there is always the risk that results will be misinterpreted.

Survival Analysis

Survival analysis is a statistical method of estimating events such as death, time to progression, complication or other occurrences over time. Theoretically, a survival analysis begins at time t=0 at which point all are alive, and decline to reach t=end, or the point of exit,

at which point all are dead. However, it is often the case that not all events will be captured within the period of observation due to some random cause, and the time to event will be unknown for some.

This can be due to several factors, namely, a subject survived beyond the end of the study and did not experience an event, a subject dropped out of the study, or follow-up information is missing from the dataset due to some other reason like an illness or relocation. This phenomenon is called censoring and if there are a lot of censored events, it can cause us to severely underestimate the survival time, if not adjusted for in the survival analysis.

This requires us to make an assumption about the likely development of the disease and extrapolate the data to a particular endpoint. Censored data is still of great value in survival analytic models.

The survival function, S(t), is the probability of a subject surviving from the origin to beyond a particular time t, and can be approximated using different methods: a Kaplan-Meier (KM) estimator or parametric method.

The KM estimator is a method commonly used in survival estimations, with the advantage of being non-parametric, i.e. not requiring assumptions about the functional form of survival distribution (61). The KM curve provides a simple step-wise visualization of survival modeled over time, and gives a clear depiction of median survival. The parametric method fits a smooth curve through the observations based on assumptions we make about the distribution.

Extrapolating Survival

In economic evaluations, standard parametric models are often fitted to KM curves using estimated parameters for long-term extrapolation. The most commonly used parametric survival functions are exponential [Equation 5], Weibull [Equation 6], log-logistic [Equation 7] and log-normal [Equation 8].

The exponential distribution is the simplest, assuming a constant hazard rate (λ) over time (t). In this sense, any death or event that is captured will be at random; it cannot adequately represent most survival data.

$$S(t) = \exp[-\lambda t] \qquad [5]$$

As a generalization of the exponential distribution, the Weibull function makes no assumption of a constant hazard rate and is therefore more broadly applicable than the exponential. The Weibull distribution is characterized by its shape (γ) and scale (λ) parameters, and can be used to model distributions where the hazard is increasing, decreasing or constant (62).

$$S(t) = \exp[-\lambda t^{\gamma}] \qquad [6]$$

The log-logistic distribution is unique in that it accommodates functions that involve both increasing and decreasing hazard rates (63). There are many applications for this, for example an infectious disease that is after some time, t, is treatable with a vaccine. Again, it has two parameters for shape (α) and scale (λ).

$$S(t) = \frac{1}{1+\lambda t^{\alpha}} \qquad [7]$$

The log-normal distribution is a transformation of the normal distribution, from which it takes into account the cumulative distribution function (φ) (64). Again, it has two parameters for shape (σ) and scale (μ).

$$S(t) = 1\text{-}\varphi\left[\frac{\ln t-\mu}{\sigma}\right] \quad [8]$$

In situations where the outcome of an intervention is gradual, parametric functions can lead to an underestimation of survival, and an alternative approach is necessary, such as the restricted cubic spline model (RCS). With RCS, polynomial functions fit particular segments of the KM, and smoothness at segment joints (knots) allows for nonlinearity between log time & log cumulative hazard, in contrast to the Weibull model (65).

Multiple factors inform the selection of the optimal parametric distribution. The Norwegian guidelines specifically state the expectation for documenting that any adjustment to observed study data is done in accordance with strategies based on those defined by Latimer (66) of an initial visual assessment, the clinical plausibility of likely disease outcome, and the lowest Akaike information criterion (AIC) and Bayesian information criterion (BIC), which are measures of the goodness of fit for estimated statistical models. The distribution selected affects the shape of the hazard function and impacts survival estimates, so making the appropriate choice is crucial.

Sensitivity Analysis

There are some methods for mitigating the uncertainty of particular inputs, one being sensitivity analysis. Using sensitivity analysis, we can determine how sensitive the results are

to changes in the parameter values. If the ICER is shown to be sensitive to changes to a specific parameter, this may indicate an area where further research is required. There are several types of sensitivity analysis: univariate sensitivity analysis, multivariate sensitivity analysis and probabilistic sensitivity analysis (PSA).

In univariate analysis, we vary only one parameter at a time and observe changes in the outcome. To isolate these changes, univariate analyses are deterministic, working with fixed values. In multivariate sensitivity analysis, we vary two or more parameters, or all parameters at a time and find a 'worst case' and 'best case' scenario based on the outcomes achieved.

Another approach commonly used is scenario analysis, which is a process of analyzing outcomes through the consideration of possible scenarios (52). The scenarios constructed represent a number of possible pathways, typically representing the base case, best case and worst case scenarios. Scenario analysis can be useful when investigating the effect of model assumptions.

PSA is based on a sense of likelihood in variation of the input parameters. It is the most informative method since it presents extreme outcomes, but also accounts for the likelihood of outcomes. It relies on a defined range of values for each variable and operates using a random number draw from each distribution to calculate the ICER. This procedure is repeated thousands of times and the results are based on what the range of true values are likely to be. The range of variation selected will have a strong impact on the interpretation of results, so it is important that appropriate values are selected from the literature or informed by expert opinion or a realistic judgement.

4. Methodology

4.1 Population

The target population used for this thesis was patients with proven stage II and stage III all risk colorectal cancer at age 65.

4.2 Intervention and Comparator

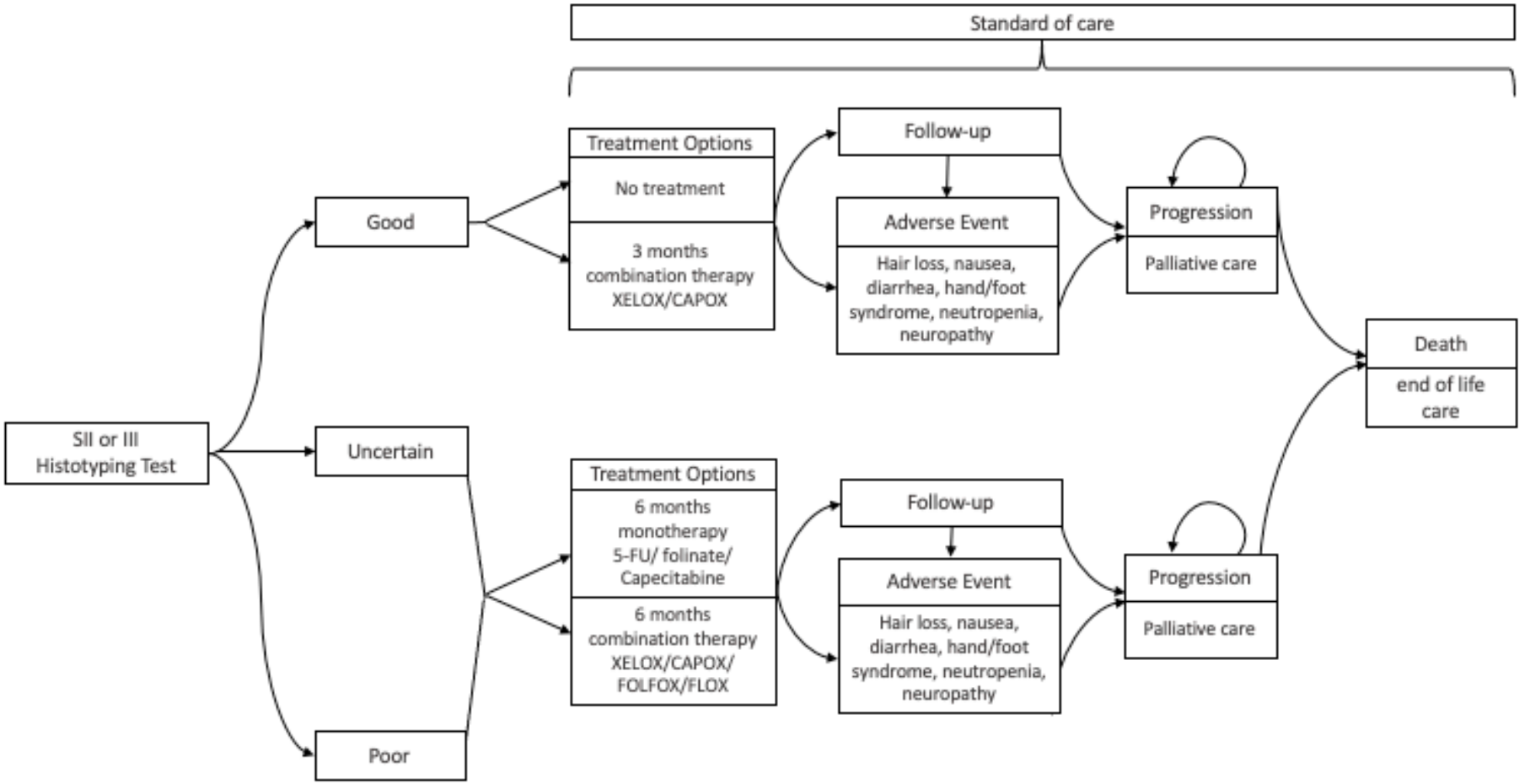
The intervention we examined as part of this case study was Histotyping in combination with the current standard of care introduced directly following surgical resection of the primary tumor and prior to the clinical decision on treatment. Patients are stratified according to risk of recurrence and were assumed to receive a corresponding arm of reduced or intensified chemotherapy based on Histotyping test results.

Histotyping is intended not to replace, but to serve as a supplement, offered as a companion prognostic service to the clinical routine. The effect of its implementation will be related to how it can alter the adjuvant chemotherapy treatment pathway assignments to reduce the health burden caused by overtreatment and undertreatment.

The standard of care of stage II and III CRC in the form of adjuvant chemotherapy, as is defined by Norwegian clinical guidelines for the treatment of CRC has been utilized as the primary comparator. For external validation of survival in each group, we used stage-specific data from Joranger et al., which modeled the clinical pathway of CRC patients by stage collected between 1993 and 2010(67). The median age of participants was 70 years,

The different pathways that result from the incorporation of Histotyping are depicted in the conceptual model (Figure 8) which was developed to demonstrate the effect of Histotyping based on input from meetings with clinicians.

Figure 8. The conceptual model of Histotyping integrated with the standard of care, reflecting the different pathways



4.3 Perspective

A healthcare perspective was applied in accordance with the Norwegian guidelines.

4.4 Outcome

The primary outcome of the early HTA was an ICER conditioned on certain characteristics of Histotyping presented by incremental costs in NOK per QALY from which the NMB was calculated using a weighted severity threshold of 495,000 NOK. Both costs and effects were discounted at an annual discount rate of 4%.

4.5 Time Horizon

The model adopted a time horizon of 10 years to capture all patients experiencing recurrence related to the primary cancer. A cycle time of 1 year was selected to study downstream consequences and costs relative to the test properties.

4.6 Model Structure

A decision-analytic partitioned survival model was developed in Microsoft® Excel 2019 to explore the cost-utility of Histotyping in stratifying stage II and III CRC patients for adjuvant chemotherapy treatment compared to the standard of care alone.

A partitioned survival model was chosen based on the presentation of the data available in the publication we drew from as a primary source (9). Partitioned survival models adequately represent the continuous risk of progression or death of patients and time elapses explicitly in the model. The model tracks a stage II and III CRC cohort using published Histotyping validation results (9) drawn from an open-label randomized controlled phase 3 trial where patients were assigned adjuvant chemotherapy treatment (33).

Stage II and III CRC were modeled separately using three partitioned survival models based on the assignment to good, uncertain and poor prognosis.

We maintained the same structure when modeling the two stages independently, defining the health states as progression-free (PF), progressed disease (PD) and death (Figure 9). Progression free is defined as the time between a primary resection and prognosis assignment and a cancer recurrence. Progressed disease is defined as the time between a recurrence and death. All patients enter the model in PF directly following a primary resection of CRC and assignment to a prognosis group and adjuvant chemotherapy treatment is begun and completed within the first cycle, assuming a cycle time of 1 year. Patients in PF will either remain in that health state or advance to PD. From PD, patients will either remain in this health state or advance to Death. Once a patient enters the health state PD, they cannot recover and reenter the PF health state. Furthermore, all patients must experience a recurrence and move through PD before moving to Death, which is a permanent health state.

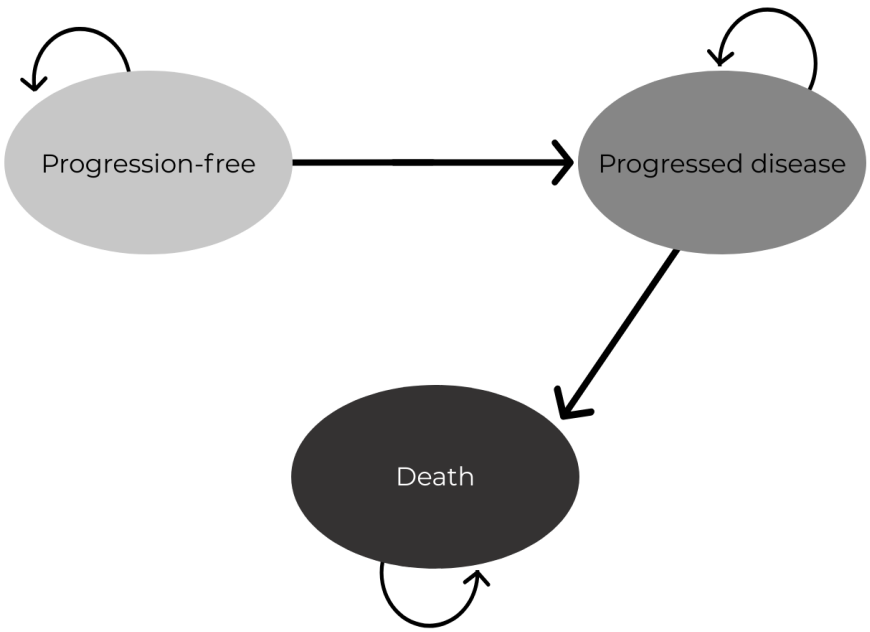


Figure 9. Model structure for partitioned survival analysis health outcomes. Each health state depends on stage (II or III) and prognosis group (good, uncertain, poor).

The starting point of the analysis was Overall Survival (OS) (9). Without having progression-free survival, we adjusted the traditional framework, beginning instead by calculating the proportion of patients entering the Death state at each cycle then those in PD, with the remainder of patients in PF. We consulted the literature for data values such as the rate of recurrence and the stage IV mortality rate.

The health state for Death was calculated using Equation 9 by subtracting the output values for each cycle derived from the survival analysis values. This was then multiplied by the cohort size.

$$Death_t = (1 - OS_t) \times cohort \quad [9]$$

For cycle 0 of PD, we utilized the stage IV mortality rate with the understanding that a recurrence in all stages is likely to advance to death to a similar degree as a stage IV recurrence. As the formulas show, progressed disease in cycle 0 is defines as the number of new deaths in cycle 1 divided by the mortality rate, resulting in 60% of deaths in cycle 1 transitioned from PD in cycle 0 [Equation 10a].

$$PD_{n=0} = \frac{Death_1}{Stage\ IV\ mortality\ rate} \quad [10a]$$

From cycle 1 onwards, we used a similar approach but calculate the cumulative mortality in cycles n+2 and n in order to smoothen the mortality over two cycles to compensate for the fluctuations caused by a small sample size. We divided the smoothened cumulative mortality by the stage IV mortality rate to calculate the number of PD patients in cycle n. In order to account for patients that remain in PD from the previous cycle, we subtracted the patients that remain in PD from the previous cycle [Equation 10b].

$$PD_n = \left(\underbrace{\frac{Death_{n+2} - Death_n}{2}}_{\text{Cumulative mortality in cycles n+2 and n}} \times \frac{1}{Stage\ IV\ mortality\ rate}\right) - \underbrace{(PD_{n-1} - Death_n)}_{\text{PD remaining from cycle n-1}} \quad [10b]$$

Finally, PF is the remaining patients in the cohort once we have accounted for PD and Death transitions away from PF [Equation 11].

$$PF_t = cohort - PD_t - Death_t \quad [11]$$

4.7 Data Inputs

Limited literature searches were conducted using combinations of search terms including those listed in Appendix Table 1 and performed using the electronic resource database available through the University of Oslo (Oria) and PubMed searches of MEDLINE, life science journals, and online books, filtered for the last 10 years in most cases. Examples of relevant search terms are: colorectal cancer, stage II and/or III, adjuvant chemotherapy, CAPOX and Capecitabine. Other model inputs were compiled through public health registries, public statistical databases, stakeholder interviews and estimates based on expert opinions.

Inputs for the model come from various sources. Wherever possible, model parameters for utilities, costs and probabilities were found in literature sources pertaining to CRC and adjuvant therapy.

Transition Probabilities

A detailed survival analysis was performed using the Kaplan-Meier curves from The Lancet Histotyping publication (9) in order to inform the selection into each prognosis group in the partitioned survival model.

Recall from section 4.6 that transitions are derived from progression free and overall survival curves in a standard partitioned survival model. In this study, OS was derived from the KM curves, which used the open-label randomized controlled QUASAR 2 trial dataset (56) for independent validation. As a result, the extrapolated values relied heavily on the choice of appropriate survival function.

Curves from the patient-level validation results (Figure 1 & 2) were digitized using the WebPlot Digitizer (68). Pseudo-patient level data was recreated based on the KM curve data and survival curves were fitted following the method described by Hoyle & Henley (69) using Microsoft® Excel 2019 and RStudio.

Our first strategy was to generate a smoothed hazard estimate in Stata (70), the results of which, pictured in Appendix Figure 1 and 2, showed us that the hazard was not constant and therefore we could rule out the use of the exponential function. Once the remaining standard panel of parametric distribution functions were applied, a determination of optimal survival function was informed by guidelines put forth by Nicholas Latimer, namely visual inspection, statistical tests, clinical validity and systematic assessment (71).

Visual assessment showed an improper fit with the KM curves and the extrapolations did not match clinical plausibility. We faced challenges of severe underestimation of survival

(Appendix Figure 3 & 4) likely due to too few observations, with little data to define the end of the KM curves.

We explored extrapolated survival using a method with restricted cubic splines. For stage II CRC survival, the number of knots for each RCS were selected on the basis of lowest AIC and BIC, with one knot being the most favorable option in each group as shown in Table 3.

Table 3. Stata OS survival output of spline models in Stage II good, uncertain and poor prognosis groups

	Stage II Good	**Stage II Uncertain**	**Stage II Poor**
AIC	96.261	28.815	57.259
BIC	103.39	32.557	62.213
coefficients (SE)			
_rcs1	1.894 (.5956)	1.858 (1.061)	2.030(.7625)
const	-16.497 (4.177)	-15.735(7.437)	-16.706 (5.345)

For stage III CRC, again, the number of knots for each RCS were selected on the basis of lowest AIC and BIC, with results shown in Table 4.

Table 4. Stata OS survival output of spline models in Stage III good, uncertain and poor prognosis groups

	Stage III Good	Stage III Uncertain	Stage III Poor
AIC	335.192	115.1578	339.4699
BIC	355.66	122.5898	349.0819
coefficients (SE)			
_rcs1	4.466 (1.268)	4.137 (1.503)	2.502 (.5714)
_rcs2	12.534 (4.410)	1.988 (.9601)	.5003 (.2394)
_rcs3	-33.513 (12.164)		
_rcs4	40.040 (15.459)		
const	-33.062 (8.372)	-29.349 (9.781)	-18.067 (3.6136)
knots			
1	6.229	5.651	5.079
2	6.647	6.786	6.878
3	7.144	7.291	7.402
4	7.333		
5	7.509		

By applying a flexible restricted cubic spline model, we could closely follow the curve in the period of observation and extrapolated values were clinically valid. Initial results showed that spline modeling did not reflect real world survival in stages II and III where survival was greatly overestimated (Figure 10 & 11).

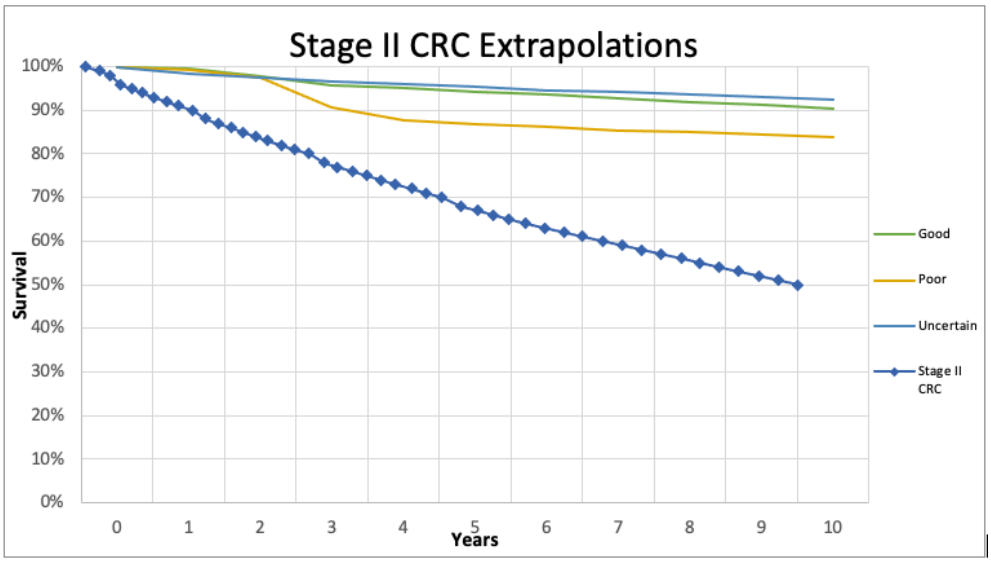


Figure 10. Stage II survival extrapolated using RCS on all years of observation. Stage II CRC survival from a Norwegian population is included for comparison

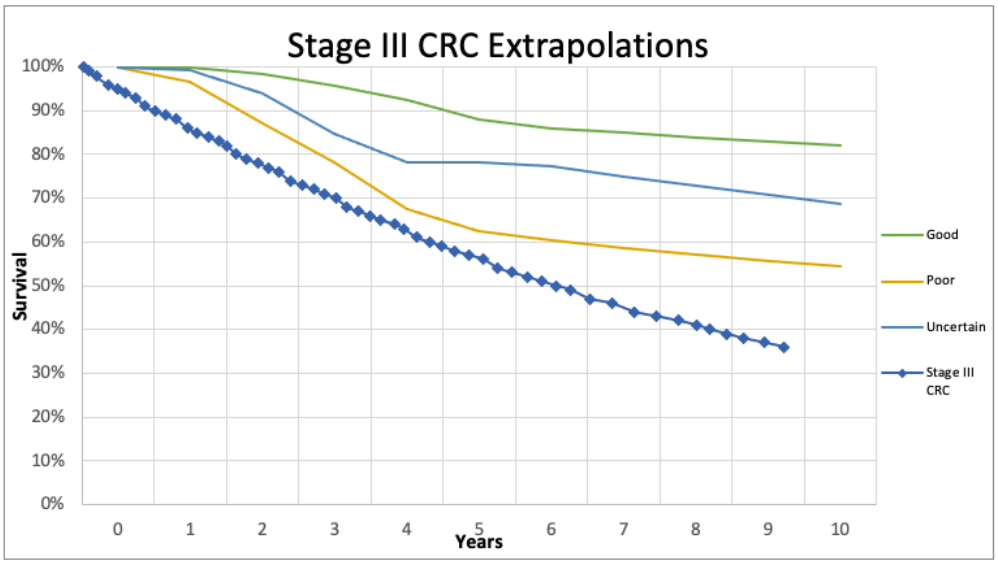


Figure 11. Stage III survival extrapolated using RCS on all years of observation. Stage III CRC survival from a Norwegian population is included for comparison

In a second attempt to apply restricted cubic spline modeling, observations were censored from the period of three years, with the justification that Histotyping sensitivity and specificity performance was drawn from this point. As depicted in Figure 12 and Figure 13, the extrapolated curves are a closer fit, but overestimated survival in the first half of the time horizon for all prognosis groups and did not reflect real world survival in the second half of the time horizon, with steep decreases and underestimated survival.

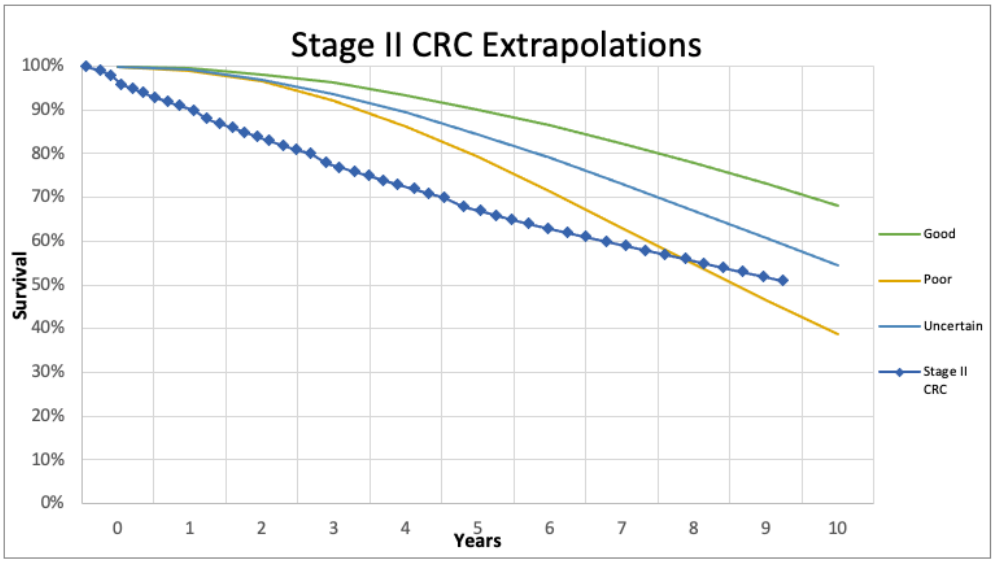


Figure 12. Stage II survival extrapolated using RCS with three years of observation. Stage II CRC survival from a Norwegian population is included for comparison

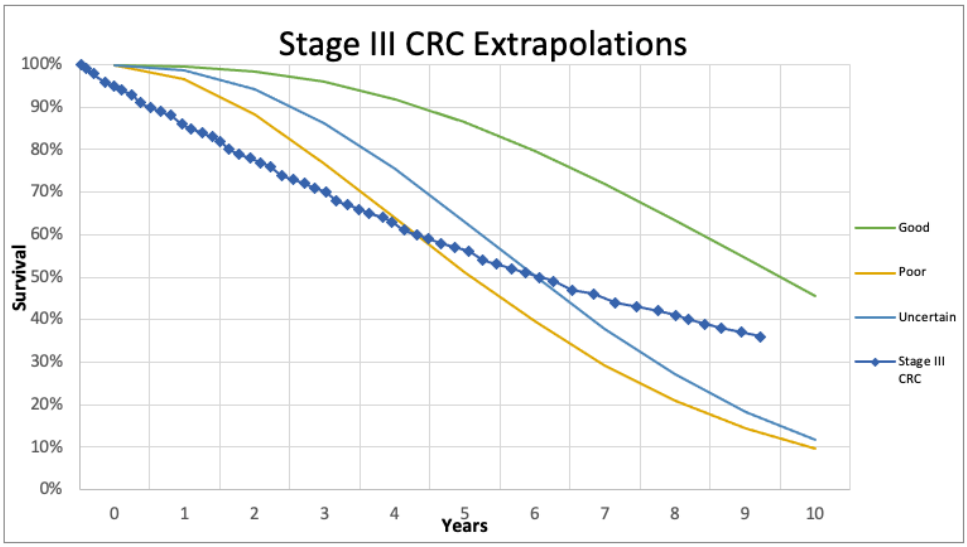


Figure 13. Stage III survival extrapolated using RCS with three years of observation Stage III CRC survival from a Norwegian population is included for comparison

In an explorative attempt to mimic real world survival, we censored stage III events after five years (Figure 14). The justification for this was that the stage III cohort was larger in the original publication with 713 patients as opposed to 397 in the stage II cohort and five years ensured that the greater number of events were captured for proper estimation of survival (9).

We noted that uncertain prognosis converged and crossed the good prognosis survival. This was unlikely to be the case in a real world scenario.

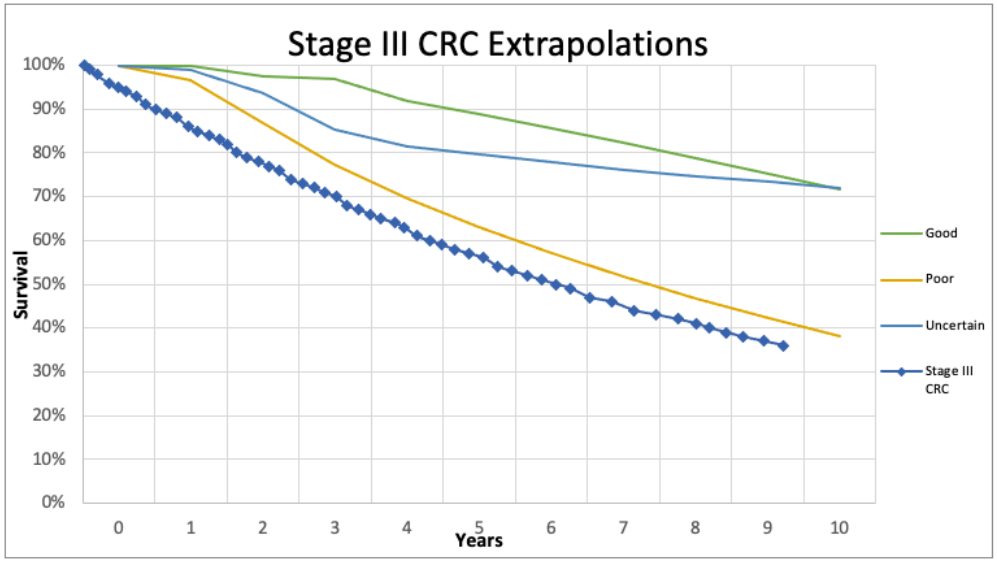


Figure 14. Stage III survival extrapolated using RCS with five years of observation Stage III CRC survival from a Norwegian population is included for comparison

In a final attempt, we kept stage II CRC individual patient level data censored after three years. Life table data (72) from the Norwegian setting replaced survival data from years four onwards, and the tangent of the uncertain curve and good prognosis curve were presumed to maintain throughout the time horizon. The poor prognosis curve assumed the trend of stage II CRC survival at the point at which the curves intersected. The extrapolated survival curve that resulted is depicted in Figure 15.

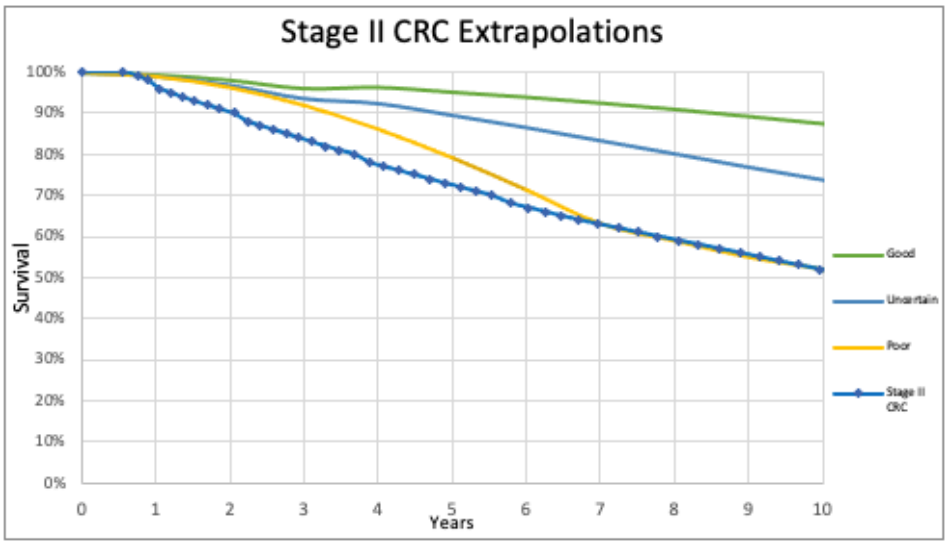


Figure 15. Stage II survival extrapolated hybrid of RCS with three years of observation and Norwegian life table data. Stage II CRC survival from a Norwegian population is included for comparison.

The final extrapolated curves for stage III maintained censoring after five years (Figure 16). The uncertain prognosis curve was adapted to maintain the same space differential from the good prognosis group as observed at year six directly following censored data. More specifically, the good prognosis group survival was 86% at year six and the uncertain prognosis group was 74% at year six, and we projected this 8% difference would be maintained for the remaining time horizon.

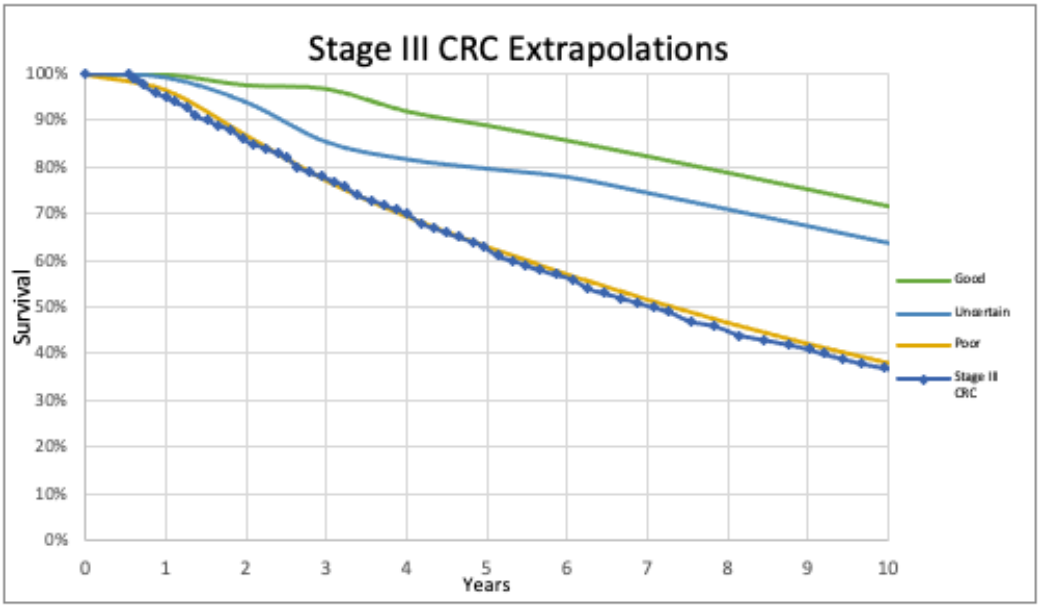


Figure 16. Stage III survival extrapolated hybrid of RCS with censored data from five years and adapted curve for uncertain prognosis. Stage III CRC survival from a Norwegian population is included for comparison.

Adverse Events

Adverse event incidences for stage III patients were derived from the SCOT study (Table 5), a phase 3 randomized controlled trial (RCT) comparing three and six-month regimens of adjuvant CAPOX in CRC patients (23).

Table 5. Incidence of adverse events for CAPOX 3-month and 6-month arms

Adverse Event Grade ≥ 3	**CAPOX 3 months**	**CAPOX 6 months**
Chemo-induced Diarrhea	12.5%	19
Hand–foot syndrome	3%	5.8%
Neutropenia	3%	6.2%
Peripheral Neuropathy	25%	58%

For stage II patients, (Table 6) AEs were derived from a systematic review of the clinical effectiveness of Capecitabine monotherapy in locally advanced and/or metastatic breast cancer patients (73). Although the cancer type differs from the population in this thesis, the

dose administered was identical to that stage II CRC patients are recommended in the standard of care. Therefore, adverse event values were assumed to be the same.

Table 6. Incidence of adverse events for Capecitabine monotherapy 6-month arm in stage II CRC (73)

Adverse Event Grade ≥ 3	**Capecitabine**
Chemo-induced Diarrhea	11%
Hand–foot syndrome	13%
Nausea and Vomiting	4%

Health Outcome

Since neither the paper by Skrede, De Raedt et al. (9), nor the QUASAR2 (33) publication offered cost or HRQoL data, we derived these values from other sources.

HRQoL utilities for stage III treatment assignment were generalized from a paper by Robles-Zurita, Boyd, Briggs et al. (22) in which a cost-effectiveness study was performed to investigate the non-inferiority of three-months versus six-months of adjuvant chemotherapy for CRC patients. Values were inferred for stage II patients based on the three-month values (22). The aforementioned study used an EQ-5D (Figure 17) to obtain HRQoL values at baseline, months 1-6 and 9,12,18 and yearly up to 8 years. These detailed treatment-related adverse event utilities allowed us to apply them cycle-wise to the calculation of QALYs in PF.

HRQoL utilities based on recurrence were applied to the PD health state QALY calculation using results from EQ-5D measure of health status in a study by Bjørnelv et al. (74).

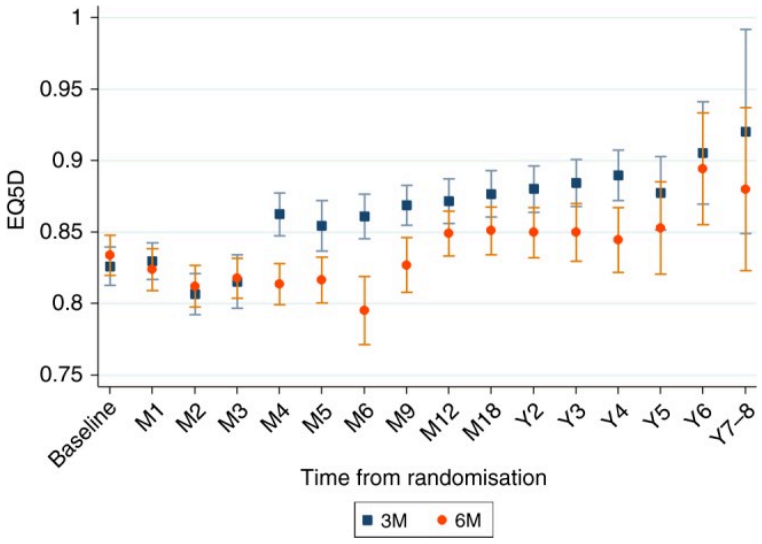


Figure 17. EQ-5D utility evolution over time by treatment arms. Average and 95% CI (22)

Costs

Based on treatment guidelines, standard therapy choices are selected for Stage II and III patients based on histological grades for low or high risk (Table 7).

Table 7. Norwegian stage II and III colorectal cancer adjuvant chemotherapy guidelines

Stage	Adjuvant Treatment <70 yrs.	Norway Recommendations	Duration
Stage II			
Low Risk	No	High MSI, no risk factors, Adjuvant therapy not recommended	-
High Risk	Yes	Monotherapy: 5-FU/folinate or Capecitabine	6 months
Stage III			
Low Risk	Yes	T1-T3, N1 XELOX/CAPOX	3 months
High Risk	Yes	T4 or N2 XELOX/CAPOX/FOLFOX/FLOX	6 months

We assumed that patients classified as uncertain would be treated as intensively as those with a poor classification and are included in the “high risk” treatment group. The basis for this assumption was that it would show the most conservative estimate of the effect of the

integration of Histotyping with the standard of care. For treatment, we assumed Capecitabine to be the monotherapy offered in stage II CRC patients and CAPOX to be the chemotherapy offered to stage III CRC patients (Table 8). These assumptions were made based on expert opinion and a base of literature that related to the target patient population.

Table 8. Chemotherapy assignment

Parameter	**Type**	**Dose**	**Histotyping**	**SoC**
Stage II Good	none	-	-	-
Stage II Uncertain	Capecitabine monotherapy	1250 mg/m2, 2x per day	6	6
Stage II Poor	Capecitabine monotherapy	1250 mg/m2, 2x per day	6	6
Stage III Good	CAPOX combo therapy	Cap: 1000 mg/m2, 2x per day Ox: Oxaliplatin 130 mg/m2 IV over 2 hours day 1	3	6
Stage III Uncertain	CAPOX combo therapy	Cap: 1000 mg/m2, 2x per day Ox: Oxaliplatin 130 mg/m2 IV over 2 hours day 1	6	6
Stage III Poor	CAPOX combo therapy	Cap: 1000 mg/m2, 2x per day Ox: Oxaliplatin 130 mg/m2 IV over 2 hours day 1	6	6

According to the literature and clinical guidelines from both Norway (11) and Europe (16), standard regimens of Capecitabine and combination therapy CAPOX were accounted for using the schedule in Appendix Table 2.

While low risk stage II CRC patients do not receive adjuvant chemotherapy, high risk patients are assigned to a 6-month regimen of monotherapy in the form of Capecitabine monotherapy. A dose of 1250 mg/m^2 is administered two times per day in the form of a tablet (2500 mg/m^2 total per day) for a period of 14 days, followed by a seven day rest period before the next three-week cycle.

Low risk stage III CRC patients receive a three month regimen of CAPOX combination therapy, in which oxaliplatin is administered intravenously over two hours on the first day of a 21-day cycle at a dosage of 130 mg/m². Also beginning on day one, patients take a dose of 1000 mg/m^2 twice per day (2000 mg/m^2 total per day) for a period of 14 days, followed by a seven day rest period before the next three-week cycle. High risk stage III CRC patients will receive the same dose intensity, but over a six month duration.

Adjuvant therapy is continued for the period defined by the Norwegian treatment guidelines, however, if patients progress, treatment administration is interrupted and patients are treated with palliative care, for which an average cost was applied based on a Norwegian cost-effectiveness study on colorectal metastases (74).

Norwegian cost data was used from the literature as well as Norwegian Directorate of Health and other public sources. Costs pertaining to all primary and secondary healthcare were included.

It was important to take treatment related adverse events into account, as these health effects have implications for both patient HRQoL as well as costs. The adverse event incidence for each adjuvant therapy regimen were used to calculate costs of treatment-related adverse events accumulated in the first year. All adverse events included were high-grade events presumed to require hospitalization, and associated costs were found in publicly available DRG lists (Table 9). Treatment-related costs including medication costs and intravenous medication administration costs were included also only for the first year, at which point adjuvant chemotherapy is completed. Upon a transition into the PD health state, costs for palliative care and recurrence were included in each subsequent cycle. Follow-up costs were

accounted for in all patients in PF and PD at each cycle from baseline to the end of the fifth year per the Norwegian guidelines (11). Costs and sources for each health state are detailed in Table 9.

Table 9. PF, PD and Death health state costs and sources

Resource (Cycles)	Cycle Dosage/ Cycle Visit	Unit Price (NOK)	Total Price (NOK)	Source	Histo-typing Arm	Comp-arator Arm
Chemotherapy - Progression Free						
Capecitabine - mono 6 months 1250 mg/m2 x2/day	63,700 mg	2,193	19,738	Felles-katalogen AS	SII uncer/poor	SII high
Capecitabine - combo 3 months 1000 mg/m2 x2/day	50,960 mg	11,113	8,773		SIII good	SIII low
Oxaliplatin – combo 3 months 130 mg/m2 over 2 hours day 1 each cycle	237 mg	2,193	44,532		SIII good	SIII low
Nurse for Oxaliplatin IV 3 months	2	528/hr.	4,224	Finansrådet		
Capecitabine – combo 6 months	50,960 mg	11,113	15,352	Felles-katalogen AS	SIII uncer/poor	SIII high
Oxaliplatin – combo 6 months 130 mg/m2 over 2 hours day 1 each cycle	237 mg	2,193	89,065		SIII uncer/poor	SIII high
Nurse for Oxaliplatin IV 6 months	2	528/hr.	8,448	Finansrådet	SIII uncer/poor	SIII high
Adverse events - Progression Free					**% patients**	
Monotherapy						
6 months						
Diarrhea		37,471	4,121	DRG 173	11%	
Hand–foot syndrome		37,471	4,871	DRG 173	13%%	
Nausea and Vomiting		64,131	1,498	DRG 172	4%	
Combination Therapy						
3 months						
Diarrhea		37,471	4,684	DRG 173	13%	
Hand–foot syndrome		37,471	1,124	DRG 173	3%	
Neutropenia		64,131	1,924	DRG 172	3%	
Peripheral Neuropathy		37,471	9,368	DRG 173	25%	
6 months						
Diarrhea		37,471	2,173	DRG 173	19%	
Hand–foot syndrome		37,471	2,323	DRG 173	6%	
Neutropenia		64,131	12,185	DRG 172	6%	
Peripheral Neuropathy		37,471	21,733	DRG 173	58%	
Follow-up Progression Free and Progressed Disease					**Health State**	
OP consult	varies by year	1,969	19,697	DRG 906A	Stage II and III PF &PD	
CT abdomen	varies by year	1,617 [a]	3,234 [a]	(75)		

Colonoscopy	varies by year	3,160	6,322	DRG 710O	Stage II and III PF &PD
CEA test	varies by year	185 [a]	1,665 [a]	(75)	
CT Lungs	varies by year	1,617 [a]	16,170 [a]	(75)	
CT Liver	varies by year	2,214 [a]	11,070 [a]	(75)	
Recurrence costs - Progressed Disease					
Recurrence y1	year 1	82,703	82,703	(67)	Stage II and III PD
Recurrence y2+	year 2+	35,399	35,399		
Palliative care	year 1-2	192,029	192,029	(74)	
Palliative care	year 2+	204,276	204,276		
End of life costs - Death					
Primary care (GP)	1	-	-	(76)	Stage II and III Death
Secondary care (inpatient/ outpatient)	1	-	-		
Home/community care (nursing/practical assistance)	1	-	-		
End of life mean	1	118,215	118,215		

a. Figures have been indexed from 2015 values to 2020

Modeling the Effect of Histotyping

Several scenarios were constructed for analysis, with the base case and standard of care detailed in Table 10. In the base case scenario, standard treatment is assigned per the Norwegian guidelines. For all scenarios, we made alterations to the base case while the standard of care remained the same.

For the standard of care, stage II low risk was assumed to be those testing positive for high MSI, which was assumed to be 15% as explained in section 2.3 (33). With Histotyping incorporated, the proportion of patients in good prognosis is increased to 66%. For stage II good prognosis and low risk patients, surgery is considered curative and they do not receive adjuvant chemotherapy.

The stage II uncertain prognosis group was treated the same as poor prognosis in all scenarios. Poor and uncertain prognosis patients in the Histotyping arm receive six months of Capecitabine monotherapy, the same as the high risk patients in the standard of care arm.

For stage III patients in the standard of care, all risk groups are assumed to receive six months of CAPOX combination therapy. With Histotyping incorporated, the proportion of patients in good prognosis is 61%. Treatment in good prognosis was defined as three months of CAPOX combination therapy, while six months was defined for uncertain and poor prognosis groups.

Table 10. Base case scenario versus the standard of care

			Type	**Dose**	**Duration (months)**
Histotyping + Standard of Care	Stage II	Good	None	-	-
		Uncertain	Capecitabine monotherapy	1250 mg/m2, 2x per day	6
		Poor	Capecitabine monotherapy	1250 mg/m2, 2x per day	6
	Stage III	Good	CAPOX combination therapy	Cap: 1000 mg/m2, 2x per day Ox: Oxaliplatin 130 mg/m2 IV over 2 hours, day 1	3
		Uncertain	CAPOX combination therapy	Cap: 1000 mg/m2, 2x per day Ox: Oxaliplatin 130 mg/m2 IV over 2 hours, day 1	6
		Poor	CAPOX combination therapy	Cap: 1000 mg/m2, 2x per day Ox: Oxaliplatin 130 mg/m2 IV over 2 hours, day 1	6
Standard of Care	Stage II	Low	None	-	-
		High	Capecitabine monotherapy	1250 mg/m2, 2x per day	6
	Stage III	All	CAPOX combination therapy	Cap: 1000 mg/m2, 2x per day Ox: Oxaliplatin 130 mg/m2 IV over 2 hours, day 1	6

In the first constructed scenario, we kept the stage III Histotyping arm uncertain prognosis group the same as they were in the base case, but the poor prognosis patients were assumed to demonstrate a need for more intensive adjuvant therapy in order to prevent future recurrence and receive an intensified regimen. In this case, we selected the standard six month regimen followed by an additional three month regimen in order to avoid the heightened toxicity of an increased dose. The cost calculations were straightforward, as it required a 50% increase in drug, administration and adverse event costs overall.

A next series of scenarios explored the risk of undertreatment as it relates to recurrence. We assumed an additional risk or recurrence for those in the stage II good prognosis group who receive no adjuvant chemotherapy, and the stage III good prognosis group who receive a reduced duration of three months instead of six months of combination therapy. As we defined earlier, good prognosis groups compose a large proportion of stage II (66%) and III (61%) patients in the Histotyping arm. Starting with 10% increased recurrence and advancing to 15%, 20% and 25%, we applied an increased relative risk of 1.1, 1.15, 1.2, and 1.25 respectively. This resulted in an increased the number of good prognosis progression free patients that transitioned from progression free to progressed at each cycle.

Sensitivity Analysis

The deterministic model serves as an estimate of expected resource use and cost values and has an inherent uncertainty. This is in large part due to the estimated parameters used as well as a number of assumptions forming the basis for calculations. To account for this uncertainty, a probabilistic sensitivity analysis was incorporated, as per ISPOR good practice guidelines for model parameter estimation and uncertainty (49).

The PSA functions as an estimate accounting for uncertainty, resulting in a presentation of likely outcomes as well as extreme outcomes. Probability distributions were defined for each variable through logical consideration of each individual parameter (Table 11). Random values were selected based on the pre-determined distributions and the standard errors inputted based on the literature and assumptions (Table 12).

Table 11: PSA parameter distribution assumptions

Parameter	Distribution	Justification
Base assumptions (cohort size, cycle length, number of cycles, follow-up schedule)	none	These values are assumed as fixed, defined by Skrede, De Raedt et al.
Anthropometric parameters	normal	Given a large enough dataset, this parameter can be assumed to follow a normal distribution
Probabilities	beta	Constraint: [0,1] Value will always fall between 0 and 1
Utilities	beta	Constraint: [0,1] No health state is worse than death in our model (Utility of 0)
Costs	gamma	Constraint: [0,1] We only account for positive costs, not negative (reimbursements)

Table 12: PSA parameter standard error assumptions

Parameters	Standard Error	Source
Vial use per patient	No sharing	Assumption, expert opinion
Tablet use per patient	No sharing	Assumption, expert opinion
Dosage over time	Stable over time	Assumption
CAPOX treatment utilities	Normalized IQR with scale factor for SE = IQR/3.92	Robles-Zurita et al. (22)
Adverse event utilities	SE = 20%	Assumption
Adverse event unit costs	SE = 20%	Assumption
End of life resource use	Translated SD with $SE = \sigma/\sqrt{n}$	Bjørnelv et al. (76)
Other healthcare resources	SE = 20%	Norwegian DRG Reference costs use an average with no SE. SE assumed to be 0.2

5. Results

5.1 Early Cost-Utility Analysis

In order to explore whether Histotyping has the potential of being cost-effective, we explored several scenarios.

The analysis of cost utility in the base case resulted in expected LYs and QALYs gained compared to the standard of care alone. Discounted outcomes are presented in Table 13.

Table 13. Early cost-effectiveness of Histotyping plus the standard of care vs. the standard of care alone in Stage II and III CRC patients with a time horizon of 10 years

	Treatment	Cost (NOK)	Incremental Costs (NOK)	QALYs	Incremental QALYs	LYs	Incremental LYs	NMB (NOK)
Stage II	Histotyping + Standard of Care	148,199	- 85,106	6.52	0.59	7.54	0.64	375,666
	Standard of Care	233,305		5.93		6.90		
Stage III	Histotyping + Standard of Care	342,576	- 91,931	5.88	0.35	6.74	0.27	267,139
	Standard of Care	434,508		5.53		6.47		

For stage II CRC, where low risk patients received no chemotherapy and high risk patients received 6 months of monotherapy, expected life-years were 7.54 for Histotyping plus the standard of care and 6.90 for the standard of care alone. In terms of QALYs, QALYs accumulated for Histotyping plus the standard of care were 6.52 QALYs and 5.93 QALYs in standard of care alone. Costs related to treatments were lower in the Histotyping integrated arm compared to the standard of care arm. As a result, the incremental costs for stage II treatment was 85,106 NOK lower than the standard of care. Based on these results, the net monetary benefit of the integration of Histotyping into the standard of care versus the standard of care alone was 375,666 NOK.

For stage III CRC, good prognosis patients in the base case received three months of CAPOX as opposed to six months of CAPOX given to uncertain and poor prognosis patients and as the standard of care. The expected life-years were 6.74 for Histotyping plus the standard of care and 6.47 for the standard of care alone. In terms of QALYs, QALYs accumulated for Histotyping plus the standard of care were 5.88 QALYs and 5.53 QALYs in standard of care alone. Costs related to treatments were lower in the Histotyping integrated arm compared to the standard of care arm in the base-case scenario. As a result, the incremental costs for stage III treatment was 91,932 NOK lower for the Histotyping arm than the standard of care. Based on these results, the net monetary benefit of the integration of Histotyping into the standard of care versus the standard of care alone was 267,139 NOK.

Disaggregated Results

Major differences in costs are most noticeable for the death related costs, which were 84,831 NOK and 62,414 NOK cheaper for Histotyping plus the standard of care in stage II and III respectively. The next greatest difference was in drug and administration costs, which were 10,146 NOK and 31,267 NOK cheaper in the Histotyping plus standard of care arm for stage II and III respectively. Stage III progressed disease maintenance costs were 702 NOK more expensive in the Histotyping plus standard of care arm. In a similar trend, follow-up costs for stage II and III patients were 2,995 NOK and 2,432 NOK more expensive respectively, for the Histotyping plus the standard of care.

Table 14. Disaggregated and undiscounted costs of Histotyping plus the standard of care vs. the standard of care alone in Stage II and III CRC patients in the base case scenario with a time horizon of 10 years (all costs in NOK)

	Treatment	Cost of AE	Drug, administration costs	PD maintenance costs	Death related costs	Follow-up costs	Total costs
Stage II	Histotyping + Standard of care	3,525	6,632	41,786	89,530	43,569	185,043
	Standard of care	8,918	16,777	47,322	174,361	40,574	287,954
	Increment	- 5,392	- 10,145	- 5,536	- 84,831	2,995	- 102,910
Stage III	Histotyping + Standard of care	25,369	73,157	58,474	214,946	41,242	413,190
	Standard of care	38,414	104,425	57,772	277,360	38,810	516,783
	Increment	- 13,044	- 31,267	702	- 62,414	2,432	- 103,592

Scenario Analysis

The alternative scenario analyses were performed to determine the impact of different input parameter values on the NMB.

The first scenario analyzed was the intensification of chemotherapy treatment for Histotyping-stratified stage III poor prognosis (Table 15). Patients were assumed to receive three additional months of adjuvant chemotherapy (CAPOX), making the total duration nine months as opposed to the six months received in the standard of care. Costs associated with Histotyping were still lower than standard of care alone, but to a lesser degree than the base case 81,442 NOK versus 91,931 NOK. No improvements to incremental QALYs or LYs were observed when compared to the base case scenario.

Table 15. Intensified chemotherapy regimen for stage III poor prognosis patients treated with Histotyping plus the standard of care vs. the standard of care alone.

	Treatment	Cost (NOK)	Incremental Costs (NOK)	QALYs	Incremental QALYs	LYs	Incremental LYs	NMB (NOK)
Stage III	Histotyping + Standard	342,576	- 81,442	5.88	0.35	6.74	0.27	256,649
	Standard of Care	424,019		5.53		6.47		

The next set of scenarios we explored were the likelihood of increased recurrence as a result of reduced adjuvant chemotherapy treatment in both stage II and III CRC patient cohorts stratified by Histotyping. Note that the stage II good prognosis and low risk groups did not receive any adjuvant chemotherapy, and that there was a significantly larger proportion of patients in the Histotyping arm cohort (66%) versus (15%) in the standard of care. For this reason, increased recurrence was calculated only in the Histotyping arm. The results are presented in Table 16 and incremental values listed are as compared to the standard of care.

Table 16. Recurrence rate increases for undertreatment scenarios in the good prognosis patients in both stage II and III treated with Histotyping plus the standard of care.

Recurrence rate	Stage	Cost (NOK)	Incremental Costs (NOK)	QALYs	Incremental QALYs	LYs	Incremental LYs	NMB (NOK)
10%	Stage II	155,523	- 77,782	7.13	1.18	7.54	0.64	662,182
	Stage III	343,589	- 90,919	5.88	0.35	6.74	0.27	265,350
15%	Stage II	159,245	- 74,059	7.12	1.16	7.54	0.64	650,595
	Stage III	349,787	- 84,720	5.87	0.34	6.74	0.27	253,883
20%	Stage II	162,911	- 70,393	7.10	1.15	7.54	0.64	638,245
	Stage III	352,100	- 82,407	5.88	0.34	6.74	0.27	249,587
25%	Stage II	166,590	- 66,715	7.08	1.13	7.54	0.64	627,586
	Stage III	354,368	- 80,139	5.86	0.33	6.74	0.27	245,351

Incremental costs, QALYs and overall NMB lost to increased recurrence is minimal with 5% increases, but more obvious when comparing a best case of a 10% increase in recurrence with the worst case of a 25% increase in recurrence. Incremental costs for 10% increased recurrence in stage II and III were 77,782 NOK and 90,919 NOK cheaper respectively than the base case scenario, but 66,715 NOK and 80,139 NOK cheaper respectively in the 25% increased recurrence group. Changes to the incremental QALYs were less noticeable than NMB, which also showed very gradual decreases with increasing recurrence.

5.2 Protocol

Recall from section 3.2 that the final step to the early HTA framework is to develop recommendations based on the results and inputs. The goal of this early HTA was to fully populate the conceptual model (section 4.2) for both the intervention and the standard of care, which relied on the literature as well as expert opinion. Meetings with a surgeon, oncologist and biotech analyst informed the following important considerations for future analysis.

PICO	• Define a sufficiently large sample size to capture the costs and effects of small incidence subgroups, i.e. stage II CRC poor prognosis • Consider possible market entry areas (service, software) • Competitors on the horizon: how fast are other technologies coming, and could anything in the foreseeable future disrupt the use of histology? • A clear treatment pathway should be defined in order for prognostics to be impactful • What are the expectation for test performance? • At which levels of sensitivity and specificity will the value of this tool be demonstrated?
Data collection	• Collect individual level data on treatment (adjuvant, adverse events, follow-up costs) • Collect health related quality of life values alongside trial • Overall survival and progression free survival should be explicitly named • If conducting a prospective trial, consider study transferability
Implementation	• How much effort will be required for this to be integrated? • Consider costs of implementation, i.e. training, and use costs • Does it currently fit within a procedural code (DRG code)? • Are there thought leaders that would endorse this? • Hospitals are conscious of funds being allocated for tests; can we demonstrate Histotyping's necessity?

6. Discussion

6.1 Main Findings

Main findings of the early cost utility analysis showed that Histotyping integrated into the standard of care has the potential of being cost-effective compared with the standard of care alone. This was evident in all scenarios examined.

Differences observed in costs and outcomes were driven by the underlying survival inputs, and the stratification of patients by risk of recurrence, where Histotyping classified a significantly higher proportion of patients in the good prognosis groups compared to the standard of care. Recall from section 4.6 that stage II and III good prognosis groups accounted for 66% and 61% of the entire cohort, respectively.

There were generally higher progressed disease maintenance costs in the Histotyping arm, which was a reflection of the greater number of patients that remained in the progressed state, rather than death, compared to the standard of care. Furthermore, follow-up costs were notably greater for Histotyping plus the standard of care. This was a reflection of a greater proportion of surviving patients in either PF or PD compared to those treated with the standard of care alone, making follow-up more costly overall.

A similar phenomenon throughout all scenario analyses was that changes were mostly noticeable in terms of cost, while QALYs and LYs gained remained similar between stage II groups and stage III groups. Noticeable decreases in incremental costs per QALY were driven by the cost of either more intensive treatment and AEs for patients with poor prognosis, or in the case of recurrence, it was the number of survivors that increased progressed disease maintenance costs that drove the results.

Cost-effectiveness was conditioned on the fact that those stratified to a good prognosis actually have a good outcome and will benefit from a reduced regiment of adjuvant chemotherapy and that the poor prognosis groups actually have a poor outcome and benefit from intensified treatment. Under these conditions, when we increased recurrence to a worst case scenario of 25% and increased the intensity of chemotherapy for stage III poor prognosis patients, Histotyping was still cost-effective.

6.2 Limitations and Future Research

Recalling the guidelines of early HTA from section 3.2, the initial step in conducting this study was narrowing the scope based on the likely point of implementation using the PICO method.

The population of stage II and III CRC patients from the validation dataset, which informed the construction of the partitioned survival model, were from a phase three RCT (77).
As QUASAR 2 was an RCT investigating adjuvant chemotherapy treatment options, all patients received adjuvant chemotherapy and it undoubtedly impacted survival and clinical follow-up in some way.

As is often the case with trial populations, there were strict inclusion criteria and it is likely the stage II and III patients were not emblematic of a general population. Furthermore, the QUASAR 2 population was from seven different countries (Australia, Austria, Czech Republic, New Zealand, Serbia, Slovenia and the UK). In the absence of country-specific data, we used QUASAR 2 as a proxy, but some variation can be expected.

The data used for the comparator population were from a model from Joranger et al. (67) that utilized prospective observational data on CRC patient survival between the years of 1993 and 2010. As some of this survival data is older than QUASAR 2 (77) data which began collection in 2005, we can expect there would be some differences in survival due to later refinement in treatment choices for adjuvant therapy. Further, the median age of patients in the QUASAR 2 study were 65, and 70 in the study by Joranger et al. (67).

To best replicate the standard of care as the comparator, we relied on the Norwegian guidelines, which clearly specify several treatment options and dose suggestions. For simplicity in the model, Capecitabine monotherapy and CAPOX combination therapy were selected for stage II and III patients, respectively. In reality, this selection is more nuanced and patient-specific.

For stage III patient stratification in the standard of care, we assumed all would be treated with six months of CAPOX combination therapy as opposed to three months of combination therapy in a subset of patients with low risk. There is certainly a distinction between high and low risk in clinical practice, however it was difficult to distinguish the groups and there were no clear pathways. In future studies, the implications of lower duration chemotherapy in stage III standard of care would be an interesting case to explore with more information.

As a second step, we made an inventory of available evidence and data on our intervention. The data used for cost and outcomes were primarily derived from the literature and Norwegian public sources such as DRGs, and statistical databases. Expert opinion strengthened the model in identifying the follow-up procedures and understanding the complexity of the patient pathway following primary resection.

Third, we drafted a conceptual model that incorporated the different scenarios that could occur as a result of stratification into a prognostic group, treatment with different intensities of adjuvant chemotherapy. Due to constraints of time and the data, we did not have a great degree of flexibility for some considerations such as population, vast differences in the type of adjuvant chemotherapy used, and relevant HRQoL data and left out disease prevalence and test performance.

Ideally, the conceptual model would have several modifications. Most importantly, we would better define the adjuvant therapy choices as in this study it was difficult to find empirical data on the likelihood of selection into the various treatment choices. For our purposes, we made the assumption of Capecitabine as monotherapy and CAPOX as the combination therapy of choice, but it would be interesting to examine the other options. Another modification we considered were the inclusion of early sensitivity and specificity inputs to investigate the changes in costs and implications for patient HRQoL and over- and undertreatment. The Histotyping test properties were reflected in the KM curves we used; however they were not explicitly taken into account in the analysis. Another point of interest would be to include more adverse events, as opposed to the severe events as we did (grade 3 and above), and to include specific treatment options for these events. In our analysis, we assumed hospitalization at different levels of care, however there are likely concomitant medications used to counteract symptoms and prevent long term effects.

Fourth, a cost utility analysis was performed to explore the influence of different scenarios on the outcomes of the intervention. In a limited capacity, we were able to construct an early HTA that demonstrated the potential for cost-effectiveness within certain conditions. As our

analysis worked primarily with costs, we are likely underestimating the benefits of the test. A headroom analysis was not performed due to a lack of feasibility.

Ideally, we would have used progression free survival data as the basis for our survival estimations as is the traditional approach for partitioned survival modeling. Progression free survival would have resulted in a more reliable estimation of the health states and may have allowed us to more accurately capture costs and benefits. We tried several approaches to estimate survival and compensate for these challenges and the flexible method of restricted cubic splines was our best option. Nevertheless, we lacked some inputs that impacted upstream and downstream resource use, and made for a less robust analysis.

6.3 Strengths

As an early HTA, the model structure and choices for time horizon and cycle time were logical. As the purpose of the model was to create a framework demonstrating the benefits of Histotyping, cycle times of one year allowed for a broad depiction of costs and effects as opposed to studying the cost and consequences of the adjuvant chemotherapy cycles in detail with three-week cycles. In the case of stage II and III CRC where patients enter the model at age of 65, a length of 10 years was appropriate to capture the effects of cancer and adjuvant chemotherapy in that timeframe. The use of a partitioned survival model was a strength in that it allowed us to work with the data we had available and enabled the creation of a simplified analysis of how patients move through the model over time and relative to their previous health states.

7. Conclusion

Histotyping as an addition to the standard of care has, under certain conditions, the potential of improving both length of life and health related quality of life compared to standard of care alone in the selection of adjuvant chemotherapy for stage II and III colorectal cancer patients. For the Histotyping arm, we defined that stage II good prognosis patients making up 66% of the cohort, as opposed to 15% of the standard of care cohort, receive no adjuvant chemotherapy. Also, stage II uncertain and poor prognosis patients receive six months of capecitabine, the same as the high risk group in the standard of care. The conditions for stage III patients in the Histotyping arm were that good prognosis patients (61%) receive three months of CAPOX combination therapy while uncertain and poor receive six months, compared with six months of CAPOX for all patients in stage III standard care. Under these conditions in the early HTA, the intervention showed good potential to be cost-effective, also when considering the trade-offs of increased progression and increased therapy intensity.

The framework contains considerable parameter uncertainty, but was intended to be updated with new data as it becomes available, and was informed by input from expert opinion and stakeholder preferences. We identified the need for more thorough analysis with detailed knowledge about altered treatment decisions and defined a protocol for future analysis. In its current form, our framework identifies conditions for cost-effectiveness and areas of use and could serve as a foundation for assessment of prospective net monetary benefit. Finally, we demonstrated the promise of Histotyping and its potential for meaningful patient implications when stratified for adjuvant chemotherapy assignment.

References

1. Wild CP WE, Stewart BW, editors (2020). World Cancer Report: Cancer Research for Cancer Prevention. Lyon,
France: International Agency for Research on Cancer.
2. GlobalData. Colorectal Cancer:
Epidemiology Forecast to 2028. 2019. Contract No.: GDHCER220-19.
3. WHO. Cancer Key Facts [Available from:
https://www.who.int/cancer/resources/keyfacts/en/.
4. Norway CRo. Cancer in Norway 2018 - Cancer incidence, mortality, survival and prevalence in
Norway. Oslo; 2019. Report No.: 0806-3621.
5. Crooke H, Kobayashi M, Mitchell B, Nwokeji E, Laurie M, Kamble S, et al. Estimating 1-and 5-year relative survival trends in colorectal cancer (CRC) in the United States: 2004 to 2014. American Society of Clinical Oncology; 2018.
6. André T, Boni C, Navarro M, Tabernero J, Hickish T, Topham C, et al. Improved overall survival with oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment in stage II or III colon cancer in the MOSAIC trial. Journal of Clinical Oncology. 2009;27(19):3109-16.
7. Copija A, Waniczek D, Witkoś A, Walkiewicz K, Nowakowska-Zajdel E. Clinical significance and prognostic relevance of microsatellite instability in sporadic colorectal cancer patients. International journal of molecular sciences. 2017;18(1):107.
8. Ijzerman MJ, Steuten LM. Early assessment of medical technologies to inform product development and market access. Applied health economics and health policy. 2011;9(5):331-47.
9. Skrede O-J, De Raedt S, Kleppe A, Hveem TS, Liestøl K, Maddison J, et al. Deep learning for prediction of colorectal cancer outcome: a discovery and validation study. The Lancet. 2020;395(10221):350-60.
10. Edge SB, Byrd DR, Carducci MA, Compton CC, Fritz A, Greene F. AJCC cancer staging manual: Springer New York; 2010.
11. Helsedirektoratet. Nasjonalt handlingsprogram med retningslinjer for diagnostikk, behandling og oppfølging av kreft i tykktarm og endetarm. 2019. Contract No.: IS-2849.
12. Osterman E, Glimelius B. Recurrence Risk After Up-to-Date Colon Cancer Staging, Surgery, and Pathology: Analysis of the Entire Swedish Population. Dis Colon Rectum. 2018;61(9):1016-25.
13. Tsikitis VL, Larson, D. W., Huebner, M., Lohse, C. M., & Thompson, P. A. . Predictors of recurrence free survival for patients with stage II and III colon cancer. BMC cancer. 2014;14(1).
14. Moertel CG, Fleming TR, Macdonald JS, Haller DG, Laurie JA, Goodman PJ, et al. Levamisole and fluorouracil for adjuvant therapy of resected colon carcinoma. New England Journal of Medicine. 1990;322(6):352-8.
15. Kemeny N. Chemotherapy for colorectal carcinoma: one small step forward, one step backward. Journal of clinical oncology: official journal of the American Society of Clinical Oncology. 1995;13(6):1287.
16. Schmoll H, Van Cutsem E, Stein A, Valentini V, Glimelius B, Haustermans K, et al. ESMO Consensus Guidelines for management of patients with colon and rectal cancer. a

personalized approach to clinical decision making. Annals of oncology. 2012;23(10):2479-516.
17. Group QC. Adjuvant chemotherapy versus observation in patients with colorectal cancer: a randomised study. The Lancet. 2007;370(9604):2020-9.
18. André T, De Gramont A, Vernerey D, Chibaudel B, Bonnetain F, Tijeras-Raballand A, et al. Adjuvant fluorouracil, leucovorin, and oxaliplatin in stage II to III colon cancer: updated 10-year survival and outcomes according to BRAF mutation and mismatch repair status of the MOSAIC study. Journal of Clinical Oncology. 2015;33(35):4176-87.
19. Tomita N, Kunieda K, Maeda A, Hamada C, Yamanaka T, Sato T, et al. Phase III randomised trial comparing 6 vs. 12-month of capecitabine as adjuvant chemotherapy for patients with stage III colon cancer: final results of the JFMC37-0801 study. British Journal of Cancer. 2019;120(7):689-96.
20. Peng J, Li W, Zhang R, Lin J, Tang J, Wen Y, et al. Safety and efficacy of a modified XELOX adjuvant regimen for patients with operated stage III colon cancer: a Chinese single-center experience. Cancer Communications. 2019;39(1):59.
21. Mizushima T, Ikeda M, Kato T, Ikeda A, Nishimura J, Hata T, et al. Postoperative XELOX therapy for patients with curatively resected high-risk stage II and stage III rectal cancer without preoperative chemoradiation: a prospective, multicenter, open-label, single-arm phase II study. BMC Cancer. 2019;19(1):929.
22. Robles-Zurita J, Boyd KA, Briggs AH, Iveson T, Kerr RS, Saunders MP, et al. SCOT: a comparison of cost-effectiveness from a large randomised phase III trial of two durations of adjuvant Oxaliplatin combination chemotherapy for colorectal cancer. British journal of cancer. 2018;119(11):1332-8.
23. Iveson TJ, Kerr RS, Saunders MP, Cassidy J, Hollander NH, Tabernero J, et al. 3 versus 6 months of adjuvant oxaliplatin-fluoropyrimidine combination therapy for colorectal cancer (SCOT): an international, randomised, phase 3, non-inferiority trial. The Lancet Oncology. 2018;19(4):562-78.
24. Tai E, Guy GP, Dunbar A, Richardson LC. Cost of cancer-related neutropenia or fever hospitalizations, United States, 2012. Journal of oncology practice. 2017;13(6):e552-e61.
25. Bates DW, Saria S, Ohno-Machado L, Shah A, Escobar G. Big data in health care: using analytics to identify and manage high-risk and high-cost patients. Health Affairs. 2014;33(7):1123-31.
26. Dixon SB, Chen Y, Yasui Y, Pui C-H, Hunger SP, Silverman LB, et al. Reduced Morbidity and Mortality in Survivors of Childhood Acute Lymphoblastic Leukemia: A Report From the Childhood Cancer Survivor Study. Journal of Clinical Oncology. 2020:JCO. 20.00493.
27. Louis DN, Perry A, Reifenberger G, Von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 World Health Organization classification of tumors of the central nervous system: a summary. Acta neuropathologica. 2016;131(6):803-20.
28. Koncina E, Haan S, Rauh S, Letellier E. Prognostic and Predictive Molecular Biomarkers for Colorectal Cancer: Updates and Challenges. Cancers (Basel). 2020;12(2).
29. Argiles G, Tabernero J, Labianca R, Hochhauser D, Salazar R, Iveson T, et al. Localised colon cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of Oncology. 2020.
30. van Helden EJ, Angus L, Menke-van der Houven van Oordt CW, Heideman DA, Boon E, van Es SC, et al. RAS and BRAF mutations in cell-free DNA are predictive for outcome of cetuximab monotherapy in patients with tissue-tested RAS wild-type advanced colorectal cancer. Molecular oncology. 2019;13(11):2361-74.

31. Sinicrope FA. DNA mismatch repair and adjuvant chemotherapy in sporadic colon cancer. Nature reviews Clinical oncology. 2010;7(3):174-7.
32. Kerr D, Gray R, Quirke P, Watson D, Yothers G, Lavery I, et al. A quantitative multigene RT-PCR assay for prediction of recurrence in stage II colon cancer: Selection of the genes in four large studies and results of the independent, prospectively designed QUASAR validation study. Journal of Clinical Oncology. 2009;27(15_suppl):4000-.
33. Kerr RS, Love S, Segelov E, Johnstone E, Falcon B, Hewett P, et al. Adjuvant capecitabine plus bevacizumab versus capecitabine alone in patients with colorectal cancer (QUASAR 2): an open-label, randomised phase 3 trial. The Lancet Oncology. 2016;17(11):1543-57.
34. DoMore! - A project for cancer patients and society [Available from: www.DoMore.no.
35. Young AM, Hobbs R, Kerr DJ. ABC of colorectal cancer: John Wiley & Sons; 2011.
36. Magnussen J VK, Saltman R. Nordic Health Care Systems: Recent Reforms And Current Policy Challenges: Recent Reforms and Current Policy Challenges: McGraw-Hill Education (UK); 2009 Sep 1.
37. National Institute for Health
Care Excellence (Great Britain). Guide to the processes of technology appraisal: National Institute for Health and Care Excellence; 2018.
38. Neumann PJ, Cohen JT, Weinstein MC. Updating cost-effectiveness—the curious resilience of the $50,000-per-QALY threshold. N Engl J Med. 2014;371(9):796-7.
39. NOU 1997:18. Prioriteringer på ny : Gjennomgang av retningslinjer for prioriteringer innen norsk helsetjeneste [Prioritizations revisited : Review of guidelines for priorities within the Norwegian health service]. Oslo: Statens forvaltningstjeneste; 1997.
40. Magnussen J, Aaserud M, Granaas T, Magelssen M, Syse A, Celius E. På ramme alvor-Alvorlighet og prioritering. Department of Health (p. 71). English summary available from: https://www regje ringe n no/conte ntass ets/d5da4 8ca5d 1a4b1 28c72 fc5da a3b4f d8/summ a ry_the_magnu ssen_repor t_on_sever ity pdf. 2015.
41. Organization WH. Health system efficiency: how to make measurement matter for policy and management: World Health Organization. Regional Office for Europe; 2016.
42. Nolte E, McKee M. Does health care save lives? Avoidable mortality revisited: The Nuffield Trust; 2004.
43. Schmid A, Cacace M, Gotze R, Rothgang H. Explaining health care system change: problem pressure and the emergence of “hybrid” health care systems. Journal of Health Politics, Policy and Law. 2010;35(4):455-86.
44. Buisman LR, Rutten-van Mölken M, Postmus D, Luime JJ, Uyl-de Groot CA, Redekop WK. The early bird catches the worm: early cost-effectiveness analysis of new medical tests. Int J Technol Assess Health Care. 2016;32(1):1-8.
45. O'Connor D, Green S, Higgins JP. Defining the review question and developing criteria for including studies. Cochrane handbook for systematic reviews of interventions: Cochrane book series. 2008:81-94.
46. MJ IJ, Koffijberg H, Fenwick E, Krahn M. Emerging Use of Early Health Technology Assessment in Medical Product Development: A Scoping Review of the Literature. Pharmacoeconomics. 2017;35(7):727-40.
47. Huygens SA, Rutten-van Mölken MP, Bekkers JA, Bogers AJ, Bouten CV, Chamuleau SA, et al. Conceptual model for early health technology assessment of current and novel heart valve interventions. Open Heart. 2016;3(2).

48. Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P, Krahn M. Conceptualizing a model: a report of the ISPOR-SMDM modeling good research practices task force–2. Medical Decision Making. 2012;32(5):678-89.
49. Briggs AH, Weinstein MC, Fenwick EA, 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. Medical decision making. 2012;32(5):722-32.
50. 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. Medical Decision Making. 2012;32(5):733-43.
51. R. Labianca BN, G. D. Beretta, S. Mosconi, M. Mandalà, A. Cervantes & D. Arnold, ESMO Guidelines Working Group. Early colon cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of Oncology. 2013;24(suppl_6):vi64-72.
52. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes: Oxford university press; 2015.
53. Johannesson M. Theory and methods of economic evaluation of health care: Springer Science & Business Media; 1996.
54. Guidelines for the submission of
documentation for single technology
assessment (STA) of pharmaceuticals. The Norwegian Medicines Agency (NOMA); 2018. p. 1-63.
55. Paulden M, Claxton K. Budget allocation and the revealed social rate of time preference for health. Health economics. 2012;21(5):612-8.
56. Goodman CS. HTA 101: introduction to health technology assessment: National Library of Medicine; 2014.
57. Black WC. The CE plane: a graphic representation of cost-effectiveness. Medical decision making. 1990;10(3):212-4.
58. Briggs A, Fenn P. Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane. Health economics. 1998;7(8):723-40.
59. Briggs A, Sculpher M. An introduction to Markov modelling for economic evaluation. Pharmacoeconomics. 1998;13(4):397-409.
60. Cranmer H, Shields GE, Bullement A. A comparison of partitioned survival analysis and state transition multi-state modelling approaches using a case study in oncology. Journal of Medical Economics. 2020;23(10):1176-85.
61. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. Journal of the American statistical association. 1958;53(282):457-81.
62. Lee ET, Wang J. Statistical methods for survival data analysis: John Wiley & Sons; 2003.
63. Mills M. Introducing survival and event history analysis: Sage; 2010.
64. Kurniasari D, Widyarini R, Antonio Y, editors. Characteristics of Hazard Rate Functions of Log-Normal Distributions. Journal of Physics: Conference Series; 2019: IOP Publishing.
65. Gibson E, Koblbauer I, Begum N, Dranitsaris G, Liew D, McEwan P, et al. Modelling the survival outcomes of immuno-oncology drugs in economic evaluations: a systematic approach to data analysis and extrapolation. Pharmacoeconomics. 2017;35(12):1257-70.

66. Latimer N. NICE DSU technical support document 14: survival analysis for economic evaluations alongside clinical trials-extrapolation with patient-level data. Sheffield: Report by the Decision Support Unit. 2011;2013.
67. Joranger P, Nesbakken A, Hoff G, Sorbye H, Oshaug A, Aas E. Modeling and Validating the Cost and Clinical Pathway of Colorectal Cancer. Medical Decision Making. 2015;35(2):255-65.
68. Rohatgi A. WebPlotDigitizer. [Online]. Available from: https://automeris.io/WebPlotDigitizer. 2020.
69. Hoyle MW, Henley W. Improved curve fits to summary survival data: application to economic evaluation of health technologies. BMC medical research methodology. 2011;11(1):139.
70. LLC S. STATA 16. 2019.
71. Latimer NR. Survival analysis for economic evaluations alongside clinical trials—extrapolation with patient-level data: inconsistencies, limitations, and a practical guide. Medical Decision Making. 2013;33(6):743-54.
72. StatBank. Statistics Norway Mortality Tables. [Online]. Available from: https://www.ssb.no/en/statbank/table/07902. 2019.
73. Jones L, Hawkins N, Westwood M, Wright K, Richardson G, Riemsma R. Systematic review of the clinical effectiveness and cost-effectiveness of capecitabine (Xeloda) for locally advanced and/or metastatic breast cancer. Health technology assessment (Winchester, England). 2004;8(5):iii-143.
74. Bjørnelv G, Dueland S, Line P, Joranger P, Fretland ÅA, Edwin B, et al. Cost-effectiveness of liver transplantation in patients with colorectal metastases confined to the liver. 2019.
75. Joranger P, Nesbakken A, Sorbye H, Hoff G, Oshaug A, Aas E. Survival and costs of colorectal cancer treatment and effects of changing treatment strategies: a model approach. Eur J Health Econ. 2020;21(3):321-34.
76. Bjørnelv GMW, Edwin B, Fretland ÅA, Deb P, Aas E. Till death do us part: the effect of marital status on health care utilization and costs at end-of-life. A register study on all colorectal cancer decedents in Norway between 2009 and 2013. BMC health services research. 2020;20(1):115.
77. Kerr RS, Love S, Segelov E, Johnstone E, Falcon B, Hewett P, et al. Adjuvant capecitabine plus bevacizumab versus capecitabine alone in patients with colorectal cancer (QUASAR 2): an open-label, randomised phase 3 trial. The Lancet Oncology. 2016;17(11):1543-57.

Appendix

Literature Search

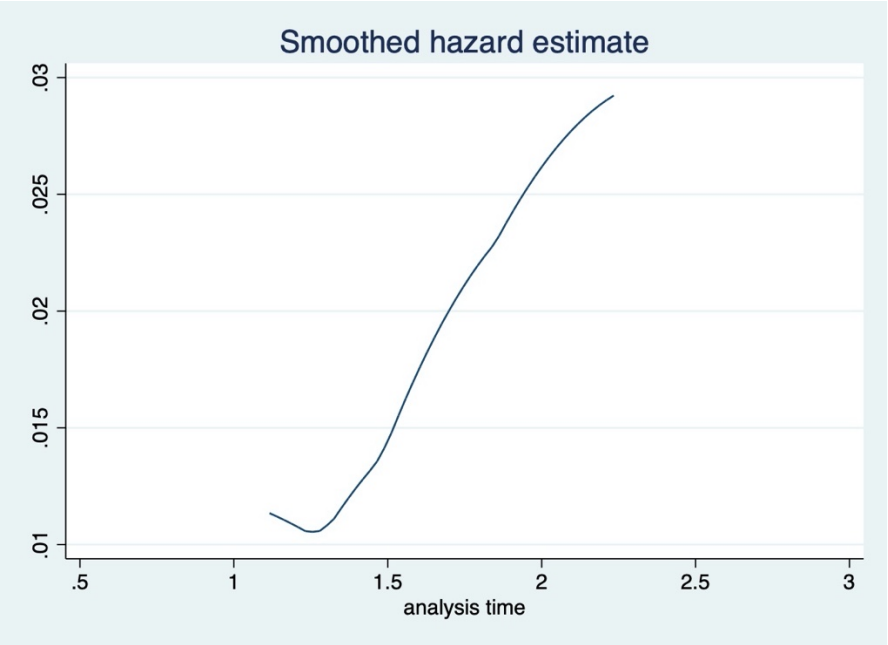
Limited literature searches were conducted using combinations of search terms including those listed in Appendix Table 1.

Appendix Table 1. Limited literature search strategy

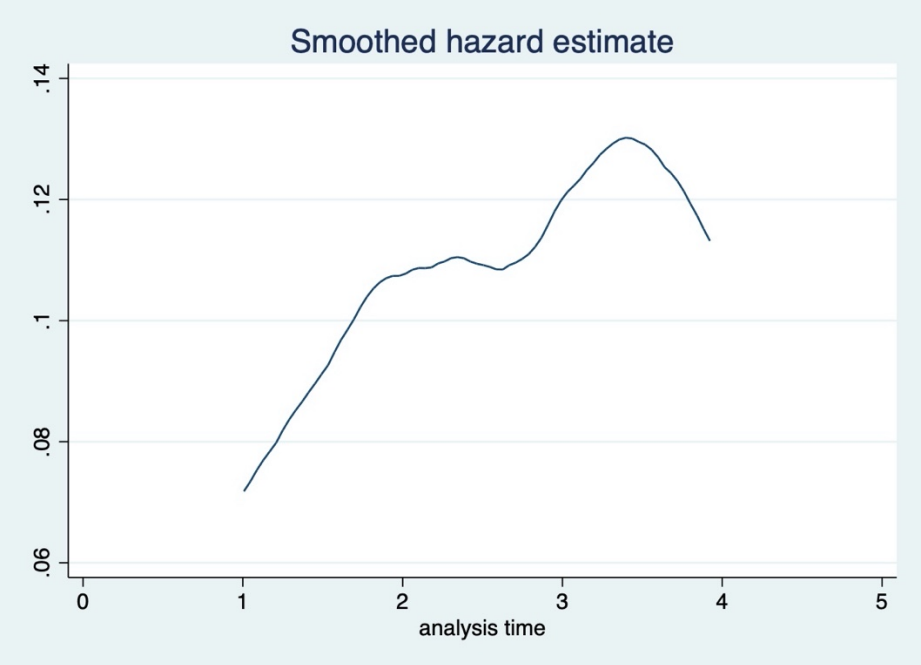
#	Searches	Results
1	(CRC) or (Colorectal cancer)	109,280
2	Adjuvant chemotherapy	32,497
3	1 and 2	4,457
4	Stage II	47,067
5	Stage III	42,505
6	1 and 2 and 4	1,073
7	((Cost-Effectiveness) OR (Economic Evaluation)) OR (Health Economic Analysis)	106,837
8	1 and 2 and 4 and 7	64
9	(Recurrence risk) OR (risk of recurrence)	-
10	7 and 9 and 1	38
11	Treatment induced and 1	2,947
12	Recurrence	-
13	Treatment-induced	6,411
14	1 and 2 and 5 and 7	30
15	(Risk stratifying) or (risk stratification)	70,833
16	Adverse events	-
17	1 and 2 and 4 and 5 and 16	125

Survival Analysis

When we began the process of survival analysis, our first step was smoothed hazard estimates to rule out the option of the exponential parametric survival function. Results are depicted for poor prognosis groups with stage II (Appendix Figure 1) and stage III (Appendix Figure 2).

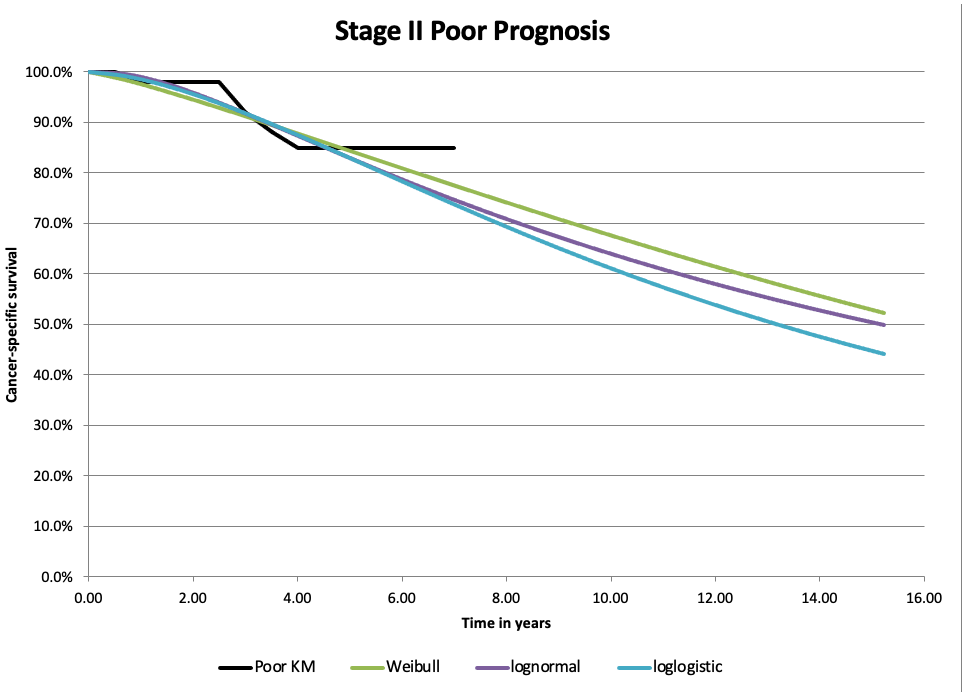


Appendix Figure 1. Smoothed hazard function for stage II poor prognosis

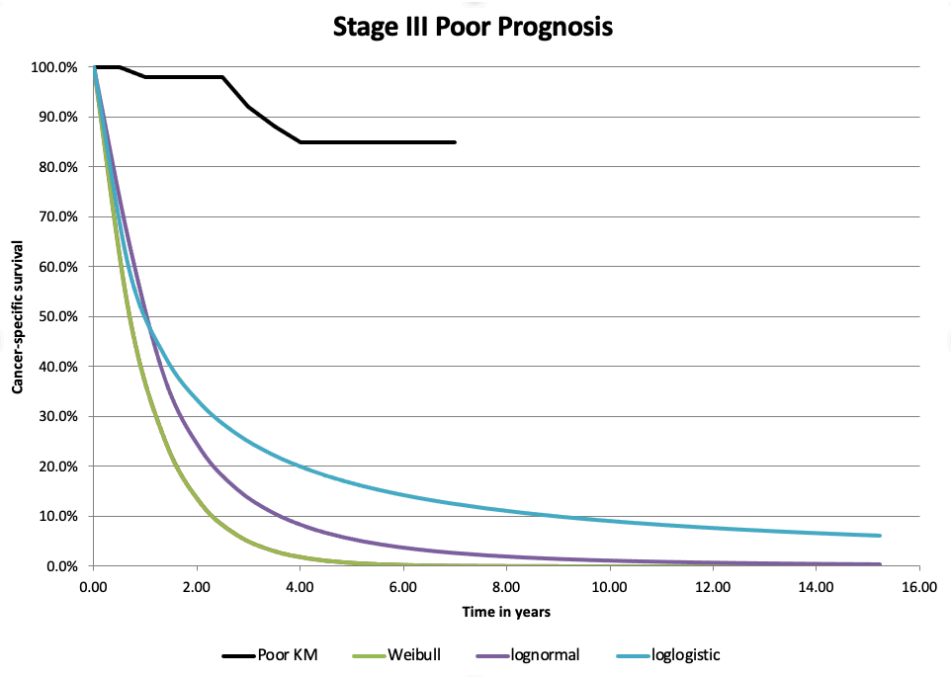


Appendix Figure 2. Smoothed hazard function for stage III poor prognosis

Visual assessment of the panel of parametric survival functions tested showed an improper fit with the KM curve. Extrapolated survival is shown to be underestimated in both poor prognosis groups for stage II (Appendix Figure 3) and stage III (Appendix Figure 4).



Appendix Figure 3. Survival was underestimated by all parametric functions when applied to the stage II poor prognosis KM



Appendix Figure 4. Survival was strongly underestimated by all parametric functions when applied to the stage III poor prognosis KM

Treatment Schedules

Using Norwegian (11) and European (16) guidelines, Capecitabine monotherapy was selected ad monotherapy for stage II patients and CAPOX selected in two duration options for treating stage III patients., standard regimens of Capecitabine and combination therapy CAPOX were accounted for using the schedule in Appendix Table 2.

Appendix Table 2. Adjuvant chemotherapy schedules by therapy type and duration

Capecitabine monotherapy single cycle : 6 month duration			
		weeks 1 2 3	
Capecitabine	1250 mg/m2, 2x per day	■ ■	x eight 3-week cycles
CAPOX single cycle : 3 month duration			
		weeks 1 2 3	
Oxaliplatin	130 mg/m2, day 1 over 2 hrs	day 1	x four 3-week cycles = 12 weeks = 3 months
Capecitabine	1000 mg/m2, 2x per day	■ ■	
CAPOX single cycle : 6 month duration			
		weeks 1 2 3	
Oxaliplatin	130 mg/m2, day 1 over 2 hrs	day 1	x eight 3-week cycles = 24 weeks = 6 months
Capecitabine	1000 mg/m2, 2x per day	■ ■	

Declaration in Lieu of Oath

Declaration in lieu of oath

With this declaration, the student confirms having written the thesis him or herself without any outside help. Others' thoughts and ideas are clearly marked as such and the master thesis has not been handed in during the course of another program and has not yet been published. Each master's thesis needs to contain such a declaration and has to be signed by the student in person. An electronic signature cannot be accepted. Exact formulation of this declaration:

"DECLARATION IN LIEU OF OATH

I hereby declare, under oath, that this master thesis has been my independent work and has not been aided with any prohibited means. I declare, to the best of my knowledge and belief, that all passages taken from published and unpublished sources or documents have been reproduced whether as original, slightly changed or in thought, have been mentioned as such at the corresponding places of the thesis, by citation, where the extent of the original quotes is indicated.

The paper has not been submitted for evaluation to another examination authority or has been published in this form or another."

SIGNATURE	
Anna Ruth Kenseth 28/10/2020	28/10-20
date and signature of student	date and signature of supervisor